Grammig, Joachim; Schaub, Eva-Maria

Conference Paper

Give me strong moments and time - Combining GMM and SMM to estimate long-run risk asset pricing models

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Financial Econometrics, No. C12-V3

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Give me strong moments and time
– Combining GMM and SMM to estimate long-run risk asset pricing models *

Joachim Grammig and Eva-Maria Schaub **

February 20, 2014

Abstract

The long-run consumption risk (LRR) model is a convincing approach towards resolving prominent asset pricing puzzles. Whilst the simulated method of moments (SMM) provides a natural framework to estimate its deep parameters, caveats concern model solubility and weak identification. We propose a two-step estimation strategy that combines GMM and SMM, and for which we elicit informative moment matches from the LRR model structure. In particular, we exploit the persistent serial correlation of consumption and the equilibrium conditions for market return and risk-free rate, as well as the model-implied predictability of the risk-free rate. We match analytical moments when possible and simulated moments when necessary and determine the crucial factors that are required for identification and reasonable estimation precision. By means of a simulation study—the first in the context of long-run risk modeling—we delineate the pitfalls associated with SMM estimation of LRR models, and we present a blueprint for successful estimation.

Key words: asset pricing, long-run risk, simulated method of moments

JEL: C58, G10, G12

*We are grateful to H. Hasseltoft for sharing his MATLAB code for the computation of the endogenous LRR parameters and to J. Krause for providing a MATLAB implementation of the CMAES algorithm. We retain the responsibility for all remaining errors.

**Joachim Grammig: University of Tübingen and Centre for Financial Research (CFR), Cologne, Germany. Address: University of Tübingen, School of Business and Economics, Mohlstrasse 36, D-72074 Tübingen, Germany. joachim.grammig@uni-tuebingen.de, +49-7071-2976009. Eva-Maria Schaub: University of Tübingen. eva-maria.schaub@uni-tuebingen.de

1 Introduction

Bansal and Yaron (2004) have introduced a dynamic asset pricing model (DAPM) that aims to resolve the prominent asset pricing puzzles by accounting for three risk factors: long-run consumption risk, short-run risk and volatility risk. Due to its far-reaching impact on the model dynamics, the first factor is the namesake for the long-run risk (LRR) model. While the LRR approach is theoretically appealing, empirical tests are impeded by an intricate model structure that involves unobserved state variables. As pointed out by Singleton (2006), the Simulated Method of Moments (SMM) should provide a convenient framework to estimate and test complex DAPMs like the LRR model.

We show that successful SMM estimation of LRR models must account for several theoretical and econometric caveats, such that the identification of the model parameters is ensured. Accordingly, the choice of moment conditions needs to be guided by a thorough understanding of the model characteristics. For that purpose we propose a two-step estimation strategy that exploits the persistent serial correlation of consumption and the equilibrium conditions for market return and risk-free rate, as well as the model-implied predictability of the risk-free rate. Simple first and second moment matches lead to weak identification of the deep model parameters. By means of a simulation study—the first in the context of long-run risk modeling—we delineate the pitfalls associated with SMM estimation of LRR models, and we present a blueprint for successful estimation.

LRR models have already been empirically assessed in previous studies using calibration and econometric estimation techniques. Bansal and Yaron (2004) perform a calibration exercise to demonstrate the ability of the LRR model to explain the equity premium. Bansal et al. (2007a) estimate a cointegrated version of the LRR model using a vector autoregressive model with stochastic volatility (SV). However,

the LRR model structure needs to be adjusted to the estimation technique to permit its application. Bansal et al. (2007b) show that the LRR approach can also account for size and value premia in the cross-section. Drechsler and Yaron (2011) calibrate an LRR model to explain the variance premium and its relationship to investor preferences. They introduce a Poisson-jump process, which allows for non-Gaussian macroeconomic shocks. Bansal and Shaliastovich (2013) advocate the LRR model as a potential solution to the bond return predictability puzzle. For that purpose, they extend the model by an inflation process; additional complexity is introduced by two different stochastic volatility processes which drive the real and the nominal side of the economy. Hasseltoft (2012) also includes inflation into the LRR framework in order to model stock and bond market jointly. Constantinides and Ghosh (2011) make use of the fact that the latent model variables can be expressed as functions of observables. This approach allows to use the Generalized Method of Moments (GMM) for estimation. Ferson et al. (2013) evaluate out-of-sample forecasts produced by a cointegrated LRR model, and find its performance to be superior to the stationary version. However, the authors impose restrictions in order to identify the deep model parameters from their reduced-form estimation, which is sufficient for forecasting, but not for the estimation of all structural parameters. Pakoš (2013) generalizes the LRR model by introducing incomplete information and a cyclical risk component in the dividend growth rate.

While the SMM approach towards estimating LRR models is natural and appealing, its concrete implementation is impeded by methodological and numerical intricacies. These obstacles have largely been ignored—or circumvented—by the previous literature. The silence on these issues is surprising, as it is well known that LRR models are inherently fragile in that the permissible parameter space—the set of parameters for which the model has a solution—has a complex topology.

For certain economically plausible parameter values the LRR model becomes insoluble, a fact that needs to be accounted for in the estimation procedure. Dividends and consumption are driven by a small but persistent latent growth component and stochastic volatility, which exacerbates the identification of the structural parameters, in particular when the data series are short. In fact, the estimation of univariate SV processes has preoccupied excellent researchers for quite some time.[1] Yet, in the LRR model, the SV process is just one element of a complex, non-linear, structural model. It seems to be a daunting task to estimate such a model, and we show that it is a losing game when using an ad-hoc choice of first and second moment matches.

We demonstrate that the one-step estimation of the deep model parameters using an ad-hoc selection of moments can mislead the researcher, as weak or non-identification is not always obvious in such a highly non-linear model. It might go unnoticed that even sophisticated optimizers converge to one of the many local minima on the rugged objective function surface. We show that identification crucially hinges on the choice of informative moment conditions, which must be tailored to the model. We advocate a two-step approach, in which we estimate separately the subset of parameters associated with the macroeconomic environment and the representative investor's preference parameters. The first step consists of a GMM estimation using analytical moment conditions resulting from the macro sub-model, and the second step is an SMM estimation that exploits the asset pricing implications of the LRR model. We emphasize that the precision of the macro parameter estimates is of utmost importance for the successful estimation of the model parameters.

A comprehensive simulation study documents the performance of our estimation strategy. Our findings constitute a call for econometric due diligence and reality checks when estimating the LRR model. We also point out that because the available

[1] Cf. e.g. Ruiz (1994), Gallant et al. (1997), Sandmann and Koopman (1998), Kim et al. (1998), Andersen et al. (1999), and Jacquier et al. (2002).

(macro) time series are relatively short, estimation precision for some of the model parameters will inevitably be moderate, which emphasizes even more the need for informative moment matches. Our two-step estimation strategy delivers credible empirical results. The caveats and solutions presented in this study are important for the estimation of other DAPMs as well.

The remainder of the paper is organized as follows. In Section 2 we briefly review the theoretical basics of LRR models. Section 3 delineates our econometric methodology. In Section 4 we present the results of a Monte Carlo simulation study to show the suitability of our approach. We conclude in Section 5.

2 Anatomy of the LRR model

In this section we describe the dynamics of the long-run risk model by Bansal and Yaron (2004). We present all key equations needed for the simulation of the full model, and hence for SMM.[2]

The LRR model is based on a non-linear four-equation VAR with two observable variables, log consumption growth g_t and log dividend growth $g_{d,t}$, and two latent variables, a growth component x_t and a stochastic variance process σ_t^2:

$$g_{t+1} = \mu_c + x_t + \sigma_t \eta_{t+1} \tag{1}$$

$$x_{t+1} = \rho x_t + \varphi_e \sigma_t e_{t+1} \tag{2}$$

$$g_{d,t+1} = \mu_d + \phi x_t + \varphi_d \sigma_t u_{t+1} \tag{3}$$

$$\sigma_{t+1}^2 = \sigma^2 + \nu_1(\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1}. \tag{4}$$

The i.i.d. innovations η_t, e_t, w_t, and u_t, are standard normally distributed and contemporaneously uncorrelated random variables. The latent fundamental drivers of the economic dynamics, x_t and σ_t^2, are assumed to be highly persistent, hence the autoregressive parameters ρ and ν_1 are usually chosen close to one in calibration excercises (cf. Bansal and Yaron, 2004). For a model simulation, the trajectories of g_t, x_t, $g_{d,t}$, and σ_t^2 represent the elementary components for all other model variables.

The representative investor, who faces these macro dynamics, has recursive preferences (cf. Kreps and Porteus, 1978; Epstein and Zin, 1989) as expressed by the utility function

$$U_t = \left[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \left(\mathbb{E}_t \left(U_{t+1}^{(1-\gamma)} \right) \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}, \tag{5}$$

[2]Detailed derivations are collected in Sections A.1–A.5 of the internet appendix, `http://tinyurl.com/lrr-internet-appendix`. These results appear somewhat dispersed in the literature, and we collect them in order to provide the interested reader with a complete picture.

where $\theta = \frac{(1-\gamma)}{\left(1-\frac{1}{\psi}\right)}$. The preference parameters δ, γ, and ψ denote subjective discount factor, relative risk aversion, and the intertemporal elasticity of substitution, respectively. The representative investor has aggregate wealth W_t and an aggregate consumption C_t. Utility maximization under the budget constraint $W_{t+1} = (W_t - C_t)R_{a,t+1}$, where $R_{a,t}$ denotes the gross return of the latent aggregate wealth portfolio, yields the basic asset pricing equation for a gross asset return $R_{i,t}$:

$$\mathbb{E}_t \left[M_{t+1} R_{i,t+1} - 1 \right] = 0, \tag{6}$$

where

$$M_{t+1} = \delta^{\theta} G_{t+1}^{-\frac{\theta}{\psi}} R_{a,t+1}^{-(1-\theta)} \tag{7}$$

denotes the stochastic discount factor, and G_t denotes gross consumption growth.

Bansal and Yaron (2004) explicitly model the log returns of the latent aggregate wealth portfolio and the observable market portfolio, $r_{a,t}$ and $r_{m,t}$, using the linear approximations suggested by Campbell and Shiller (1988):[3]

$$r_{a,t+1} = \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{t+1} \tag{8}$$

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1}, \tag{9}$$

where z_t denotes the log price-consumption ratio and $z_{m,t}$ the log price-dividend ratio. Furthermore,

$$\kappa_1 = \frac{\exp(\bar{z})}{1+\exp(\bar{z})} \qquad \kappa_{1,m} = \frac{\exp(\bar{z}_m)}{1+\exp(\bar{z}_m)} \tag{10}$$

$$\kappa_0 = \ln(1+\exp(\bar{z})) - \kappa_1 \bar{z} \quad \kappa_{0,m} = \ln(1+\exp(\bar{z}_m)) - \kappa_1 \bar{z}_m, \tag{11}$$

[3]Detailed derivations of Equations (8)–(11) are given in the internet appendix in Section A.1.

where $\bar{z}$ and $\bar{z}_m$ denote the means of z_t and $z_{m,t}$. Bansal and Yaron (2004) model the latent log price-consumption ratio and the observable log price-dividend ratio as

$$z_t = A_0 + A_1 x_t + A_2 \sigma_t^2 \tag{12}$$

$$z_{m,t} = A_{0,m} + A_{1,m} x_t + A_{2,m} \sigma_t^2. \tag{13}$$

The A-parameters in (12) and (13) have to be determined by an analytical solution of the model, which we present in detail in the internet appendix. This amounts to pricing the gross returns of aggregate wealth portfolio and market portfolio, $R_{a,t}$ and $R_{m,t}$, using (6). The solutions are given by:

$$A_1 = \frac{1 - \frac{1}{\psi}}{1 - \kappa_1 \rho} \tag{14}$$

$$A_2 = \frac{1}{2} \frac{\left(\theta - \frac{\theta}{\psi}\right)^2 + (\theta A_1 \kappa_1 \varphi_e)^2}{\theta(1 - \kappa_1 \nu_1)} \tag{15}$$

$$A_0 = \frac{1}{1 - \kappa_1} \left[\ln \delta + \left(1 - \frac{1}{\psi}\right) \mu_c + \kappa_0 + \kappa_1 A_2 \sigma^2 (1 - \nu_1) + \frac{\theta}{2} (\kappa_1 A_2 \sigma_w)^2 \right] \tag{16}$$

$$A_{1,m} = \frac{\phi - \frac{1}{\psi}}{1 - \kappa_{1,m} \rho} \tag{17}$$

$$\begin{aligned} A_{2,m} = & \frac{(1-\theta)(1 - \kappa_1 \nu_1) A_2}{(1 - \kappa_{1,m} \nu_1)} \\ & + \frac{\frac{1}{2}[(-\frac{\theta}{\psi} + \theta - 1)^2 + ((\kappa_{1,m} A_{1,m} \varphi_e) - ((1-\theta)\kappa_1 A_1 \varphi_e))^2 + \varphi_d^2]}{(1 - \kappa_{1,m} \nu_1)} \end{aligned} \tag{18}$$

$$\begin{aligned} A_{0,m} = & \frac{1}{(1 - \kappa_{1,m})} \Bigg[\theta \ln \delta - \frac{\theta}{\psi} \mu_c + (\theta - 1) \Big[\kappa_0 + \kappa_1 A_0 + \kappa_1 A_2 (1 - \nu_1) \sigma^2 - A_0 + \mu_c \Big] \\ & + \kappa_{0,m} + \kappa_{1,m} A_{2,m} \sigma^2 (1 - \nu_1) + \mu_d + \frac{1}{2} \left[(\theta - 1) \kappa_1 A_2 + \kappa_{1,m} A_{2,m} \right]^2 \sigma_w^2 \Bigg] \end{aligned} \tag{19}$$

The A-parameters given by Equations (14)–(19) depend on κ_0, κ_1, $\kappa_{0,m}$, and $\kappa_{1,m}$ from Equations (10) and (11), which in turn depend on $\bar{z}$ and $\bar{z}_m$. As a consequence, the κ-parameters, and hence the A-parameters, are endogenous.

For SMM estimation, we need to generate series for z_t, $z_{m,t}$, $r_{a,t}$, $r_{m,t}$, and for that purpose we need to solve the model, i.e. find $\bar{z}$ and $\bar{z}_m$ that determine the endogenous κ- and A-parameters such that Equations (10)–(19) are fulfilled.

As we outline below, our estimation approach exploits the implications of the LRR model for the log risk-free rate $r_{f,t}$, and we therefore need to simulate time series of the risk-free rate. To obtain the LRR model-implied expression for $r_{f,t}$, we price the risk-free asset using Equation (6) and obtain

$$r_{f,t} = -\theta \ln(\delta) + \frac{\theta}{\psi}\left[\mu_c + x_t\right] + (1-\theta)\mathbb{E}_t(r_{a,t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1}), \tag{20}$$

where m_t is the logarithm of the stochastic discount factor M_t and

$$\begin{aligned}\mathbb{E}_t(r_{a,t+1}) &= \kappa_0 + \kappa_1\left[A_0 + A_1\rho x_t + A_2(\sigma^2 + \nu_1(\sigma_t^2 - \sigma^2))\right] \qquad (21)\\ &\quad - A_0 - A_1 x_t - A_2\sigma_t^2 + \mu_c + x_t,\\ \mathrm{Var}_t\left(m_{t+1}\right) &= \left(\frac{\theta}{\psi} + 1 - \theta\right)^2 \sigma_t^2 + \left[(1-\theta)\kappa_1 A_1 \varphi_e\right]^2 \sigma_t^2 \qquad (22)\\ &\quad + \left[(1-\theta)\kappa_1 A_2\right]^2 \sigma_w^2.\end{aligned}$$

The derivation of (20)–(22) can be found in Section A.4 of the internet appendix.

Using these equations for the economic environment, the financial variables, and the expressions for the endogenous parameters, the LRR model can be simulated, given the macro parameters $\boldsymbol{\xi}_M = (\mu_c, \mu_d, \rho, \sigma, \varphi_e, \phi, \varphi_d, \nu_1, \sigma_w)'$, and the preference parameters $\boldsymbol{\xi}_P = (\delta, \gamma, \psi)'$. The next section describes our two-step estimation approach, which combines GMM for the first-step estimation of the macro parameters with an SMM estimation of $\boldsymbol{\xi}_P$ in the second step.

3 Econometric Methodology

3.1 Matching Moments for the LRR Model: choices and caveats

Singleton (2006) suggests to estimate DAPMs like the one in the previous section by SMM. The method is perfectly suited to deal with non-linearity, latent variables, and endogenous parameters, those complexity-driving features of the LRR model. Its application amounts to selecting measurable functions $\boldsymbol{g}(\cdot)$ of the economic or financial system variables and their model-implied expected values $\mathbb{E}\left[\boldsymbol{g}(\boldsymbol{q}_t;\boldsymbol{\xi})\right]$, where $\boldsymbol{q}_t$ holds all relevant observable and latent system variables and $\boldsymbol{\xi}$ contains the model parameters (or a subset of them). Observations for $\boldsymbol{g}(\cdot)$ are collected in the vector $\boldsymbol{g}_t^*$. A match of sample moments with population moments yields

$$\boldsymbol{G}_T(\boldsymbol{\xi}) = \left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{g}_t^* - \mathbb{E}\left[\boldsymbol{g}(\boldsymbol{q}_t;\boldsymbol{\xi})\right]\right]. \tag{23}$$

SMM is applied when the expectations cannot be expressed analytically as functions of $\boldsymbol{\xi}$, but need to be simulated such that the moment matches now read:

$$\boldsymbol{G}_T(\boldsymbol{\xi}) = \left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{g}_t^* - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)}\boldsymbol{g}\left(\boldsymbol{q}_s;\boldsymbol{\xi}\right)\right], \tag{24}$$

where T and $\mathcal{T}(T)$ denote the sample size and the simulated sample size, respectively. To obtain $\boldsymbol{q}_s$ for $s = 1,\ldots,\mathcal{T}(T)$, we need to simulate the LRR model using the equations and results outlined in the previous section. By choosing a large size for the simulated sample, a good approximation to population moments can be ensured. GMM estimates, using (23), or SMM estimates, using (24), then result from

$$\hat{\boldsymbol{\xi}}_T = \underset{\boldsymbol{\xi}\in\boldsymbol{\Theta}}{\operatorname{argmin}}\ \boldsymbol{G}_T(\boldsymbol{\xi})'\,\boldsymbol{W}_T\,\boldsymbol{G}_T(\boldsymbol{\xi}), \tag{25}$$

where $\boldsymbol{W}_T$ is a symmetric and positive definite weighting matrix.

The LRR model being a complex, highly non-linear model, it is an appealing idea to estimate it by matching some selected first and second moments. However, the identification of the deep parameters is likely to require information that is not or only weakly reflected in the first and second moments. Instead, the key model characteristics need to be translated into informative moment matches.

This insight emerges from an attempt to estimate the twelve LRR parameters in $\boldsymbol{\xi}_M$ and $\boldsymbol{\xi}_P$ simultaneously, using a set of moment matches adapted from Hasseltoft (2012). Table 1 shows that it includes ten first and second moments, two autocovariances, and two moments based on the prediction relationship between past log price-dividend ratio and future consumption growth volatility.

[Insert Table 1 about here]

The estimation is performed on simulated data of lenght $T = 1000$ and $T = 100000$, which are generated by a parameterized LRR model with structural parameters given in Table 6. To minimize the SMM objective we use $\boldsymbol{W}_T = \boldsymbol{I}$ and $\mathcal{T}(T) = 10^6$, and we employ one of the most sophisticated optimization algorithms currently available, the Covariance Matrix Adaptation Evolution Strategy (CMAES) algorithm developed by Hansen and Ostermeier (2001). To start the optimization of the SMM objective, we use three different, but not very dissimiliar starting values. Initial set s_1 uses the true parameters, which we slightly change for initial set s_2; the initial values in set s_3 are more away from the true parameters, but still economically quite plausible. Panel A of Table 2 documents a disturbing result that raises doubts whether the ad-hoc moment matches in Table 1 are useful to estimate the LRR parameters. Using the different initial values the optimization algorithm terminates at different parameter values (some vastly different), which nevertheless all fullfill the convergence criteria. We purposefully choose quite strong convergence criteria,

hence the CMAES algorithm does not stop prematurely, but takes its time to search the surface of the SMM objective. Panel A of Table 3 shows that the problem lingers even for a large sample size of $T = 100000$.[4] The conclusion from these results is obvious. You can obtain different (convenient) parameter estimates by choosing different starting values. The ad-hoc moment matches that invoke the augmented first two moments are of course valid, but they are too weak to help identify the structural parameters, even in large samples. Using alternative weighting matrices, such as an estimate of the efficient weighting matrix or the inverse covariance matrix of the GMM residuals, does not solve the problem.[5]

[Insert Tables 2–3 about here]

Panels B of Tables 2 and 3 show that the GMM estimation strategy pursued by Constantinides and Ghosh (2011) is prone to similar problems as the SMM estimation using the ad-hoc moment matches. Here, too, the optimization algorithm that minimizes the GMM objective converges to different values when started from different initial values. Hence, the same question looms: can the moment matches ensure the identification of the LRR model parameters? The alarming result of our simulation exercise may occur for real data, too: that the optimizer stops at values close to plausibly chosen starting values. If the neighborhood of that point happens to be well-defined, asymptotic inference may yield favorably small standard errors for plausible, but utterly arbitrary estimates. This caveat is aggravated by the small sample sizes used in empirical studies.

[4]How rugged the surface of the objective function is becomes obvious when using a gradient-based quasi-Newton algorithm. Irrespective of using starting values close to or remote from the true parameter values, the optimization algorithm only impalpably moves away from the starting point and converges immediately, despite strict convergence criteria.

[5]Using the optimization algorithms available in MATLAB's (Global) Optimiziation Toolboxes, such as the Nelder-Mead simplex, Simulated Annealing, Genetic Algorithm, and Pattern Search produces the same problem. The CMAES algorithm is superior, as it is designed to deal with very rugged objective functions. This is a desirable feature, but it comes at the cost of computation time.

The highly non-linear LRR model structure precludes an analytical check for non-identification. However, we can provide numerical evidence. If a set of theoretical moments chosen for a match is to be useful for the identification of a model parameter, it is a necessary condition that one or more moments respond to a change in that parameter. Table 4 shows a moment sensitivity matrix for the ad-hoc first and second moments, which displays each moment's percentage change as a reaction to a 50% c.p. change in one LRR model parameter.

[Insert Table 4 about here]

There is little or no sensitivity of the moments to changes in the SV parameters. The largest change provoked by a 50% change in ν_1 and σ_w is a 4%, or respectively, a 3% change in the mean market excess return. Given that this is a very weak reaction of this moment which responds to a much larger extent to many other model parameters, it is doubtful whether the SV parameters are identified. Considering the efforts invested in SV estimation that are documented in the related literature, this does not come as a surprise. This raises the question whether parameter estimation is only hampered by the presence of stochastic volatility. Repeating the optimization with SV turned off, however, delivers the same results: global optimization is infeasible, even with sophisticated optimizers.

We conclude that model identification and successful estimation hinges on the choice of more informative moment matches. The next two subsections outline a two-step estimation approach, in which we carefully exploit the properties of the LRR model in order to be able to identify the structural parameters through informative moment matches. Since the macro dynamics in (1)–(4) do not depend on the preference parameters, we focus on estimating the macro parameters by GMM in the first step. Using the estimated macro parameters, the preference parameters are estimated by SMM in the second step. It turns out that the optimization prob-

lems in each step become well-defined and using sophisticated optimizers becomes unnecessary.

3.2 Analytic moment conditions

For the estimation of the parameters of the dynamic system (1)–(4) using two observable series g_t and $g_{d,t}$ only, we need to choose moment matches that are able to capture their characteristic features implied by the LRR model. Naturally, the first and second moments should be matched. It turns out that they can be analytically expressed as functions of a subset of $\boldsymbol{\xi}_M$ that does not contain the SV parameters. We can then use the following moment matches for GMM:

$$\boldsymbol{G}_T^{(M),1}(\boldsymbol{\xi}_{M_0}) = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} g_t-\mu_c \\ g_{d,t}-\mu_d \\ g_t^2-\mu_c^2+\frac{\varphi_e^2\sigma^2}{1-\rho^2}+\sigma^2 \\ g_{d,t}^2-\mu_d^2+\phi^2\frac{\varphi_e^2\sigma^2}{1-\rho^2}+\varphi_d^2\sigma^2 \\ g_{d,t}g_t-\mu_c\mu_d+\phi\frac{\varphi_e^2\sigma^2}{1-\rho^2} \end{bmatrix}, \tag{26}$$

where $\boldsymbol{\xi}_{M_0} = (\mu_c, \mu_d, \rho, \sigma, \varphi_e, \phi, \varphi_d)'$.

The defining feature of the LRR model is the latent growth component x_t, a small but allegedly very persistent driver of consumption and dividend growth. If such a component is present, the observed consumption and dividend growth series exhibit small positive serial correlations, which, however, prevail over many lags. Figure 1 provides an illustration.

[Insert Figure 1 about here]

It seems obvious that this key model feature can only be captured by using moment matches for auto-moments of a high lag order. We therefore consider the following

moment matches that involve the (cross) auto-moments of consumption and dividend growth:

$$\boldsymbol{G}_T^{(M),2}(\boldsymbol{\xi}_{M_0}) = \begin{bmatrix} \frac{1}{T-1}\sum_{t=1}^{T-1} g_{t+1}g_t - \mu_c^2 + \rho\frac{\varphi_e^2\sigma^2}{1-\rho^2} \\ \vdots \\ \frac{1}{T-L_1}\sum_{t=1}^{T-L_1} g_{t+L_1}g_t - \mu_c^2 + \rho^{L_1}\frac{\varphi_e^2\sigma^2}{1-\rho^2} \\ \frac{1}{T-1}\sum_{t=1}^{T-1} g_{d,t+1}g_{d,t} - \mu_d^2 + \phi^2\rho\frac{\varphi_e^2\sigma^2}{1-\rho^2} \\ \vdots \\ \frac{1}{T-L_2}\sum_{t=1}^{T-L_2} g_{d,t+L_2}g_{d,t} - \mu_d^2 + \phi^2\rho^{L_2}\frac{\varphi_e^2\sigma^2}{1-\rho^2} \\ \frac{1}{T-1}\sum_{t=1}^{T-1} g_{d,t+1}g_t - \mu_c\mu_d + \phi\rho\frac{\varphi_e^2\sigma^2}{1-\rho^2} \\ \vdots \\ \frac{1}{T-L_3}\sum_{t=1}^{T-L_3} g_{d,t+L_3}g_t - \mu_c\mu_d + \phi\rho^{L_3}\frac{\varphi_e^2\sigma^2}{1-\rho^2} \end{bmatrix}, \tag{27}$$

where L_1, L_2, and L_3 denote the maximum lag length for the respective (cross) auto-moments. Bansal and Yaron's (2004) reasoning suggests that the number of lags should be large in order to capture the persistence of the series induced by x_t.

The theoretical moments that we match in (27) are derived analytically and written as functions of the parameters. Hence, a simulation-based estimation of $\boldsymbol{\xi}_{M_0}$ is unnecessary. We perform a standard GMM estimation by stacking $\boldsymbol{G}_T^{(M),1}(\boldsymbol{\xi}_{M_0})$ and $\boldsymbol{G}_T^{(M),2}(\boldsymbol{\xi}_{M_0})$ and using these moment matches in the GMM objective function in (25).

The moment conditions in (26) and (27) are valid irrespective of the SV parameter values ν_1 and σ_w. Hence, the GMM estimation of $\boldsymbol{\xi}_{M_0}$ can be performed regardless of whether SV prevails or not. Of course, this also implies that these moment matches cannot identify the SV parameters. We revisit this issue in the next section.

3.3 Theory-based moment conditions

We now turn to the estimation of the preference parameters $\boldsymbol{\xi}_P$, taking the macro parameters $\boldsymbol{\xi}_{M_0}$ as given. There are three observable financial variables, the log market return r_m, the log risk-free rate r_f, and the log price-dividend ratio z_m. The LRR model implies that they depend on the preference parameters, and hence represent useful candidates for moment matches. The macro variables and their associated moments cannot contribute to the identification of the preference parameters since they do not depend on $\boldsymbol{\xi}_P$.

We consider two approaches to estimate $\boldsymbol{\xi}_P$. The first uses an ad-hoc moment match of the first two moments of the three observable time series $r_{m,t}$, $r_{f,t}$, and $z_{m,t}$. Yet again, we suspect that those "agnostic" moment matches may not suffice to identify the deep model parameters. In a second approach, we therefore continue to pursue our philosophy to transfer key model characteristics into informative moment conditions. For that purpose, we exploit asset pricing and prediction relationships implied by the LRR model.

For both the ad-hoc and the theory-based approach we need to simulate the population moments to be matched with the data. The moments of the observable financial variables cannot be expressed analytically as a function of the preference parameters, since the LRR model needs to be solved numerically for the endogenous parameters (cf. Section 2), such that we resort to SMM.[6]

[6]The observant reader will recall that the moment conditions in (28) were shown to be insensitive to ν_1 and σ_w, which is why we do not try to use them for the estimation of the SV parameters. We will resolve this issue below.

The ad-hoc approach uses the following first and second moment matches:

$$\boldsymbol{G}_T^{(P,A)}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} r_{m,t} - r_{f,t} - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} r_{m,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) - r_{f,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) \\ \left(r_{m,t} - r_{f,t}\right)^2 - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} \left(r_{m,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) - r_{f,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P)\right)^2 \\ r_{f,t} - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} r_{f,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) \\ r_{f,t}^2 - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} r_{f,s}^2(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) \\ z_{m,t} - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} z_{m,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) \\ z_{m,t}^2 - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} z_{m,s}^2(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) \end{bmatrix}. \tag{28}$$

For the theory-based approach we exploit the pricing implications of the LRR model for the excess return of the market portfolio and the risk-free rate. Using the basic asset pricing equation (6) to price the risk-free rate yields

$$\mathbb{E}_t\left(M_{t+1}\right) = \mathbb{E}_t\left(\frac{1}{R_{f,t+1}}\right), \tag{29}$$

where M_{t+1} is given in Equation (7). Applying the law of total expectation leads to the unconditional moment constraint

$$\mathbb{E}\left(M\right) = \mu_M = \mathbb{E}\left(\frac{1}{R_f}\right). \tag{30}$$

Since μ_M cannot be expressed analytically as a function of the parameters, we match the mean of the simulated SDF with the sample mean of the inverse gross risk-free rate. This entails the following moment match:

$$\boldsymbol{G}_T^{(P,T),1}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} \frac{1}{R_{f,t}} - \mu_M \\ \mu_M - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} M_s(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) \end{bmatrix}. \tag{31}$$

Pricing the market excess return $(R_m - R_f)$ using Equation (6) implies:

$$\mathbb{E}_t\left(R_{m,t+1} - R_{f,t+1}\right) = -\frac{\mathbb{E}_t\left[\left(M_{t+1} - \mathbb{E}_t(M_{t+1})\right)\left(R_{m,t+1} - R_{f,t+1}\right)\right]}{\mathbb{E}_t(M_{t+1})}. \tag{32}$$

Applying the law of total expectation yields the unconditional moment constraint

$$\mathbb{E}\left(R_m - R_f\right) = -\frac{\mathbb{E}\left[\left(M - \mu_M\right)\left(R_m - R_f\right)\right]}{\mu_M}, \tag{33}$$

such that we can use the match of sample and simulated moment

$$\begin{aligned} \boldsymbol{G}_T^{(P,T),2}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) = \frac{1}{T}\sum_{t=1}^{T}\left(R_{m,t} - R_{f,t}\right) \\ + \frac{\frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)}\left[R_{m,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) - R_{f,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P)\right]\left[M_s(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) - \mu_M\right]}{\mu_M}. \end{aligned} \tag{34}$$

The set of test assets could be extended by including managed portfolios as suggested by Cochrane (1996). Given a vector of instruments available at t, $\boldsymbol{Z}_t$, pricing the managed portfolio payoffs $(R_{m,t+1} - R_{f,t+1})\boldsymbol{Z}_t$ using Equation (6) implies:

$$\mathbb{E}\left[(R_{m,t+1} - R_{f,t+1})\boldsymbol{Z}_t\right] = -\frac{\mathbb{E}\left[(M_{t+1} - \mu_M)(R_{m,t+1} - R_{f,t+1})\boldsymbol{Z}_t\right]}{\mu_M}.$$

Note that in order to use these moment conditions for SMM one is limited in the choice of instruments. $\boldsymbol{Z}_t$ needs to be determined within the LRR model since the instruments have to be simulated, too.

Furthermore, we would like to make sure that the unconditional Sharpe-Ratio of the market portfolio,

$$\frac{\mathbb{E}\left(R_m - R_f\right)}{\sqrt{\mathbb{E}\left[\left(R_m - R_f\right)^2\right] - \left[\mathbb{E}\left(R_m - R_f\right)\right]^2}}, \tag{35}$$

as a key statistic for the risk-return trade-off implied by the LRR model, is matched. Since the means of market excess return and the risk-free rate appear already in the moment matches in (31) and (34), the only remaining moment to be matched is the expected value of the squared market excess return. Hence, we use the second moment of the market excess return to match the market volatility, viz:

$$\boldsymbol{G}_T^{(P,T),3}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) = \frac{1}{T}\sum_{t=1}^{T}(R_{m,t} - R_{f,t})^2 - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)}\left(R_{m,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) - R_{f,s}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P)\right)^2. \tag{36}$$

Campbell and Shiller (1988) point out that there is a prediction relation implied by the linear approximations in (8) and (9), namely that the log price-dividend ratio predicts future discount rates.[7] This predictive relationship can be exploited by matching the slope parameter of a regression of the risk-free rate on past values of the log price-dividend ratio[8], which entails matching the first and the second moment of $z_{m,t}$, too:

$$\boldsymbol{G}_T^{(P,T),4}(\boldsymbol{\xi}_M, \boldsymbol{\xi}_P) = \begin{bmatrix} \frac{\frac{1}{T-1}\sum_{t=1}^{T-1}\left[R_{f,t+1} - \frac{1}{T-1}\sum_{t=1}^{T-1} R_{f,t+1}\right] z_{m,t}}{\frac{1}{T}\sum_{t=1}^{T}\left(z_{m,t} - \frac{1}{T}\sum_{t=1}^{T} z_{m,t}\right)^2} - \frac{\frac{1}{\mathcal{T}(T)-1}\sum_{s=1}^{\mathcal{T}(T)-1}\left[z_{m,s}(\boldsymbol{\xi}_M,\boldsymbol{\xi}_P) - \mu_{1,z_m}\right] R_{f,s+1}(\boldsymbol{\xi}_M,\boldsymbol{\xi}_P)}{\mu_{2,z_m} - \mu_{1,z_m}^2} \\ \frac{1}{T}\sum_{t=1}^{T} z_{m,t} - \mu_{1,z_m} \\ \frac{1}{T}\sum_{t=1}^{T} z_{m,t}^2 - \mu_{2,z_m} \\ \mu_{1,z_m} - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} z_{m,s}(\boldsymbol{\xi}_M,\boldsymbol{\xi}_P) \\ \mu_{2,z_m} - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} z_{m,s}^2(\boldsymbol{\xi}_M,\boldsymbol{\xi}_P) \end{bmatrix}. \tag{37}$$

[7] A simulation exercise shows that the predictive power of $z_{m,t}$ for $R_{f,t+1}$ is indeed strong: the R^2 of a one-step predictive regression is 95%. The simulation is again based on the parameter values given in Table 6 and a sample size of 10^6 observations.

[8] Note that a predictive relationship between price-dividend ratio and future excess returns is not a feature of the LRR model. Hence, $z_{m,t}$ cannot be used as an instrument for the construction of a managed portfolio or in a predictive regression.

Stacking the components (31), (34), (36), and (37) yields a moment match of six theory-motivated moments, which will be used in the objective function in (25).[9]

[Insert Table 5 about here]

The moment sensitivity analysis in Table 5, shows which of the six moment matches provides information about which parameter values. All simulated moments react strongly to a 10% change in the subjective discount factor δ. For both γ and ψ there is one moment condition that reacts sizeably to a parameter change, whereas all other moments do not. Most information about γ is contained in the moment condition (33), which captures the LRR model's pricing implication for the market excess return. This is reflected in a 10% change in the simulated moment in (34) in response to a 10% change in γ. The other simulated moments only respond weakly. The identification of ψ is mainly provided by the slope parameter of the predictive regression of $R_{f,t+1}$ on $z_{m,t}$. The corresponding simulated moment reacts to a 10% change in ψ by an increase of 14%, while the other moments only respond by 4% or less. This "prediction moment" is not sensitive to a change in γ, which helps to disentangle risk aversion and intertemporal elasticity of substitution.

The analytic moment conditions discussed in Section 3.2 can only be used to identify and estimate the unconditional variance σ^2, but not the SV parameters ν_1 and σ_w. The estimation of a latent volatility process has been the topic of a developed econometric literature. The methodological caveats discussed there are aggravated in the LRR model, in which stochastic consumption volatility is just one ingredient of a non-linear multiple-equation model. We have also seen that the ad-hoc moment conditions in (28) are not useful to identify the SV parameters, and we do not claim that our theory-based moment conditions do a better job. These

[9] As in the application of Parker and Julliard (2005), it has to be ensured that the auxiliary parameters μ_M, μ_{1,z_m}, and μ_{2,z_m} in (31) and (37) are exactly matched. This can be achieved by inserting the simulated means into the respective moment conditions.

moment conditions are based on unconditional moment constraints and stochastic volatility is about changing conditional variances. Furthermore, the available data series will be relatively short and thus not informative about the dynamics of the latent consumption volatility process. Alternative approaches to identify ν_1 and σ_w, like exploiting the autocorrelation in the squared market returns or fourth moments of returns, are also not informative enough.

Instead of looking for more sophisticated ways to estimate the SV parameters, we try to simplify the problem. Table 4 shows that even large changes in ν_1 and σ_w merely have a small impact on the unconditional equity premium. Hence, if the primary interest is not the estimation of the SV parameters and the evolution of the conditional risk premium but the estimation of the preference parameters and the model-implied risk premium, an alternative estimation strategy could "concentrate out" the SV parameters. In a simulation of the model in the course of SMM estimation, the stochastic volatility σ_t^2 is replaced by its unconditional forecast, namely σ^2, which is estimated in the first step. We conjecture that the unconditional moments of the simulated financial variables (and the measurable functions of them used for moment matching) would not be greatly affected when σ_t^2 is replaced by its unconditional forecast σ^2. It may be the case that this strategy reduces efficiency, but on the other hand it may also provide more robust results, as the SV parameters may be poorly identified by weak moment conditions and/or a small sample size. In Section 4.3 we assess the feasibility of our estimation strategy in a simulation study. By comparing the estimation results assuming that the SV parameters are known and the estimation results based on the strategy that concentrates out ν_1 and σ_w, we can quantify the loss in efficiency.

The suggested two-step estimation strategy implies that standard asymptotic inference for $\hat{\boldsymbol{\xi}}_P$ cannot be applied due to the use of the macro parameter estimates $\hat{\boldsymbol{\xi}}_{M_0}$. We recommend using a parametric bootstrap instead.

3.4 Computational caveats

The LRR anatomy outlined in Section 2 implies that for every iteration in the optimization process the endogenous parameters have to be computed anew. This is achieved by solving for the mean of z_t and $z_{m,t}$, such that the squared difference between the respective mean hypothesized by the solver and the resulting model-implied mean is equal to zero. Hence, the endogenous parameters are implied by the roots of two functions.

[Insert Figure 2 about here]

The reason why this affects SMM estimation is illustrated in Figure 2, which shows a plot of these two functions. Using the parameter values in Table 6 yields the roots $\bar{z}^* = 6.24$ and $\bar{z}_m^* = 5.49$ (cf. upper panels of Figure 2). The lower panel shows that a change of these parameters within a plausible range may yield an unsolvable model: for example, changing the value of the risk aversion parameter from $\gamma = 10$ to $\gamma = 4$ and the mean of dividend growth from $\mu_d = 0.0015$ to $\mu_d = 0.0035$, leaving all other parameters unchanged, implies that one of the two functions does not have a root. Hence, the endogenous parameters could not be computed in the course of SMM estimation. Model insolubility is not limited to implausible parameter values and the permissible range of values for each parameter changes with the values in the remaining parameters. This inherent "fragility" of the LRR model exacerbates parameter estimation.

A practical way to account for such a problem would be a penalty term that keeps the optimizer away from those unfavorable parameter combinations. However,

this leads to a discontinuity in the surface of the objective function, which might pose a challenge to any optimizer, in particular for gradient-based methods, which are widely used in applied econometrics. We therefore recommend to use a robust optimization algorithm, such as the Nelder Mead simplex.

4 Simulation study

4.1 General setup

We evaluate our methodological approach in an extensive simulation study. For that purpose, we generate data by simulating series of varying length T for g, g_d, r_m, r_f, and z_m as implied by the LRR model. The values of the true macro parameters $\boldsymbol{\xi}_M$ and preference parameters $\boldsymbol{\xi}_P$ are those calibrated by Bansal and Yaron (2004). They are listed in Table 6.

[Insert Table 6 about here]

The length of the simulated series varies between $T = 1000$, $T = 2000$, $T = 5000$, and $T = 100000$. Assuming a monthly sampling frequency, $T = 1000$ is equivalent to about 83 years, which is a reasonable size for a real-world application. The longer series should illustrate the behavior of the estimates for a growing sample size, and $T = 100000$ serves as a reality check of the estimation method, since the parameter estimates should be precise and unbiased. Parameter estimation is performed on the simulated data as described in the previous sections. To ensure logical consistency of the parameter estimates, we restrict $\hat{\varphi}_e$, $\hat{\phi}$, and $\hat{\varphi}_d$ to positive values. $\hat{\rho}$, $\hat{\mu}_c$, and $\hat{\mu}_d$ have to take on values between 0 and 1.

In Section 3.1 we saw that SMM estimation based on the moment conditions in Table 1 yields unreliable results, and we have emphasized the danger of reporting sugar-coated estimates that result from a convergence to a point near plausibly chosen starting values. It is crucial to avoid that fallacy in our simulation study. Prior to engaging in a large-scale simulation study, we have therefore carefully tested both the analytic moment conditions and the theory-based moment matches used for the two-step estimation procedure. For that purpose we started the GMM/SMM optimization algorithms from different, also far-off, starting values on a variety of test

data to make sure that the optimization algorithm converges to identical values. Panels C of Tables 2 and 3 provide a numerical example and a comparison with the estimation strategies that produced the problematic results discussed in Section 3.1. Using the same data as for the Panel A and B estimations, our combined GMM/SMM procedure yields the same estimates regardless of the chosen initial values. The GMM and SMM objective functions within the two-step estimation procedure are well-defined, and the moment matches ensure identification of the deep parameters.

To assess the precision of the estimates, we use 400 replications for each T. To economize computation time, we choose not to use a penalty term in the simulation study, but instead drop replications for which the optimizer stops at an unsolvable model. When applying the method to real data, a penalty term should be used. The next subsection discusses the quality of the first-step GMM estimates for $\boldsymbol{\xi}_{M_0}$, in 4.3 we present the SMM estimation results for $\boldsymbol{\xi}_P$. To ensure a good approximation of the population moments for SMM, we use $\mathcal{T}(T) = 10^6$.

4.2 GMM estimation of the macro parameters

For each time series length T, we perform eight GMM estimations of $\boldsymbol{\xi}_{M_0}$, using the identity matrix for $\boldsymbol{W}_T$ and the moment matches listed in Table 7. The number of moment conditions ranges from exact identification (7 mc) to heavy overidentification (185 mc). Each moment set includes the first and second moment matches of (26) and a varying number of (cross) auto-moments from (27). The maximum lag length is $L_1 = L_2 = L_3 = 60$, meaning that we account for autocovariances up to five years, assuming a monthly frequency. As pointed out previously, identification of the macro parameters should benefit from taking into account long lags for the autocovariances because the drivers of the LRR macro dynamics are slow-moving

processes. We use the Nelder-Mead simplex algorithm for optimization of the GMM objective.

To ensure the feasibility of our study, in which we conduct $400 \times 8 \times 4$ estimations, we use one set of starting values only. The initial robustness checks make us confident that starting value dependence is not an issue for the moment matches we advocate. Nevertheless we choose, quite on purpose, starting values at some distance from the true parameters.[10] A poor starting point makes the problem harder for the optimization algorithm, and hence more time-consuming, but it also prevents the danger of reporting sugar-coated results.

Tables 8–11 report means and standard deviations of the macro parameter estimates computed across the 400 replications. Figures 3–9 illustrate the estimation quality using kernel densities.

[Insert Table 7 about here]

[Insert Tables 8 through 11 about here]

[Insert Figures 3 through 9 about here]

We can see that estimation precision varies across parameters. For ρ, φ_e, and ϕ, both a relatively large sample size ($T > 2000$) and informative moment conditions are needed for good results (cf. Tables 9 and 10), whereas μ_c, σ, and φ_d are less difficult to estimate (cf. Tables 8, 10, and 11). The finite sample simulation evidence reported here indicates that the favorably small asymptotic standard errors reported in empirical estimations of LRR models should be taken with a grain of salt. These applications use a much smaller sample size.

Naturally, the precision of the parameter estimates increases with the sample size. However, for the critical parameters ρ, φ_e, and ϕ, the accuracy also improves

[10] We use $\boldsymbol{\xi}_{M_0} = (0.018, 0.018, 0.881, 0.082, 0.003, 7.389, 7.389)'$ as starting values.

with the number of moment conditions, i.e. with the maximum lag length allowed for the autocovariances. The ability to estimate ρ precisely is crucial because the persistent growth component x_t is the defining feature of the LRR model, which determines the dynamics of macroeconomic and financial variables. Large sample sizes in the range of $T = 5000$ are unattainable in real-world applications, such that the availability of moment conditions that prove to be informative also in small samples is of utmost importance.

For small samples, a moderate increase of the number of moment conditions (up to 35 mc, say) does not lead to an improvement in the parameter estimates. This applies in particular to the critical parameters ρ, φ_e, and ϕ. On the other hand, even an increase of the maximum lag length from three (113 mc) to five years (185 mc) has a beneficial effect. The same conclusion is drawn from consulting the kernel density plots in Figures 3–9, where four different moment sets—7 mc, 35 mc, 113 mc, and 185 mc—are compared. For the parameters which can be estimated relatively easily, the difference between 113 mc and 185 mc is small. However, the larger moment sets offer a considerable improvement for ρ, ϕ, and φ_e. These results indicate that one should choose the largest moment set in order to obtain precise estimates, also and in particular for small sample sizes.

In order to formally determine how many lags are informative, we propose to select the moment set that minimizes the Bayes-Schwarz information criterion for GMM suggested by Andrews (1999):

$$\text{GMM-BIC} = J_T - (|c| - p_b) \ln T, \tag{38}$$

where

$$J_T = T\, \boldsymbol{G}_T(\hat{\boldsymbol{\xi}}_{M_0}) \left[\widehat{\text{Avar}}(\boldsymbol{G}_T(\hat{\boldsymbol{\xi}}_{M_0})) \right]^+ \boldsymbol{G}_T(\hat{\boldsymbol{\xi}}_{M_0}), \tag{39}$$

is Hansen's (1982) J-statistic. $\left[\widehat{\text{Avar}}(\boldsymbol{G}_T(\hat{\boldsymbol{\xi}}_{M_0}))\right]^+$ denotes the pseudo-inverse of the estimated asymptotic covariance matrix of $\boldsymbol{G}_T(\hat{\boldsymbol{\xi}}_{M_0})$, $|c|$ stands for the number of moment conditions, and p_b stands for the number of estimated parameters.

[Insert Table 12 about here]

Table 12 shows that in the vast majority of replications the GMM-BIC selects the largest moment set (185 mc), confirming our previous finding that a high number of lags is indeed informative, which should be exploited to enhance estimation precision.

Replications for which the optimization algorithm did not converge or produced economically implausible results are not included in Tables 8–11 and the kernel densities in Figures 3–9. We consider a result to be implausible if one of the parameter estimates differs from the true parameter value by a factor of ten or more.

In an application using a single data set one would try to tackle those problematic data using the remedies of applied econometrics: using different (and more favorable) starting values, probing alternative optimization algorithms and tuning the algorithm's parameters. However, such a "clinical" handling of the problematic simulated data sets would hamper the feasibility of a simulation study.

[Insert Table 13 about here]

Nevertheless, the analysis of the number of problematic cases across moment matches and sample sizes is quite informative for our purposes. Table 13 shows that the number of successful estimations tends to be smaller for shorter time series and more parsimonious moment sets. It is not surprising that estimation problems are exacerbated in small samples. We want to estimate the parameters of highly persistent latent processes. Because of the slow convergence of sample to population moments this is a difficult endeavor. However, Table 13 shows that the problems can also be mitigated by including more informative moment conditions. For any sample size

T, the use of more lags for the auto-moments—that is, allegedly more informative moment matches—increases the number of successfully estimated replications. For example, for the parsimonious 7 mc we have to discard 152 cases for T=1000, for 185 mc only 79 are discarded. Hence, accounting for remote lags in the moment matches does not only improve estimation precision but also increases the probability of a successful estimation. The exclusion of the problematic replications implies a sample selection effect that strengthens our conclusions even more. The larger moment sets facilitate the computation of estimates also for some of those problematic replications, for which the parsimonious moment sets fail. Since the more difficult replications will yield worse estimates, the superiority of the larger moment sets is even understated in the kernel densities and Tables 8–11.

4.3 SMM estimation of the preference parameters

We now turn to the SMM estimation of the preference parameters $\boldsymbol{\xi}_P = (\delta, \gamma, \psi)'$. For that purpose we compare the theory-based moment conditions that we advocate in Section 3.3 to the ad-hoc moment matches from Equation (28). SV is present in the simulated data, but for estimation we concentrate out the SV parameters as described in Section 3.3. To ensure robust optimization results, we perform an initial grid search of reasonable ranges for the three preference parameters and use the parameter combination that minimizes the SMM objective function as starting values for the Nelder-Mead simplex.

We contrast the results from an estimation that takes the true macro parameters as given to those using the macro parameters estimated in the first step. Using the true macro parameters allows to assess the quality of the moment set selected for the second estimation step independently of the effect of potentially imprecise

first-step estimates. The estimated macro parameters are based on the moment set that produced the most favorable results in the first estimation step (185 mc).

[Insert Figures 10 through 12 about here]

[Insert Table 14 about here]

Table 14 displays means and standard deviations of the SMM parameter estimates using the true macro parameters. Figures 10–12 illustrate the estimation quality by kernel densities. We can see that the time preference δ can be estimated most precisely using both ad-hoc and theory-based moment matches. Parameter standard deviations and biases are low, even for small T. By contrast, the estimation of risk aversion γ and intertemporal elasticity of substitution ψ are more challenging. For that purpose, the theory-based moment matches are clearly more suitable. The superiority of the theory-motivated moment conditions is particularly pronunced when it comes to the estimation of ψ.

Estimation quality is not hampered by concentrating out stochastic volatility. Using the unconditional volatility instead of σ_t^2 when simulating moments only impalpably changes the results, as shown in Table 15. The variation in the parameter estimates for $\boldsymbol{\xi}_P$ is even lower in most cases when the SV parameters are concentrated out.

[Insert Table 15 about here]

The right-hand side panel of Table 14 reports the estimation results using the estimated macro parameters. Compared to the results based on the true macro parameters, standard deviation and bias of the preference parameter estimates are large, even with the best available macro parameter estimates. This result demonstrates the vital importance of precise macro parameter estimates as an input for

the second estimation step. The theory-based moments alone, although informative, do not suffice to deliver good estimates of the preference parameters. For the results reported in Table 16 we have raised the bar for the quality of the macro parameter estimates by restricting the deviations from the true values to less than 100%, which enhances the precision of the preference parameter estimates.

[Insert Table 16 about here]

[Insert Table 17 about here]

Again, the reported results are based on successful estimations only. The numbers of replications for which the preference parameters could be successfully estimated are given in Table 17. Replications were dropped if the algorithm converged to a parameter set for which the model cannot be solved. Furthermore, if one or more parameter exceeded the hundredfold of the true parameter value, the replication was discarded due to economic implausibility. For the simulation studies based on the estimated macro parameters, a successful first-stage estimation is required. Since an estimation of the preference parameters is doomed to fail when one of the underlying macro parameter estimates exceeds the tenfold of the true parameter in absolute value, such replications were skipped on the second stage. This is particularly relevant for the simulation studies with small T where the number of replications is thereby considerably reduced.

Means and standard deviations reported in Table 14 show that the theory-motivated moments outperform the ad-hoc moment matches, irrespective of the underlying macro parameters. The theory-based moments ensure identification, they yield smaller standard deviations when operating both on true and estimated macro parameters and usually produce a smaller bias.

We saw that the accuracy of the estimated macro parameters $\hat{\boldsymbol{\xi}}_{M_0}$ is crucial for the precise estimation of the preference parameters $\boldsymbol{\xi}_P$. In order to increase the

accuracy of $\hat{\boldsymbol{\xi}}_{M_0}$, we use an efficiency-enhancing weighting for GMM estimation. Table 18 compares the results of a GMM estimation of the macro parameters using the identity matrix to a two-stage efficient weighting approach using the inverse covariance matrix of the GMM residuals $\boldsymbol{u}_t$, $\boldsymbol{W}_T = \left[\widehat{\mathrm{Var}}(\boldsymbol{u}_t)\right]^{-1}$.

[Insert Table 18 about here]

It turns out that the efficiency gains seem to be limited. Only for ϕ, which determines the dynamics of dividend growth, there is a non-negligible efficiency gain. However, ϕ is an important parameter that drives the macro dynamics. We therefore explore the impact of using these presumably more efficient macro parameter estimates in the estimation of $\boldsymbol{\xi}_P$.

Table 19 displays the estimation results relying on macro estimates produced with $\boldsymbol{W}_T = \boldsymbol{I}$ and the results from $\boldsymbol{W}_T = \left[\widehat{\mathrm{Var}}(\boldsymbol{u}_t)\right]^{-1}$. Both use the theory-based moment matches to estimate the preference parameters.

[Insert Table 19 about here]

We can see that the results particularly improve for small samples and the parameters which are difficult to estimate. Standard deviations and biases get smaller for all estimates but one. This indicates that a minor improvement in the estimation of the macro dynamics can have a substantial impact. However, the larger improvements of the macro parameters—which subsequently lead to better results for the preference parameters—result from the well-chosen moment conditions rather than from just applying an efficient weighting scheme. Efficient weighting is no panacea that compensates for ill-conceived moment conditions.

5 Conclusion

Estimating an asset pricing model that features two latent processes as fundamental economic drivers, and a pricing kernel that depends on unobservable variables is a demanding job for financial econometrics. The bar is raised if such a model must be solved every time it is computed at new values of the model parameters, and if it is not unlikely that the solution does not exist. Add as a final obstacle the fact that the number of observations available for empirical analyses is small, and you have collected the hurdles for empirical tests of long-run risk asset pricing models.

SMM is an estimation technique that is designed to cope with such methodological challenges. It combines a compelling estimation philosophy—matching sample moments and their model-implied counterparts—with computational feasibility: the model-implied moments need not be analytically expressed as functions of the parameters, but can be approximated by sample means of simulated model series.

While SMM is thus appealing for our purpose, some empirically important questions have not been addressed, and our study aims to close that gap. Are the moments selected for matching really informative enough to identify those deep model parameters, which describe the dynamics of latent processes and investor preferences? Non-identification may hide itself well in such a complex model structure. We saw powerful optimizers go astray on an objective function surface with myriads of local minima implied by weakly identifying moment conditions. This caveat calls for due diligence when transferring the key model characteristics into informative moment matches. An agnostic match of some low-order moments is clearly insufficient. And even if meaningful theory-rooted and practically useable moment conditions can be found, what is the sample size that is required to deliver precise estimates?

We tackle these issues by proposing a combined GMM/SMM two-step estimation strategy, in which we elicit moment conditions that reflect the key features of the LRR model. We first focus on the parameters that drive the macroeconomic dynamics and deal with investor preference parameters in a second step, using the first-step estimates as input and exploiting the asset pricing equations and predictive relations implied by the LRR framework. The question of how large the number of observations has to be for a successful estimation is addressed in an extensive simulation study.

The theoretical moments that we use in the first estimation step can be analytically expressed as functions of the macro parameters, such that GMM estimation becomes feasible. These moment matches are valid in the presence of stochastic consumption volatility, but they cannot identify the SV parameters. They can, however, identify the unconditional consumption volatility, which is required in the second estimation step. The properties of the latent persistent growth component, that defining feature of the LRR model, are captured by including remote lags of (cross-) autocovariances of consumption and dividend growth. Our estimation strategy considerably improves the ability to estimate the parameters of the latent consumption growth component.

Given the notorious difficulties related to estimating stochastic volatility processes, we propose to concentrate out the SV parameters in the second (SMM) estimation step. We do not preclude that SV prevails in the data, but we replace time-varying stochastic volatility by the first-step unconditional volatility estimate when computing the simulated moments in the second step. Unless you are interested in conditional pricing implications, estimating the parameters of interest is feasible without loss of precision.

Using the theory-based moment conditions, which we propose, delivers precise estimates for subjective discount factor, relative risk aversion and the intertemporal elasticity of substitution also for smaller samples. Given the complexity of the LRR asset pricing equations, this is an encouraging result. SMM lives up to the promise of being able to deliver good estimates in a difficult setup, provided that informative moment matches are used.

The caveat is that the estimates of the macro parameters, which are used for the second estimation step, have to be of high quality. To achieve that quality, both informative first-stage moment matches and relatively large sample sizes are mandatory. This finding adds grains of salt to empirical applications, which inevitably have to work with small samples. It may sound like a truism, but in order estimate a complex DAPM like the LRR model it is indispensable to have informative data (long time series) and strong moment matches. We saw that short data series may turn out to be uninformative, such that first-step estimation either fails or the estimates are very imprecise. But if data and/or estimation quality of the economic parameters is poor, you cannot expect too much from the second-step estimation of the preference parameters. Refraining from estimating them in the first place is then the scientifically honest decision. Our two-step approach is therefore a reality check for applied work.

Let us conclude. In this study we have tried to elicit the key features out of the LRR model into meaningful and strong moment matches, and we have discussed the benefits and limitations of our estimation strategy. We believe that subsequent research will most fruitfully be invested in increasing the quality of the macro parameter estimates. And, yes, time has to pass.

References

Andersen, T. G., Chung, H.-J., Sørensen, B. E., 1999. Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. Journal of Econometrics 91 (1), 61–87.

Andrews, D. W. K., 1999. Consistent moment selection procedures for generalized method of moments estimation. Econometrica 67 (3), 543–563.

Bansal, R., Gallant, A. R., Tauchen, G., 2007a. Rational pessimism, rational exuberance, and asset pricing models. The Review of Economic Studies 74 (4), 1005–1033.

Bansal, R., Kiku, D., Yaron, A., 2007b. Risks for the long run: Estimation and inference, Working Paper.

Bansal, R., Shaliastovich, I., 2013. A long-run risks explanation of predictability puzzles in bond and currency markets. The Review of Financial Studies 26 (1), 1–33.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. The Journal of Finance 59 (4), 1481–1509.

Campbell, J. Y., Shiller, R. J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. The Review of Financial Studies 1 (3), 195–228.

Cochrane, J. H., 1996. A cross-sectional test of an investment-based asset pricing model. Journal of Political Economy 104 (3), 572–621.

Constantinides, G. M., Ghosh, A., 2011. Asset pricing tests with long-run risks in consumption growth. The Review of Asset Pricing Studies 1 (1), 96–136.

Drechsler, I., Yaron, A., 2011. What's vol got to do with it. The Review of Financial Studies 24 (1), 1–45.

Epstein, L. G., Zin, S. E., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. Econometrica 57 (4), 937–969.

Ferson, W., Nallareddy, S., Xie, B., 2013. The "out-of-sample" performance of long run risk models. Journal of Financial Economics 107 (3), 537–556.

Gallant, A., Hsieh, D., Tauchen, G., 1997. Estimation of stochastic volatility models with diagnostics. Journal of Econometrics 81 (1), 159–192.

Hansen, L. P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50 (4), 1029–1054.

Hansen, N., Ostermeier, A., 2001. Completely derandomized self-adaptation in evolution strategies. Evolutionary Computation 9 (2), 159–195.

Hasseltoft, H., 2012. Stocks, bonds, and long-run consumption risks. Journal of Financial and Quantitative Analysis 47, 309–332.

Jacquier, E., Polson, N. G., Rossi, P. E., 2002. Bayesian analysis of stochastic volatility models. Journal of Business & Economic Statistics 20 (1), 69–87.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: Likelihood inference and comparison with ARCH models. The Review of Economic Studies 65 (3), 361–393.

Kreps, D. M., Porteus, E. L., 1978. Temporal resolution of uncertainty and dynamic choice theory. Econometrica 46 (1), 185–200.

Pakoš, M., 2013. Long-run risk and hidden growth persistence. Journal of Economic Dynamics and Control 37 (9), 1911–1928.

Parker, J. A., Julliard, C., 2005. Consumption risk and the cross section of expected returns. Journal of Political Economy 113 (1), 185–222.

Ruiz, E., 1994. Quasi-maximum likelihood estimation of stochastic volatility models. Journal of Econometrics 63 (1), 289–306.

Sandmann, G., Koopman, S. J., 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood. Journal of Econometrics 87 (2), 271–301.

Singleton, K. J., 2006. Empirical Dynamic Asset Pricing. Princeton University Press.

Tables and Figures

Table 1: Moments used for moment sensitivity analysis
The table lists the moments used in a first attempt for the estimation of the structural parameters of the long-run risk model. Three types of moment conditions are applied. Crucial moments describing the macroeconomic environment are matched, asset pricing properties are replicated, and prediction relationships are exploited. The macro moments are obtained from the fundamental processes which drive the economy. The formulas for the asset pricing moments emerge as a result of the linear approximations derived on the basis of the macroeconomic environment. ξ_{t+1}, the residual of an AR(1) process for log consumption growth, is obtained by regressing g_{t+1} on g_t.

Macro moments	
Mean of log consumption growth	$\mathbb{E}(g_t)$
Mean of log dividend growth	$\mathbb{E}(g_{d,t})$
Mean of squared log consumption growth	$\mathbb{E}(g_t^2)$
Mean of squared log dividend growth	$\mathbb{E}(g_{d,t}^2)$
Mean of the product of log consumption growth with its first lag	$\mathbb{E}(g_t\ g_{t+1})$
Mean of the product of log consumption growth with its second lag	$\mathbb{E}(g_t\ g_{t+2})$
Asset Pricing moments	
Mean excess return of the market portfolio (in logs)	$\mathbb{E}(r_{m,t} - r_{f,t})$
Mean log risk-free rate	$\mathbb{E}(r_{f,t})$
Mean log price-dividend ratio	$\mathbb{E}(z_{m,t})$
Mean squared excess return of the market portfolio (in logs)	$\mathbb{E}[(r_{m,t} - r_{f,t})^2]$
Mean squared log risk-free rate	$\mathbb{E}[r_{f,t}^2]$
Mean squared log price-dividend ratio	$\mathbb{E}[z_{m,t}^2]$
Prediction moments	
Mean squared residual of $AR(1)$ for log consumption growth	$\mathbb{E}(\xi_{t+1}^2)$
Mean product of squared residual and log price-dividend ratio	$\mathbb{E}(\xi_{t+1}^2\ z_{m,t})$

Table 2: Parameter estimates and objective for different methods and different starting values, short data set

This table compares the results obtained from different estimation approaches using three sets of starting values. We generate data with $T = 1000$ using the parameter values in Table 6 and subsequently try to recover the deep model parameters using different estimation approaches. Panel A reports the results from an SMM estimation using the moment conditions from Table 1, Panel B holds the results from the GMM estimation strategy proposed by Constantinides and Ghosh (2011), and Panel C holds the results from our two-step approach. All results use the identity matrix for weighting. The starting value set $\boldsymbol{s_1}$ is the true parameter set, the other sets of starting values $\boldsymbol{s_2} = (0.0015, 0.0015, 0.95, 0.05, 0.95, 0.00001, 0.008, 3, 5, 0.99, 12, 1.2)'$ and $\boldsymbol{s_3} = (0.001, 0.001, 0.8, 0.05, 0.5, 0.00001, 0.008, 3, 5, 0.96, 15, 2)'$ differ from the true values. Only for the two-step estimation approach, the Nelder-Mead simplex suffices to obtain reliable and starting value-independent results. For the other approaches, a more powerful optimization algorithm (CMAES) had to be used. The two-step estimation approach delivers starting value-independent results. For the other methods, this could not be achieved, even with a sophisticated algorithm designed for global optimization and allowance for a multiple of the computation time.

	Panel A			**Panel B**			**Panel C**		
	$\boldsymbol{s_1}$	$\boldsymbol{s_2}$	$\boldsymbol{s_3}$	$\boldsymbol{s_1}$	$\boldsymbol{s_2}$	$\boldsymbol{s_3}$	$\boldsymbol{s_1}$	$\boldsymbol{s_2}$	$\boldsymbol{s_3}$
$\boldsymbol{\mu_c}$	0.0025	0.0021	0.0021	0.0011	0.0024	0.0004	0.0024	0.0024	0.0024
$\boldsymbol{\mu_d}$	0.0025	0.0026	0.0005	0.0036	0.0026	$2 \cdot 10^{-6}$	0.0026	0.0026	0.0026
$\boldsymbol{\rho}$	0.9782	0.9601	0.0799	0.9991	0.9092	0.7321	0.9765	0.9765	0.9765
$\boldsymbol{\varphi_e}$	0.0405	0.0784	0.1749	0.0094	0.0957	0.1762	0.0324	0.0324	0.0324
$\boldsymbol{\nu_1}$	0.9845	0.9660	0.9772	0.9477	0.9957	0.5739			
$\boldsymbol{\sigma_w}$	$4.0 \cdot 10^{-6}$	$6.7 \cdot 10^{-6}$	$2.3 \cdot 10^{-5}$	$3.6 \cdot 10^{-5}$	$4.0 \cdot 10^{-6}$	$8.3 \cdot 10^{-8}$			
$\boldsymbol{\sigma}$	0.0100	0.0050	0.0261	0.0091	0.0077	0.0071	0.0078	0.0078	0.0078
$\boldsymbol{\phi}$	2.5382	4.3642	6.5399	2.7077	3.1668	0.0420	4.1545	4.1545	4.1545
$\boldsymbol{\varphi_d}$	3.3072	4.6125	0.6058	3.6421	4.3360	4.7927	4.2913	4.2913	4.2913
$\boldsymbol{\delta}$	0.9976	0.9978	0.9919	0.9433	0.9959	-0.1375	0.9986	0.9986	0.9986
γ	10.0745	12.1557	12.1288	11.8520	16.3428	10.5208	17.9922	17.9922	17.9922
ψ	1.8988	2.0425	4.8776	6.8827	4.4990	0.3847	1.1891	1.1891	1.1891
objective	$9.7 \cdot 10^{-8}$	$3.8 \cdot 10^{-7}$	$3.4 \cdot 10^{-5}$	$1.5 \cdot 10^{-5}$	$9.1 \cdot 10^{-11}$	0.0024437	$7.2 \cdot 10^{-8}$	$7.2 \cdot 10^{-8}$	$7.2 \cdot 10^{-8}$
							$2.0 \cdot 10^{-7}$	$2.0 \cdot 10^{-7}$	$2.0 \cdot 10^{-7}$
obj(start)	0.1296	74.6042	408.7485	0.0037	0.0304	6.5822	$2.1 \cdot 10^{-6}$	$2.3 \cdot 10^{-6}$	$4.8 \cdot 10^{-6}$
							33.8966	87.8011	408.4100

Table 3: Parameter estimates and objective for different methods and different starting values, long data set

This table compares the results obtained from different estimation approaches using three sets of starting values. We generate data with $T = 100000$ using the parameter values in Table 6 and subsequently try to recover the deep model parameters using different estimation approaches. Panel A reports the results from an SMM estimation using the moment conditions from Table 1, Panel B holds the results from the GMM estimation strategy proposed by Constantinides and Ghosh (2011), and Panel C holds the results from our two-step approach. All results use the identity matrix for weighting. The starting value set $\boldsymbol{s_1}$ is the true parameter set, the other sets of starting values $\boldsymbol{s_2} = (0.0015, 0.0015, 0.95, 0.05, 0.95, 0.00001, 0.008, 3, 5, 0.99, 12, 1.2)'$ and $\boldsymbol{s_3} = (0.001, 0.001, 0.8, 0.05, 0.5, 0.00001, 0.008, 3, 5, 0.96, 15, 2)'$ differ from the true values. Only for the two-step estimation approach, the Nelder-Mead simplex suffices to obtain reliable and starting value-independent results. For the other approaches, a more powerful optimization algorithm (CMAES) had to be used. The two-step estimation approach delivers starting value-independent results. For the other methods, this could not be achieved, even with a sophisticated algorithm designed for global optimization and allowance for a multiple of the computation time. A star indicates that the algorithm did not converge given the maximum number of iterations.

	Panel A			**Panel B**			**Panel C**		
	s_1^*	s_2	s_3^*	s_1^*	s_2	s_3	s_1	s_2	s_3
μ_c	0.0015	0.0010	0.0005	0.0015	0.0015	0.0003	0.0015	0.0015	0.0015
μ_d	0.0017	0.0020	0.0005	0.0017	0.0017	0.0036	0.0017	0.0017	0.0017
ρ	0.9786	0.9618	0.8044	0.9993	0.9993	0.7305	0.9813	0.9813	0.9813
φ_e	0.0454	0.0658	0.1324	0.0083	0.0083	0.1778	0.0426	0.0426	0.0426
ν_1	0.9864	0.9645	0.7387	0.9830	0.9743	0.0325			
σ_w	$2.1 \cdot 10^{-6}$	$1.1 \cdot 10^{-5}$	$4.4 \cdot 10^{-6}$	$1.5 \cdot 10^{-7}$	$2.2 \cdot 10^{-9}$	$4.5 \cdot 10^{-6}$			
σ	0.0078	0.0059	0.0117	0.0078	0.0078	0.0007	0.0078	0.0078	0.0078
ϕ	2.9786	3.7873	6.0687	2.9246	2.9317	27.9402	2.6930	2.6930	2.6930
φ_d	4.4856	2.6125	2.6915	4.4867	4.4971	51.0524	4.4919	4.4919	4.4919
δ	0.9980	0.9975	0.9969	0.9977	0.9926	0.9916	0.9979	0.9979	0.9979
γ	10.1883	11.7574	14.7956	10.4048	8.7461	9.1228	10.4593	10.4593	10.4593
ψ	1.4974	2.1615	1.0517	0.9483	0.8227	-0.0796	1.5677	1.5677	1.5677
objective	$4.4 \cdot 10^{-13}$	$1.7 \cdot 10^{-6}$	$7.2 \cdot 10^{-6}$	$5.9 \cdot 10^{-6}$	$6.0 \cdot 10^{-6}$	$5.5 \cdot 10^{-4}$	$1.1 \cdot 10^{-9}$	$1.1 \cdot 10^{-9}$	$1.1 \cdot 10^{-9}$
							$3.4 \cdot 10^{-8}$	$3.4 \cdot 10^{-8}$	$3.4 \cdot 10^{-8}$
obj(start)	$3.3 \cdot 10^{-4}$	68.8180	395.0515	$4.1 \cdot 10^{-5}$	0.0102	6.9203	$2.7 \cdot 10^{-8}$	$1.6 \cdot 10^{-7}$	$8.3 \cdot 10^{-7}$
							0.3797	94.3059	398.5025

Table 4: An analysis of moment sensitivity

This table holds the moment sensitivity matrix computed from the moments given in Table 1. The moments are computed based on a simulated sample size of $5 \cdot 10^6$ observations, based on the parameter values from Table 6. The moment sensitivity in this table is computed as the relative change of a moment when one given parameter c.p. decreases by 50%. Each column of the table displays the sensitivity of all moments to a change of that size in the parameter given in the column header.

	μ_c	μ_d	ρ	φ_e	ν_1	σ_w	σ	ϕ	φ_d	δ	γ	ψ
$\mathbb{E}(g)$	-0.50	0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00
$\mathbb{E}(g_d)$	0.00	-0.50	0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	0.00
$\mathbb{E}(g^2)$	-0.03	0.00	-0.04	-0.03	0.00	0.00	-0.67	0.00	0.00	0.00	0.00	0.00
$\mathbb{E}(g_d^2)$	0.00	-0.00	-0.02	-0.02	0.00	0.00	-0.69	-0.02	-0.73	0.00	0.00	0.00
$\mathbb{E}(g_{t+1} \cdot g_t)$	-0.34	0.00	-0.53	-0.41	0.00	0.00	-0.39	0.00	0.00	0.00	0.00	0.00
$\mathbb{E}(g_{t+2} \cdot g_t)$	-0.34	0.00	-0.53	-0.41	0.00	0.00	-0.38	0.00	0.00	0.00	0.00	0.00
$\mathbb{E}(r_m - r_f)$	0.01	-0.03	-1.18	-0.86	-0.04	-0.03	-0.66	-0.69	0.12	-1.18	-0.69	-0.38
$\mathbb{E}(r_f)$	-0.23	0.00	0.17	0.13	0.01	0.01	0.27	0.00	0.00	$> 10^2$	0.21	0.70
$\mathbb{E}(z_m)$	0.02	-0.03	0.43	0.20	0.01	0.00	0.10	0.16	-0.02	-1.00	0.12	-0.01
$\mathbb{E}((r_m - r_f)^2)$	0.02	-0.02	-0.45	-0.32	-0.01	-0.01	-0.66	-0.39	-0.42	-0.46	0.06	-0.23
$\mathbb{E}(r_f^2)$	-0.32	0.00	0.08	0.05	0.00	0.00	0.32	0.00	0.00	$> 10^4$	0.36	2.12
$\mathbb{E}(z_m^2)$	0.04	-0.05	1.06	0.43	0.01	0.01	0.21	0.34	-0.04	-1.00	0.25	-0.02
$\mathbb{E}(\xi^2)$	0.00	0.00	-0.04	-0.03	0.00	0.00	-0.70	0.00	0.00	0.00	0.00	0.00
$\mathbb{E}(\xi_{t+1}^2 \cdot z_{m,t})$	0.02	-0.03	0.38	0.16	0.01	0.01	-0.67	0.16	-0.02	-1.00	0.12	-0.01

Table 5: Moment sensitivity to parameters for theory-based moments
This table holds the moment sensitivity matrix for the theory-based moments using the simulated part of the moment match. The moments are computed from a simulated dataset with a sample size of 10^6 observations, based on the parameters from Table 6. The moment sensitivity in this table is computed as the relative change of a moment when one given parameter c.p. decreases by 10%. Each column of the table displays the sensitivity of all moments to a change of that size in the parameter given in the column header.

	$\boldsymbol{\delta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\psi}$
$\mathbb{E}(M)$	-0.10	-0.00	-0.00
$\frac{-\mathrm{Cov}(R_m - R_f, M)}{\mathbb{E}(M)}$	-0.97	-0.10	-0.04
$\mathbb{E}\left[(R_m - R_f)^2\right]$	-0.32	0.01	-0.03
$\frac{\mathrm{Cov}(R_{f,t+1}, z_{m,t})}{\mathrm{Var}(z_m)}$	4.28	-0.01	0.14
$\mathbb{E}(z_m)$	-0.60	0.02	-0.00
$\mathbb{E}(z_m^2)$	-0.84	0.04	-0.00

Table 6: True parameter values
This table holds the parameter values calibrated by Bansal and Yaron (2004). These values are used as true parameter values for the simulation of the LRR model.

μ_c	0.0015	σ	0.0078
μ_d	0.0015	ϕ	3.0
ρ	0.9790	φ_d	4.5
φ_e	0.0440	δ	0.998
ν_1	0.9870	γ	10.0
σ_w	$2.3 \cdot 10^{-6}$	ψ	1.5

Table 7: Moment conditions used for GMM estimation of macro parameters
For GMM estimation of $\boldsymbol{\xi}_{M_0}$, the basic set of first and second moment conditions in (26) is always included. The maximum lag lengths of the (cross) auto-moments in (27) vary according to the scheme below.

moment set	$\boldsymbol{L_1}$	$\boldsymbol{L_2}$	$\boldsymbol{L_3}$
7 mc	2	0	0
15 mc	5	5	0
20 mc	5	5	5
35 mc	10	10	10
87 mc	36	36	10
113 mc	36	36	36
149 mc	48	48	48
185 mc	60	60	60

Table 8: Means and standard deviations of GMM estimates $\hat{\mu}_c$ and $\hat{\mu}_d$

This table holds the means and the standard deviations of the GMM parameter estimates. 400 data sets were simulated for each sample size T and the estimation was done with 8 different sets of moment conditions, ranging from 7 to 185 moment conditions. The table shows that the precision of the estimates increases both with the length of the data series and with the number of moment conditions.

	7mc	**15mc**	**20mc**	**35mc**	**87mc**	**113mc**	**149mc**	**185mc**
$\boldsymbol{\mu_c}$**=0.0015**								
T=1000	0.001588	0.001557	0.001566	0.001539	0.001549	0.001557	0.001552	0.001566
	(0.000501)	(0.000529)	(0.000548)	(0.000556)	(0.000536)	(0.000534)	(0.000525)	(0.000520)
T=2000	0.001546	0.001530	0.001529	0.001532	0.001531	0.001521	0.001524	0.001524
	(0.000359)	(0.000381)	(0.000378)	(0.000371)	(0.000366)	(0.000368)	(0.000373)	(0.000367)
T=5000	0.001522	0.001523	0.001529	0.001524	0.001519	0.001517	0.001519	0.001522
	(0.000239)	(0.000240)	(0.000237)	(0.000237)	(0.000243)	(0.000242)	(0.000242)	(0.000238)
T=100000	0.001496	0.001495	0.001496	0.001496	0.001496	0.001496	0.001496	0.001496
	(0.000058)	(0.000058)	(0.000058)	(0.000059)	(0.000059)	(0.000059)	(0.000059)	(0.000059)
$\boldsymbol{\mu_d}$**=0.0015**								
T=1000	0.002084	0.001948	0.002033	0.002012	0.001949	0.001982	0.001947	0.002017
	(0.001446)	(0.001559)	(0.001604)	(0.001602)	(0.001485)	(0.001505)	(0.001486)	(0.001467)
T=2000	0.001749	0.001661	0.001685	0.001730	0.001677	0.001651	0.001669	0.001665
	(0.001124)	(0.001172)	(0.001190)	(0.001179)	(0.001153)	(0.001139)	(0.001158)	(0.001149)
T=5000	0.001590	0.001601	0.001616	0.001604	0.001574	0.001576	0.001577	0.001589
	(0.000816)	(0.000804)	(0.000801)	(0.000805)	(0.000832)	(0.000836)	(0.000835)	(0.000830)
T=100000	0.001482	0.001481	0.001481	0.001481	0.001482	0.001482	0.001482	0.001482
	(0.000191)	(0.000192)	(0.000192)	(0.000192)	(0.000192)	(0.000192)	(0.000192)	(0.000192)

Table 9: Means and standard deviations of GMM estimates $\hat{\rho}$ and $\hat{\varphi}_e$

This table holds the means and the standard deviations of the GMM parameter estimates. 400 data sets were simulated for each sample size T and the estimation was done with 8 different sets of moment conditions, ranging from 7 to 185 moment conditions. The table shows that the precision of the estimates increases both with the length of the data series and with the number of moment conditions.

	7mc	**15mc**	**20mc**	**35mc**	**87mc**	**113mc**	**149mc**	**185mc**
ρ=0.9790								
T=1000	0.808712	0.751218	0.788771	0.803248	0.917071	0.926602	0.925134	0.914304
	(0.293789)	(0.325286)	(0.290278)	(0.292826)	(0.156420)	(0.150617)	(0.145199)	(0.170581)
T=2000	0.843687	0.803634	0.807005	0.837530	0.949981	0.954998	0.952202	0.954707
	(0.238097)	(0.259096)	(0.274239)	(0.263833)	(0.103534)	(0.095277)	(0.107220)	(0.087877)
T=5000	0.890665	0.879996	0.898257	0.933842	0.971759	0.972920	0.972324	0.972734
	(0.155095)	(0.181472)	(0.148819)	(0.121116)	(0.049922)	(0.047260)	(0.047247)	(0.027517)
T=100000	0.931789	0.969324	0.971714	0.977699	0.978944	0.978968	0.978865	0.978726
	(0.099321)	(0.033394)	(0.030141)	(0.015070)	(0.004005)	(0.003863)	(0.003003)	(0.002683)
φ_e=0.0440								
T=1000	0.075223	0.065644	0.071262	0.068932	0.053135	0.052875	0.056154	0.056177
	(0.099367)	(0.089732)	(0.088866)	(0.081661)	(0.055733)	(0.054500)	(0.051267)	(0.047299)
T=2000	0.079019	0.065668	0.071638	0.068225	0.046782	0.045572	0.050383	0.053195
	(0.090658)	(0.074406)	(0.080417)	(0.072669)	(0.043367)	(0.042948)	(0.044188)	(0.044524)
T=5000	0.080681	0.054945	0.062062	0.055617	0.042828	0.043298	0.045803	0.047415
	(0.081350)	(0.059491)	(0.066043)	(0.053241)	(0.039532)	(0.034905)	(0.033780)	(0.028952)
T=100000	0.062456	0.040432	0.040156	0.041713	0.043773	0.043746	0.043974	0.044204
	(0.052848)	(0.034373)	(0.034101)	(0.020359)	(0.004770)	(0.005509)	(0.004320)	(0.004015)

Table 10: Means and standard deviations of GMM estimates $\hat{\sigma}$ and $\hat{\phi}$

This table holds the means and the standard deviations of the GMM parameter estimates. 400 data sets were simulated for each sample size T and the estimation was done with 8 different sets of moment conditions, ranging from 7 to 185 moment conditions. The table shows that the precision of the estimates increases both with the length of the data series and with the number of moment conditions.

	7mc	**15mc**	**20mc**	**35mc**	**87mc**	**113mc**	**149mc**	**185mc**
σ=0.0078								
T=1000	0.007772	0.007791	0.007799	0.007775	0.007789	0.007770	0.007784	0.007778
	(0.000412)	(0.000431)	(0.000410)	(0.000464)	(0.000404)	(0.000459)	(0.000411)	(0.000404)
T=2000	0.007788	0.007824	0.007804	0.007766	0.007799	0.007798	0.007790	0.007781
	(0.000296)	(0.000284)	(0.000296)	(0.000493)	(0.000356)	(0.000331)	(0.000355)	(0.000384)
T=5000	0.007785	0.007834	0.007800	0.007779	0.007814	0.007818	0.007814	0.007815
	(0.000198)	(0.000182)	(0.000187)	(0.000386)	(0.000184)	(0.000177)	(0.000179)	(0.000179)
T=100000	0.007791	0.007803	0.007800	0.007801	0.007803	0.007803	0.007803	0.007802
	(0.000048)	(0.000043)	(0.000044)	(0.000043)	(0.000041)	(0.000042)	(0.000042)	(0.000042)
ϕ=3.0								
T=1000	4.079648	5.322702	5.377577	5.060300	4.193287	3.875414	4.036691	4.118246
	(5.334688)	(5.634923)	(5.253479)	(5.231113)	(4.212258)	(3.460940)	(3.689799)	(3.976272)
T=2000	3.878558	5.367465	5.066246	4.353708	3.594153	3.436140	3.307435	3.357487
	(4.334499)	(4.839307)	(4.745072)	(4.097956)	(2.747727)	(2.523522)	(1.996180)	(1.942996)
T=5000	2.983722	4.753321	3.837608	3.372612	3.116605	3.095433	3.104853	3.111128
	(2.576938)	(4.281023)	(2.570627)	(1.651467)	(0.905312)	(0.828071)	(0.823260)	(0.759768)
T=100000	2.973181	3.116188	3.038033	3.011684	3.010016	3.012849	3.012190	3.010695
	(1.616907)	(0.934635)	(0.290304)	(0.190327)	(0.148244)	(0.144897)	(0.142191)	(0.142076)

Table 11: Means and standard deviations of GMM estimates $\hat{\varphi}_d$

This table holds the means and the standard deviations of the GMM parameter estimates. 400 data sets were simulated for each sample size T and the estimation was done with 8 different sets of moment conditions, ranging from 7 to 185 moment conditions. The table shows that the precision of the estimates increases both with the length of the data series and with the number of moment conditions.

	7mc	**15mc**	**20mc**	**35mc**	**87mc**	**113mc**	**149mc**	**185mc**
φ_d**=4.5**								
T=1000	4.442414	4.470252	4.461027	4.494593	4.491056	4.495483	4.483134	4.489017
	(0.272011)	(0.200707)	(0.198498)	(0.297396)	(0.174139)	(0.251202)	(0.167098)	(0.168620)
T=2000	4.468676	4.448006	4.464787	4.581143	4.503892	4.503805	4.509860	4.516286
	(0.192713)	(0.160857)	(0.165845)	(1.643127)	(0.321805)	(0.221512)	(0.284563)	(0.332658)
T=5000	4.504565	4.461806	4.486654	4.561332	4.495095	4.493135	4.495772	4.495031
	(0.116551)	(0.121951)	(0.101916)	(1.214736)	(0.077883)	(0.072184)	(0.072824)	(0.073611)
T=100000	4.505292	4.496131	4.498794	4.498771	4.498175	4.498054	4.498214	4.498434
	(0.044705)	(0.020153)	(0.019967)	(0.018085)	(0.016623)	(0.016878)	(0.016946)	(0.016911)

Table 12: Moment matches selected by GMM-BIC

This table evaluates the performance of the different moment sets by comparing the GMM-BIC. We count the number of replications in which a given moment set minimizes the BIC criterion for each T. Replications with implausible parameter estimates, failed computation of the GMM-BIC, or for which the optimization algorithm did not converge, were dropped. We consider an estimate implausible if it is ten times bigger than the true parameter value. Furthermore, a replication is discarded if the J-statistic lies above the 99.999% quantile of the respective χ^2-distribution, which leads to an implausibly high value of the GMM-BIC. The table shows that the highest number of moment conditions minimizes the information criterion in the majority of the cases. Hence, we find that we should include a high number of lags for autocorrelations and cross-correlations.

	7mc	**15mc**	**20mc**	**35mc**	**87mc**	**113mc**	**149mc**	**185mc**
T=1000	0	1	5	12	11	13	31	255
T=2000	0	1	2	6	1	12	25	308
T=5000	0	1	0	0	0	3	24	364
T=100000	0	0	0	0	0	0	0	400

Table 13: Successful estimations for the macro parameters

This table gives the number of successful estimations for each simulation study of the macro submodel. The total number of replications is 400. Results of a replication are dropped if a parameter estimate is larger than ten times the true value or if the algorithm did not converge.

	7mc	**15mc**	**20mc**	**35mc**	**87mc**	**113mc**	**149mc**	**185mc**
T=1000	248	227	257	281	317	326	325	321
T=2000	325	267	317	324	355	363	358	365
T=5000	375	341	360	369	392	393	394	389
T=100000	397	394	397	399	400	399	399	400

Table 14: Means and standard deviations of the SMM estimates $\hat{\delta}$, $\hat{\gamma}$, and $\hat{\psi}$
This table displays the results for the preference parameters obtained from two different moment sets, the ad-hoc chosen moments and the theory-based moments. For both moment sets the estimation was performed once based on the true macro parameters and once based on the estimated macro parameters from the first stage. The table shows that the precision of the estimates tremendously depends on the quality of the macro parameters.

	true macro parameters		estimated macro parameters	
	ad-hoc	**theory-based**	**ad-hoc**	**theory-based**
δ=0.9980				
T=1000	0.9981	0.9980	0.9955	0.9965
	(0.0008)	(0.0006)	(0.0047)	(0.0021)
T=2000	0.9980	0.9980	0.9966	0.9972
	(0.0006)	(0.0004)	(0.0027)	(0.0019)
T=5000	0.9979	0.9980	0.9979	0.9978
	(0.0004)	(0.0003)	(0.0011)	(0.0007)
T=100000	0.9979	0.9980	0.9980	0.9980
	(0.0002)	(0.0001)	(0.0005)	(0.0002)
γ=10				
T=1000	10.5409	10.3399	34.4836	26.5381
	(1.4146)	(1.1108)	(82.5210)	(28.4748)
T=2000	10.3125	10.2983	19.5936	16.4719
	(1.0566)	(0.7999)	(38.9336)	(14.8120)
T=5000	10.2542	10.3380	11.7276	12.7821
	(0.7275)	(0.5097)	(6.9328)	(5.7417)
T=100000	10.1376	10.3287	10.3326	10.3929
	(0.3114)	(0.1112)	(0.6271)	(0.6996)
ψ=1.5				
T=1000	1.8402	1.5171	4.0093	3.1949
	(1.7972)	(0.0542)	(9.5282)	(3.6131)
T=2000	1.9827	1.5149	3.3048	2.5063
	(4.9717)	(0.0256)	(5.4362)	(1.8937)
T=5000	1.7570	1.5154	3.2752	1.9462
	(1.3764)	(0.0428)	(7.4588)	(1.5177)
T=100000	1.7964	1.5129	1.7586	1.5279
	(2.2075)	(0.0017)	(0.9212)	(0.1533)

Table 15: Means and standard deviations of $\hat{\xi}_P$ with SV concentrated out or using true ν_1 and σ_w

This table shows that it does not make a difference whether we concentrate out ν_1 and σ_w or use their true values when we estimate the preference parameters. The simulated data have SV present. The estimation is based on the theory-motivated moment conditions and the true macro parameters.

	SV concentrated out	**true values for ν_1, σ_w**
δ=0.9980		
T=1000	0.9980	0.9982
	(0.0006)	(0.0008)
T=2000	0.9980	0.9983
	(0.0004)	(0.0007)
T=5000	0.9980	0.9984
	(0.0003)	(0.0005)
T=100000	0.9980	0.9983
	(0.0001)	(0.0001)
γ=10		
T=1000	10.3399	10.3238
	(1.1108)	(1.3043)
T=2000	10.2983	10.4376
	(0.7999)	(1.0647)
T=5000	10.3380	10.6081
	(0.5097)	(0.7407)
T=100000	10.3287	10.5481
	(0.1112)	(0.1607)
ψ=1.5		
T=1000	1.5171	1.5260
	(0.0542)	(0.0923)
T=2000	1.5149	1.5195
	(0.0256)	(0.0591)
T=5000	1.5154	1.5125
	(0.0428)	(0.0141)
T=100000	1.5129	1.5140
	(0.0017)	(0.0022)

Table 16: Means and standard deviations of the SMM estimates $\hat{\delta}$, $\hat{\gamma}$, and $\hat{\psi}$ using precise first-step estimates $\hat{\boldsymbol{\xi}}_{M_0}$

This table displays the results for the preference parameters obtained from two different moment sets, the ad-hoc chosen moments and the theory-based moments. For both moment sets the estimation of the preference parameters is only performed for reasonably precisely estimated macro parameters $\hat{\boldsymbol{\xi}}_{M_0}$ from the first stage. The benchmark is that the estimates $\hat{\varphi}_e$, $\hat{\phi}$, and $\hat{\varphi}_d$ do not deviate by more than 100% from their true values. The table shows that the precision of the estimates is enhanced by ensuring the quality of the macro parameters, compared to the results based on estimated macro parameters displayed in Table 14.

	ad-hoc	**theory-based**
δ=0.9980		
T=1000	0.9963	0.9969
	(0.0024)	(0.0018)
T=2000	0.9967	0.9972
	(0.0026)	(0.0014)
T=5000	0.9979	0.9978
	(0.0011)	(0.0006)
T=100000	0.9980	0.9980
	(0.0005)	(0.0002)
γ=10		
T=1000	14.5026	17.6116
	(12.1657)	(17.3300)
T=2000	15.2075	13.7967
	(33.0646)	(10.7466)
T=5000	11.2477	12.2792
	(3.3315)	(4.1390)
T=100000	10.3326	10.3929
	(0.6271)	(0.6996)
ψ=1.5		
T=1000	3.1976	2.3595
	(9.1846)	(3.4152)
T=2000	3.0002	1.9782
	(4.7453)	(1.0916)
T=5000	3.3112	1.7258
	(7.5025)	(0.6467)
T=100000	1.7586	1.5279
	(0.9212)	(0.1533)

Table 17: Successful second-step SMM estimations for the preference parameters

This table gives the number of successfully estimated replications for each simulation study of the asset pricing model. The maximum possible number of successful estimations for $\boldsymbol{\xi}_P$ when the estimated macro parameters are used is limited by the respective number of successfully estimated replications from the macro model, given in the last column of Table 13. In addition, results are dropped if the resulting preference parameter estimates are larger than the hundredfold of their true values or if the algorithm did not converge.

	true macro parameters		estimated macro parameters	
	ad-hoc	**theory-based**	**ad-hoc**	**theory-based**
T=1000	382	333	159	240
T=2000	392	368	177	281
T=5000	393	382	226	329
T=100000	392	399	396	392

Table 18: Means and standard deviations of first-step GMM estimates using identity matrix and "efficient" weighting scheme

The results from a two-stage version of the GMM efficient weighting are compared to first-stage GMM using the identity matrix. The "efficient" weighting matrix is obtained as the inverse variance matrix of the GMM residuals from a first-stage GMM estimation. We use the 185 mc described in Table 7.

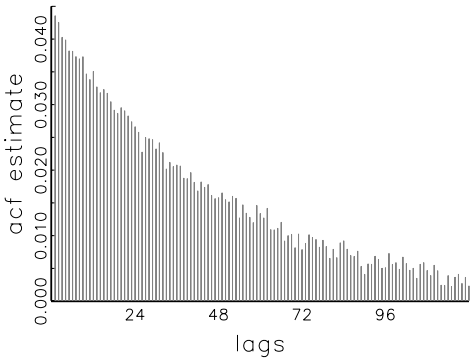
	$\boldsymbol{W}_T = \boldsymbol{I}$	$\boldsymbol{W}_T = \left[\widehat{\mathrm{Var}}(\boldsymbol{u}_t)\right]^{-1}$	$\boldsymbol{W}_T = \boldsymbol{I}$	$\boldsymbol{W}_T = \left[\widehat{\mathrm{Var}}(\boldsymbol{u}_t)\right]^{-1}$
	$\boldsymbol{\mu_c}$**=0.0015**		$\boldsymbol{\mu_d}$**=0.0015**	
T=1000	0.001566	0.001431	0.002017	0.001924
	(0.000520)	(0.000569)	(0.001467)	(0.001532)
T=2000	0.001524	0.001465	0.001665	0.001628
	(0.000367)	(0.000382)	(0.001149)	(0.001195)
T=5000	0.001522	0.001490	0.001589	0.001557
	(0.000238)	(0.000243)	(0.000830)	(0.000858)
T=100000	0.001496	0.001490	0.001482	0.001475
	(0.000059)	(0.000095)	(0.000192)	(0.000215)
	$\boldsymbol{\rho}$**=0.9790**		$\boldsymbol{\varphi_e}$**=0.0440**	
T=1000	0.914304	0.926508	0.056177	0.058211
	(0.170581)	(0.172008)	(0.047299)	(0.049801)
T=2000	0.954707	0.961988	0.053195	0.048828
	(0.087877)	(0.096559)	(0.044524)	(0.029814)
T=5000	0.972734	0.976753	0.047415	0.045632
	(0.027517)	(0.009447)	(0.028952)	(0.010915)
T=100000	0.978726	0.978759	0.044204	0.044211
	(0.002683)	(0.001637)	(0.004015)	(0.001899)
	$\boldsymbol{\sigma}$**=0.0078**		$\boldsymbol{\phi}$**=3.0**	
T=1000	0.007778	0.006866	4.118246	3.279867
	(0.000404)	(0.000396)	(3.976272)	(2.524031)
T=2000	0.007781	0.007327	3.357487	3.065498
	(0.000384)	(0.000264)	(1.942996)	(1.493208)
T=5000	0.007815	0.007612	3.111128	2.964123
	(0.000179)	(0.000172)	(0.759768)	(0.604832)
T=100000	0.007802	0.007791	3.010695	2.997759
	(0.000042)	(0.000041)	(0.142076)	(0.184094)
	$\boldsymbol{\varphi_d}$**=4.5**			
T=1000	4.489017	4.749044		
	(0.168620)	(0.224397)		
T=2000	4.516286	4.627509		
	(0.332658)	(0.135163)		
T=5000	4.495031	4.549365		
	(0.073611)	(0.072046)		
T=100000	4.498434	4.501541		
	(0.016911)	(0.016699)		

Table 19: Means and standard deviations of the preference parameter estimates using an efficient weighting matrix estimate in the estimation of the macro parameters

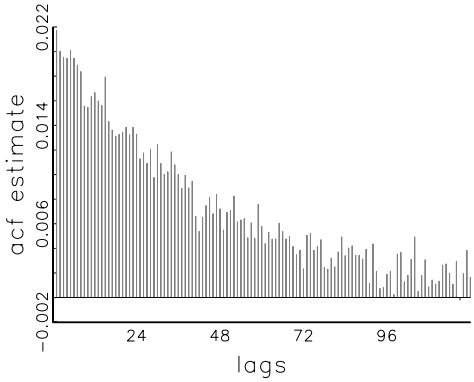
This table displays the results from the preference parameters, estimated based on two different underlying sets of estimated macro parameters: in the left column, the macro parameters are estimated via a first-stage GMM approach, in the right column the macro parameters are estimated using a two-stage efficient weighting matrix. The moment set used for the estimation of the preference parameters is the theory-motivated moment set. The table shows that a higher precision of a relevant macro parameter estimate helps to improve the results of the asset pricing model estimation.

	$\boldsymbol{W}_T = \boldsymbol{I}$	$\boldsymbol{W}_T = \left[\widehat{\mathrm{Var}}(\boldsymbol{u}_t)\right]^{-1}$	$\boldsymbol{W}_T = \boldsymbol{I}$	$\boldsymbol{W}_T = \left[\widehat{\mathrm{Var}}(\boldsymbol{u}_t)\right]^{-1}$
	δ=0.9980		**γ=10**	
T=1000	0.9965	0.9970	26.5381	19.8183
	(0.0021)	(0.0024)	(28.4748)	(23.9993)
T=2000	0.9972	0.9974	16.4719	14.5322
	(0.0019)	(0.0013)	(14.8120)	(14.4060)
T=5000	0.9978	0.9978	12.7821	12.0499
	(0.0007)	(0.0007)	(5.7417)	(4.7351)
T=100000	0.9980	0.9980	10.3929	10.3809
	(0.0002)	(0.0001)	(0.6996)	(0.6776)
	ψ=1.5			
T=1000	3.1949	2.8801		
	(3.6131)	(2.9754)		
T=2000	2.5063	2.1048		
	(1.8937)	(1.4821)		
T=5000	1.9462	1.7309		
	(1.5177)	(0.6011)		
T=100000	1.5279	1.5248		
	(0.1533)	(0.0976)		

Figure 1: Autocorrelograms of simulated consumption and dividend growth
These autocorrelograms illustrate the persistence of the growth processes defining the macroeconomy. The graphs are based on a model simulation based on the parameter values used by Bansal and Yaron (2004) listed in Table 6 and a sample size of 10^6 observations. The abscissa spans a time interval of 10 years, the half-life of both autocorrelations is about three years.



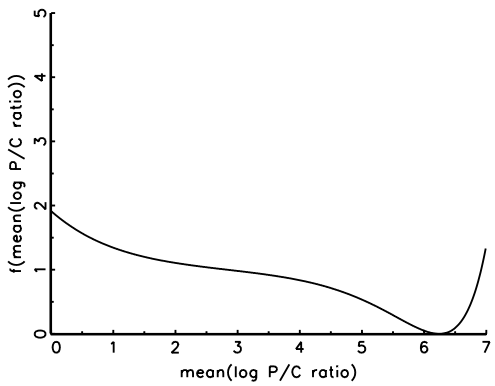
log consumption growth



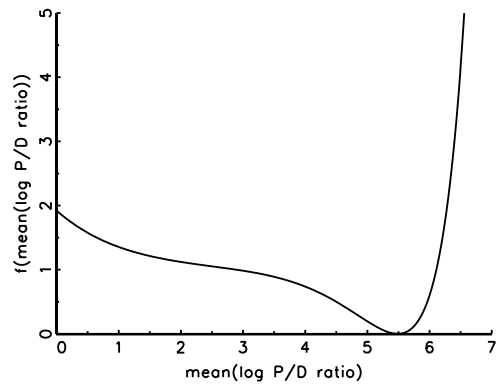
log dividend growth

Figure 2: (Non-) existence of the solution for the endogenous LRR model parameters

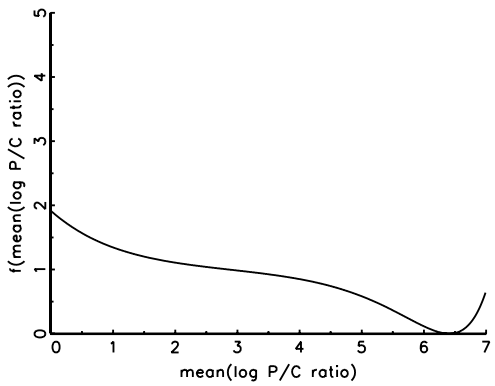
Finding the roots of the squared deviations between hypothetical and model-implied mean is required to solve for the endogenous parameters. If the deviation functions do not both have a root the model cannot be solved. These cases need to be prevented in simulations of the model.



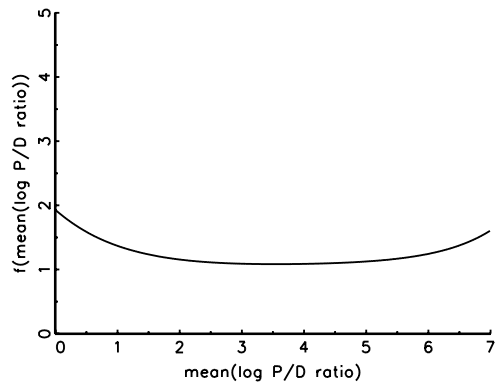
root for $\bar{z}$ for $\gamma = 10$, $\mu_d = 0.0015$



root for $\bar{z}_m$ for $\gamma = 10$, $\mu_d = 0.0015$



root for $\bar{z}$ for $\gamma = 4$, $\mu_d = 0.0035$



root for $\bar{z}_m$ for $\gamma = 4$, $\mu_d = 0.0035$

Figure 3: Kernel densities for $\hat{\mu}_c$

The figure displays the kernel densities for $\hat{\mu}_c$ resulting from different moment sets. In order to account for the boundedness of the parameters, we use beta kernels for the parameters between 0 and 1. The vertical lines indicate the position of the true parameter. The information for the estimate $\hat{\mu}_c$ is mainly contained in $\mathbb{E}(g_t)$, which is matched in all moment sets. Therefore, the estimation precision of $\hat{\mu}_c$ varies little across moment sets.

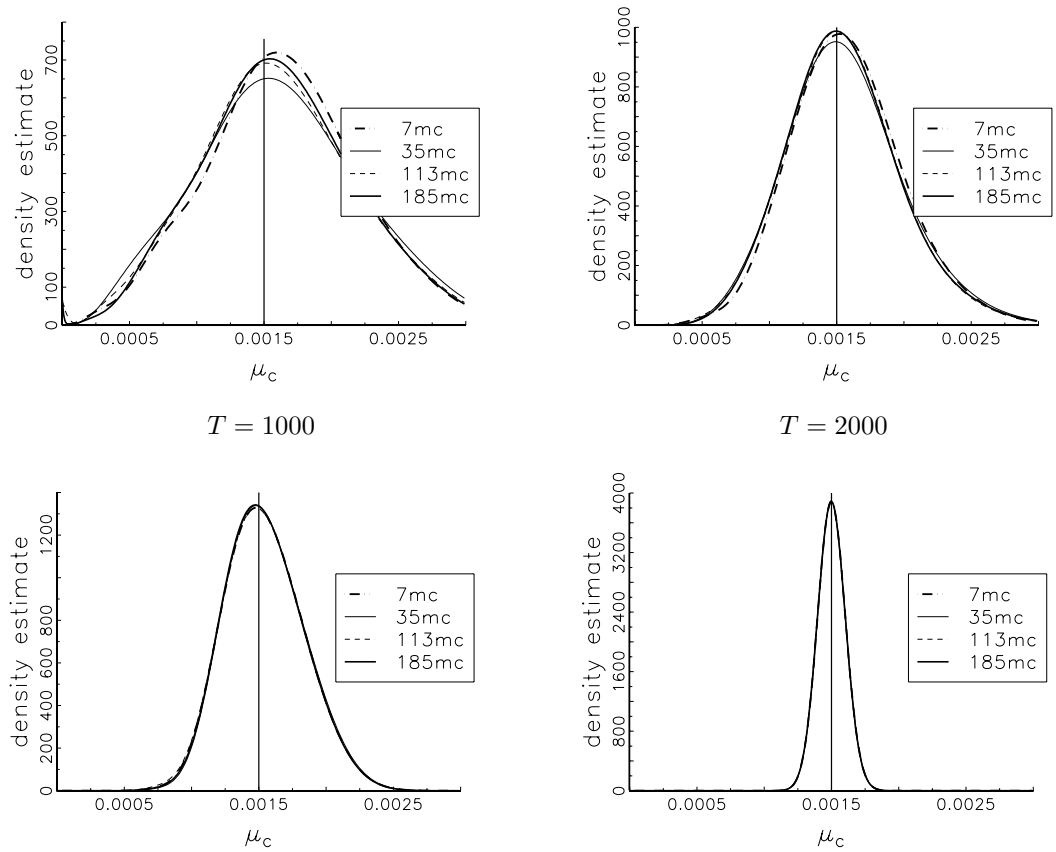


Figure 4: Kernel densities for $\hat{\mu}_d$

The figure displays the kernel densities for $\hat{\mu}_d$ resulting from different moment sets. In order to account for the boundedness of the parameters, we use beta kernels for the parameters between 0 and 1. The vertical lines indicate the position of the true parameter. The information for the estimate $\hat{\mu}_d$ is mainly contained in $\mathbb{E}(g_{d,t})$, which is matched in all moment sets. Therefore, the estimation precision of $\hat{\mu}_d$ varies little across moment sets.

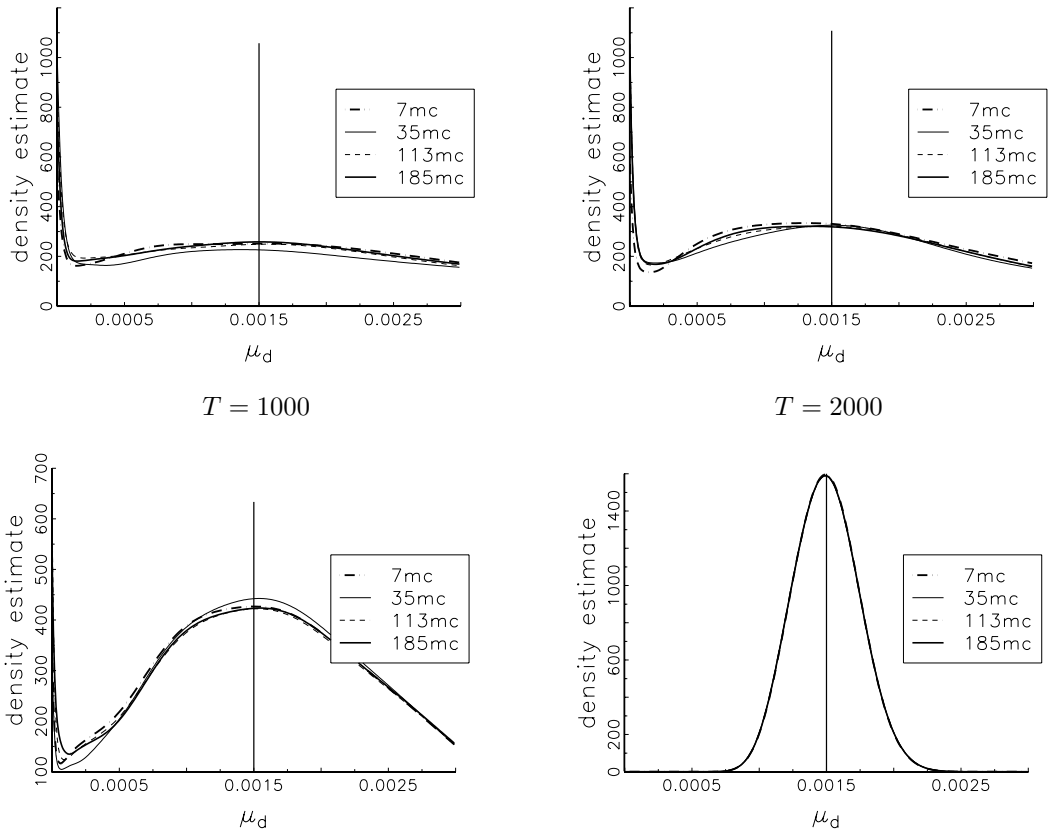


Figure 5: Kernel densities for $\hat{\rho}$

The figure displays the kernel densities for $\hat{\rho}$ resulting from different moment sets. In order to account for the boundedness of the parameters, we use beta kernels for the parameters between 0 and 1. The vertical lines indicate the position of the true parameter. It stands out that the large moment sets are clearly superior to the smaller moment sets, hence, matching autocovariances for long lags enhances the precision of $\hat{\rho}$.

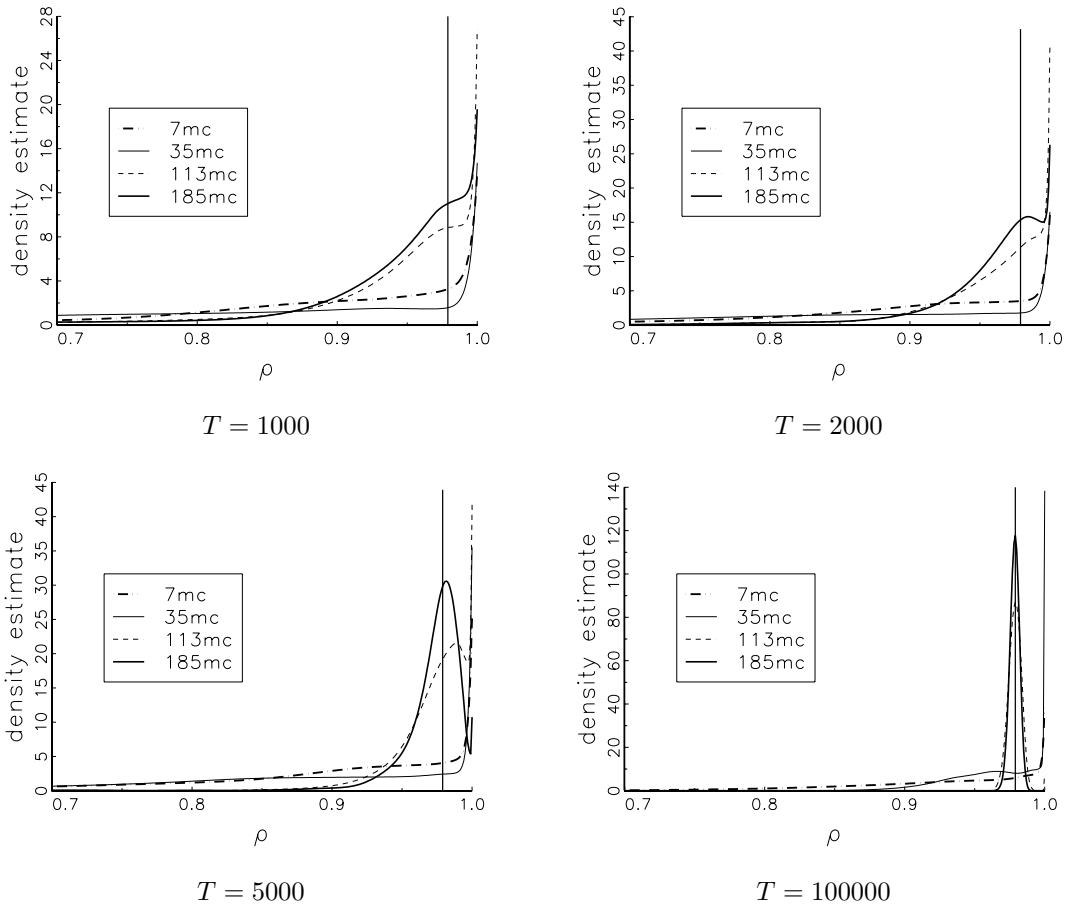


Figure 6: Kernel densities for $\hat{\varphi}_e$
The figure displays the kernel densities for $\hat{\varphi}_e$ resulting from different moment sets. The vertical lines indicate the position of the true parameter. It stands out that the large moment sets are clearly superior to the smaller moment sets, hence, matching autocovariances for long lags enhances the precision of $\hat{\varphi}_e$.

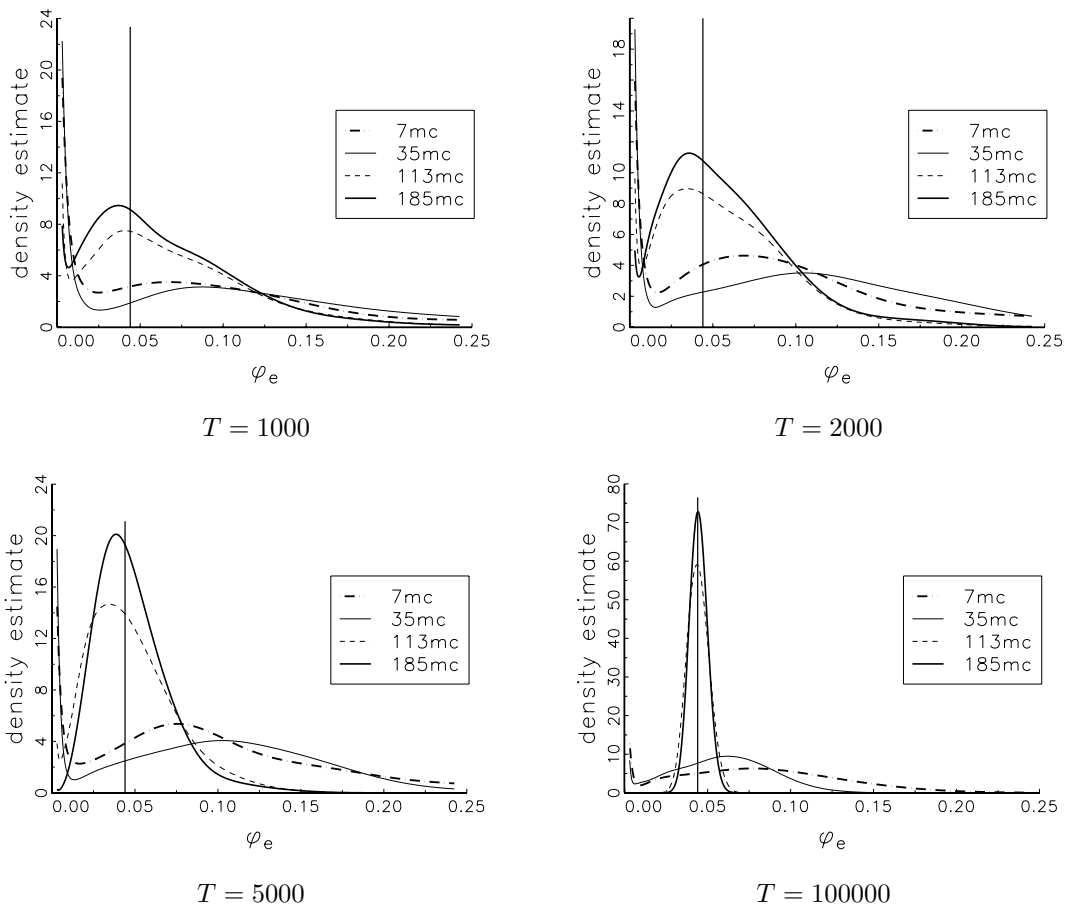


Figure 7: Kernel densities for $\hat{\sigma}$

The figure displays the kernel densities for $\hat{\sigma}$ resulting from different moment sets. The vertical lines indicate the position of the true parameter. The information for the estimate of the unconditional variance $\hat{\sigma}$ is mainly contained in $\mathbb{E}(g_t^2)$ and $\mathbb{E}(g_{d,t}^2)$, which are matched in all moment sets. Therefore, the estimation precision of $\hat{\sigma}$ varies little across moment sets.

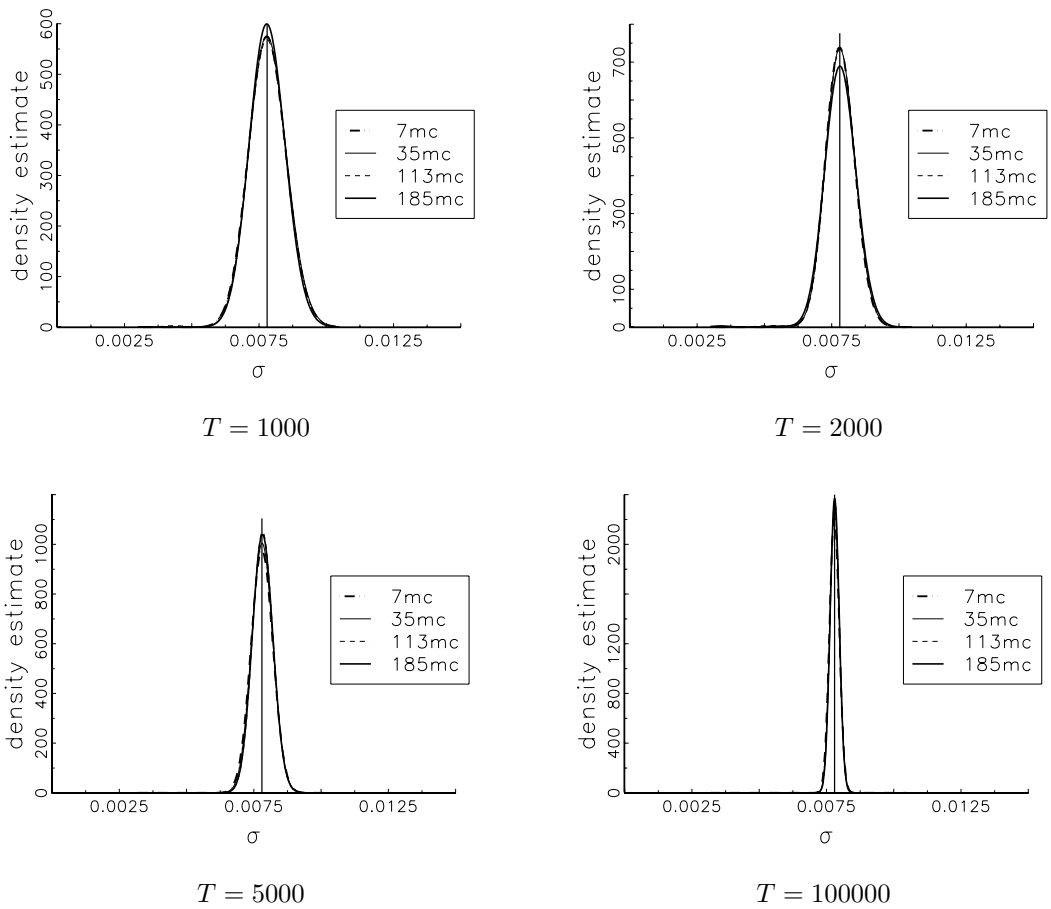
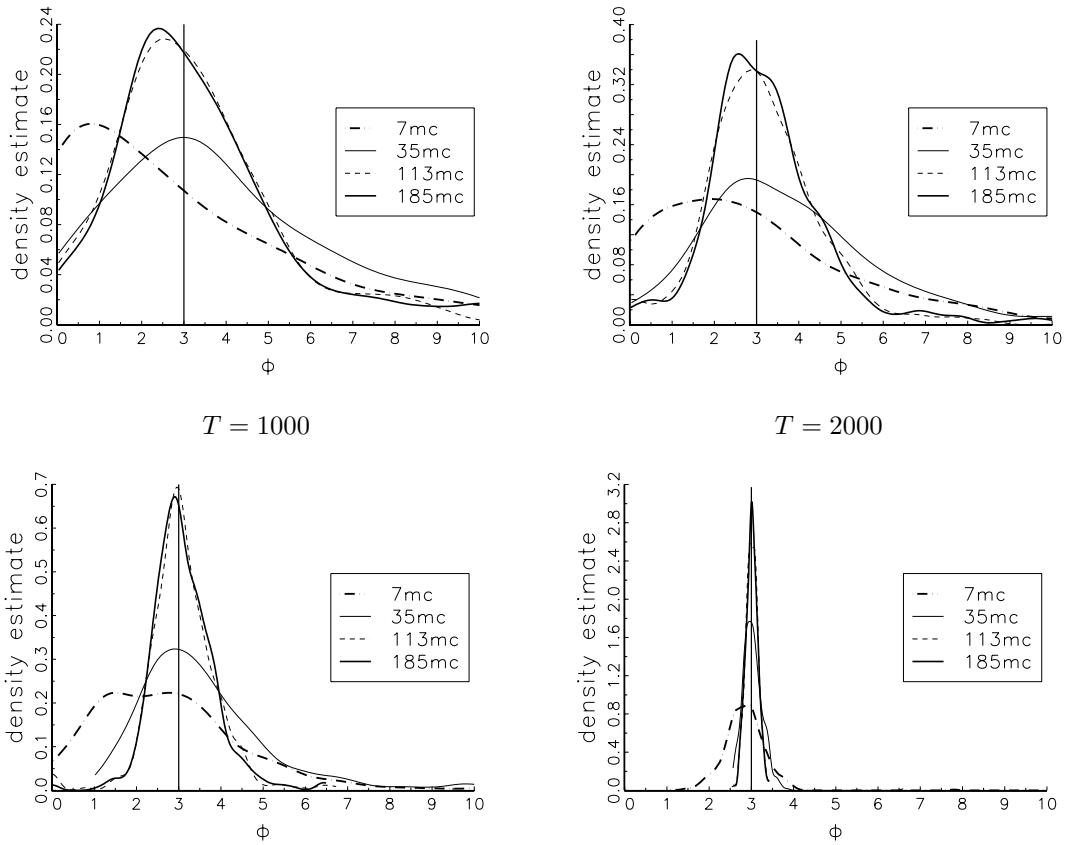


Figure 8: Kernel densities for $\hat{\phi}$

The figure displays the kernel densities for $\hat{\phi}$ resulting from different moment sets. The vertical lines indicate the position of the true parameter. It stands out that the large moment sets are clearly superior to the smaller moment sets, hence, matching autocovariances for long lags enhances the precision of $\hat{\phi}$.

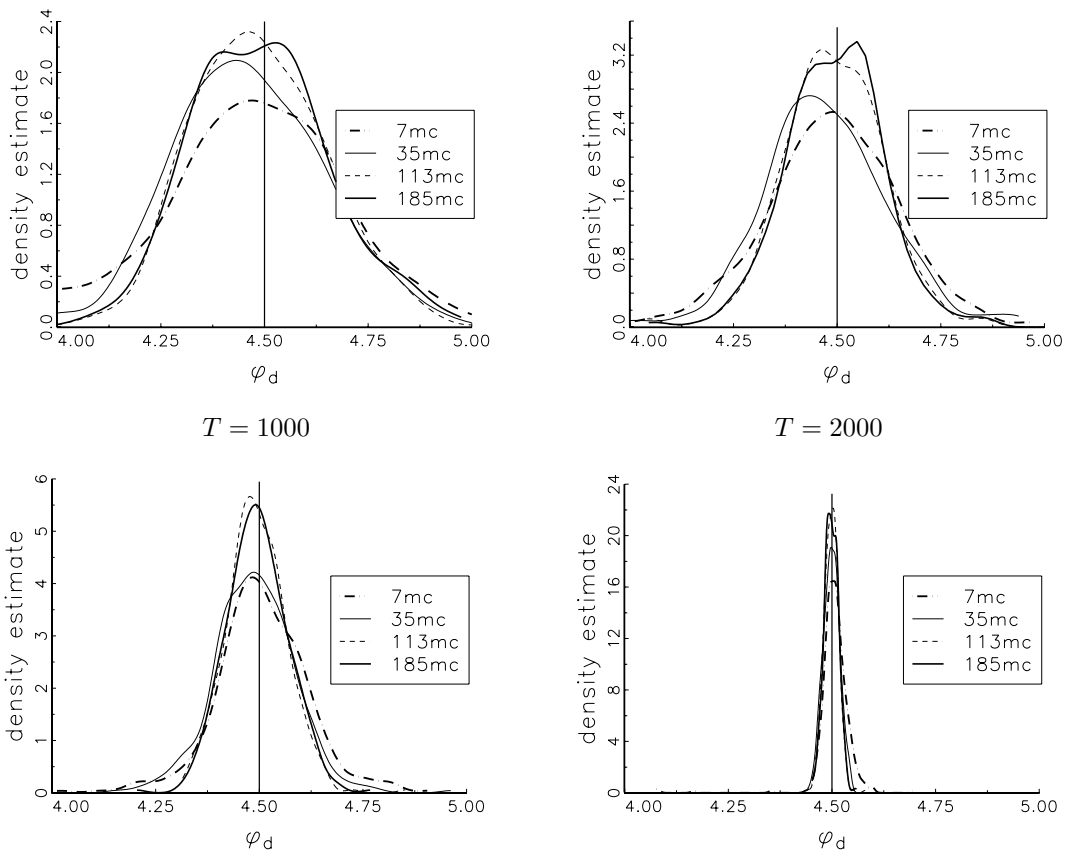


$T = 1000$ $T = 2000$

$T = 5000$ $T = 100000$

Figure 9: Kernel densities for $\hat{\varphi}_d$

The figure displays the kernel densities for $\hat{\varphi}_d$ resulting from different moment sets. The vertical lines indicate the position of the true parameter. Only $\mathbb{E}(g_{d,t}^2)$, which is matched in all moment sets, contains information for the estimate $\hat{\varphi}_d$. Therefore, the estimation precision of $\hat{\varphi}_d$ varies little across moment sets.



$T = 1000$ $T = 2000$

$T = 5000$ $T = 100000$

Figure 10: Kernel densities for $\hat{\delta}$

The figure displays the kernel densities for $\hat{\delta}$ resulting from different moment sets, using the true macro parameters. The vertical lines indicate the position of the true parameter. The theory-motivated moments are superior to the ad-hoc moments when it comes to estimating the subjective discount factor δ.

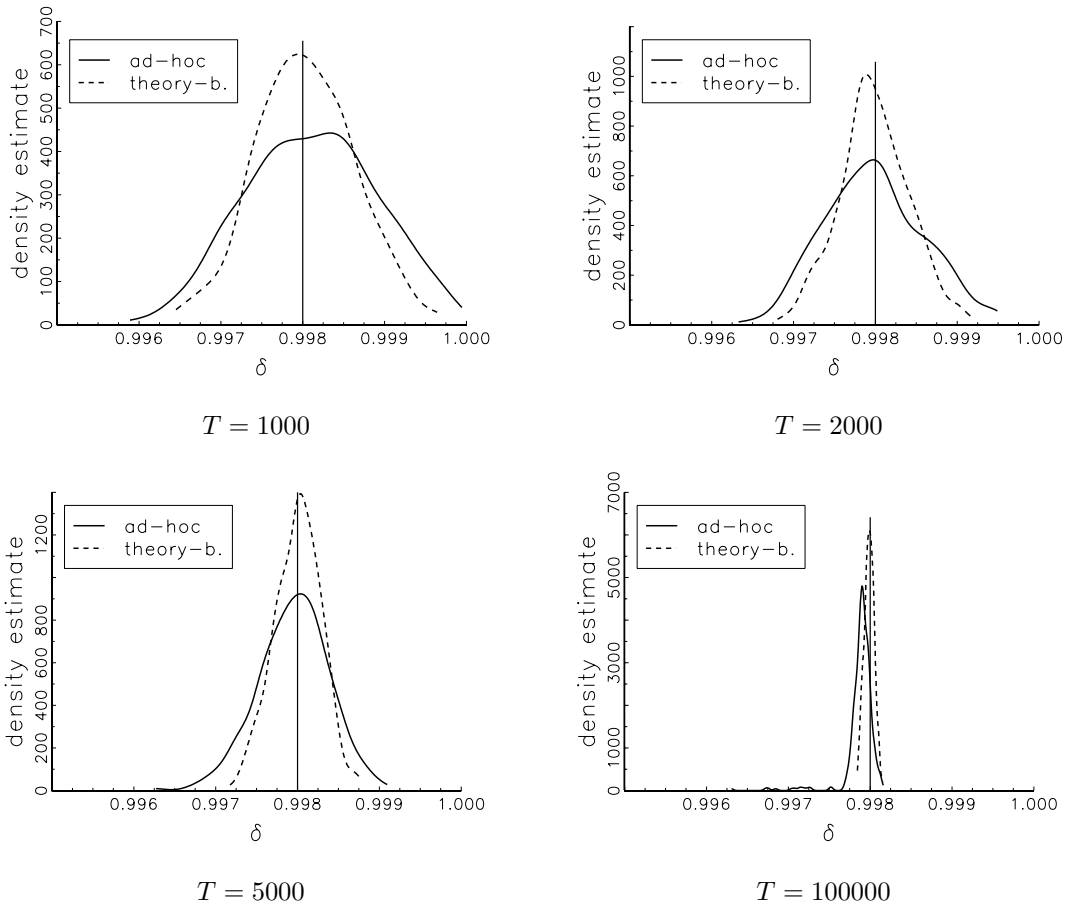


Figure 11: Kernel densities for $\hat{\gamma}$

The figure displays the kernel densities for $\hat{\gamma}$ resulting from different moment sets, using the true macro parameters. The vertical lines indicate the position of the true parameter. The theory-motivated moments are superior to the ad-hoc moments when it comes to estimating the risk aversion γ.

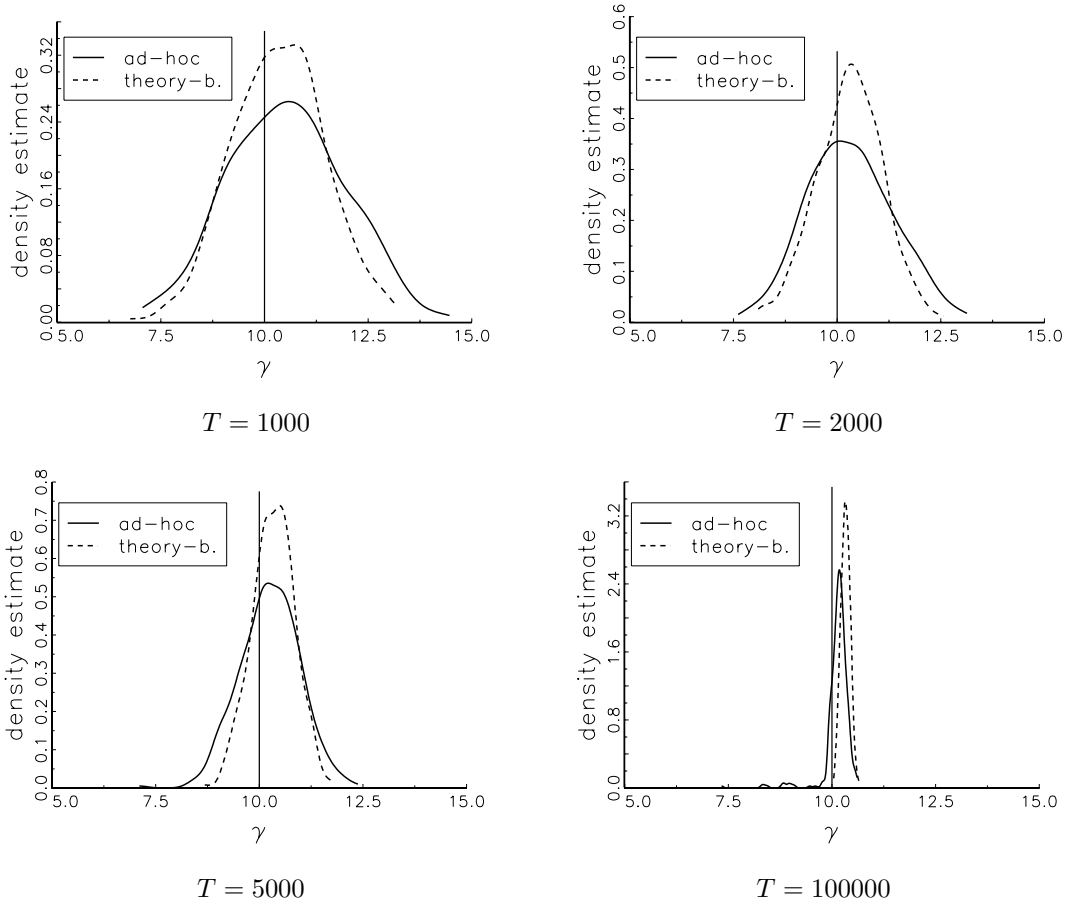


Figure 12: Kernel densities for $\hat{\psi}$

The figure displays the kernel densities for $\hat{\psi}$ resulting from different moment sets, using the true macro parameters. The vertical lines indicate the position of the true parameter. The theory-motivated moments are clearly superior to the ad-hoc moments when it comes to estimating the intertemporal elasticity of substitution ψ. When using the ad-hoc moments, ψ only seems to be weakly identified.