Econometric Analysis of Long-Run Risk in Empirical Asset Pricing

Dissertation zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von
Eva-Maria Küchlin (geb. Schaub)
aus Lohr am Main

Tübingen
2016

Tag der mündlichen Prüfung:	24.04.2017
Dekan:	Professor Dr. rer. soc. Josef Schmid
1. Gutachter:	Professor Dr. rer. pol. Joachim Grammig
2. Gutachter:	Professor Dr.-Ing. Rainer Schöbel

Preface

"Do not follow where the path may lead.
Go instead where there is no path
and leave a trail."
Ralph Waldo Emerson

Research is a journey into the unknown. To discover new scientific insights, the traveler must go off the beaten path and into the rough terrain that holds the promise of new discovery. The path is certainly nonlinear, often paved with obstacles, occasionally lengthy, and always challenging. Yet sometimes, when the wayfarer has finally cut a swath through the thicket, all he finds is a hint indicating where to go next. The rough terrain I try to negotiate in this work is the long-run risk asset pricing model, which has a complex, nonlinear structure, and is inaccessible by means of standard econometric methods. My path passes through the maze of model equations, leads past previous econometric approaches, constantly tries to keep a safe distance from the model's precipice, and finally leads to the discovery of viable estimation strategies. I hope to leave a small trail for future travelers that can serve as a beginning of new journeys through the fascinating world of consumption-based asset pricing.

My thesis could not have been completed without the support of many. I would like to take this opportunity to express my gratitude to all of them. First of all, I am

grateful to my supervisor Joachim Grammig for his guidance and constant support. Working on a joint research project with him was a great experience that helped me to find my way in the academic world and I would like to thank him for his encouragement and advice. I am also much obliged to my co-advisor Rainer Schöbel for constructive discussions and his interest in my research. I would like to thank him for kindly agreeing to be the second referee of my thesis.

This work has greatly benefited from helpful comments and suggestions of seminar and conference participants at various venues, amongst others the Econometric Society European Meetings in 2014 and 2016, and the Conference on Indirect Estimation Methods in Finance and Econometrics in Constance, 2014. In particular, I would like to acknowledge valuable comments by Ron Gallant, Roxana Halbleib, Enrique Sentana, and George Tauchen.

I would like to thank my present and former colleagues Johannes Bleher, Thomas Dimpfl, Tobias Langen, Franziska Peter, and Jantje Sönksen for sharing thoughts and many coffee breaks, and Sylvia Bürger for kindly managing all administrative matters. Thomas Dimpfl deserves special thanks for the administration of the IT infrastructure, which was indispensable for my research. I also gratefully acknowledge the use of the high-performance computing resources provided by the bwHPC initiative of the state of Baden-Württemberg, Germany.

I would like to thank my family for their unconditional support and confidence. I am much obliged to my father-in-law for indefatigable proofreading throughout my academic career. Finally, I am immensely indebted to my wonderful husband, whose optimism and unfailing encouragement has been an invaluable source of strength to me. Without him, I would neither have started nor finished this work.

Contents

Chapter 1

Introduction

Research in financial economics has endeavored to explain asset pricing puzzles for decades. Most efforts are dedicated to the equity premium puzzle, a term coined by Mehra and Prescott (1985) for the lack of a theoretical explanation for the extraordinarily high risk premium paid by risky assets in the postwar U.S. financial market. A popular theoretical approach that promises to resolve this and other asset pricing puzzles is the long-run risk (LRR) asset pricing model proposed by Bansal and Yaron (2004), a model that is intricate in nature and thus challenging to analyze with econometric techniques. This study is concerned with the econometric analysis of the LRR model, encompassing obstacles to the estimation, identification issues, and an empirical evaluation. For that purpose, different econometric methods are applied to the theoretical model, including the generalized method of moments (GMM), the simulated method of moments (SMM), indirect inference estimation, and maximum likelihood (ML) estimation that relies on filtering techniques.

This study is based on three separate working papers concerned with the estimation of the long-run risk asset pricing model. In the first paper, entitled "Give me strong moments and time: Combining GMM and SMM to estimate the long-run risk

asset pricing model," Joachim Grammig and myself suggest a two-step GMM/SMM approach to estimate the LRR model (cf. Grammig and Küchlin, 2016a). The second paper, "Estimating the long-run risk asset pricing model with a two-step indirect inference approach," also joint work with Joachim Grammig, presents an indirect inference estimation strategy that is more parsimonious and allows to estimate the model parameters at a frequency higher than that of the data (cf. Grammig and Küchlin, 2016b). In a third paper, "Filtering methods for the estimation of the long-run risk asset pricing model," I suggest a maximum likelihood estimation approach that promises efficiency gains and finally allows to estimate the full set of LRR model parameters (cf. Küchlin, 2016). This thesis presents all studies in a unified manner. Derivations and additional results collected in Web appendices to the aforementioned papers are included to provide further details.

Chapter 2 reviews the related literature and describes the LRR model in detail. First, the macroeconomic part of the model is introduced, which is driven by two latent variables that emerge as the key sources of risk in the economy; subsequently, the asset pricing implications resulting from the macroeconomic variables are examined, thereby highlighting the recursive LRR model structure. The solvability of the model for its endogenous parameters is addressed, and a calibration provides intuition for the role of each parameter in the model. Appendix A collects the analytical derivations of various LRR model components.

In Chapter 3, identification issues implied by the LRR model are revealed by scrutinizing estimation strategies presented in the literature. The insights gained from this analysis warrant the conclusion that the estimation strategy should be consistent with the recursive model structure, implying a two-step approach that reflects the dependencies between the economic processes. In light of these findings, we suggest a moment-based two-step estimation strategy that exploits analytical

moments where possible and simulated moments where necessary. The availability of analytical moment expressions permits the use of GMM in the first estimation step, whereas the model endogeneity precludes closed-form expressions for the moments of the financial variables, thus calling for the use of simulated moments. The two-step GMM/SMM estimation strategy thus combines the advantages of both methods regarding computational cost and feasibility. A Monte Carlo study and an empirical study using quarterly U.S. data illustrate the validity and precision, as well as the limits, of the estimation approach. A key finding of this study is that the precise estimation of the long-run risk component in the LRR model requires the inclusion of a large number of auto-moments in the estimation. This is an issue for any empirical application that relies on a rather limited data set.

A more parsimonious estimation approach is developed in Chapter 4. Adhering to the concept of two-step estimation, an indirect inference estimation strategy is suggested: in each step, tailor-made auxiliary models are used to consecutively estimate the parameters that determine the macroeconomy and the financial market, where the auxiliary models are designed to capture the salient features of the respective model part. In contrast to the estimation strategy of Chapter 3, the two-step indirect inference approach is entirely simulation-based and thus allows for more flexibility regarding the frequencies of the model and the data. As a result, the model can be estimated on a monthly basis from quarterly data in the empirical application, which allows to emulate an economically plausible decision frequency of the representative investor. Both the Monte Carlo study and the empirical application to quarterly U.S. data corroborate that the estimation precision is low, given the currently available scope of data.

Compared to moment-based estimation methods, maximum likelihood estimation is typically more efficient, as it takes the complete distribution of the model

variables into account, as opposed to focusing only on isolated properties of the distribution, such as moments. Chapter 5 introduces a maximum likelihood-based estimation approach that aims to improve estimation precision. The use of filtering methods permits the application of maximum likelihood, despite the presence of latent variables. The proposed three-step method allows estimation of the full set of LRR model parameters and thus overcomes the lack of identification of the parameters that characterize the fluctuating economic uncertainty, an issue that could not be resolved by either of the estimation strategies presented in Chapters 3 and 4. A Monte Carlo study demonstrates the efficiency gains and establishes the viability of the suggested method. Subsequently, an empirical application is conducted on monthly U.S. data, which provides evidence for a rather risk-averse investor, even though long-run risk is accounted for.

The main results of all studies are reviewed and summarized in Chapter 6.

Chapter 2

Asset pricing with long-run risk

2.1 Literature review

The beginnings of consumption-based asset pricing are founded on the Capital Asset Pricing model (CAPM) (cf. Sharpe, 1964; Lintner, 1965; Mossin, 1966), and in particular on the idea of the Intertemporal Capital Asset Pricing Model (ICAPM) by Merton (1973), which states that the expected return of an asset is determined by its covariance with the market portfolio and a set of state variables that describe the investment opportunity set. Rubinstein (1976) presents a discrete-time approach that is consistent with Merton's ICAPM, while Breeden (1979) generalizes the ICAPM concept in continuous time by replacing the multiple betas of Merton's (1973) model by a single beta that relates to the return's covariance with aggregate consumption. The discrete-time model published by Rubinstein (1976) and the continuous-time model by Breeden (1979) establish the consumption-based asset pricing paradigm. An exposition of the consumption-based model is provided in the textbook by Cochrane (2005).

Although theoretically appealing, the consumption-based asset pricing model (CBM) was soon found to be incompatible with empirical data. Numerous studies produced disappointing empirical results, giving rise to several asset pricing puzzles, among others the equity premium puzzle (cf. Mehra and Prescott, 1985), and the risk-free rate puzzle (cf. Weil, 1989). A comprehensive overview of the empirical developments in the consumption-based asset pricing literature can be found in Campbell (2003) and Breeden, Litzenberger, and Jia (2014). Empirical estimations and tests of the CBM turned out to yield implausible values for the investor's preference parameters, in particular for the risk aversion parameter (cf. Cochrane, 1996, who reports relative risk aversion estimates above 100). Campbell and Cochrane (2000) explain the empirical failure of the CBM with the crucial role of conditioning information, which is unavailable for empirical applications. The subsequent attempts to resurrect the consumption-based asset pricing paradigm are numerous.

One main strand of literature focuses on data-related issues, such as the suitability of the commonly used U.S. consumption data, measurement problems, the unavailability of the investor's information set, or rare disasters that could have occurred, but are not realized in the data. Building on the findings of Campbell and Cochrane (2000), Lettau and Ludvigson (2001) achieve a better empirical performance by performing a conditional estimation of the linear CBM, thereby allowing for time-varying risk premia. As a conditioning variable, they propose the log consumption-wealth ratio, which allegedly captures the investor's information set in a more comprehensive way than the previously suggested conditioning variables, namely the dividend-price ratio or the term spread. Parker and Julliard (2005) find that while contemporaneous consumption risk can only explain a small fraction of cross-sectional variation in asset prices, their measure of the so-called "ultimate risk to consumption" considerably improves the empirical performance of

the consumption-based model. By aggregating consumption growth over several periods, they also mitigate typical shortcomings of consumption data, such as measurement error and adjustment costs. In the same vein, Yogo (2006) also focuses on consumption data issues and suggests a model that separates consumption of durable and non-durable goods. In the broader sense, also the rare disasters literature founded by Rietz (1988) and Barro (2006, 2009) explains the equity premium puzzle with data-related issues. They argue that rare, but disastrous contractions of consumption are anticipated by consumers and thus incorporated in asset prices, even though no such disaster may have realized in the observed sample. Jagannathan and Wang (2007) discover that a consumption measure computed between the fourth quarters of each year substantially improves the empirical performance of the linear CBM, indicating that consumers might adjust their decisions rather infrequently at the end of each calendar year. Savov (2011) shows that consumption data from the National Income and Product Accounts (NIPA) are too smooth to justify large risk premia and argues that using data on garbage growth, which exhibits more variation and is more strongly related to stock returns, can solve the equity premium puzzle.

The other main stream of literature comprises structural asset pricing models that extend the CBM to larger model frameworks and thereby try to explain the notorious asset pricing puzzles. In particular, the external habit model of Campbell and Cochrane (1999) and the long-run risk model proposed by Bansal and Yaron (2004) constitute the principal competing approaches in this area. Both models allow to match asset market phenomena by using a larger number of parameters than previous models. Campbell and Cochrane (1999) include persistent habits into the model, which imply slow-moving countercyclical risk premia, a feature that serves to improve the empirical performance substantially. Drawing on the psychologi-

cal concept that well-being is typically evaluated in comparison to a reference level instead of in absolute values, this countercyclical effect is achieved by the following mechanism: an economic downturn causes consumption to shrink towards the investor's habit level, which in turn increases relative risk aversion, and thereby expected risk premia. In contrast, the long-run risk model by Bansal and Yaron (BY, 2004) focuses on the macroeconomic sources of risk that the representative investor must face. In particular, changes in growth expectations of consumption (long-run consumption risk) and in the fluctuating economic uncertainty (volatility risk) drive the decisions of the representative investor and thereby serve to explain asset pricing puzzles, such as the large equity premium. As opposed to the habit model, in which the time-varying risk premia are obtained by a variation in risk aversion, the long-run risk model involves time-varying risk. Combined with short-run consumption risk as an additional risk factor, long-run consumption risk and volatility risk are the main ingredients of the stochastic discount factor that prices all assets in the LRR model. Due to its far-reaching impact on model dynamics, the first source of risk provides the name for the long-run risk asset pricing model.

In their seminal paper, BY perform a calibration that demonstrates the ability of the LRR model to explain the equity premium. The LRR approach is theoretically appealing because the calibrated model matches numerous features of financial markets with a plausible theoretical framework based on macroeconomic risk. Therefore, the model has been extraordinarily popular and its properties have been studied in several articles. In response to a comment by Bui (2007), Bansal, Kiku, and Yaron (2007b) present a slightly modified version of the model with an alternative calibration to improve the model's forecasting implications, in particular to reduce the implausibly high predictability of consumption growth. The majority of the following studies, however, did not adopt this modification, but continued to use the

original model as proposed in 2004. Drechsler and Yaron (2011) perform another calibration of a generalized LRR model including jumps to explain the variance premium and its relationship to investor preferences. Bansal and Shaliastovich (2013) advocate the LRR approach as a solution to the bond return predictability puzzle. Bansal, Kiku, and Yaron (2012a) and Beeler and Campbell (2012) disagree on the consistency of LRR model calibrations with empirical data. The discussion includes different opinions on the slope of the yield curve, which is negative in BY's calibration but too large in absolute value according to Beeler and Campbell (2012); moreover, opinions differ with respect to predictability issues, and, in particular, regarding the size of the intertemporal elasticity of substitution (IES) parameter. While Bansal et al. (2012a) argue that the IES must be larger than 1, Beeler and Campbell (2012) point out the discrepancy between the weak response of consumption growth to changes in the risk-free rate, implying an IES clearly smaller than 1, and the strong negative effect of increasing consumption volatility on stock returns, implying an IES larger than 1. Ferson, Nallareddy, and Xie (2013) evaluate out-of-sample forecasts of a cointegrated version of the LRR model and find the performance to be superior to the stationary model.

Calibrations can provide helpful insights into the ability of the LRR model to replicate certain features of the data, however, they involve a confirmation bias. A critical appraisal of the model can only be obtained by econometric analysis. Empirical tests of the LRR model are impeded by its complex model structure, which precludes the use of standard econometric techniques. A comprehensive econometric analysis and a profound empirical evaluation is therefore the goal of the present thesis.

This research contributes to a literature that empirically assesses the LRR model. Bansal, Gallant, and Tauchen (2007a) conduct the first econometric analysis of the

LRR model, in which they also compare its empirical performance to that of Campbell and Cochrane's (1999) habit model. While they find that in many aspects both models are similarly suited to explain and replicate the stylized facts of the data, their analysis of the models' dynamics over longer horizons speaks in favor of the LRR model. For their estimation, Bansal et al. (2007a) use the efficient method of moments (EMM) to estimate a cointegrated LRR model variant. However, even using EMM, some important structural LRR model parameters, among them the intertemporal elasticity of substitution, could not be estimated and had to be calibrated instead. This notable result indicates some unresolved identification issues. Interestingly, some subsequent empirical studies that rely on less sophisticated estimation techniques do report estimates of all LRR model parameters, sometimes with remarkable precision.

Aldrich and Gallant (2011) present the first Bayesian estimation of the LRR model. Hasseltoft (2012) includes inflation in the LRR framework to model stock and bond markets jointly. He uses the simulated method of moments for the estimation of all parameters, except the subjective discount factor, which is calibrated to a value very close to 1. Constantinides and Ghosh (2011) show how to express the latent model variables as functions of observables, which in turn permits the use of the generalized method of moments. The same analytical inversion is exploited by Bansal, Kiku, and Yaron (2012b), who derive analytical expressions to aggregate the moments used in their GMM estimation, permitting an estimation of the model dynamics at a monthly frequency. In a recent paper, Calvet and Czellar (2015) estimate a simplified version of the LRR model using an exactly identifying auxiliary model within an indirect inference estimation approach. They also report estimates of all LRR model parameters, but their simplification, which greatly facilitates the model simulation, is not benign and comes at the cost of a built-in inconsistency.

The out-of-sample analysis by Ferson et al. (2013) focuses on forecasting, for which it suffices to estimate the model in a reduced form without identifying all of the structural parameters.

The empirical analysis of the LRR model is impeded by methodological and numerical intricacies. Such obstacles have not been explicitly debated in previous literature, which is surprising, because it is well known that the model structure is inherently fragile: for certain economically plausible parameter values, the model becomes unsolvable, and the estimation procedure must account for that problem. Moreover, dividends and consumption in the LRR framework are driven by a small, but persistent latent growth component and stochastic volatility (SV), which exacerbates the estimation of the structural parameters, especially when the data series are short. The estimation of SV models has preoccupied econometric research for some time, see e.g. Ruiz (1994), Gallant, Hsieh, and Tauchen (1997), Sandmann and Koopman (1998), Kim, Shepard, and Chib (1998), Andersen, Chung, and Sørensen (1999), and Jacquier, Polson, and Rossi (2002). In the LRR model, the SV process is just one component of a complex system.

The econometric analysis of the LRR model is challenging, as identification problems are not obvious in the highly nonlinear structure of the model. Those issues can easily be overlooked when an optimization algorithm converges to one of many local minima on a rugged objective function surface. In the following chapters, identification matters are discussed in depth, using as examples different estimation methods that have previously been applied to the LRR model. The recurrent theme of this thesis will be the necessity to adhere to the recursive model structure in the estimation process, which implies multi-step estimation strategies. In the subsequent section, the LRR model will be described in detail.

2.2 LRR model anatomy

To review the LRR model in its original formulation by BY, the recursive model structure is described in two consecutive sections: the first delineates the macroeconomic dynamics; the second details the asset pricing implications of the model. Having introduced the elementary components of the model, the issue of model solvability is discussed, which is of vital importance for a successful estimation of the representative investor's preference parameters. The presentation of the model highlights the intricacies of the model structure, which complicate generating simulated data, and thus the estimation by simulation-based methods. Detailed derivations of various model equations are collected in Appendices A.1–A.5.

2.2.1 Time series macro dynamics

The LRR macroeconomy is described by a nonlinear vector-autoregression with two observable variables, log consumption growth g_t and log dividend growth $g_{d,t}$, as well as two latent variables, a small and persistent growth component x_t and a stochastic variance σ_t^2:

$$g_{t+1} = \mu_c + x_t + \sigma_t \eta_{t+1}, \tag{2.1}$$

$$x_{t+1} = \rho x_t + \varphi_e \sigma_t e_{t+1}, \tag{2.2}$$

$$g_{d,t+1} = \mu_d + \phi x_t + \varphi_d \sigma_t u_{t+1}, \tag{2.3}$$

$$\sigma_{t+1}^2 = \sigma^2 + \nu_1(\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1}. \tag{2.4}$$

The i.i.d. innovations η_t, e_t, u_t, and w_t are standard normally distributed, contemporaneously uncorrelated random variables. The latent processes are assumed to be highly persistent, such that ρ and ν_1 are chosen to be close to 1 in calibration

exercises. The parameters that describe the macro dynamics of the LRR model are collected in the vector $\boldsymbol{\xi}^M = (\mu_c, \mu_d, \rho, \sigma, \varphi_e, \phi, \varphi_d, \nu_1, \sigma_w)'$.

It is important to observe that the LRR model is inherently recursive: The variables on the left-hand sides of Equations (2.1)-(2.4) are elementary components for all other (financial) model variables. When LRR model-implied data are required for simulation-based estimation, it is necessary to generate time series of g_t, x_t, $g_{d,t}$, and σ_t^2, before simulating financial variables such as asset returns and price-dividend ratios.

2.2.2 Asset pricing relations

The representative LRR investor who faces the macro dynamics in Equations (2.1)-(2.4) is assumed to have recursive preferences (cf. Epstein and Zin, 1989), as expressed by the utility function

$$U_t = \left[(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\left(\mathbb{E}_t\left(U_{t+1}^{(1-\gamma)}\right)\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}, \tag{2.5}$$

where C_t is aggregate consumption, and $\theta = \frac{(1-\gamma)}{\left(1-\frac{1}{\psi}\right)}$. The three preference parameters, collected in the vector $\boldsymbol{\xi}^P = (\delta, \gamma, \psi)'$, denote the subjective discount factor, relative risk aversion (RRA), and intertemporal elasticity of substitution, respectively. The representative investor has aggregate wealth W and maximizes utility under the budget constraint $W_{t+1} = (W_t - C_t)R_{a,t+1}$. The gross return of the aggregate wealth portfolio, R_a, constitutes a claim to aggregate consumption. From the first order condition of this optimization problem, we obtain the pricing equation for a gross asset return R_i,

$$\mathbb{E}_t\left[M_{t+1}R_{i,t+1} - 1\right] = 0, \tag{2.6}$$

where

$$M_{t+1} = \delta^{\theta} G_{t+1}^{-\frac{\theta}{\psi}} R_{a,t+1}^{-(1-\theta)} \tag{2.7}$$

is the stochastic discount factor (SDF), and G denotes gross consumption growth.

Drawing on the linear approximations suggested by Campbell and Shiller (1988), BY use the following expressions for r_a, the log return of the aggregate wealth portfolio, and r_m, the log return of the market portfolio, which constitutes a claim to the dividend stream:

$$r_{a,t+1} = \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{t+1}, \tag{2.8}$$

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1}, \tag{2.9}$$

where z is the log price-consumption ratio of the latent wealth portfolio, and z_m is the log price-dividend ratio of the observable market portfolio. Furthermore,

$$\kappa_1 = \frac{\exp(\bar{z})}{1+\exp(\bar{z})}, \qquad \kappa_{1,m} = \frac{\exp(\bar{z}_m)}{1+\exp(\bar{z}_m)}, \tag{2.10}$$

$$\kappa_0 = \ln(1+\exp(\bar{z})) - \kappa_1 \bar{z}, \text{ and } \quad \kappa_{0,m} = \ln(1+\exp(\bar{z}_m)) - \kappa_1 \bar{z}_m, \tag{2.11}$$

where $\bar{z}$ and $\bar{z}_m$ denote the means of z and z_m. The derivations of Equations (2.8)–(2.11) can be found in Appendix A.1. The latent log P/C ratio z and the observable log P/D ratio z_m are assumed to evolve as:

$$z_t = A_0 + A_1 x_t + A_2 \sigma_t^2, \tag{2.12}$$

$$z_{m,t} = A_{0,m} + A_{1,m} x_t + A_{2,m} \sigma_t^2. \tag{2.13}$$

The A-coefficients in Equations (2.12) and (2.13) must be determined by an analytical solution of the model. Pricing the gross return of the aggregate wealth portfolio using Equation (2.6), as outlined in Appendix A.2, leads to the expressions

$$A_1 = \frac{1 - \frac{1}{\psi}}{1 - \kappa_1 \rho}, \tag{2.14}$$

$$A_2 = \frac{1}{2} \frac{\left(\theta - \frac{\theta}{\psi}\right)^2 + (\theta A_1 \kappa_1 \varphi_e)^2}{\theta[1 - \kappa_1 \nu_1]}, \quad \text{and} \tag{2.15}$$

$$A_0 = \frac{1}{1 - \kappa_1} \left[\ln \delta + \left(1 - \frac{1}{\psi}\right) \mu_c + \kappa_0 + \kappa_1 A_2 \sigma^2 (1 - \nu_1) + \frac{\theta}{2} (\kappa_1 A_2 \sigma_w)^2 \right]. \tag{2.16}$$

Subsequently pricing the gross return to the market portfolio, as shown in Appendix A.3, yields

$$A_{1,m} = \frac{\phi - \frac{1}{\psi}}{1 - \kappa_{1,m} \rho}, \tag{2.17}$$

$$A_{2,m} = \frac{(1 - \theta)(1 - \kappa_1 \nu_1) A_2 + \frac{1}{2}[\lambda_{m,\eta}^2 + (\beta_{m,e} - \lambda_{m,e})^2 + \varphi_d^2]}{(1 - \kappa_{1,m} \nu_1)}, \quad \text{and} \tag{2.18}$$

$$\begin{aligned} A_{0,m} = \frac{1}{(1 - \kappa_{1,m})} \Bigg[& \theta \ln \delta - \frac{\theta}{\psi} \mu_c + (\theta - 1) \Bigg[\kappa_0 + \kappa_1 A_0 + \kappa_1 A_2 (1 - \nu_1) \sigma^2 \\ & - A_0 + \mu_c \Bigg] + \kappa_{0,m} + \kappa_{1,m} A_{2,m} \sigma^2 (1 - \nu_1) + \mu_d \\ & + \frac{1}{2} \left[(\theta - 1) \kappa_1 A_2 + \kappa_{1,m} A_{2,m} \right]^2 \sigma_w^2 \Bigg]. \end{aligned} \tag{2.19}$$

To obtain the LRR model-implied expression for the log risk-free rate, r_f, the same procedure is applied. Pricing the risk-free return using Equation (2.6) yields

$$r_{f,t} = -\theta \ln(\delta) + \frac{\theta}{\psi} [\mu_c + x_t] + (1 - \theta) \mathbb{E}_t(r_{a,t+1}) - \frac{1}{2} \mathrm{Var}_t(m_{t+1}), \tag{2.20}$$

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

$$\begin{aligned}\mathbb{E}_t(r_{a,t+1}) &= \kappa_0 + \kappa_1 \left[A_0 + A_1\rho x_t + A_2(\sigma^2 + \nu_1(\sigma_t^2 - \sigma^2))\right] \\ &\quad - A_0 - A_1 x_t - A_2\sigma_t^2 + \mu_c + x_t, \quad \text{and}\end{aligned} \tag{2.21}$$

$$\begin{aligned}\mathrm{Var}_t\left(m_{t+1}\right) &= \left(\frac{\theta}{\psi} + 1 - \theta\right)^2 \sigma_t^2 + \left[(1-\theta)\kappa_1 A_1 \varphi_e\right]^2 \sigma_t^2 \\ &\quad + \left[(1-\theta)\kappa_1 A_2\right]^2 \sigma_w^2.\end{aligned} \tag{2.22}$$

The detailed derivation is provided in Appendix A.4.

2.2.3 Model solvability

The analytical solution of the model yields expressions for the A-coefficients that depend on the model parameters in $\boldsymbol{\xi}^M$ and $\boldsymbol{\xi}^P$ (cf. Equations (2.14)–(2.19)), but also on the κ-parameters in Equations (2.10) and (2.11), which in turn depend on $\bar{z}$ and $\bar{z}_m$. As a consequence, the κ-parameters, and thus the A-coefficients, are endogenously determined.

To estimate the LRR model by simulation-based methods, model-implied series of z, z_m, r_a, and r_m must be generated. For that purpose, a numerical solution of the model is required. To that end, we determine $\bar{z}$ and $\bar{z}_m$ such that Equations (2.10)–(2.19) are fulfilled. This can be achieved by numerically solving for the means of z and z_m, such that the squared differences between the hypothesized means and the resulting model-implied means are equal to 0. The endogenous parameters are thus implied by the roots of two functions f_1 and f_2:

$$f_1(\bar{z}, \boldsymbol{\xi}^M, \boldsymbol{\xi}^P) = \left[\bar{z} - A_0(\bar{z}, \boldsymbol{\xi}^M, \boldsymbol{\xi}^P) - A_2(\bar{z}, \boldsymbol{\xi}^M, \boldsymbol{\xi}^P)\sigma^2\right]^2, \tag{2.23}$$

$$f_2(\bar{z}, \bar{z}_m, \boldsymbol{\xi}^M, \boldsymbol{\xi}^P) = \left[\bar{z}_m - A_{0,m}(\bar{z}, \bar{z}_m, \boldsymbol{\xi}^M, \boldsymbol{\xi}^P) - A_{2,m}(\bar{z}, \bar{z}_m, \boldsymbol{\xi}^M, \boldsymbol{\xi}^P)\sigma^2\right]^2. \tag{2.24}$$

The upper panels of Figure 2.1 show a plot of $f_1(\bar{z})$ and $f_2(\bar{z}_m)$ and their roots using the LRR parameter values calibrated by BY (see Table 2.1). The lower panels show that changing these parameters within a plausible range can easily yield an unsolvable model. Whether the model is solvable or not, and thus whether LRR model-implied data can be generated in the first place, entirely depends on the values of the structural parameters in $\boldsymbol{\xi}^M$ and $\boldsymbol{\xi}^P$. This fragility of the LRR model poses a challenge for any econometric analysis, for which—for both theoretical and numerical reasons—one must rely on a certain regularity of the admissible parameter space.

2.2.4 Simulation of LRR model-implied data

Simulation of LRR model-implied data is frequently performed throughout all chapters of this study, as it is required for calibration, simulation-based estimation, Monte Carlo assessment of the estimation strategies, and for bootstrap inference. For a given set of structural parameter values for $\boldsymbol{\xi}^M$ and $\boldsymbol{\xi}^P$, the first step is to simulate data for the latent macro variables σ_t^2 and x_t and the observable macro variables g and g_d; then, in a second step, time series of the financial variables z, r_a, z_m, r_m, and r_f can be obtained.

For a desired sample size S, the simulation of the macro variables involves drawing 4 independent series of standard normally distributed random variables of length $(S + L)$ to obtain series of realizations of the i.i.d. innovations η_t, e_t, u_t, and w_t in Equations (2.1)–(2.4). L is the number of observations of a "burn-in" period, which is discarded to mitigate the impact of the choice of starting values on the autoregressive processes. For all simulations, $L = 100$ is used.

When generating data for the latent processes σ_t^2 and x_t, the unconditional expectations are used as starting values for the forward-iteration of Equations (2.4)

and (2.2), i.e. $\sigma_0^2 = \sigma^2$ and $x_0 = 0$. Incidental negative values of σ_t^2 are replaced by 0. Subsequently, the series for g and g_d can be simulated using Equations (2.1) and (2.3).

Based on the simulated macro series, we can simulate data for the financial variables. For that purpose, the LRR model must be solved for the endogenous means $\bar{z}$ and $\bar{z}_m$, such that Equations (2.10)–(2.19) are fulfilled. The means $\bar{z}$ and $\bar{z}_m$ can then be used to obtain the values of the κ- and A-parameters.

Numerically solving the equation $f_1(\bar{z}, \boldsymbol{\xi}^M, \boldsymbol{\xi}^P) = 0$ for the mean of the log P/C ratio ($\bar{z}$) yields values for κ_1 and κ_0, as well as A_1, A_2, and A_0, which are computed in this order. The observations for z and r_a are then obtained by using Equations (2.12) and (2.8). Using the results from the solution for $\bar{z}$, the second part of the model solution $f_2(\bar{z}, \bar{z}_m, \boldsymbol{\xi}^M, \boldsymbol{\xi}^P) = 0$ can be performed at this point to obtain the mean of the log P/D ratio ($\bar{z}_m$), and thereby the values for the endogenous parameters $\kappa_{1,m}$ and $\kappa_{0,m}$, as well as $A_{1,m}$, $A_{2,m}$, and $A_{0,m}$. Having solved the entire model, the time series of z_m and r_m can be computed using Equations (2.13) and (2.9). Finally, a series of LRR model-implied log risk-free rates r_f is obtained from Equation (2.20).

2.3 LRR model calibration and implications

The first calibration of the LRR model in its original form, as presented in Section 2.2, was performed by BY. Their choice of parameter values is listed in Table 2.1. They calibrate the model on a monthly basis, thereby assuming a monthly decision frequency of the representative investor. Before an econometric analysis can be attempted, it is instructive to gain insights into the LRR model structure and the role of each parameter in the system, as the parameter values determine the model's ability to reproduce the stylized facts of financial market data.

The unconditional means of log consumption and dividend growth are specified by μ_c=μ_d=0.0015 on a monthly level, implying annual growth rates of 1.8%. The growth expectations for consumption and dividends are governed by the latent growth component x_t, which enters the time-t conditional expectation of consumption growth, $\mu_c + x_t$, and dividend growth, $\mu_d + \phi x_t$, respectively. Since its autoregressive parameter is calibrated to ρ=0.979, the latent growth component is assumed to be highly persistent, implying persistent growth expectations for the macroeconomy. The time-t conditional variances of the growth processes and the latent growth component are uniformly driven by the stochastic variance process σ_t^2, which has an unconditional mean of σ^2=0.0078^2, an autoregressive parameter ν_1=0.987, and a volatility parameter σ_w=2.3e-06. Thus, the fluctuating economic uncertainty represented by this process is assumed to be highly persistent with a rather low volatility. Consequently, the economy tends to remain in its current state of volatility, whether it is high in a crisis period or low in moderate economic conditions. By scaling the size of the innovations to the latent growth component by φ_e=0.044, while scaling the innovations to consumption and dividend growth by a factor of 1 and φ_d=4.5, respectively, the predictable part of consumption and dividend growth is kept small. The discrepancy between the scaling parameters for shocks to consumption and to dividend growth implies a considerably more volatile growth process for dividends as compared to consumption. In the same vein, the leverage parameter ϕ=3 translates positive (negative) growth expectations for consumption to even larger (worse) growth expectations for dividends. For illustration purposes, a simulated set of macro data is displayed in Figure 2.2. The parameters correspond to the BY calibration and the sample size equals S=10^3.

The LRR investor prefers present to future consumption by a subjective discount factor of δ=0.998. Risk aversion and intertemporal elasticity of substitution are

disentangled by using the utility function suggested by Epstein and Zin (1989) in Equation (2.5). Since the risk aversion parameter is chosen to be γ=10 and the IES parameter is calibrated to ψ=1.5, the utility is clearly distinct from a time-separable power utility function, which would imply equality of the risk aversion and the reciprocal of the IES parameter (for a detailed discussion of the relationship between risk aversion and intertemporal elasticity of substitution in asset pricing, see e.g. Campbell, 1993). Figure 2.3 shows a simulated set of financial data resulting from the BY calibration and from the macro series in Figure 2.2.

To simulate the financial data series, the model is numerically solved for its endogenous parameters. The BY calibration implies the following model solution: the mean of the log price-consumption ratio is given by $\bar{z} = 6.24$, which entails (in the order of computability) κ_1=0.9981, κ_0=0.0141, A_1=14.55, A_2=-470.27, and A_0=6.27. The endogenous mean of the log price-dividend ratio is obtained as $\bar{z}_m = 5.49$, which implies $\kappa_{1,m}$=0.9959, $\kappa_{0,m}$=0.0267, $A_{1,m}$=93.22, $A_{2,m}$=-2397.8, and $A_{0,m}$=5.63. The signs of the A-coefficients have important implications for the relationships between sources of risk and risk premia.

Expected returns conditional on time-t information in the LRR model are inversely related to the asset return's conditional covariance with the stochastic discount factor, or equivalently, with the SDF's innovations in excess of its time-t conditional expectation. As derived in Appendix A.5, the time-t expected risk premium for asset i is given by

$$-\mathrm{Cov}_t\left[m_{t+1} - \mathbb{E}_t(m_{t+1}), r_{i,t+1} - \mathbb{E}_t(r_{i,t+1})\right] - \frac{1}{2}\mathrm{Var}_t(r_{i,t+1}). \tag{2.25}$$

Leaving aside the asset-specific variance term, the components of the covariance reveal three macroeconomic sources of risk that are priced in the LRR model. They can be deduced from the expression of the log SDF innovations

$$m_{t+1} - \mathbb{E}_t(m_{t+1}) = \lambda_{m,\eta}\sigma_t\eta_{t+1} - \lambda_{m,e}\sigma_t e_{t+1} - \lambda_{m,w}\sigma_w w_{t+1}, \tag{2.26}$$

as derived in Equation (A-14): long-run consumption risk represented by e_{t+1}, short-run consumption risk due to η_{t+1}, and volatility risk related to w_{t+1}. Given the calibration and the model solution, we can infer the signs of the coefficients associated with the sources of risk.

As $\lambda_{m,\eta}$ and $\lambda_{m,w}$ are negative and $\lambda_{m,e}$ is positive, while the stochastic volatility σ_t can safely be assumed to be positive, a positive covariance with shocks to consumption growth or long-run growth expectations (η_{t+1} or e_{t+1}) bears a positive risk premium, while assets with a positive covariance with volatility risk w_{t+1} carry a negative risk premium. Thus, the BY calibration has plausible implications regarding the risk compensation scheme: assets that tend to have low returns in states of the economy in which growth or growth expectations are low, or in which the volatility is high, must pay a higher risk premium than assets with opposite properties.

BY emphasize the importance of the IES parameter ψ being larger than 1. Since there is an ongoing debate on this issue in the literature, this matter is worth to be assessed in depth. Important implications of the LRR model are determined by the relationship between the values of ψ and γ captured by θ. BY calibrate the risk aversion and the IES such that the resulting θ is negative. This choice ensures a plausible pricing scheme (cf. SDF in Equation (2.7)) because it establishes a negative relationship between the marginal rate of substitution and the return to

the aggregate wealth portfolio. For that matter, it is not necessary to have $\theta < 0$, but $\theta < 1$ is sufficient to maintain sensible implications of the SDF. This condition, however, is crucial for the economic implications of the model.[1] If γ does not exceed $1/\psi$, the requirement $\theta < 1$ is met. Given $\gamma = 10$, it would therefore be sufficient to restrict ψ to values larger than 0.1 to prevent an implausible SDF.

An IES larger than 1 entails that the substitution effect dominates the wealth effect. Considering Equation (2.14) shows that $\psi > 1$ is required for A_1 to be positive, as both ρ and κ_1 are close to but smaller than 1. A value smaller than 1 would imply a negative relationship between growth expectations and the log price-consumption ratio. Thus, a rise in growth expectations would prompt the representative agent to invest less into the aggregate wealth portfolio, thereby causing its price to fall. Furthermore, Equation (2.15) implies that $\theta < 0$, and thus $\psi > 1$, ensures that the coefficient A_2 takes negative values.[2] BY assert that this is necessary to match the negative correlation between consumption volatility and the log price-dividend ratio, a feature of the data. It should be mentioned, however, that a negative A_2 is rather required to obtain a negative correlation between economic uncertainty and the log price-consumption ratio. For the price-dividend ratio, it is the sign of $A_{2,m}$ that matters. Equation (2.18) shows that the sign cannot be easily determined by analytical considerations. Numerical analysis demonstrates that for the BY calibration, the IES can be lowered as far as $\psi = 0.36$ before the sign flips from negative to positive.

[1] Consider an asset that covaries positively with the SDF and thus should bear a negative risk premium. If θ was larger than 1, the SDF would be positively correlated with the return to the aggregate wealth portfolio. In turn, this would imply that we should expect assets that exhibit a positive covariance with the aggregate wealth portfolio to have a negative risk premium. This contradicts the economic basics of risk compensation, as an asset with pro-cyclical payoffs should carry a positive risk premium.

[2] A negative θ is obtained by choosing $\psi > 1$, given that $\gamma > 1$, i.e. that the investor's risk aversion is not extraordinarily small, which will be assumed throughout the following considerations.

The analytical considerations show that an IES larger than 1 is required for a negative correlation between economic uncertainty and the log price-consumption ratio and for a powerful substitution effect that dominates the wealth effect. For economic plausibility of the SDF, however, it is sufficient to ensure that $\gamma < 1/\psi$, which does not necessarily imply that the IES must be larger than 1. Neither is $\psi > 1$ necessary to attain a negative correlation between consumption volatility and the log price-dividend ratio.

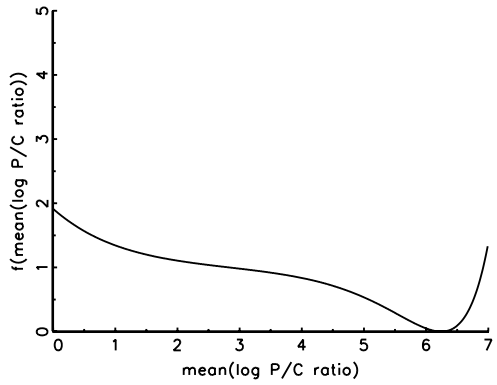
The importance of an intertemporal elasticity of substitution larger than 1 can be better understood when subjecting the BY calibration to a univariate variation in the IES parameter. However, the complex nonlinear expressions involved in the model solution preclude an analytical assessment of signs, let alone magnitudes implied by a variation in ψ. A simulation exercise with $T{=}10^5$ can help to reveal the resulting effects. Due to the model structure, a change in ψ leaves the macro variables unaffected. As the most important goal of the LRR model is to match the features of the data on the equity premium and the risk-free rate, Figure 2.4 illustrates the role of the IES in the annualized magnitudes of the equity premium, the risk-free rate, and the volatilities of the market portfolio and the riskless asset. Panel (a) shows that to obtain a sizeable equity premium, a large IES is required, which is partly due to the impact of ψ on the risk-free rate, as illustrated in Panel (b). Also, the desired low variation in the risk-free rate crucially hinges on $\psi > 1$ according to Panel (d), while the value of $\psi = 1.5$ is shown in Panel (c) to imply a market volatility similar to that observed in the data. This analysis shows that, given the remainder of the calibrated parameters, an IES larger than 1 is indispensable for the ability of the LRR model to resolve the equity premium and risk-free rate puzzle.

Table 2.1: LRR model parameter values calibrated by Bansal and Yaron (2004)

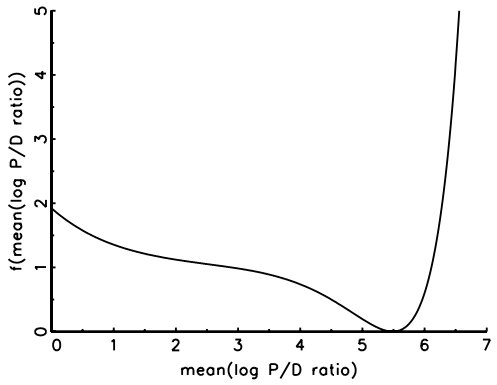
μ_c	μ_d	ρ	φ_e	ν_1	σ_w	σ	ϕ	φ_d	δ	γ	ψ
0.0015	0.0015	0.979	0.044	0.987	2.3e-06	0.0078	3	4.5	0.998	10	1.5

Figure 2.1: Existence of the solution for the endogenous LRR model parameters

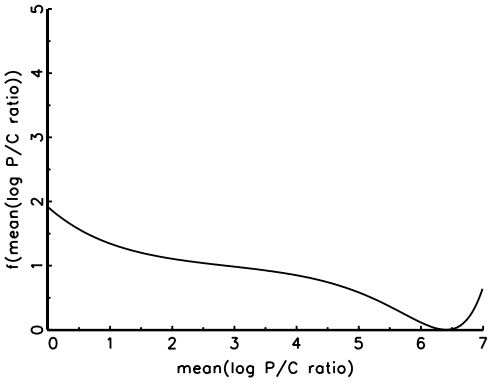
The figure displays the functions $f_1(\bar{z})$ and $f_2(\bar{z}_m)$ in Equations (2.23) and (2.24). Solving for the endogenous parameters amounts to finding the roots of f_1 and f_2. If those functions do not both have a root, the LRR model cannot be solved. The upper panels show a plot of $f_1(\bar{z})$ and $f_2(\bar{z}_m)$ based on the LRR parameter values chosen by Bansal and Yaron (2004) for their calibration of the LRR model (see Table 2.1). The lower panels show that a change of these parameters within a plausible range may yield an unsolvable model: 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.



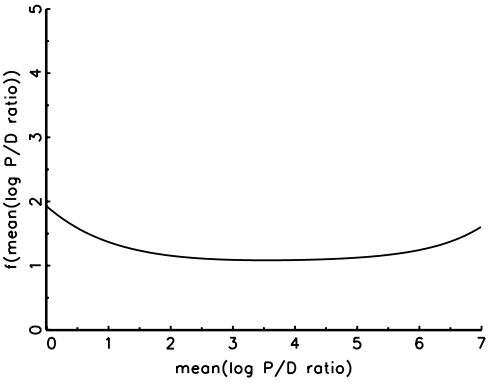
(a) root exists for $\gamma = 10$, $\mu_d = 0.0015$



(b) root exists for $\gamma = 10$, $\mu_d = 0.0015$

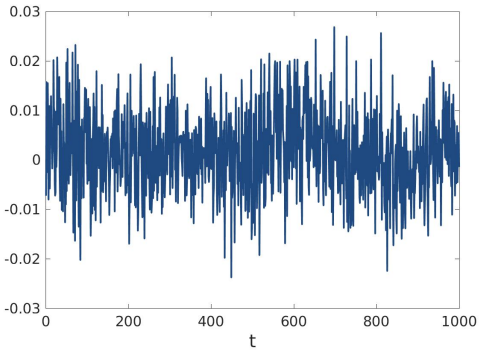


(c) root exists for $\gamma = 4$, $\mu_d = 0.0035$

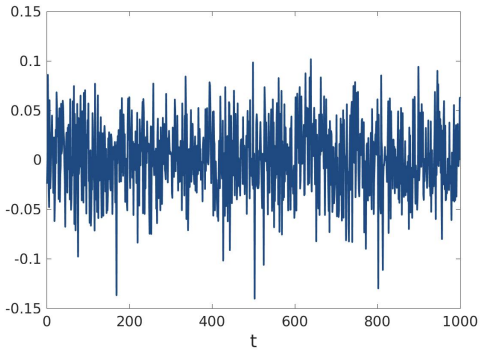


(d) no root exists for $\gamma = 4$, $\mu_d = 0.0035$

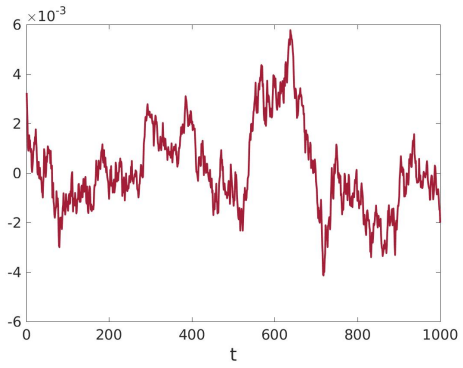
Figure 2.2: Simulated macro data series
The figure displays a set of simulated macro data series obtained from the BY calibration using a sample size of T=10^3.



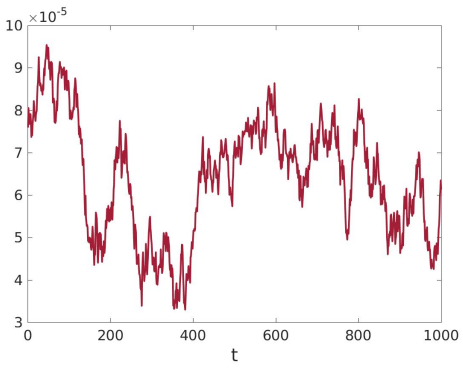
(a) log consumption growth g_t



(b) log dividend growth $g_{d,t}$

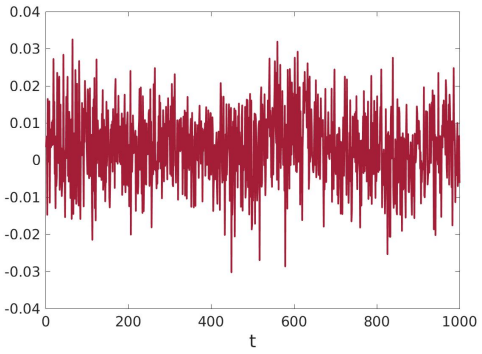


(c) latent growth component x_t

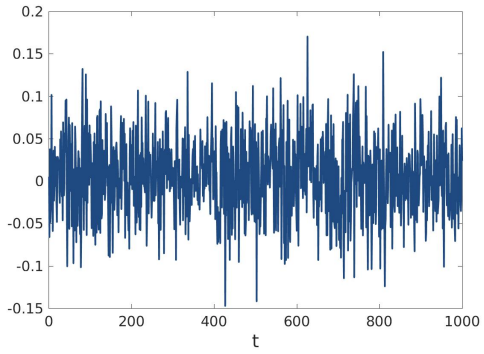


(d) latent stochastic variance σ_t^2

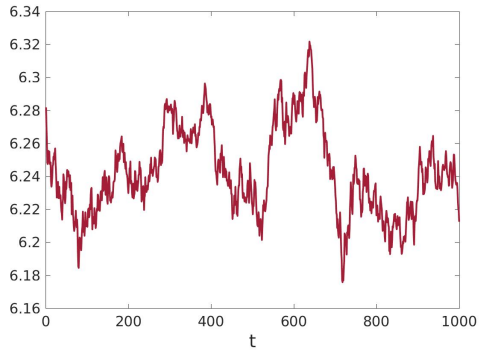
Figure 2.3: Simulated financial data series
The figure displays a set of simulated financial data series obtained from the BY calibration using a sample size of T=10^3.



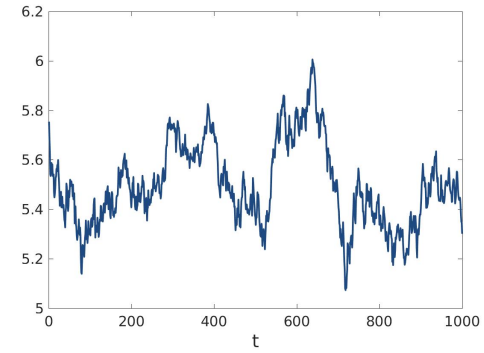
(a) log aggregate wealth return $r_{a,t}$



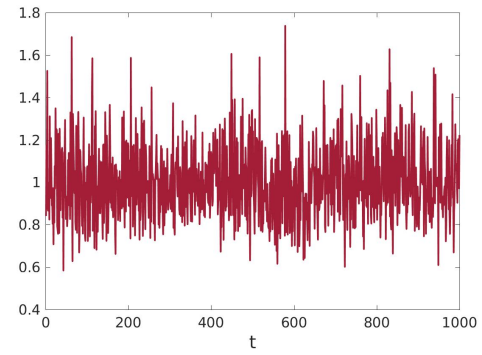
(b) log market return $r_{m,t}$



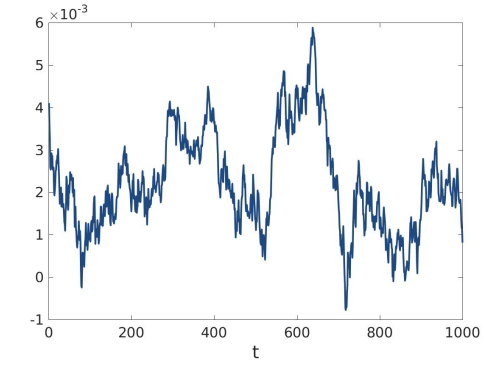
(c) log price-consumption ratio z_t



(d) log price-dividend ratio $z_{m,t}$

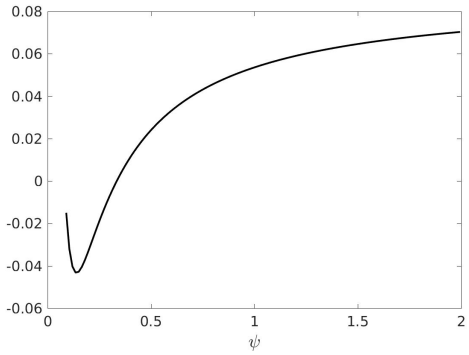


(e) stochastic discount factor M_t

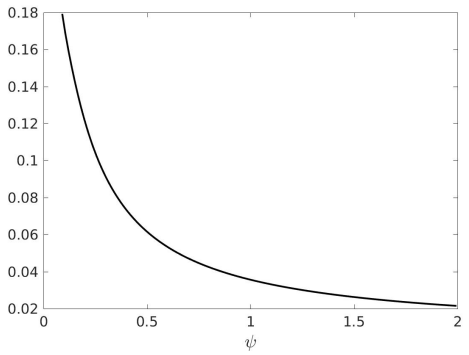


(f) log risk-free rate $r_{f,t}$

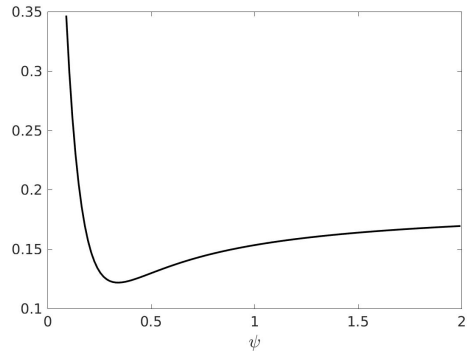
Figure 2.4: Variation in ψ
The figure displays the variation in the key stylized facts of financial market data produced by the BY calibration for different values of the IES. The sample moments are computed from a simulated data set of size of T=10^5. Model solvability is not an issue throughout the resulting parameter sets.



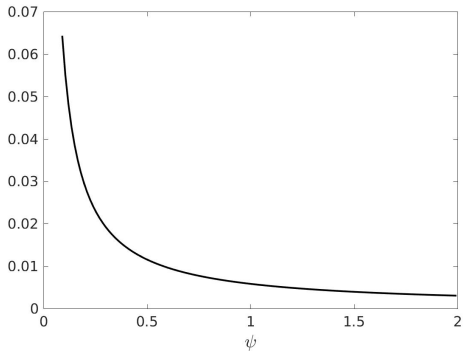
(a) $\hat{\mathbb{E}}(R_{m,t} - R_{f,t})$



(b) $\hat{\mathbb{E}}(r_{f,t})$



(c) $\hat{\sigma}(r_{m,t})$



(d) $\hat{\sigma}(r_{f,t})$

Chapter 3

A two-step GMM/SMM estimation of the long-run risk model

3.1 Introduction

The long-run risk model outlined in Chapter 2 resolves prominent puzzles of financial economics by accounting for long-run consumption risk and long-run volatility risk: shocks to growth expectations or macroeconomic uncertainty are assumed to have long-lasting effects on the economy, thus causing the investor to demand considerable compensation for holding risky assets. Empirical tests of the LRR approach are complicated by various features of the model, such as latent variables and endogenous parameters, which preclude the use of standard econometric techniques.

With this study, we show that any empirical analysis of the LRR model must overcome theoretical and econometric caveats related to model solvability and identification. To reveal the roots of the identification issues, we implement two moment-

based approaches suggested in the literature. We discuss the GMM approach suggested by Constantinides and Ghosh (2011), which relies on an analytical model inversion, and the SMM approach used by Hasseltoft (2012). A moment sensitivity analysis helps to illustrate shortcomings of the extant moment-based LRR model estimation strategies. We propose a two-step, generalized/simulated method of moments estimation strategy that exploits the recursive LRR model structure to disentangle the moment conditions associated with the macroeconomic and financial system variables. In each step, we motivate theory-based moment matches derived from the equilibrium conditions for the market return and risk-free rate and the LRR model-implied time series properties of consumption and dividend growth. With a Monte Carlo study and an empirical application, we explore the feasibility and estimation precision of a reliable econometric analysis of the long-run risk asset pricing model.

We argue that estimating the LRR parameters in one step by using an ad hoc choice of first and second moment matches does not constitute a sound econometric analysis of the LRR model. Identification problems are not obvious in the highly nonlinear model structure, and it might go unnoticed that even sophisticated optimizers converge to a local minimum on the rugged objective function surface. We provide evidence that the identification of the deep LRR model parameters, and thus the ability to produce reliable estimation results, hinges on carefully thought-out moment matches that must reflect the recursive LRR model structure. We advocate a two-step estimation approach, in which we estimate the parameters associated with the macroeconomic environment of the LRR model separately from the representative investor's preference parameters. The first step consists of a GMM estimation that uses moment conditions derived from the LRR macro dynamics; the second step is an SMM estimation that exploits the asset pricing and predictive

relationships implied by the LRR model. We show that the precision of the macro parameter estimates is of crucial importance for the successful estimation of the preference parameters. An exhaustive Monte Carlo study documents the performance of our proposed two-step estimation strategy, which is then applied to empirical data. Our findings constitute a call for econometric due diligence, reality checks, and some degree of modesty when estimating a complex dynamic asset pricing model like the LRR model. The available low-frequency macro time series are short, such that the estimation precision for some model parameters will inevitably be limited, emphasizing even more the need for informative moment matches.

One of the advantages of our theory-based identification strategy is that we can contrast the empirical results with the theoretical implications of the LRR model, and thereby assess their validity. We find that Andrews' (1999) moment selection criterion indicates the usefulness of precisely those moment matches that should be informative from a theoretical perspective, which can be regarded as implicit support for the LRR model. Moreover, the economically plausible and precise second-step estimate of the subjective discount factor indicates that the LRR model can help to resolve the interest rate puzzle. The second-step estimate of the intertemporal elasticity of substitution (IES) is greater than 1, which corroborates the long-run risk perspective on asset pricing. However, we also estimate a large coefficient of relative risk aversion, which suggests that Campbell and Cochrane's (1999) caveat that high risk aversion may be unavoidable in the class of identical-agent models also applies to the LRR model.

The remainder of the chapter is organized as follows: Section 3.2 details our two-step methodology. In Section 3.3, we present the results of a Monte Carlo study that assesses the suitability of our approach, before discussing the empirical results in Section 3.4. We conclude in Section 3.5.

3.2 Econometric methodology

3.2.1 Matching moments for GMM/SMM estimation of the LRR model

The presence of two latent processes, the highly nonlinear expressions for the equilibrium conditions for asset prices, along with the need to solve the model whenever evaluated at structural parameter values chosen from a fragmentary admissible parameter space preclude the use of standard econometric methods to analyze the LRR model. Singleton (2006) advocates the simulated method of moments, arguing that it is well suited for dealing with the complexity-driving features of the LRR model.

Adopting Singleton's (2006) notation, we define an m-dimensional observation function $\boldsymbol{g}(\boldsymbol{q}_t;\boldsymbol{\xi})$, where the p-dimensional vector $\boldsymbol{\xi} = \left(\boldsymbol{\xi}^{M'}, \boldsymbol{\xi}^{P'}\right)'$ collects the model parameters, and where $\boldsymbol{q}_t$ contains macroeconomic and financial model variables. In the present application, the observation function can consist of powers of consumption and dividend growth, market equity premium, risk-free rate, model-implied pricing errors, and so on. Matching sample moments of the observed series $\boldsymbol{g}_t^* \equiv \boldsymbol{g}(\boldsymbol{q}_t;\boldsymbol{\xi}_0)$, where $\boldsymbol{\xi}_0$ denotes the true parameter vector, with population moments yields:

$$\boldsymbol{G}_T(\boldsymbol{\xi}) = \mathbb{E}_T(\boldsymbol{g}_t^*) - \mathbb{E}\left[\boldsymbol{g}(\boldsymbol{q}_t;\boldsymbol{\xi})\right]. \tag{3.1}$$

We use Hansen's (1982) notation, $\mathbb{E}_T(\cdot) \equiv \frac{1}{T}\sum_1^T(\cdot)$, where T denotes the sample size. We resort to SMM if the population moments cannot be expressed analytically as functions of $\boldsymbol{\xi}$, yet can be simulated. Then,

$$\boldsymbol{G}_T(\boldsymbol{\xi}) = \mathbb{E}_T(\boldsymbol{g}_t^*) - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} \boldsymbol{g}\left(\boldsymbol{q}_s;\boldsymbol{\xi}\right), \tag{3.2}$$

where $\mathcal{T}(T)$ denotes the size of the simulated sample after discarding an initial portion that is left out to mitigate the transient effects of the initial conditions. To obtain $\boldsymbol{q}_s$ for $s = 1, \ldots, \mathcal{T}(T)$, we simulate LRR model-implied data following the blueprint outlined in Section 2.2.4. Then, GMM estimates, using Equation (3.1), or SMM estimates, using Equation (3.2), are obtained from

$$\hat{\boldsymbol{\xi}}_T = \underset{\boldsymbol{\xi} \in \boldsymbol{\Theta}}{\operatorname{argmin}} \, \boldsymbol{G}_T(\boldsymbol{\xi})' \, \boldsymbol{W}_T \, \boldsymbol{G}_T(\boldsymbol{\xi}) \equiv \underset{\boldsymbol{\xi} \in \boldsymbol{\Theta}}{\operatorname{argmin}} \, Q_T(\boldsymbol{\xi}), \tag{3.3}$$

where $\boldsymbol{W}_T$ is a symmetric and positive semi-definite distance matrix and $\boldsymbol{\Theta} \subset \mathbb{R}^p$ denotes the admissible parameter space.

3.2.2 Caveats

GMM and SMM are versatile tools. The theoretical conditions for consistency and asymptotic normality have been rigorously researched,[1] the range of applications is wide, and simulation-based moment matching facilitates an empirical analysis when standard econometric methods fail. Yet the key question in any application is, "Which moments should be matched?" It might seem appealing to try to estimate the LRR model parameters by matching some moments of macro and financial system variables ad hoc, but it is not obvious that these moment matches would support the identification of the structural parameters. As Hall (2005) notes, failures in identification may become apparent only when the estimation is attempted.

[1] Canonical references include Hansen (1982) for GMM and Duffie and Singleton (1993) for SMM; in addition, excellent synopses are provided by Hall (2005) and Singleton (2006). Briefly, consistency requires that a uniform law of large numbers (ULLN) applies to $\mathbb{E}_T(\boldsymbol{g}(\boldsymbol{q}_t; \boldsymbol{\xi}))$, such that it converges uniformly over $\boldsymbol{\Theta}$ to $\mathbb{E}\left[\boldsymbol{g}(\boldsymbol{q}_t; \boldsymbol{\xi})\right]$. This demand ensures that $Q_T(\boldsymbol{\xi})$ converges uniformly to the limit function $Q_0(\boldsymbol{\xi}) = [\mathbb{E}(\boldsymbol{g}_t^*) - \mathbb{E}\left(\boldsymbol{g}(\boldsymbol{q}_t; \boldsymbol{\xi})\right]' \, \boldsymbol{W} \, [\mathbb{E}(\boldsymbol{g}_t^*) - \mathbb{E}\left(\boldsymbol{g}(\boldsymbol{q}_t; \boldsymbol{\xi})\right]$, where $\boldsymbol{W}$ is the probability limit of $\boldsymbol{W}_T$. Intuitively, we assume that the data are not too fat-tailed to justify the assumption that a ULLN applies. The criterion for global identification is that Q_0 is uniquely minimized (i.e., is equal to 0) at $\boldsymbol{\xi}_0$. A necessary but not sufficient condition for global identification is that the rank of $\mathbb{E}\left[\frac{\partial \boldsymbol{g}(\boldsymbol{q}_t; \boldsymbol{\xi}_0)}{\partial \boldsymbol{\xi}'}\right]$ is equal to p (local identification criterion).

We take heed of Hall's warning and conduct an experiment using simulated LRR model-implied data, and attempt to estimate the p=12 LRR parameters in $\boldsymbol{\xi}^M$ and $\boldsymbol{\xi}^P$. We use two sets of moment matches, adapted from studies by Constantinides and Ghosh (2011) and Hasseltoft (2012), respectively. The latter uses ten first and second moments of the observable macro and financial LRR model variables, along with two auto-moments, and two moments based on an empirically motivated prediction relationship between the log price-dividend ratio and the squared future shocks to consumption growth (see Panel A of Table 3.1). Constantinides and Ghosh (2011) instead use nine macro moments along with six unconditional asset pricing moments associated with the market portfolio return, the risk-free rate, and four managed portfolios (see Panel B of Table 3.1).

Estimation problems may arise due to a small and uninformative sample, but their persistence in a very large sample indicates identification failure. We therefore perform the estimations on simulated data with T=100k observations. These data are generated from an LRR model, for which we use the parameter values of Bansal and Yaron's (2004) calibration (see Table 2.1) as true values. We initially use the identity matrix for $\boldsymbol{W}_T$ and $\mathcal{T}(T)$=10^6, after dropping the first 100 values.

Previous studies hint at optimization problems with the estimation of the LRR model, which is indicated by the use of sophisticated optimization algorithms.[2] We therefore also employ an advanced optimization technique, the covariance matrix adaptation evolution strategy (CMAES) developed by Hansen and Ostermeier (2001), which is specifically designed to deal with difficult objective functions. We start each optimization of the objective function in Equation (3.3) at three different, but not very dissimilar initial values. The starting value vector $\boldsymbol{\xi}_{s_1}$ corresponds to

[2] Hasseltoft (2012) uses simulated annealing, whereas Constantinides and Ghosh (2011) apply the differential evolution algorithm. These algorithms promise to find the global minimum of a rugged objective function much better than the gradient-based methods usually employed for econometric analysis.

the true parameter vector $\boldsymbol{\xi}_0$, which we only slightly change for $\boldsymbol{\xi}_{s_2}$. The starting value vector $\boldsymbol{\xi}_{s_3}$ is somewhat further away from $\boldsymbol{\xi}_0$ but still perfectly reasonable.

Panels A and B of Table 3.2 show that when started from these different initial values, the CMAES either converges to different parameter values or cannot meet the convergence criteria within a reasonable time (3 million function evaluations). The same problem occurs using other optimization algorithms—inter alia simulated annealing, genetic algorithm, and pattern search—available in MATLAB's global optimization toolbox. An alternative distance matrix, such as an estimate of the efficient weighting matrix, does not resolve the problem either. These results raise doubts about whether the moment matches in Table 3.1 can identify all structural model parameters. Recall that Bansal et al. (2007a), who use EMM, refrain from estimating all LRR parameters and instead resort to calibrating some key model parameters, such as the IES.

3.2.3 Moment sensitivity

The LRR model structure precludes analytic identification checks, but we can provide numerical evidence. The local identification criterion requires that the rank of the sensitivity matrix $\mathbb{E}\left[\frac{\partial \boldsymbol{g}(\boldsymbol{q}_t;\boldsymbol{\xi}_0)}{\partial \boldsymbol{\xi}'}\right]$ must be equal to p. It is not possible to calculate all of these population moments analytically, but they can be simulated. Using a simulated sample size of $\mathcal{T}(T)=10^7$ and the parameter values calibrated by Bansal and Yaron (2004) for $\boldsymbol{\xi}_0$, we find the rank condition fulfilled for both sets of moment conditions in Table 3.2.

Local identification is necessary (albeit not sufficient) to ensure global identification, but the analysis of moment sensitivity is instructive beyond checking the rank condition. It reveals which moments are useful to identify which parameter. Intuitively, if none of the moments responds to a change of a model parameter, then

the selected moment matches cannot be useful to identify that parameter. In applications, the moment sensitivity may not be zero but rather might be very small, in which case the rank condition may be met, but the identifying information provided by the moment matches is weak. In contrast, if a moment responds to many parameters, it is not clear which, if any, model parameter can be identified by it.

Table 3.3 displays the percentage change of some of the macro and financial moments in Table 3.1, along with some higher-order (cross) auto-moments, in response to a 50% c.p. decrease of one of the LRR model parameters. The computation is performed at the true parameter values, for which we draw on Bansal and Yaron's calibration.

Due to the recursive LRR model structure, the macro moments must be insensitive to changes in the preference parameters. Accordingly, only the financial moment matches can help to identify the subjective discount factor δ, the RRA coefficient γ, and the IES ψ. Among the macro moments, only the fourth moments of consumption and dividend growth respond to a change in the SV parameters ν_1 and σ_w, but the sensitivity is weak. The financial moments also are insensitive to a change in these parameters. The largest response to a 50% decrease of ν_1 (σ_w) is a 4% (3%) decrease in the expected market excess return. These are small responses by a moment that is very sensitive to almost every other model parameter. These findings raise doubts whether the SV parameters can be identified by matching the moments in Table 3.3. Estimation attempts based on the moments in Table 3.1 could not identify all LRR model parameters in one step.[3]

We also observe that the financial moments are quite sensitive to the preference parameters, which only financial moments can identify. However, Table 3.3 also

[3] Parameter estimation using the moments in Table 3.1 is not only hampered by the presence of stochastic volatility, though. Repeating the estimation procedure without SV in the data generating process (i.e. setting $\nu_1 = 0$ and $\sigma_w = 0$) delivers the same result. Reliable optimization is thus infeasible, even when using sophisticated algorithms.

shows that the financial moments respond strongly to changes in the macro parameters (except the SV parameters ν_1 and σ_w). When minimizing the objective function in Equation (3.3), it is inevitable that financial moment matches interfere with the estimation of the macro parameters. We conjecture that the stark sensitivity of financial moments to the macro parameters may hamper their ability to identify the preference parameters, thus causing the aforementioned estimation problems. It is not obvious that the financial moment matches should have the additional task of promoting the identification of the macro parameters.

In the next sections, we describe a way to achieve reliable results. The key insights are that the moment matches must reflect the recursive structure of the LRR model, and the LRR model characteristics must be incorporated in informative moment matches. Our conclusion from the moment sensitivity analysis is that we should be much more considerate in choosing which moment match to use to identify which model parameter.

3.2.4 Disentangling moment matches

The recursive structure of the LRR model implies moment matches that involve the variables g and g_d only and that therefore only depend on $\boldsymbol{\xi}^M$. We denote those macro moment matches by $\boldsymbol{G}_T^M(\boldsymbol{\xi}^M)$. Other moment matches involve financial variables (e.g. market return, risk-free rate); by the LRR model design, they depend on both $\boldsymbol{\xi}^M$ and $\boldsymbol{\xi}^P$. We denote those moment matches by $\boldsymbol{G}_T^P(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)$. The optimization problem in Equation (3.3) entails setting linear combinations of $\boldsymbol{G}_T^M$ and $\boldsymbol{G}_T^P$ to 0. Using $\boldsymbol{G}_T(\boldsymbol{\xi}) = \left(\boldsymbol{G}_T^{M\prime}, \boldsymbol{G}_T^{P\prime}\right)'$ and properly partitioning the distance ma-

trix $\boldsymbol{W}_T$, we can write the first order conditions for the optimization problem in Equation (3.3) as:

$$\underbrace{\begin{bmatrix} \frac{\partial \boldsymbol{G}_T^M(\hat{\boldsymbol{\xi}}^M)'}{\partial \boldsymbol{\xi}^M} & \frac{\partial \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M,\hat{\boldsymbol{\xi}}^P)'}{\partial \boldsymbol{\xi}^M} \\ \mathbf{0} & \frac{\partial \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M,\hat{\boldsymbol{\xi}}^P)'}{\partial \boldsymbol{\xi}^P} \end{bmatrix}}_{\frac{\partial \boldsymbol{G}_T(\hat{\boldsymbol{\xi}})'}{\partial \boldsymbol{\xi}}} \times \underbrace{\begin{bmatrix} \boldsymbol{W}_T^M & \boldsymbol{W}_T^{12} \\ \boldsymbol{W}_T^{21} & \boldsymbol{W}_T^P \end{bmatrix}}_{\boldsymbol{W}_T} \times \underbrace{\begin{bmatrix} \boldsymbol{G}_T^M(\hat{\boldsymbol{\xi}}^M) \\ \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M,\hat{\boldsymbol{\xi}}^P) \end{bmatrix}}_{\boldsymbol{G}_T(\hat{\boldsymbol{\xi}})} = \mathbf{0}. \tag{3.4}$$

Equation (3.4) reveals how the estimation procedure intertwines the financial and macro moment matches, in particular that the financial moment matches $\boldsymbol{G}_T^P$ interfere with the estimation of the macro parameters $\boldsymbol{\xi}^M$. Using a distance matrix with non-zero elements off its main diagonal generates the most complex mix of moment matches, but macro and financial moment matches remain entangled even when using $\boldsymbol{W}_T = \boldsymbol{I}$. In this case, Equation (3.4) becomes:

$$\frac{\partial \boldsymbol{G}_T^M(\hat{\boldsymbol{\xi}}^M)'}{\partial \boldsymbol{\xi}^M} \boldsymbol{G}_T^M(\hat{\boldsymbol{\xi}}^M) + \frac{\partial \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M,\hat{\boldsymbol{\xi}}^P)'}{\partial \boldsymbol{\xi}^M} \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M,\hat{\boldsymbol{\xi}}^P) = \mathbf{0}, \tag{3.5}$$

$$\frac{\partial \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M,\hat{\boldsymbol{\xi}}^P)'}{\partial \boldsymbol{\xi}^P} \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M,\hat{\boldsymbol{\xi}}^P) = \mathbf{0}. \tag{3.6}$$

The analysis of moment sensitivity suggests that the entanglement of macro and financial moment matches might be the reason for the aforementioned estimation problems. Due to the recursive nature of the LRR model, only the financial moment matches $\boldsymbol{G}_T^P$ can be useful to identify the preference parameters. Yet Equations (3.4) and (3.5) show that by minimizing the GMM objective function in Equation (3.3), the financial moment matches cannot help but interfere in the estimation of the macro parameters.

These considerations lead us to conclude that macro and financial moment matches should be disentangled. By disentangling, we mean that only linear com-

binations of the macro moment matches $\boldsymbol{G}_T^M$ should be set to 0 when estimating $\boldsymbol{\xi}^M$, such that the term $\frac{\partial \boldsymbol{G}_T^{P'}}{\partial \boldsymbol{\xi}^M}\boldsymbol{G}_T^P$ should not be present in Equation (3.5). However, no positive semi-definite and symmetric matrix $\boldsymbol{W}_T$ can accomplish this task. Disentangling macro and financial moment matches is not possible when parameter estimates result from a minimization of the GMM/SMM objective function in Equation (3.3). As a consequence, the restrictions implied by this estimation procedure must be lifted.

We formalize these considerations by conceiving of the estimation procedure as a generic GMM problem. By generic GMM, we mean that the parameter estimates are obtained by setting linear combinations of the moment matches to 0, i.e. $\boldsymbol{a}_T(\boldsymbol{\xi})\boldsymbol{G}_T(\boldsymbol{\xi}) \stackrel{!}{=} 0$, but not necessarily $\boldsymbol{a}_T(\boldsymbol{\xi}) = \frac{\partial \boldsymbol{G}_T(\boldsymbol{\xi})'}{\partial \boldsymbol{\xi}}\boldsymbol{W}_T$, as in Equation (3.4). The desired disentangling of moment matches can be achieved by estimating $\boldsymbol{\xi}^M$ and $\boldsymbol{\xi}^P$ by solving:

$$\underbrace{\begin{bmatrix} \frac{\partial \boldsymbol{G}_T^M(\boldsymbol{\xi}^M)'}{\partial \boldsymbol{\xi}^M}\boldsymbol{W}_T^M & \boldsymbol{0} \\ \boldsymbol{0} & \frac{\partial \boldsymbol{G}_T^P(\boldsymbol{\xi}^M,\boldsymbol{\xi}^P)'}{\partial \boldsymbol{\xi}^P}\boldsymbol{W}_T^P \end{bmatrix}}_{\boldsymbol{a}_T(\boldsymbol{\xi})} \times \begin{bmatrix} \boldsymbol{G}_T^M(\boldsymbol{\xi}^M) \\ \boldsymbol{G}_T^P(\boldsymbol{\xi}^M,\boldsymbol{\xi}^P) \end{bmatrix} \stackrel{!}{=} \boldsymbol{0}, \tag{3.7}$$

where $\boldsymbol{W}_T^M$ and $\boldsymbol{W}_T^P$ are symmetric and positive semi-definite matrices. The resulting estimates $\hat{\boldsymbol{\xi}}^M$ thus obey:

$$\frac{\partial \boldsymbol{G}_T^M(\hat{\boldsymbol{\xi}}^M)'}{\partial \boldsymbol{\xi}^M}\, \boldsymbol{W}_T^M\, \boldsymbol{G}_T^M(\hat{\boldsymbol{\xi}}^M) = \boldsymbol{0}, \tag{3.8}$$

which corresponds to the first order conditions of the problem:

$$\hat{\boldsymbol{\xi}}^M = \underset{\boldsymbol{\xi}^M \in \boldsymbol{\Theta}^M}{\operatorname{argmin}}\, \boldsymbol{G}_T^M(\boldsymbol{\xi}^M)'\, \boldsymbol{W}_T^M\, \boldsymbol{G}_T^M(\boldsymbol{\xi}^M), \tag{3.9}$$

where $\boldsymbol{\Theta}^M$ denotes the admissible parameter space of the macro parameters. Equation (3.7) also implies that:

$$\frac{\partial \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M, \hat{\boldsymbol{\xi}}^P)'}{\partial \boldsymbol{\xi}^P} \boldsymbol{W}_T^P \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M, \hat{\boldsymbol{\xi}}^P) = \mathbf{0}, \tag{3.10}$$

which corresponds to the first order conditions of the problem:

$$\hat{\boldsymbol{\xi}}^P = \underset{\boldsymbol{\xi}^P \in \boldsymbol{\Theta}^P}{\operatorname{argmin}} \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M, \boldsymbol{\xi}^P)' \boldsymbol{W}_T^P \boldsymbol{G}_T^P(\hat{\boldsymbol{\xi}}^M, \boldsymbol{\xi}^P), \tag{3.11}$$

where $\boldsymbol{\Theta}^P$ denotes the admissible parameter space for $\boldsymbol{\xi}^P$.

LRR parameter estimates that are based on disentangled macro and financial moment matches can thus be obtained by a two-step estimation procedure. Because the procedure is equivalent to the generic GMM problem in Equation (3.7), asymptotic inference on $\boldsymbol{G}_T(\hat{\boldsymbol{\xi}})$ and $\hat{\boldsymbol{\xi}}$ applies (cf. Hansen, 1982, Theorem 3.1 and Lemma 4.1). Alternatively, we can exploit the parametric nature of the LRR model to obtain inference through a parametric bootstrap simulation. We employ a simulation-based estimation method, so the necessary ingredients are readily available. We explain the bootstrap procedure in detail in Section 3.4.

3.2.5 Macro moment matches: motivation

The consequence of the idea of disentangling macro and financial moments is that the estimation of the macro parameters must rely exclusively on moment matches that involve the two observable time series, consumption growth g and dividend growth g_d. The upside, besides providing stability, is that the moments of these variables can be represented analytically as a function of $\boldsymbol{\xi}^M$, which allows for the use of GMM instead of simulation-based estimation methods. In particular, we can

write the five moment matches that involve the first two moments and the cross moment of g and g_d as

$$\boldsymbol{G}_T^{M_1}(\boldsymbol{\xi}^{M_*}) = \begin{bmatrix} \mathbb{E}_T(g_t) - \mathbb{E}(g_t; \boldsymbol{\xi}^{M_*}) \\ \mathbb{E}_T(g_{d,t}) - \mathbb{E}(g_{d,t}; \boldsymbol{\xi}^{M_*}) \\ \mathbb{E}_T(g_t^2) - \mathbb{E}(g_t^2; \boldsymbol{\xi}^{M_*}) \\ \mathbb{E}_T(g_{d,t}^2) - \mathbb{E}(g_{d,t}^2; \boldsymbol{\xi}^{M_*}) \\ \mathbb{E}_T(g_{d,t} \cdot g_t) - \mathbb{E}(g_{d,t} \cdot g_t; \boldsymbol{\xi}^{M_*}) \end{bmatrix} = \begin{bmatrix} \mathbb{E}_T(g_t) - \mu_c \\ \mathbb{E}_T(g_{d,t}) - \mu_d \\ \mathbb{E}_T(g_t^2) - \mu_c^2 - \frac{\varphi_e^2 \sigma^2}{1-\rho^2} - \sigma^2 \\ \mathbb{E}_T(g_{d,t}^2) - \mu_d^2 - \phi^2 \frac{\varphi_e^2 \sigma^2}{1-\rho^2} - \varphi_d^2 \sigma^2 \\ \mathbb{E}_T(g_{d,t} \cdot g_t) - \mu_c \mu_d - \phi \frac{\varphi_e^2 \sigma^2}{1-\rho^2} \end{bmatrix}, \tag{3.12}$$

where $\boldsymbol{\xi}^{M_*} = (\mu_c, \mu_d, \rho, \sigma, \varphi_e, \phi, \varphi_d)'$. Moreover, we can use the following moment matches that involve auto-moments and cross auto-moments,

$$\boldsymbol{G}_T^{M_2}(\boldsymbol{\xi}^{M_*}) = \begin{bmatrix} \mathbb{E}_{T-1}(g_{t+1} \cdot g_t) - \mathbb{E}(g_{t+1} \cdot g_t; \boldsymbol{\xi}^{M_*}) \\ \vdots \\ \mathbb{E}_{T-L_1}(g_{t+L_1} \cdot g_t) - \mathbb{E}(g_{t+L_1} \cdot g_t; \boldsymbol{\xi}^{M_*}) \\ \mathbb{E}_{T-1}(g_{d,t+1} \cdot g_{d,t}) - \mathbb{E}(g_{d,t+1} \cdot g_{d,t}; \boldsymbol{\xi}^{M_*}) \\ \vdots \\ \mathbb{E}_{T-L_2}(g_{d,t+L_2} \cdot g_{d,t}) - \mathbb{E}(g_{d,t+L_2} \cdot g_{d,t}; \boldsymbol{\xi}^{M_*}) \\ \mathbb{E}_{T-1}(g_{d,t+1} \cdot g_t) - \mathbb{E}(g_{d,t+1} \cdot g_t; \boldsymbol{\xi}^{M_*}) \\ \vdots \\ \mathbb{E}_{T-L_3}(g_{d,t+L_3} \cdot g_t) - \mathbb{E}(g_{d,t+L_3} \cdot g_t; \boldsymbol{\xi}^{M_*}) \end{bmatrix} = \begin{bmatrix} \mathbb{E}_{T-1}(g_{t+1} \cdot g_t) - \mu_c^2 - \rho \frac{\varphi_e^2 \sigma^2}{1-\rho^2} \\ \vdots \\ \mathbb{E}_{T-L_1}(g_{t+L_1} \cdot g_t) - \mu_c^2 - \rho^{L_1} \frac{\varphi_e^2 \sigma^2}{1-\rho^2} \\ \mathbb{E}_{T-1}(g_{d,t+1} \cdot g_{d,t}) - \mu_d^2 - \phi^2 \rho \frac{\varphi_e^2 \sigma^2}{1-\rho^2} \\ \vdots \\ \mathbb{E}_{T-L_2}(g_{d,t+L_2} \cdot g_{d,t}) - \mu_d^2 - \phi^2 \rho^{L_2} \frac{\varphi_e^2 \sigma^2}{1-\rho^2} \\ \mathbb{E}_{T-1}(g_{d,t+1} \cdot g_t) - \mu_c \mu_d - \phi \rho \frac{\varphi_e^2 \sigma^2}{1-\rho^2} \\ \vdots \\ \mathbb{E}_{T-L_3}(g_{d,t+L_3} \cdot g_t) - \mu_c \mu_d - \phi \rho^{L_3} \frac{\varphi_e^2 \sigma^2}{1-\rho^2} \end{bmatrix}, \tag{3.13}$$

where L_1, L_2, and L_3 denote the maximum lag orders for the respective (cross) auto-moments.

The moments in Equations (3.12) and (3.13) do not depend on the SV parameters ν_1 and σ_w and thus cannot be used to identify those parameters. However, this fact also implies that the estimation of the macro parameters in $\boldsymbol{\xi}^{M_*}$ can be performed

without having to account for SV. Moment matches that depend on ν_1 and σ_w instead must involve fourth moments of consumption and dividend growth,

$$\boldsymbol{G}_T^{M_3}(\boldsymbol{\xi}^M) = \begin{bmatrix} \mathbb{E}_T(g_t^4) - \mathbb{E}\left(g_t^4;\, \boldsymbol{\xi}^M\right) \\ \mathbb{E}_T(g_{d,t}^4) - \mathbb{E}\left(g_{d,t}^4;\, \boldsymbol{\xi}^M\right) \end{bmatrix}. \tag{3.14}$$

The detailed expressions for the fourth moments of g and g_d, provided in Equations (3.27) and (3.28) in Appendix 3.A, show that the moments in Equation (3.14) depend on all the macro parameters. Accordingly, the LRR model features another layer of recursiveness in terms of moments. In the spirit of disentangling moments, it would be possible to estimate the parameters in $\boldsymbol{\xi}^{M_*}$ upfront—using the moment matches in Equations (3.12) and (3.13)—and then focus on estimating ν_1 and σ_w, by matching exactly the moments in Equation (3.14).

3.2.6 Macro moment matches: sensitivity analysis

Are the moment matches in Equations (3.12) and (3.13) informative for the identification of the macro parameters in $\boldsymbol{\xi}^{M_*}$? The analysis of moment sensitivity in Table 3.3 provides some guidance for answering this question. First, it is obvious that matching the first moments of consumption and dividend growth helps to identify μ_c and μ_d. The conspicuous sensitivity of the second moment of consumption growth to the unconditional volatility σ (and little else) indicates that the corresponding moment match will be helpful to identify σ. Moreover, the second moment of dividend growth is very responsive to φ_d and also sensitive to σ, which is arguably well identified. Therefore, the moment match invoking g_d^2 should ensure the identification of φ_d.

The moment sensitivity analysis indicates that the identification of the remaining three parameters, ρ, φ_e, and ϕ is more intricate, because both the cross moment of g and g_d and their (cross) auto-moments respond very strongly to at least two of these parameters. The three parameters are closely linked to the latent process x_t, so their identification is challenging. Consider first the autoregressive parameter ρ, which induces small, highly persistent serial correlations in both consumption and dividend growth (see Figure 3.1 for an illustration). Intuitively, the identification of ρ should be facilitated by the auto-moments of consumption and dividend growth, which constitute informative moment matches over long lags.

Table 3.3 shows that the auto-moments of consumption and dividend growth are quite responsive to ρ. We also observe that the auto-moment sensitivity toward ρ depends on the lag, which in turn should allow to identify this parameter. Furthermore, it requires lags of a relatively high order before the sensitivity of the auto-moments of consumption growth toward ρ changes notably. The sensitivity of the first and second auto-moments is virtually the same, which corroborates the notion that the information to identify ρ must come from moment matches that involve higher-order auto-moments.

The identification of φ_e and ϕ is supported by the fact that the sensitivity patterns of the (cross) auto-moments to these parameters are somewhat dissimilar: First, the auto-moments of consumption growth are unrelated to ϕ. Second, the sensitivities towards ρ, φ_e, and ϕ of the auto-moments of dividend growth on the one hand and the cross auto-moments on the other hand differ from one another. Ultimately, these diverging responses contribute to the identification of φ_e and ϕ.

3.2.7 Financial moment matches: motivation

The observable financial variables in the LRR model are the market return R_m, the risk-free rate R_f, and the log price-dividend ratio z_m. These variables represent candidates for financial moment matches. To motivate the first moment match, we use Equation (2.6) to price the risk-free rate, which yields

$$\mathbb{E}_t\left(M_{t+1}\right) = \frac{1}{R_{f,t+1}}. \tag{3.15}$$

Applying the law of total expectation (LTE) leads to the unconditional moment condition

$$\mathbb{E}\left(M_t\right) = \mu_M = \mathbb{E}\left(\frac{1}{R_{f,t}}\right). \tag{3.16}$$

Because $\mathbb{E}(M_t)$ 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, that is:

$$\boldsymbol{g}_T^{P_1}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) = \begin{bmatrix} \mathbb{E}_T\left[\frac{1}{R_{f,t}}\right] - \mu_M \\ \mu_M - \frac{1}{\mathcal{T}(T)} \sum_{s=1}^{\mathcal{T}(T)} M_s(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) \end{bmatrix}. \tag{3.17}$$

Another moment match results from pricing the market excess return $(R_m - R_f)$ using Equation (2.6), applying the LTE, and rearranging terms to obtain the following moment condition:

$$\mathbb{E}\left(R_{m,t} - R_{f,t}\right) = -\frac{\mathbb{E}\left[\left(M_t - \mu_M\right)\left(R_{m,t} - R_{f,t}\right)\right]}{\mu_M}. \tag{3.18}$$

Thus, we can use the following match of sample and simulated moments:

$$\begin{aligned} \boldsymbol{G}_T^{P_2}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) &= \mathbb{E}_T\left(R_{m,t} - R_{f,t}\right) \\ &+ \frac{\frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)}\left[R_{m,s}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) - R_{f,s}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)\right]\left[M_s(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) - \mu_M\right]}{\mu_M}. \end{aligned} \tag{3.19}$$

For a third moment match, we consider the unconditional Sharpe ratio of the market portfolio, which is a key statistic for the risk-return trade-off implied by the LRR model. The means of the market excess return and the risk-free rate are implicitly accounted for in Equations (3.17) and (3.19), so the remaining moment to be matched is the expected value of the squared market excess return:

$$\begin{aligned} \boldsymbol{G}_T^{P_3}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) &= \mathbb{E}_T\left(R_{m,t} - R_{f,t}\right)^2 \\ &- \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)}\left(R_{m,s}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) - R_{f,s}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)\right)^2. \end{aligned} \tag{3.20}$$

Our final financial moment matches are derived from a prediction relation pointed out by Campbell and Shiller (1988), who argue that the linear approximations in Equations (2.8) and (2.9) imply that the log price-dividend ratio predicts future discount rates.[4] We use this predictive relationship to match the slope parameter of

[4] A simulation experiment reveals that the predictive power of $z_{m,t}$ for $R_{f,t+1}$ is quite strong: The R^2 of a one-step predictive regression is 95%. The simulation is based on the parameter values given in Table 2.1 and a simulated sample size of 10^6 observations.

a regression of the risk-free rate on past values of the log price-dividend ratio, which also entails matching the first and second moments of $z_{m,t}$:

$$\boldsymbol{G}_T^{P_4}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) = \begin{bmatrix} \frac{\mathbb{E}_{T-1}\left[\left(R_{f,t+1} - \mathbb{E}_{T-1} R_{f,t+1}\right) z_{m,t}\right]}{\mathbb{E}_T\left(z_{m,t} - \mathbb{E}_T z_{m,t}\right)^2} \\ -\frac{\frac{1}{\mathcal{T}(T)-1}\sum_{s=1}^{\mathcal{T}(T)-1}\left[z_{m,s}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) - \mu_{z_m}\right] R_{f,s+1}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)}{\mu'_{z_m} - \mu_{z_m}^2} \\ \mathbb{E}_T(z_{m,t}) - \mu_{z_m} \\ \mathbb{E}_T(z_{m,t}^2) - \mu'_{z_m} \\ \mu_{z_m} - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} z_{m,s}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) \\ \mu'_{z_m} - \frac{1}{\mathcal{T}(T)}\sum_{s=1}^{\mathcal{T}(T)} z_{m,s}^2(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) \end{bmatrix}. \tag{3.21}$$

The stacked financial moment matches $\boldsymbol{G}_T^P = \left(\boldsymbol{G}_T^{P_1\prime}, \boldsymbol{G}_T^{P_2}, \boldsymbol{G}_T^{P_3}, \boldsymbol{G}_T^{P_4\prime}\right)'$ are then used for the SMM objective function in Equation (3.11). As pointed out by Parker and Julliard (2005), the auxiliary parameters μ_M, μ_{z_m}, and μ'_{z_m} in Equations (3.17) and (3.21) must be exactly matched.

3.2.8 Financial moment matches: sensitivity analysis

The moment sensitivities in Table 3.4 indicate which of the financial moment matches provides information about which preference parameter. All financial moments respond strongly to a 10% change in the subjective discount factor δ. For both the RRA coefficient γ and the IES ψ, one of the moments responds sizeably to a specific parameter change, whereas the sensitivity of the other moments is low. Most information about γ is contained in the LRR model's pricing implication for the market excess return, which is reflected in the 10% decrease in the simulated moment in Equation (3.19) in response to a 10% decrease in γ. The other moments are not particularly sensitive to γ. The identification of ψ mainly results from the slope parameter of the predictive regression of $R_{f,t+1}$ on $z_{m,t}$. The corresponding simulated moment responds to a 10% decrease in ψ with a 14% increase; the other

moments change by 4% or less. The prediction moment is not sensitive to a change in γ, which thus helps to disentangle risk aversion and the intertemporal elasticity of substitution.

3.2.9 Treatment of stochastic volatility

The estimation of stochastic volatility models is a topic of substantial econometric discussion. The methodological challenges are aggravated in the present context, because SV is just one ingredient of a complex dynamic asset pricing model. The analysis in Section 3.2.3 shows that there is no sensitivity of the macro moments in Table 3.1 to the SV parameters ν_1 and σ_w, and we have seen that the analytical moment matches presented in Section 3.2.5 can be used to identify the unconditional variance σ^2 but not ν_1 or σ_w. Moreover, the financial moment matches in Table 3.3 are also unresponsive to the SV parameters, and the theory-based moment matches presented in the previous section cannot be expected to do a better job: They are based on unconditional moments, whereas stochastic volatility pertains to changing conditional variances.

As mentioned previously, we could consider using the fourth moments of dividend and consumption growth to identify ν_1 and σ_w. However, the moment sensitivity is too weak to claim that a reliable estimation would be possible based on fourth moment matches. We draw this conclusion from a simulation experiment, in which we attempt to estimate the SV parameters ν_1 and σ_w, assuming the true values of all other macro parameters are known, using a very large simulated sample, and Bansal and Yaron's (2004) calibrated parameters. The estimation procedure, which amounts to exactly matching the fourth moments of g and g_d, yields wildly fluctuating estimates of ν_1 and σ_w across different simulated samples. This result

indicates that the fourth moments are not sufficiently informative to identify the SV parameters.[5]

Instead of looking for more sophisticated ways to estimate the SV parameters, we propose a simplification. If the primary goal is not the estimation of the SV parameters and the evolution of the conditional risk premium, but rather the estimation of the preference parameters and determining whether plausible estimates can explain the unconditional equity premium, an alternative estimation strategy is to concentrate out the SV parameters. By concentrating out, we mean that in a model simulation in the course of the SMM estimation, we replace the stochastic volatility σ_t^2 with its unconditional expected value, $\mathbb{E}(\sigma_t^2) = \sigma^2$. The estimation of the macro parameters $\boldsymbol{\xi}^{M_*}$ is thus performed by using $\boldsymbol{G}_T^M = \left(\boldsymbol{G}_T^{M_1\prime}, \boldsymbol{G}_T^{M_2\prime}\right)'$ in the GMM objective function in Equation (3.9), which yields an estimate of the unconditional stochastic volatility σ. We conjecture that the unconditional simulated moments of the financial variables are not greatly affected when σ_t^2 is replaced by σ^2. Concentrating out SV may reduce efficiency, yet it also could enhance robustness, because the SV parameters may be poorly identified by weak moment conditions and/or a small sample size.

3.3 Monte Carlo study

3.3.1 Design

We test the two-step estimation approach with an extensive Monte Carlo study. For that purpose, we generate LRR model-implied series of g, g_d, r_m, r_f, and z_m using

[5] Relaxing the rule to disentangle moment matches, we also used financial moment matches, such as the auto-moments of the squared market return or the fourth moments of returns, to obtain estimates of ν_1 and σ_w. However, the results did not improve. As indicated by the moment sensitivity analysis in Table 3.3, the financial moments are not very sensitive to the SV parameters either.

as true LRR parameter values those calibrated by BY. We consider lengths of the simulated series of T=1k, 5k, and 100k. The T=100k case provides a reality check whether the two-step strategy works and that the moment conditions can identify the structural model parameters. Assuming a monthly sampling frequency, T=1k represents a large but not unreasonable sample size for a real-world application. Using T=5k illustrates the behavior of the estimates for a growing sample size.

We restrict the estimates $\hat{\varphi}_e$, $\hat{\phi}$, and $\hat{\varphi}_d$ to positive values, whereas $\hat{\rho}$, $\hat{\mu}_c$, and $\hat{\mu}_d$ must take values between 0 and 1. For that purpose, we use exponential and logit transforms of the unrestricted auxiliary parameters. To attain a high level of accuracy of the simulated moments, we use $\mathcal{T}(T) = 10^6$. We use the Nelder-Mead simplex (NM) algorithm to minimize the objective functions in each step. The NM method is less sophisticated than the optimizers used in Section 3.2.2 and previous literature. However, as we shall see, the two-step estimation procedure does not require an elaborate optimization algorithm.

To assess the estimation precision, we generate 400 replications for each sample size. In Section 2 we pointed out the fragility of the LRR model, which may become unsolvable when certain parameter combinations are probed during SMM estimation. A practical solution would be a penalty term that moves the optimizer away from unfavorable parameter combinations. To economize computation time in the simulation study, we chose not to use a penalty term but instead to drop the replications for which the optimizer terminated with an unsolvable model.

In Section 3.2.2 we emphasized the hazard of reporting overly optimistic estimates that result from a false convergence to a point near the plausibly chosen starting values. Prior to engaging in a large-scale Monte Carlo study, we therefore carefully pre-tested the two-step estimation procedure and started the optimizations from different parameter values, using a variety of test data to ensure that the NM

algorithm converges to the same values. Panel C in Table 3.2 shows the results obtained using a particular set of moment matches specified in detail below. Using the same data as for the failed estimation attempts in Panels A and B, the two-step procedure yields identical estimates, irrespective of the initial values chosen.

3.3.2 Monte Carlo results: first-step estimates

We focus on four different sets of moment matches to estimate the seven macro parameters in $\boldsymbol{\xi}^{M_*}$, as described in Section 3.2.5. Table 3.5 shows that the number of moments m used for the GMM estimation ranges from exact identification (m=7) to ample over-identification (m=185).[6] All four setups include the five moment matches of Equation (3.12), and then add increasing numbers of auto-moments selected from Equation (3.13). The maximum lag order is L_1=L_2=L_3=60 (m=185), meaning that we use auto-moments up to five years, assuming a monthly frequency. The intermediate cases use L_1=L_2=L_3=10 (m=35) and L_1=L_2=L_3=36 (m=113). We obtain the first-stage GMM macro estimates by using $\boldsymbol{W}_T^M = \boldsymbol{I}_m$ in Equation (3.9). To check whether an asymptotically efficient weighting scheme is beneficial in smaller samples, we also compute second-stage GMM estimates, based on the distance matrix

$$\boldsymbol{W}_T^M = \left[\mathrm{Var}_T\left(\boldsymbol{g}_t^{M_*} - \mathbb{E}[\boldsymbol{g}^M(\boldsymbol{q}_t; \hat{\boldsymbol{\xi}}^{M_*(1)})]\right)\right]^{-1}, \tag{3.22}$$

where $\hat{\boldsymbol{\xi}}^{M_*(1)}$ is the first-stage GMM estimate, $\boldsymbol{g}^M$ is the observation function pertaining to the respective macro moment match, and $\mathrm{Var}_T(\cdot)$ denotes a sample variance-covariance matrix. Asymptotically efficient GMM estimation should use a distance matrix $\boldsymbol{W}_T = \hat{\boldsymbol{S}}^{-1} \underset{p}{\to} \boldsymbol{S}^{-1}$ in Equation (3.3), where $\boldsymbol{S} = \lim_{T\to\infty} \mathrm{Var}(T^{-1/2}\boldsymbol{G}_T(\boldsymbol{\xi}_0))$. We experimented with alternative estimators of $\boldsymbol{S}$ that account for serial correla-

[6] To produce the results in Panel C of Table 3.2, we use the m=185 variant for the first-step estimation and the theory-based financial moment matches $\mathbf{G}_T^P$ for the second-step estimation.

tion in $\boldsymbol{g}_t^{M*}$, but the estimate in Equation (3.22) delivered the best results in finite samples.

We intentionally choose starting values located at some distance from the true parameters.[7] Poor initial values make the problem harder for the optimization algorithm, and the Monte Carlo study more time-consuming, but they also prevent the threat of overly optimistic results. In light of the aforementioned results, we seek to avoid this fallacy at all costs. Any replications for which the optimization algorithm failed or that produced implausible estimates are excluded from Table 3.6 and Figure 3.2.[8] The number of successful estimations, which we report in Panel H of Table 3.6, is itself an interesting statistic, because it indicates how well the respective moment matches define the optimization problem. Table 3.6 contains the means and standard deviations of the macro parameter estimates computed across successful replications, and Figure 3.2 illustrates the results.

The T=100k results show that the GMM estimation strategy works, and that the macro moment matches can identify the macro parameters in $\boldsymbol{\xi}^{M*}$. The bias in the estimates vanishes, and the standard deviation shrinks; estimation failure is a rare event. There is a notable exception though: The bias and standard deviation of $\hat{\rho}$ and $\hat{\varphi}_e$ remain considerably large for m=7; the bias of $\hat{\phi}$ is small, but the standard deviation is not. The moment sensitivity analysis in Section 3.2.6 already has suggested that these three parameters, associated with the latent growth component x_t, may prove difficult to estimate. Note that the four sets of moment matches only differ with respect to the number of auto-moments. The m=7 variant uses

[7] The starting values are μ_c=0.018, μ_d=0.018, ρ=0.881, σ=0.082,φ_e=0.003, ϕ=7.389, and φ_d=7.389.

[8] An estimation result is considered implausible if one of the parameter values to which the NM algorithm converges differs from the true parameter by a factor of 10 or more. In an empirical application, a treatment of problematic data could use different starting values and optimization algorithms, and tune the algorithm's parameters. However, such a clinical approach is impractical in a large-scale simulation study.

just the first two auto-moments of consumption growth, and the simulation results indicate that this is not enough: auto-moment matches that involve higher lags of auto-moments are required to identify ρ, φ_e, and ϕ.

The T=5k results corroborate the benefits of exploiting the information contained in higher-order auto-moments, but now sampling error takes effect. As might be expected from the 100k results, the parameters associated with x_t prove hard to estimate, but the results can be improved using higher auto-moment matches. The improvement is most striking for the autoregressive parameter ρ. A comparison of the m=7 and m=35 results on the one hand, and the m=113 and m=185 results on the other hand, shows that estimation precision increases with the use of higher auto-moments. Figure 3.2 also illustrates the substantial advancement from m=7 and m=35 to m=113, whereas a further enhancement due to the use of m=185 is more marginal. Asymptotically efficient weighting is particularly useful to hone the estimation results for φ_e, though only in combination with higher auto-moment matches. Generally, using an asymptotically efficient distance matrix cannot replace the use of higher auto-moments.

The estimation precision is good for μ_c, σ, and φ_d, confirming the conjecture that the moment matches in Equation (3.12) should identify these parameters quite well. The mean of dividend growth μ_d proves hard to estimate, because the dividend growth series is volatile. Using auto-moments is no remedy here.

The T=1k results confirm these conclusions, although sampling error becomes more of an issue, as does the increasing number of failed estimations. Estimation precision is reduced in particular for the critical parameters ρ, φ_e, and ϕ. However, the usage of higher auto-moments again can mitigate these problems. Estimation precision improves when moving from m=7 to m=113, and the number of failed replications decreases. A comparison of m=35 with m=113 shows a substantial

improvement. The effect of increasing the number of auto-moments further, e.g. from m=113 to m=185, is less pronounced.

The T=1k results indicate that the favorably small asymptotic standard errors reported in the empirical estimations of LRR models should be taken with a grain of salt. These applications use much smaller sample sizes. The simulation results show that estimation precision with the currently available sample size must be limited. Using the information contained in higher auto-moments is beneficial, but time is a constraining factor. The available consumption and dividend time series are relatively short, creating the familiar trade-off between efficiency (allowing for a high lag order) and robustness. The improvement of estimation quality from m=35 (max. lag: <1 year) to m=113 (max. lag: 3 years) is considerable, but the incremental benefits of using m=185 (max. lag: 5 years) may be offset by picking up noise from the data.

3.3.3 Monte Carlo results: second-step estimates

Second-step SMM estimation of the preference parameters δ, γ, and ψ is based on the six theory-based financial moment matches in $\boldsymbol{G}_T^{P_1}$, $\boldsymbol{G}_T^{P_2}$, $\boldsymbol{G}_T^{P_3}$, and $\boldsymbol{G}_T^{P_4}$, along with $\boldsymbol{W}_T^P = \boldsymbol{I}$ in the SMM objective function in Equation (3.11). For comparison, we also use the six ad hoc financial moment matches in Panel A-2 of Table 3.1. The input from the first step is the vector of macro parameter estimates resulting from m=185. In each replication, we perform an initial grid search over reasonable ranges of the three preference parameters, and use the parameter combination that yields the smallest SMM objective function as starting values for the optimization.

We also investigate the hypothetical case in which the true macro parameters are available, which enables us to assess the quality of the financial moment matches independently of the effect of the potentially imprecise first-step macro estimates.

For that purpose, we assume either that all macro parameters in $\boldsymbol{\xi}^M$ are known or, alternatively, that only the subset $\boldsymbol{\xi}^{M_*}$ is known. In the first case, we can compute σ_t^2 when simulating the financial moments in the second step. If we know only $\boldsymbol{\xi}^{M_*}$, we instead use the unconditional mean σ^2 to simulate moments. A comparison of the resulting estimates then allows us to gauge the efficiency loss implied by concentrating out SV.

Panel A of Table 3.7 displays the means and standard deviations of the SMM estimates for the preference parameters that use the true macro parameters for the simulation of moments. The first column reports the results based on the ad hoc financial moment matches, and the second contains the results using the theory-based moment matches—both are obtained by concentrating out SV. The third column of Panel A shows the results for theory-based moment matches if we were to assume that all macro parameters were known. Figure 3.3 illustrates and compares the estimation precision using kernel density estimates.

The estimation quality delivered by SMM is very good. The bias and standard deviation of the preference parameter estimates are small; the density estimates center around the true parameters. The subjective discount factor δ can be estimated most precisely, but the estimates of the relative risk aversion parameter γ are also quite accurate. The estimation quality delivered by the theory-based moment matches outperforms the ad hoc moment matches, most prominently for the estimate of the IES coefficient, $\hat{\psi}$. The theory-based moment matches are thus particularly useful for disentangling risk aversion and intertemporal substitution elasticity. Comparing the second and third columns of Panel A in Table 3.7, we find that the estimation quality of the preference parameters is barely affected by concentrating out SV.

Panel B of Table 3.7 in turn displays the means and standard deviations of the SMM preference parameter estimates that use the first-stage GMM macro parameter estimates for the simulation of the financial moments. We observe that the theory-based moments again outperform the ad hoc moment matches. Moreover, the estimation precision for the subjective discount factor is not greatly impaired; the parameter standard deviations and bias remain small for T=1k too. Estimating the relative risk aversion γ and the IES ψ based on the estimated macro parameters poses a greater challenge. Compared with the estimates that use the true macro parameters, bias and standard deviation increase considerably. The kernel densities in Figure 3.4 retain their modes at the true values, but there are probability masses allocated in the right tails, which indicates that some large estimates of γ and ψ are responsible for the increase in the standard deviation and bias. These results emphasize the importance of using precise macro parameter estimates for the SMM estimation of the preference parameters.

To improve the quality of the first-step input, we consider two strategies. First, we use the second-stage instead of first-stage GMM macro estimates. Second, we raise the bar for the quality of the first-step estimates and discard those that do not fulfill these requirements. We summarize the effects of both strategies in Table 3.8 and Figure 3.4.

A comparison of Panel A of Table 3.8 with Panel B of Table 3.7 shows the benefits of using the second-stage GMM macro estimates. Estimation precision improves particularly for T=1k. The kernel plots in Figure 3.4 show that the likelihood of severe overestimation of γ and ψ also decreases. Note, however, that the major improvement of the macro parameter estimates results from using a sufficient number of auto-moment matches, rather than from applying an efficient weighting scheme.

Using an estimate of the optimal weighting matrix thus cannot cannot replace a careful choice of moment matches.

Panel B of Table 3.8 and Figure 3.4 illustrate the effect of raising the bar for the quality of first-step macro estimates before entering the second estimation step. Raising the bar means that we discard a replication if one of the macro parameter estimates is more than twice its true value. As a consequence, the quality of the second-step preference parameter estimates improves further. Figure 3.4 shows that the likelihood of outlier estimates diminishes and the kernel densities center more closely around the true parameters. Of course, this procedure is only applicable in a simulation experiment. An empirical application demands a judgment call, based on the quality of the first-step macro estimates: If the macro parameter estimates are implausible or too imprecise, then the researcher should refrain from moving on to the second estimation step.

3.4 Empirical application

3.4.1 Data

We use the data collected by Beeler and Campbell (2012) to conduct an empirical application of the two-step estimation strategy. Their data contain time series of U.S. consumption growth, the return of a market portfolio proxy with the corresponding P/D ratio and dividend growth, as well as a risk-free rate proxy. The data comprise T=247 observations at a quarterly frequency, spanning the time period 1947Q2–2008Q4. Figure 3.5 displays and describes the data.

Two issues should be taken into account when attempting to estimate the LRR model on these data. First, the calibrated LRR model parameters in our previous analyses correspond to a monthly decision frequency. However, empirical analy-

sis often must rely on data sampled at a quarterly (or lower) frequency. Second, dividend payments occur irregularly in time. The quarterly dividend growth series depicted in Panel (b) of Figure 3.5 is therefore quite erratic. It exhibits a strong, negative first-order autocorrelation that is not allowed for in Equation (2.3). Dividend growth is less volatile at an annual frequency, but in that case the number of observations is quite small. We therefore perform our analysis using the quarterly Beeler and Campbell (2012) data, and we follow Hasseltoft (2012) in taking the average of the current period's log dividend growth and that of the previous three quarters to obtain the smoothed dividend growth series in Panel (c) of Figure 3.5.

3.4.2 First-step estimation results

In Section 3.3.2 we found that the use of higher auto-moment matches is important to ensure good first-step macro parameter estimates. This conclusion is based on simulated data, but if the LRR model is a valid description of real-world data generating processes, we should expect that it holds true for the empirical data too. To investigate this question, we rely on the GMM Bayes-Schwarz information criterion (GMM-BIC) introduced by Andrews (1999):

$$\text{GMM-BIC} = J_T - (m - p) \ln T. \tag{3.23}$$

The GMM-BIC is based on the J-statistic, which for the first-step estimation of the macro parameters reads

$$J_T = T\, \boldsymbol{G}_T(\hat{\boldsymbol{\xi}}^{M_*}) \left[\widehat{\text{Avar}}(\boldsymbol{G}_T(\hat{\boldsymbol{\xi}}^{M_*}))\right]^+ \boldsymbol{G}_T(\hat{\boldsymbol{\xi}}^{M_*}), \tag{3.24}$$

where $^+$ denotes the Moore-Penrose inverse. The use of more moments increases J_T, because matching the sample moments with theoretical moments becomes more

difficult. This increase is counterbalanced by subtracting a term whose value rises with the sample size and the degree of overidentification. When choosing between alternative sets of moment matches, the one that yields the smallest GMM-BIC is preferable.

In Panel A of Table 3.9 we report the first-step macro parameter estimates that result from using various sets of moment matches for GMM. The entries in the table are sorted in ascending order of GMM-BIC values. All variants use the five first and second moment matches in Equation (3.12), combined with the (cross) auto-moment matches in Equation (3.13) for various maximum lag lengths L_1, L_2, and L_3. The GMM-BIC points to L_1=L_2=L_3=12, implying that auto-moment matches for up to 3 years are informative, in implicit support of the LRR model. The Monte Carlo results thus emphasize the necessity to exploit higher-order auto-moment matches to identify parameters that pertain to the latent component x_t; the GMM-BIC prompts us to do precisely that.

For the preferred set of moment matches, we also report the second-stage GMM estimates and the standard errors based on asymptotic GMM inference, as well as the bootstrap standard errors.[9] The parametric bootstrap simulation consists of generating LRR model-implied data, with the point estimates as true parameter values. The simulated samples contain the same number of observations as the empirical data. The GMM estimation then can be performed on the simulated series, and the sample drawing and estimation is repeated 250 times. Bootstrap standard errors result from computing the standard deviation across the successful bootstrap replications.[10]

[9] Asymptotic standard errors are computed assuming no serial correlation of the GMM residuals, $\boldsymbol{g}_t^* - \mathbb{E}\left[\boldsymbol{g}(\boldsymbol{q}_t; \boldsymbol{\xi}_0)\right]$.

[10] Applying the same criteria as in the Monte Carlo study, replications in which the optimization fails are discarded.

The point estimates for the moment matches preferred by the GMM-BIC are plausible (cf. Panel A of Table 3.9). The bootstrap standard errors for some parameters are close to the asymptotic standard errors, whereas for others, they are larger. Therefore, the finite sample approximation applied to asymptotic theory may produce somewhat overoptimistic results. The estimation precision varies across macro parameters, similar to the way it does for the simulated data. Moreover, second-stage GMM improves the estimation precision also in a small sample. The findings using simulated LRR model data thus arise as well when we use empirical data, a result that lends support to the LRR paradigm.

3.4.3 Second-step estimation results

We use the second-stage GMM parameter estimates of the GMM-BIC-preferred specification to estimate the preference parameters δ, γ, and ψ in the second (SMM) step. The SMM estimation relies on six theory-based moment matches, as motivated in Section 3.2.7.

It turns out that the minimization of the SMM objective function using empirical data is more challenging than the benign first-step GMM problem. In the Monte Carlo study, the known true parameter values provided a reference point for choosing starting values for the optimization, as well as a gauge of the plausibility of the estimates. The empirical analysis has no such anchor. We therefore employ a computer-intensive procedure to find the minimum of the SMM objective function. The minimization starts from a grid of 100 different parameter combinations, and then selects the parameters that pertain to the smallest of the 100 minima.[11]

[11] Numerical issues, and model solvability in particular, cannot be expected to be mitigated when using empirical instead of simulated data, especially when the sample size is small. However, the numerical problems discussed here should not be confounded with the identification problem investigated in Section 3.2.2. Estimation failure on a large sample of LRR model-generated data indicates identification problems, which can be resolved by using well-thought-out moment matches,

The optimization algorithms that we employ (NM and CMAES) converge from many different initial values to the same overall minimum, but not from all 100 grid points. We therefore caution against taking the "global optimizer" qualifier, as attributed to algorithms like CMAES and simulated annealing, literally. The optimization-from-grid strategy increases the reliability of the reported estimates, but it has a price. With limited computer resources, researchers would likely have to wait a considerable amount of time until they could obtain the final estimates. The benefit though is that they could be more confident that they have indeed found the overall minimum of the SMM objective function.

Standard errors for the preference parameter estimates result from extending the previously described bootstrap approach. To that end, we simulate LRR model-implied macro and financial series using the first- and second-step parameter estimates. The SMM estimation then can be performed on the simulated data, and the simulation/estimation steps are repeated 250 times. Standard errors are obtained by computing the standard deviation across the bootstrap estimates. Although computationally burdensome, we again recommend performing the optimization in each bootstrap replication from a grid of initial values, to prevent spurious convergence that would distort the bootstrap standard errors. The bootstrap simulation can rely on "pseudo-true" parameters—that is, the empirical estimates—to help reduce the number of grid points.

Panel B of Table 3.9 reports the SMM preference parameter point estimates and the bootstrap standard errors. The estimates of δ, ψ, and γ are, from an economic point of view, arguably the most interesting. There is an ongoing debate about whether the large U.S. market equity premium and small T-bill rate can be reconciled with reasonable investor time and risk preferences. Generating model-

as shown in the Monte Carlo study. In the empirical application, in contrast, we deal with numerical problems that occur due the small sample size.

implied data assuming plausible parameter values and comparing selected simulated and sample moments, i.e. calibration, is one way to confront the model with empirical facts. Econometric analysis instead seeks to test empirically and potentially refute a model, instead of seeking confirmatory evidence. Moreover, it gives an idea about how informative the data are, that is, which range of parameter values is compatible with the data.

When interpreting the second-step estimation results, we find it instructive to contrast them with those reported by Yogo (2006), who also relies on recursive preferences in a consumption-based asset pricing framework. Whereas the LRR model focuses on the long-run properties of aggregate consumption growth, Yogo's (2006) idea is to disentangle durable and non-durable consumption.

Assuming expected utility maximization and a power utility function often requires an implausible, negative rate of time preference (δ>1) to explain both the small average T-bill rate and the large market equity premium simultaneously. In contrast, the estimate of the subjective discount factor reported in Panel B of Table 3.9, $\hat{\delta} = 0.985$, is perfectly reasonable and also quite precise (s.e.($\hat{\delta}$)=0.0017). Yogo (2006), who also reports estimates of δ smaller than 1, interprets this result as evidence that the recursive utility specification, which is an integral part of the LRR framework, helps resolve the risk-free rate puzzle.

Bansal and Yaron (2004) emphasize that the ability of the LRR model to resolve the equity premium puzzle hinges on an IES that is greater than 1. Only then does an intertemporal substitution effect dominate the income effect, and the LRR story unfolds. The point estimate reported in Panel B of Table 3.9, $\hat{\psi}$=1.11, is therefore in accord with the LRR paradigm. In contrast, Yogo (2006) reports very small IES estimates, such as $\hat{\psi}$ between 0.023 and 0.024, and s.e.($\hat{\psi}$) of 0.002–0.009. It should be noted though that the range of the IES values supported by the data also includes

values of ψ that are less than 1 (s.e.($\hat{\psi}$)=0.88). However, we already knew from the Monte Carlo study that the estimation precision must be limited in a small sample.

The estimate of the RRA coefficient reported in Panel B of Table 3.9 is large ($\hat{\gamma} = 218.5$) and comparable, in terms of size and precision, to the estimates reported by Yogo (2006).[12] Previous attempts at an econometric estimation of the LRR model have reported considerably smaller RRA estimates, but as we have seen, these one-step estimation results need to be taken with a grain of salt. Our results suggest instead that Campbell and Cochrane's (1999, p. 243) caveat applies to the LRR model too: "High risk aversion is inescapable (or at least has not yet been escaped) in the class of identical-agent models that are consistent with the equity premium facts."

3.5 Conclusion

Econometric analyses and empirical tests of the long-run risk asset pricing model are difficult. It is a demanding task to estimate a model that features two latent processes as fundamental economic drivers, a pricing kernel that depends on unobservable variables, and that must be solved every time it is computed for new parameter values. As an estimation technique, SMM is designed to cope with such methodological challenges, but some important questions have not been addressed in prior literature, and our study seeks to close that gap: Are the moments selected for matching informative enough to identify the structural model parameters that describe the dynamics of latent processes and investor preferences? Identification problems are not obvious in such a complex model structure. And even if theory-

[12] The estimates of γ range from 174.5 to 205.9 (see Table II on p. 552 in Yogo (2006)). The standard errors for $\hat{\gamma}$ reported by Yogo (2006) range from 11.8–49.9, which is also comparable with our estimate, s.e.($\hat{\gamma}$)=12.0.

based, and practically useable moment conditions can be found, what sample size is required to deliver precise estimates? The number of observations available for empirical analysis is relatively small.

We tackle these issues by proposing a two-step estimation strategy, in which we elicit moment matches that reflect the key features of the LRR model. Most importantly, we argue that the recursive LRR model structure must be reflected in the estimation strategy, meaning that macro and financial moment matches need to be disentangled. We therefore first estimate the parameters that drive the macroeconomic dynamics, and then exploit asset pricing and predictive relations implied by the LRR framework for the estimation of the preference parameters. The 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. The properties of the latent persistent growth component, the defining feature of the LRR model, are captured by including higher-order auto-moments of consumption and dividend growth. Considering the notorious difficulty associated with estimating stochastic volatility processes, and doubtful identification, we propose to concentrate out the SV parameters in the second estimation step. We do not preclude the potential prevalence of SV in the data, but we replace time-varying stochastic volatility with the first-step unconditional volatility estimate when computing the simulated moments in the second step, in which we estimate the investor preference parameters by SMM. Using theory-based financial moment matches, SMM delivers precise estimates for the subjective discount factor, relative risk aversion, and the intertemporal elasticity of substitution, even for smaller samples, if the first-step input is of high quality. Considering the complexity of the LRR asset pricing equations, this result is encouraging.

The empirical application of our two-step GMM/SMM estimation method provides support for the LRR asset pricing paradigm, but it also challenges some previously reported results. In particular, the estimate of the subjective discount factor is both plausible and precise, which indicates that the LRR model can help to resolve the interest rate puzzle. The estimate of the intertemporal elasticity of substitution is less precise, but the point estimate is greater than 1. An IES greater than 1 is the cornerstone of the LRR paradigm. Previous empirical studies have either reported very small IES estimates or the IES has not been estimated at all, and instead fixed to a convenient value. Finally, our estimate of the coefficient of relative risk aversion indicates that the conclusion that consumption-based asset pricing models with a representative agent require a high level of risk aversion also holds true for the LRR model.

For an accurate estimation of the preference parameters, the estimates of the macro parameters used for the second estimation step must be of good quality. Therefore, both informative first-step moment matches and a relatively large sample size are required. To estimate a complex dynamic asset pricing model like the LRR model, informative and strong moment matches are indispensable. If the estimation quality of the macroeconomic parameters is poor, researchers cannot expect much from the second-step estimation of the preference parameters. Our two-step approach thus constitutes a reality check for applied work.

Fruitful extensions in subsequent research could seek to increase the quality of the macro parameter estimates. Time must pass before the confidence bounds can narrow, but strong and well-thought-out moment matches will help applied research in the meantime.

3.A Appendix

SV parameter estimation has been attempted by using higher-order moments. Constantinides and Ghosh (2011) suggest the use of the variances of g_{t+1}^2 and $g_{d,t+1}^2$:

$$\begin{aligned}
\mathrm{Var}\left(g_{t+1}^2\right) =& \frac{3\varphi_e^4\sigma_w^2(1+\nu_1\rho^2)}{(1-\rho^4)(1-\nu_1^2)(1-\nu_1\rho^2)} + \frac{1}{1-\rho^4}\left[2\varphi_e^4\sigma^4 + \frac{4\rho^2\varphi_e^4\sigma^4}{1-\rho^2}\right] + \frac{3\sigma_w^2}{1-\nu_1^2} \\
&+ 4\mu_c^2\frac{\varphi_e^2\sigma^2}{1-\rho^2} + \frac{6\varphi_e^2\sigma_w^2\nu_1}{(1-\nu_1^2)(1-\nu_1\rho^2)} + \frac{4\varphi_e^2\sigma^4}{1-\rho^2} + 2\sigma^4 + 4\mu_c^2\sigma^2, \quad (3.25) \\
\mathrm{Var}\left(g_{d,t+1}^2\right) =& \phi^4\left[\frac{3\varphi_e^4\sigma_w^2(1+\nu_1\rho^2)}{(1-\rho^4)(1-\nu_1^2)(1-\nu_1\rho^2)} + \frac{1}{1-\rho^4}\left(2\varphi_e^4\sigma^4 + \frac{4\rho^2\varphi_e^4\sigma^4}{1-\rho^2}\right)\right] \\
&+ \frac{3\sigma_w^2}{1-\nu_1^2}\varphi_d^4 + 4\mu_d^2\frac{\varphi_e^2\sigma^2}{1-\rho^2}\phi^2 + \frac{6\varphi_e^2\sigma_w^2\nu_1}{(1-\nu_1^2)(1-\nu_1\rho^2)}\phi^2\varphi_d^2 \\
&+ \frac{4\varphi_e^2\sigma^4}{1-\rho^2}\phi^2\varphi_d^2 + 2\sigma^4\varphi_d^4 + 4\mu_d^2\varphi_d^2\sigma^2. \quad (3.26)
\end{aligned}$$

These expressions are error-corrected versions of the formulas reported in Constantinides and Ghosh (2011). The fourth moments of g and g_d are then given by:

$$\begin{aligned}
\mathbb{E}\left(g_{t+1}^4\right) &= \mathrm{Var}\left(g_{t+1}^2\right) + \left(\mathbb{E}(g_{t+1}^2)\right)^2 \\
&= \mathrm{Var}\left(g_{t+1}^2\right) + \left(\mu_c^2 + \frac{\varphi_e^2\sigma^2}{1-\rho^2} + \sigma^2\right)^2, \quad (3.27) \\
\mathbb{E}\left(g_{d,t+1}^4\right) &= \mathrm{Var}\left(g_{d,t+1}^2\right) + \left(\mathbb{E}(g_{d,t+1}^2)\right)^2 \\
&= \mathrm{Var}\left(g_{d,t+1}^2\right) + \left(\mu_d^2 + \phi^2\frac{\varphi_e^2\sigma^2}{1-\rho^2} + \varphi_d^2\sigma^2\right)^2. \quad (3.28)
\end{aligned}$$

Table 3.1: Moments used for SMM and GMM estimations of the LRR model. The table lists the moments for two approaches to estimating the parameters of the LRR model. The moments in Panel A are used for an SMM estimation approach, adapted from Hasseltoft (2012). The moments in Panel B are used for the GMM estimation approach by Constantinides and Ghosh (2011). Panel 1 lists the moments related to macroeconomic LRR variables, Panel 2 contains moments related to financial LRR variables, and Panel 3 lists two moments that result from an empirically motivated prediction relationship. Finally, ζ_{t+1} is the residual of an AR(1) process for log consumption growth, obtained by regressing g_{t+1} on g_t.

Panel A: SMM approach following Hasseltoft (2012)	**Panel B: GMM approach** following Constantinides and Ghosh (2011)
Panel 1: Macro moments	
$\mathbb{E}(g_t)$	$\mathbb{E}(g_t)$
$\mathbb{E}(g_{d,t})$	$\mathbb{E}(g_{d,t})$
$\mathbb{E}(g_t^2)$	$\mathbb{E}(g_t^2)$
$\mathbb{E}(g_{d,t}^2)$	$\mathbb{E}(g_{d,t}^2)$
$\mathbb{E}(g_{t+1} \cdot g_t)$	$\mathbb{E}(g_{t+1} \cdot g_t)$
$\mathbb{E}(g_{t+2} \cdot g_t)$	$\mathbb{E}(g_{d,t+1} \cdot g_{d,t})$
	$\mathbb{E}(g_t \cdot g_{d,t})$
	$\mathbb{E}(g_t^4)$
	$\mathbb{E}(g_{d,t}^4)$
Panel 2: Financial moments	
$\mathbb{E}(r_{m,t} - r_{f,t})$	$\mathbb{E}(M_{t+1} \cdot R_{m,t+1})$
$\mathbb{E}(r_{f,t})$	$\mathbb{E}(M_{t+1} \cdot R_{f,t+1})$
$\mathbb{E}(z_{m,t})$	$\mathbb{E}(M_{t+1} \cdot R_{m,t+1} \cdot r_{f,t})$
$\mathbb{E}[(r_{m,t} - r_{f,t})^2]$	$\mathbb{E}(M_{t+1} \cdot R_{m,t+1} \cdot z_{m,t})$
$\mathbb{E}[r_{f,t}^2]$	$\mathbb{E}(M_{t+1} \cdot R_{f,t+1} \cdot r_{f,t})$
$\mathbb{E}[z_{m,t}^2]$	$\mathbb{E}(M_{t+1} \cdot R_{f,t+1} \cdot z_{m,t})$
Panel 3: Prediction moments	
$\mathbb{E}(\zeta_{t+1}^2)$	
$\mathbb{E}(\zeta_{t+1}^2 \cdot z_{m,t})$	

Table 3.2: Comparison of alternative estimation strategies.
We simulate data of length T=100k, generated by the LRR model, with Bansal and Yaron's (2004) calibrated parameters as true values $\boldsymbol{\xi}_0$. Panel A reports the results from an SMM estimation for which we match the moments in Panel A of Table 3.1; Panel B provides the results obtained from adopting the GMM estimation strategy proposed by Constantinides and Ghosh (2011), based on the moments listed in Panel B of Table 3.1; and Panel C shows the estimates from the two-step approach described in Section 3.2.4, using m=185 moment matches for the first-step estimation and the theory-based financial moment matches $\mathbf{G}_T^P$ for the second-step estimation. The results in Panels A and B were obtained using the CMAES algorithm. The Panel C results were obtained using the Nelder-Mead simplex algorithm. In all estimations, $\boldsymbol{W}_T$=$\boldsymbol{I}$. The starting values in $\boldsymbol{\xi}_{s_1}$ correspond to the true values $\boldsymbol{\xi}_0$ given in Table 2.1, $\boldsymbol{\xi}_{s_2} = (\mu_c = 0.0015, \mu_d = 0.0015, \rho = 0.95, \varphi_e = 0.05, \nu_1 = 0.95, \sigma_w = 0.00001, \sigma = 0.008, \phi = 3, \varphi_d = 5, \delta = 0.99, \gamma = 12, \psi = 1.2)'$, and $\boldsymbol{\xi}_{s_3} = (\mu_c = 0.001, \mu_d = 0.001, \rho = 0.8, \varphi_e = 0.05, \nu_1 = 0.5, \sigma_w = 0.00001, \sigma = 0.008, \phi = 3, \varphi_d = 5, \delta = 0.96, \gamma = 15, \psi = 2)'$. The table also reports the values of the objective functions at the initial and final parameter values. The convergence criteria for objective function and parameters are $1 \cdot 10^{-15}$ and $1 \cdot 10^{-9}$. A star indicates that the algorithm did not converge, given the maximum number of function evaluations, which was limited to $3 \cdot 10^6$.

	Panel A			**Panel B**			**Panel C**		
	$\boldsymbol{\xi_{s_1} = \xi_0}$	$\boldsymbol{\xi_{s_2}}$	$\boldsymbol{\xi_{s_3}}$	$\boldsymbol{\xi_{s_1} = \xi_0}$	$\boldsymbol{\xi_{s_2}}$	$\boldsymbol{\xi_{s_3}}$	$\boldsymbol{\xi_{s_1} = \xi_0}$	$\boldsymbol{\xi_{s_2}}$	$\boldsymbol{\xi_{s_3}}$
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	0.0038	0.0017	0.0017	0.0017
ρ	0.9786	0.9618	0.8044	0.9994	0.9994	0.7287	0.9813	0.9813	0.9813
φ_e	0.0454	0.0658	0.1324	0.0080	0.0076	0.1787	0.0426	0.0426	0.0426
ν_1	0.9864	0.9645	0.7387	0.9873	0.9740	0.0336			
σ_w	$2.1 \cdot 10^{-6}$	$1.1 \cdot 10^{-5}$	$4.4 \cdot 10^{-6}$	$2.0 \cdot 10^{-7}$	$3.4 \cdot 10^{-9}$	$4.2 \cdot 10^{-6}$			
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	2.9254	2.9314	29.0190	2.6930	2.6930	2.6930
φ_d	4.4856	2.6125	2.6915	4.4875	4.4962	53.1272	4.4919	4.4919	4.4919
δ	0.9980	0.9975	0.9969	0.9977	0.9916	0.9911	0.9979	0.9979	0.9979
γ	10.1883	11.7574	14.7956	10.4576	9.1161	9.1177	10.4593	10.4593	10.4593
ψ	1.4974	2.1615	1.0517	0.9522	0.8443	-0.0758	1.5677	1.5677	1.5677
obj. fct. (final)	$4.4 \cdot 10^{-13}$*	$1.7 \cdot 10^{-6}$	$7.2 \cdot 10^{-6}$*	$5.9 \cdot 10^{-6}$	$5.9 \cdot 10^{-6}$	$5.6 \cdot 10^{-4}$	$1.1 \cdot 10^{-9}$ $3.4 \cdot 10^{-8}$	$1.1 \cdot 10^{-9}$ $3.4 \cdot 10^{-8}$	$1.1 \cdot 10^{-9}$ $3.4 \cdot 10^{-8}$
obj. fct. (initial)	$3.3 \cdot 10^{-4}$	68.8180	395.0515	$4.1 \cdot 10^{-5}$	0.0102	6.9203	$2.7 \cdot 10^{-8}$ 0.3797	$1.6 \cdot 10^{-7}$ 94.3059	$8.3 \cdot 10^{-7}$ 398.5025

Table 3.3: Moment sensitivity.
The table shows the sensitivity of various moments from Table 3.1 and selected higher-order (cross) auto-moments to changes in the LRR model parameters. The moments are simulated on the basis of a sample size of 10^8 observations, using Bansal and Yaron's calibrated values from Table 2.1. The moment sensitivity is computed as the percentage change of a moment when a given parameter is decreased by 50% c.p. The columns of the table show the sensitivity of all moments to a change of the parameter in the column header.

	μ_c	μ_d	ρ	φ_e	σ	ϕ	φ_d	ν_1	σ_w	δ	γ	ψ
Panel A: Macro moments												
$\mathbb{E}(g)$	-50	0	0	0	0	0	0	0	0	0	0	0
$\mathbb{E}(g_d)$	0	-50	0	0	0	0	0	0	0	0	0	0
$\mathbb{E}(g^2)$	-3	0	-4	-3	-67	0	0	0	0	0	0	0
$\mathbb{E}(g_d^2)$	0	0	-2	-2	-69	-2	-73	0	0	0	0	0
$\mathbb{E}(g_{t+1} \cdot g_t)$	-34	0	-54	-41	-38	0	0	0	0	0	0	0
$\mathbb{E}(g_{t+2} \cdot g_t)$	-34	0	-54	-41	-38	0	0	0	0	0	0	0
$\mathbb{E}(g_{t+12} \cdot g_t)$	-38	0	-49	-37	-34	0	0	0	0	0	0	0
$\mathbb{E}(g_{t+36} \cdot g_t)$	-47	0	-37	-28	-26	0	0	0	0	0	0	0
$\mathbb{E}(g_{d,t+1} \cdot g_{d,t})$	0	-6	-89	-69	-64	-69	0	0	0	0	0	0
$\mathbb{E}(g_{d,t+2} \cdot g_{d,t})$	0	-6	-90	-69	-64	-69	0	0	0	0	0	0
$\mathbb{E}(g_{d,t+12} \cdot g_{d,t})$	0	-8	-90	-67	-62	-67	0	0	0	0	0	0
$\mathbb{E}(g_{d,t+36} \cdot g_{d,t})$	0	-12	-84	-63	-58	-63	0	0	0	0	0	0
$\mathbb{E}(g \cdot g_d)$	-10	-10	-75	-59	-55	-39	0	0	0	0	0	0
$\mathbb{E}(g_{t+1} \cdot g_{d,t})$	-11	-11	-76	-59	-55	-39	0	0	0	0	0	0
$\mathbb{E}(g_{t+2} \cdot g_{d,t})$	-11	-11	-78	-59	-54	-39	0	0	0	0	0	0
$\mathbb{E}(g_{t+12} \cdot g_{d,t})$	-13	-13	-75	-56	-52	-37	0	0	0	0	0	0
$\mathbb{E}(g_{t+36} \cdot g_{d,t})$	-18	-18	-64	-48	-44	-32	0	0	0	0	0	0
$\mathbb{E}(g^4)$	-5	0	-8	-6	-87	0	0	-5	-4	0	0	0
$\mathbb{E}(g_d^4)$	0	0	-4	-3	-88	-3	-93	-5	-4	0	0	0
Panel B: Financial moments												
$\mathbb{E}(r_m - r_f)$	1	-3	-118	-86	-66	-69	12	-4	-3	-118	-69	-38
$\mathbb{E}(r_f)$	-23	0	17	13	27	0	0	1	1	$> 10^4$	21	70
$\mathbb{E}(z_m)$	2	-3	43	20	10	16	-2	1	0	-100	12	-1
$\mathbb{E}((r_m - r_f)^2)$	2	-2	-45	-32	-66	-39	-42	-1	-1	-46	6	-23
$\mathbb{E}(r_f^2)$	-32	0	8	5	32	0	0	1	0	$> 10^6$	36	212
$\mathbb{E}(z_m^2)$	4	-5	106	43	21	34	-4	1	1	-100	25	-2
Panel C: Prediction moments												
$\mathbb{E}(\xi^2)$	0	0	-4	-3	-70	0	0	0	0	0	0	0
$\mathbb{E}(\xi_{t+1}^2 \cdot z_{m,t})$	2	-3	38	16	-67	16	-2	1	1	-100	12	-1

Table 3.4: Moment sensitivity to parameters for theory-based moments.
The table displays the sensitivity of the theory-based financial moments to changes in the preference parameters using the simulated part of the moment match. The moments are computed from a simulated sample of size of 10^6 observations, based on the parameters from Table 2.1. The moment sensitivity in this table is computed as the percentage 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.

	δ	γ	ψ
$\mathbb{E}(M)$	-10	0	0
$\frac{-\mathrm{Cov}(R_m - R_f, M)}{\mathbb{E}(M)}$	-97	-10	-4
$\mathbb{E}\left[(R_m - R_f)^2\right]$	-32	1	-3
$\frac{\mathrm{Cov}(R_{f,t+1}, z_{m,t})}{\mathrm{Var}(z_m)}$	428	-1	14
$\mathbb{E}(z_m)$	-60	2	0
$\mathbb{E}(z_m^2)$	-84	4	0

Table 3.5: Moment matches used for GMM estimation of macro parameters.
For GMM estimation of $\boldsymbol{\xi}^{M_*}$, the basic set of first and second moment matches in Equation (3.12) is always included. The maximum lag lengths of the (cross) auto-moments in Equation (3.13) vary according to the scheme below.

moment set	L_1	L_2	L_3
$m = 7$	2	0	0
$m = 35$	10	10	10
$m = 113$	36	36	36
$m = 185$	60	60	60

Table 3.6: Means and standard deviations of first-step GMM estimates.
The table reports the means and the standard deviations of the GMM macro parameter estimates. For each sample size T, 400 data sets were simulated, and the estimation was performed using different moment sets, ranging from 7 to 185 moment conditions (cf. Table 3.5).

	First-stage GMM estimation				**Second-stage GMM estimation**		
	$m = 7$	$m = 35$	$m = 113$	$m = 185$	$m = 35$	$m = 113$	$m = 185$
				Panel A: $\mu_c = 0.0015$			
T=1k	0.0016	0.0015	0.0016	0.0016	0.0016	0.0015	0.0014
	(0.0005)	(0.0006)	(0.0005)	(0.0005)	(0.0006)	(0.0005)	(0.0006)
T=5k	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
T=100k	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
				Panel B: $\mu_d = 0.0015$			
T=1k	0.0021	0.0020	0.0020	0.0020	0.0021	0.0020	0.0019
	(0.0014)	(0.0016)	(0.0015)	(0.0015)	(0.0016)	(0.0015)	(0.0015)
T=5k	0.0016	0.0016	0.0016	0.0016	0.0016	0.0015	0.0016
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0009)
T=100k	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
				Panel C: $\rho = 0.979$			
T=1k	0.809	0.803	0.927	0.914	0.834	0.940	0.927
	(0.294)	(0.293)	(0.151)	(0.171)	(0.292)	(0.134)	(0.172)
T=5k	0.891	0.934	0.973	0.973	0.968	0.976	0.977
	(0.155)	(0.121)	(0.047)	(0.028)	(0.082)	(0.041)	(0.009)
T=100k	0.932	0.978	0.979	0.979	0.980	0.979	0.979
	(0.099)	(0.015)	(0.004)	(0.003)	(0.009)	(0.002)	(0.002)
				Panel D: $\varphi_e = 0.044$			
T=1k	0.0752	0.0689	0.0529	0.0562	0.0568	0.0516	0.0582
	(0.0994)	(0.0817)	(0.0545)	(0.0473)	(0.0717)	(0.0558)	(0.0498)
T=5k	0.0807	0.0556	0.0433	0.0474	0.0368	0.0426	0.0456
	(0.0813)	(0.0532)	(0.0349)	(0.0290)	(0.0393)	(0.0233)	(0.0109)
T=100k	0.0625	0.0417	0.0437	0.0442	0.0402	0.0441	0.0442
	(0.0528)	(0.0204)	(0.0055)	(0.0040)	(0.0150)	(0.0022)	(0.0019)
				Panel E: $\sigma = 0.0078$			
T=1k	0.0078	0.0078	0.0078	0.0078	0.0076	0.0072	0.0069
	(0.0004)	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
T=5k	0.0078	0.0078	0.0078	0.0078	0.0078	0.0077	0.0076
	(0.0002)	(0.0004)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
T=100k	0.0078	0.0078	0.0078	0.0078	0.0078	0.0078	0.0078
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Table 3.6 – continued

	first-stage GMM estimation				**Second-stage GMM estimation**		
	$m = 7$	$m = 35$	$m = 113$	$m = 185$	$m = 35$	$m = 113$	$m = 185$
				Panel F: $\phi = 3.0$			
T=1k	4.08	5.06	3.88	4.12	4.15	3.45	3.28
	(5.33)	(5.23)	(3.46)	(3.98)	(4.34)	(2.95)	(2.52)
T=5k	2.98	3.37	3.10	3.11	3.02	2.99	2.96
	(2.58)	(1.65)	(0.83)	(0.76)	(0.63)	(0.57)	(0.60)
T=100k	2.97	3.01	3.01	3.01	3.00	3.01	3.00
	(1.62)	(0.19)	(0.14)	(0.14)	(0.12)	(0.11)	(0.18)
				Panel G: $\varphi_d = 4.5$			
T=1k	4.44	4.49	4.50	4.49	4.53	4.64	4.75
	(0.27)	(0.30)	(0.25)	(0.17)	(0.18)	(0.20)	(0.22)
T=5k	4.50	4.56	4.49	4.50	4.51	4.53	4.55
	(0.12)	(1.21)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
T=100k	4.51	4.50	4.50	4.50	4.50	4.50	4.50
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
				Panel H: Successful estimations			
T=1k	248	281	326	321	290	321	326
T=5k	375	369	393	389	380	399	393
T=100k	397	399	399	400	398	400	399

Table 3.7: Means and standard deviations of the second-step SMM estimates. Panel A reports the means and standard deviations of the SMM preference parameter estimates that use the true macro parameters for the moment simulation. The first column of Panel A reports the results based on the ad hoc moment matches in Panel A-2 of Table 3.1, and the second column reports the results for the theory-based moment matches $\boldsymbol{G}_T^P = (\boldsymbol{G}_T^{P_1\prime}, \boldsymbol{G}_T^{P_2}, \boldsymbol{G}_T^{P_3}, \boldsymbol{G}_T^{P_4\prime})'$. In both cases, SMM estimates are obtained by concentrating out stochastic volatility, that is, using $\sigma^2 = \mathbb{E}(\sigma_t^2)$ for σ_t^2. The third column of Panel A contains the SMM estimation results using the theory-based financial moment matches and assuming the complete vector of macro parameters is known, and thus using σ_t^2 when simulating the theory-based financial moments. Panel B reports the means and standard deviations of the SMM preference parameter estimates that use the estimated macro parameters for the simulation of moments.

	Panel A			**Panel B**	
	True macro parameters			**Estimated macro parameters**	
	ad hoc	**theory-based**	**+ SV known**	**ad hoc**	**theory-based**
$\underline{\delta = 0.998}$					
T=1k	0.9981	0.9980	0.9981	0.9955	0.9965
	(0.0008)	(0.0006)	(0.0006)	(0.0047)	(0.0021)
T=5k	0.9979	0.9980	0.9980	0.9979	0.9978
	(0.0004)	(0.0003)	(0.0003)	(0.0011)	(0.0007)
T=100k	0.9979	0.9980	0.9980	0.9980	0.9980
	(0.0002)	(0.0001)	(0.0001)	(0.0005)	(0.0002)
$\underline{\gamma = 10}$					
T=1k	10.5	10.3	10.1	34.5	26.5
	(1.4)	(1.1)	(1.0)	(82.5)	(28.5)
T=5k	10.3	10.3	10.0	11.7	12.8
	(0.7)	(0.5)	(0.5)	(6.9)	(5.7)
T=100k	10.1	10.3	10.0	10.3	10.4
	(0.3)	(0.1)	(0.1)	(0.6)	(0.7)
$\underline{\psi = 1.5}$					
T=1k	1.84	1.52	1.49	4.01	3.19
	(1.80)	(0.05)	(0.06)	(9.53)	(3.61)
T=5k	1.76	1.52	1.50	3.28	1.95
	(1.38)	(0.04)	(0.04)	(7.46)	(1.52)
T=100k	1.80	1.51	1.50	1.76	1.53
	(2.21)	(0.00)	(0.00)	(0.92)	(0.15)

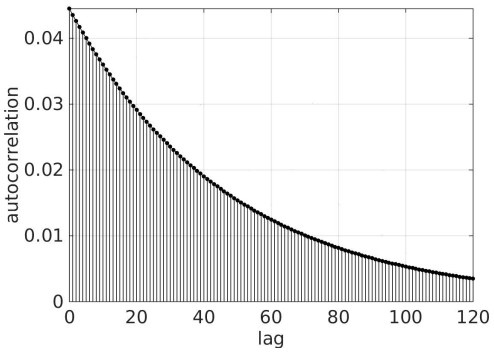
Table 3.8: Means and standard deviations of the preference parameter estimates based on second-stage GMM macro parameter estimates.
Panel A shows the SMM estimation results for the preference parameters, based on the second-stage GMM macro parameter estimates. Panel B reveals the incremental effect of raising the bar for the quality of the first-step estimates before entering the second step. For that purpose, we discard any replication for which one of the macro parameter estimates is more than twice its true value.

	Panel A **Second-stage GMM**	**Panel B** **Second-stage GMM + select.**
$\underline{\delta = 0.998}$		
T=1k	0.9970	0.9971
	(0.0024)	(0.0016)
T=5k	0.9978	0.9978
	(0.0007)	(0.0006)
T=100k	0.9980	0.9980
	(0.0001)	(0.0001)
$\underline{\gamma = 10}$		
T=1k	19.8	15.3
	(24.0)	(9.7)
T=5k	12.0	12.0
	(4.7)	(4.5)
T=100k	10.4	10.4
	(0.7)	(0.7)
$\underline{\psi = 1.5}$		
T=1k	2.88	2.10
	(2.98)	(1.19)
T=5k	1.73	1.71
	(0.60)	(0.47)
T=100k	1.52	1.52
	(0.10)	(0.10)

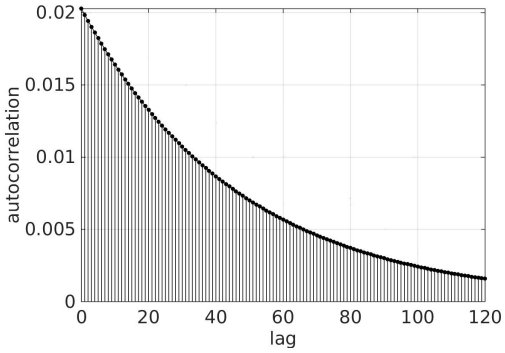
Table 3.9: Parameter estimates based on quarterly empirical data, 1947Q2 to 2008Q4.
The table reports the parameter estimates based on Beeler and Campbell's (2012) data, obtained by applying the two-step GMM/SMM estimation strategy. For the first-step estimation of the macro parameters, we use the basic set of first and second moment matches in Equation (3.12) combined with various maximum lag lengths (L_1, L_2, L_3) of the (cross) auto-moments in Equation (3.13). The second-step estimation of the preference parameters uses the six theory-based moment matches described in Section 3.2.7. The values in parentheses are asymptotic standard errors, assuming no serial correlation of the GMM residuals. The values in brackets are bootstrap standard errors. The table is sorted in ascending order by the GMM-BIC proposed by Andrews (1999).

	Panel A								**Panel B**		
	First-step results								**Second-step results**		
$\mathbf{L_1, L_2, L_3}$	μ_c	μ_d	ρ	φ_e	σ	ϕ	φ_d	**GMM-BIC**	δ	γ	ψ
12,12,12	0.0055	0.0035	0.7984	0.1227	0.0040	17.58	1.14		0.9852	218.5	1.11
(2^{nd} stage)	[0.0003]	[0.0024]	[0.1143]	[0.0614]	[0.0004]	[12.53]	[0.64]		[0.0017]	[12.0]	[0.88]
	(0.0001)	(0.0006)	(0.0450)	(0.0455)	(0.0003)	(6.42)	(0.89)				
12,12,12	0.0051	0.0051	0.7002	0.0851	0.0049	29.01	2.00	-152.1			
	[0.0003]	[0.0026]	[0.1181]	[0.0782]	[0.0006]	[41.37]	[1.40]				
	(0.0003)	(0.0012)	(0.0823)	(0.0826)	(0.0003)	(26.30)	(0.67)				
12,12,0	0.0051	0.0051	0.7010	0.1154	0.0049	21.47	2.02	-92.1			
12,0,0	0.0051	0.0051	0.7474	0.5096	0.0039	1.38	4.89	-50.5			
4,4,4	0.0051	0.0051	0.7098	0.0909	0.0049	26.41	2.10	-38.6			
4,4,0	0.0051	0.0051	0.7107	0.1158	0.0049	20.78	2.11	-18.5			
2,2,2	0.0051	0.0051	0.8761	0.0664	0.0049	21.41	2.71	-15.2			
4,0,0	0.0051	0.0051	0.7343	0.5459	0.0038	1.30	4.98	-10.1			

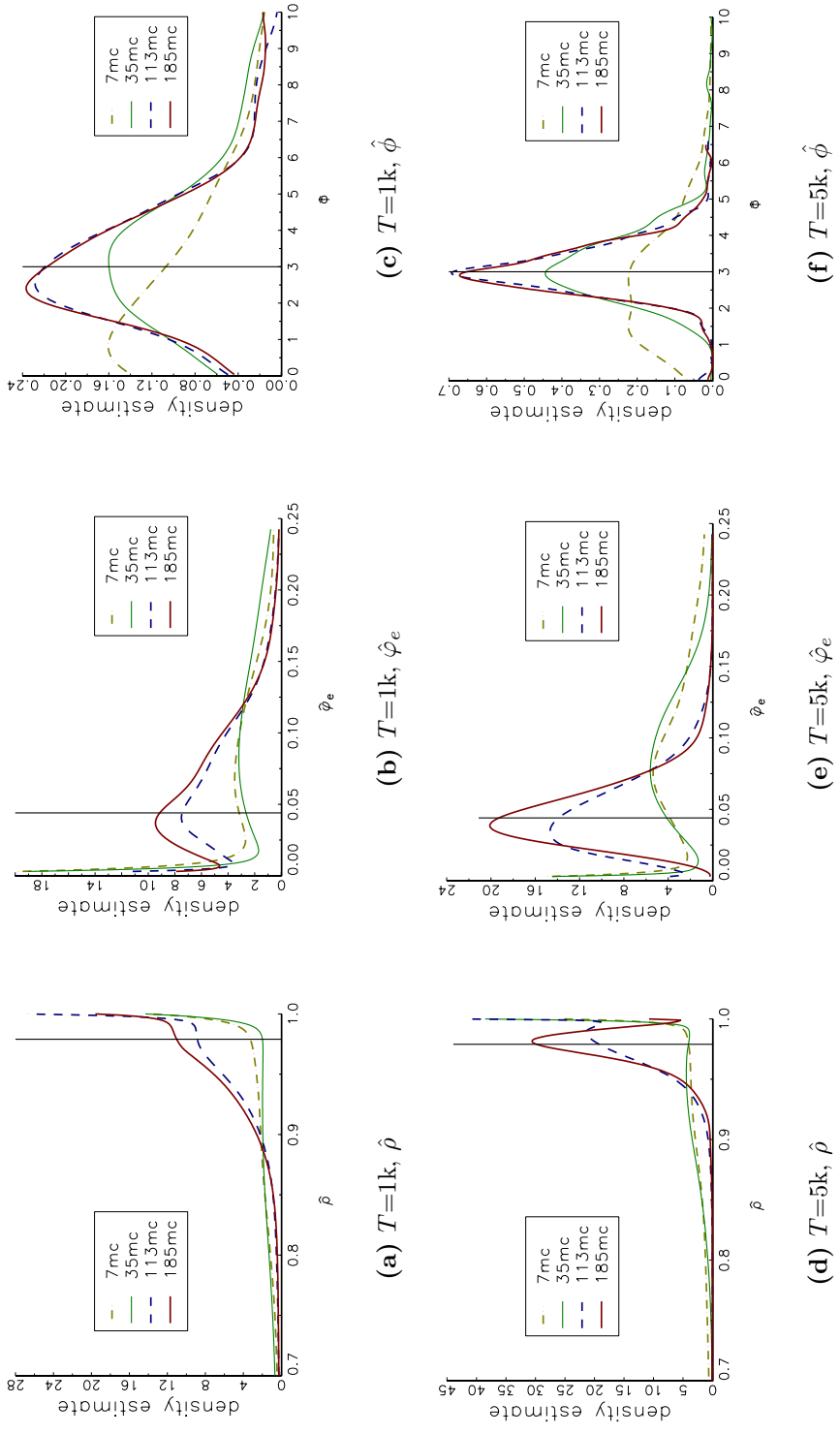
Figure 3.1: Autocorrelograms of consumption and dividend growth.
The figure shows the small but persistent autocorrelations of consumption growth (Panel a) and dividend growth (Panel b) as implied by the LRR model. The graphs display the autocorrelations from lag 1 to 120 based on the LRR parameter values calibrated by Bansal and Yaron (2004), as listed in Table 2.1. These values correspond to a monthly decision frequency of the agent, such that the abscissa spans 10 years.



(a) log consumption growth



(b) log dividend growth

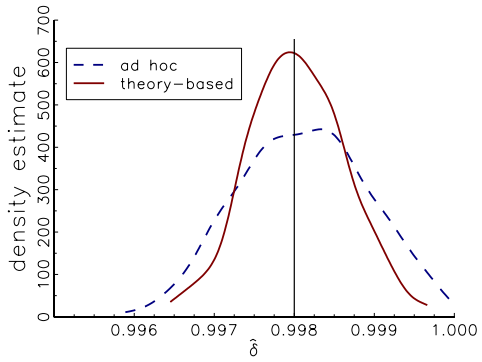


(a) T=1k, $\hat{\rho}$ **(b)** T=1k, $\hat{\varphi}_e$ **(c)** T=1k, $\hat{\phi}$

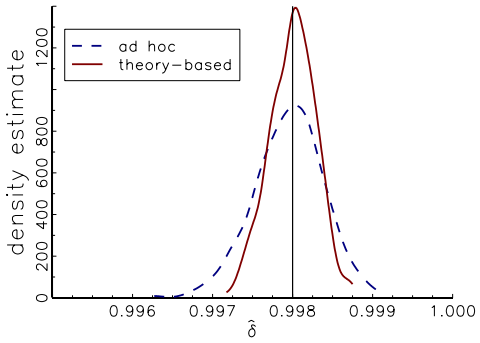
(d) T=5k, $\hat{\rho}$ **(e)** T=5k, $\hat{\varphi}_e$ **(f)** T=5k, $\hat{\phi}$

Figure 3.2: Kernel densities: $\hat{\rho}$, $\hat{\varphi}_e$, and $\hat{\phi}$.
Panels (a)–(f) display kernel densities for estimates of the macro parameters ρ, φ_e, and ϕ. The estimation is based on four different sets of moment matches, $m = 7$, $m = 35$, $m = 113$, $m = 185$ (see Table 3.5) and two time series lengths, T=1k and T=5k. To account for the boundedness of the parameter space, we use the beta kernel proposed by Chen (1999) with the bandwidth selector by Silverman (1986) adjusted for variable kernels. The vertical lines indicate the positions of the true parameters.

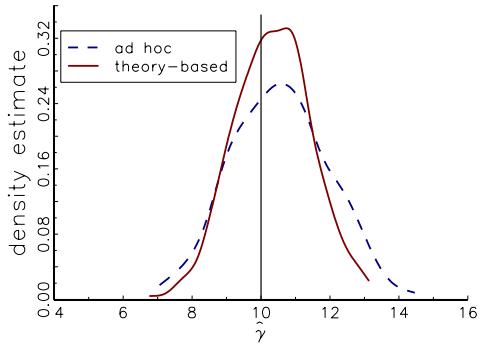
Figure 3.3: Kernel densities: $\hat{\delta}$, $\hat{\gamma}$, and $\hat{\psi}$ (true macro parameters).
Panels (a)–(f) display kernel densities for preference parameter estimates that result from using theory-based (solid) and ad hoc (dashes) financial moment matches. SMM estimation is based on the knowledge of the true macro parameters; SV is concentrated out when simulating moments. The vertical lines indicate the positions of the true parameters. A Gaussian kernel with bandwidth as proposed by Silverman (1986) is used.



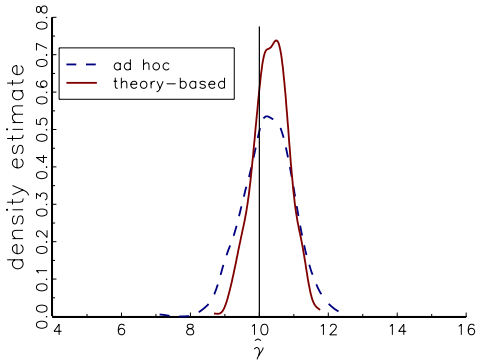
(a) T=1k, $\hat{\delta}$



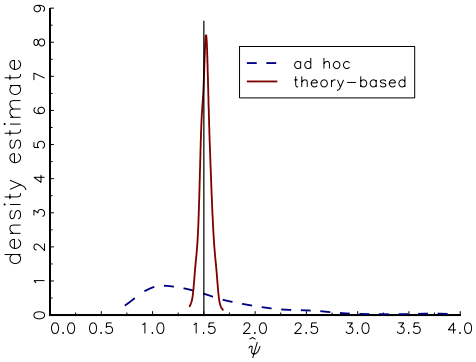
(b) T=5k, $\hat{\delta}$



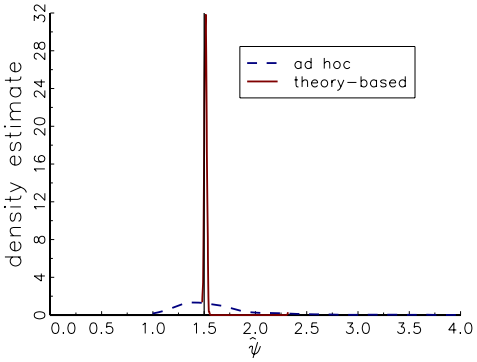
(c) T=1k, $\hat{\gamma}$



(d) T=5k, $\hat{\gamma}$

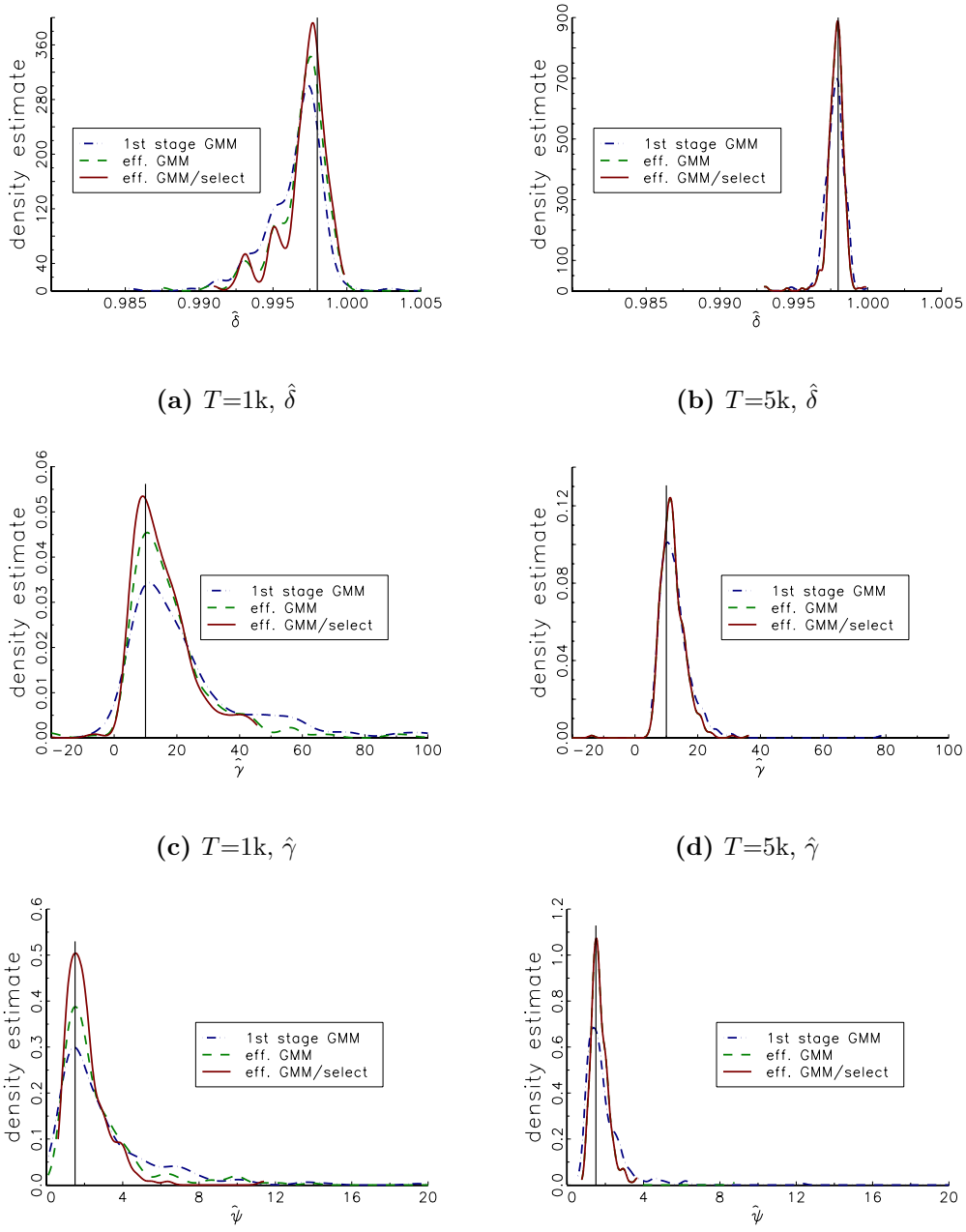


(e) T=1k, $\hat{\psi}$



(f) T=5k, $\hat{\psi}$

Figure 3.4: Kernel densities: $\hat{\delta}$, $\hat{\gamma}$, and $\hat{\psi}$ (estimated macro parameters). Panels (a)–(f) display three kernel densities for preference parameter estimates that result from using estimated macro parameters and theory-based financial moment matches. The first (dashes-dots) uses first-step macro GMM estimates based on the m=185 moment set and an identity weighting matrix for GMM (1st stage GMM). The second (dashes) instead uses an estimate of the efficient GMM weighting matrix (eff. GMM) when estimating the macro parameters. The third (solid line) also uses efficient weighting but applies a more restrictive selection criterion: the second-step SMM estimation of the preference parameters is not performed if one of the first-step macro estimates is more than twice as large as the true parameter value (eff. GMM/select). The vertical lines indicate the positions of the true parameters. A Gaussian kernel with the bandwidth proposed by Silverman (1986) is used.

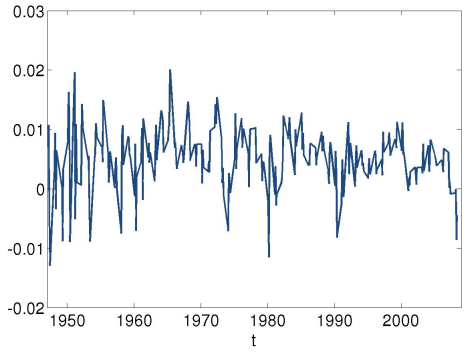


(a) T=1k, $\hat{\delta}$ **(b)** T=5k, $\hat{\delta}$

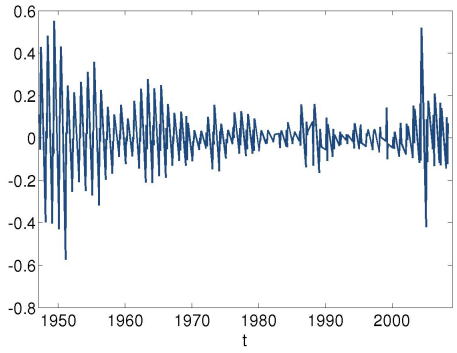
(c) T=1k, $\hat{\gamma}$ **(d)** T=5k, $\hat{\gamma}$

(e) T=1k, $\hat{\psi}$ **(f)** T=5k, $\hat{\psi}$

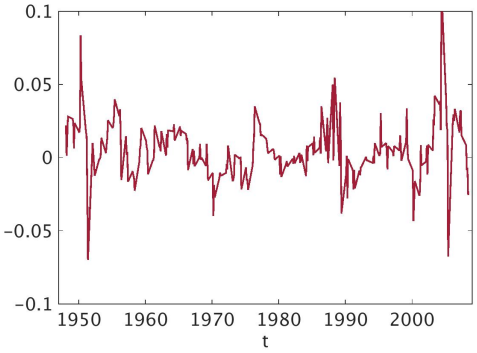
Figure 3.5: Data used for the empirical application.
Panels (a)-(f) display the time series used for the empirical application. The data come from Beeler and Campbell (2012), and they span the time period 1947Q2 to 2008Q4. Consumption growth is computed on the basis of U.S. real consumption of non-durable goods and services. The market portfolio return, dividend growth, and the price-dividend ratio are calculated for the CRSP value-weighted market portfolio. Conversions into real terms are performed using the consumer price index. The proxy for the ex ante risk-free rate is obtained from a forecast of the ex post real rate of three-month Treasury bills.



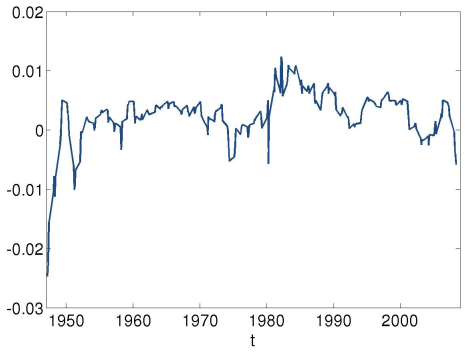
(a) log consumption growth



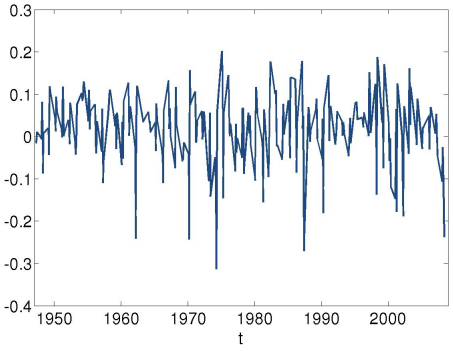
(b) log dividend growth



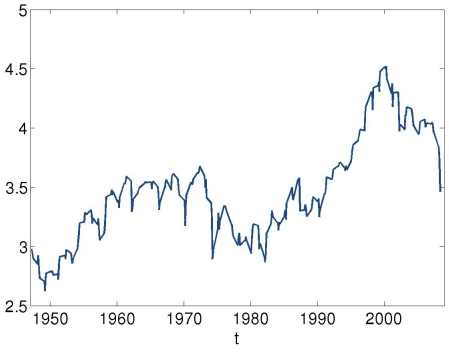
(c) log dividend growth 4 quarters avg.



(d) log risk-free rate



(e) log return market portfolio



(f) log P/D ratio

Chapter 4

Indirect inference estimation of the long-run risk model

4.1 Introduction

Allowing for long-run consumption risk in the pricing kernel holds the promise to resolve prominent asset pricing puzzles and helps restore the nexus of real economy and financial markets. Numerical calibrations show that by taking long-run risk into account, the considerable U.S. postwar equity premium can indeed be explained by a consumption-based asset pricing model that assumes plausible values for the representative agent's time preference, risk aversion, and propensity for intertemporal substitution.

The long-run risk approach, as described in Chapter 2, is theoretically appealing and the calibration results are encouraging. However, the estimation of the structural model parameters, the assessment of the estimation precision, as well as model specification tests, i.e. econometric analysis beyond calibration, are quite challenging. We propose a two-step indirect inference strategy for the estimation of the LRR

asset pricing model that avoids the drawbacks of previous approaches analyzed in detail in Chapter 3. Due to its flexibility, the indirect inference methodology allows for a more parsimonious estimation approach compared to the two-step GMM/SMM estimation strategy presented in the previous chapter. Since indirect inference estimation is a simulation-based approach, the model dynamics can be estimated at an arbitrary frequency, irrespective of the frequency of the data. The two-step approach even allows for different frequencies of macroeconomic and financial data.

Calvet and Czellar (2015) also use an indirect inference approach to estimate a version of the LRR model, in which the endogeneity is removed, such that a model solution in the course of the estimation is no longer required. Instead, the means of z_t and $z_{m,t}$, which should be endogenously determined, are set to fixed values $\bar{z}^*$ and $\bar{z}_m^*$. Although this choice reduces computation time in the estimation process, the simplification comes at the cost of a non-negligible built-in inconsistency. When simulating the LRR model using $\bar{z}^*$ and $\bar{z}_m^*$, the means of the simulated z_t and $z_{m,t}$ series will be different from the fixed values.[1] The exactly identifying auxiliary model used by Calvet and Czellar (2015) is complex and global optimization is both computationally expensive and difficult to ensure. Bearing in mind the identification issues encountered in one-step estimations of the LRR model, we suggest a two-step approach instead, for which the auxiliary models are rapidly and reliably estimated.

Recognizing the inherent recursive structure of the LRR model, the two steps separate the estimation of the macroeconomic dynamics from that of the investor preference parameters, which is the key to obtain reliable estimates of the structural model parameters. Instead of working with a single auxiliary model, which would have the difficult task to capture all important model features, each estimation step

[1] For example, using the LRR model parameter values calibrated by BY, and $\bar{z}^* = 6.96$ and $\bar{z}_m^* = 5.95$, as chosen by Calvet and Czellar (2015), to simulate LRR model-implied data series with T=100k, we obtain a sample mean of the log P/C ratio equal to 5.87 and a sample mean of the log P/D ratio equal to 5.19. These differences are large in economic terms.

uses a specific auxiliary model tailored to account for the time series properties and asset pricing implications of the LRR model, respectively. The two-step indirect inference approach allows for different frequencies of the macroeconomic and financial data, which do not have to coincide with the LRR model-implied decision frequency of the representative investor. For the auxiliary model in the first estimation step, we adopt the heterogeneous autoregressive (HAR) specification proposed by Corsi (2009). The HAR approach allows for the use of past information over long horizons in a parsimonious way. This favorable feature avoids using a large number of higher-order autocovariances of consumption and dividend growth to extract the information about the small predictable growth component. The representative investor's preference parameters are estimated in the second step, for which we exploit the asset pricing implications of the LRR model. The two-step estimation strategy implies that standard theory of asymptotic inference is not applicable, such that we rely on a bootstrap method that makes use of the parametric nature of the LRR model instead. A Monte Carlo study documents the feasibility of the two-step indirect inference estimation strategy and reveals the estimation precision that can be expected using a sample size as is currently available for empirical analysis. The results emphasize that the quality of the macro parameter estimates is crucial to deliver precise preference parameter estimates. In an empirical application, we obtain estimates of the macro parameters that support the notion of a small persistent growth component, which is a crucial ingredient of the LRR asset pricing approach. The point estimates of the parameters that describe the investor's subjective time preference and risk aversion are economically plausible, while the estimate of the intertemporal elasticity of substitution (IES) is less than 1. However, the data are also compatible with an IES>1, which is a necessary condition for the ability of the LRR model to account for the prominent asset pricing puzzles. The confidence

intervals indicate that the estimation precision will inevitably be limited by the relatively short low-frequency macroeconomic data series. The empirical evidence in favor of the LRR model is therefore less conclusive than suggested by some previous studies.

The remainder of the chapter is organized as follows. Section 4.2 explains the two-step indirect inference estimation strategy. Section 4.3 provides the results of a Monte Carlo study. Section 4.4 describes the data. Section 4.5 presents empirical results. Section 4.6 contains concluding remarks.

4.2 Econometric methodology

4.2.1 Motivation and notation

This section outlines a two-step indirect inference estimation strategy that separates the estimation of the macro parameters $\boldsymbol{\xi}^M$ from that of the preference parameters $\boldsymbol{\xi}^P$. The approach allows for different sampling frequencies in each estimation step, which may also differ from the LRR model-implied decision frequency of the representative agent. To formalize the exposition, we use a notation that draws on the seminal papers by Gourieroux, Monfort, and Renault (1993) and Smith (1993).

The LRR model as presented in Section 2.2 implies a vector stochastic process for consumption and dividend growth $\boldsymbol{y}_M^{(d)}(\boldsymbol{\xi}^M) \equiv \{\boldsymbol{y}_{M,s}^{(d)}(\boldsymbol{\xi}^M), s \geq 1\}$, where $\boldsymbol{y}_{M,s}^{(d)} = \left(g_s^{(d)}, g_{d,s}^{(d)}\right)'$. The superscript (d) indicates that the sampling frequency of the process corresponds to the decision frequency. Moreover, the LRR model implies a vector stochastic process for the return of the market portfolio, risk-free rate, and P/D ratio, $\boldsymbol{y}_P^{(d)}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) \equiv \{\boldsymbol{y}_{P,s}^{(d)}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P), s \geq 1\}$, where $\boldsymbol{y}_{P,s}^{(d)} = \left(r_{m,s}^{(d)}, r_{f,s}^{(d)}, z_{m,s}^{(d)}\right)'$. Both processes are assumed to be stationary and ergodic for any $\boldsymbol{\xi}^M \in \boldsymbol{\Theta}^M \subset \mathcal{R}^9$ and $\boldsymbol{\xi}^P \in \boldsymbol{\Theta}^P \subset \mathcal{R}^3$, respectively. It may be necessary to consider a time aggre-

gation of the two model-implied processes to a lower frequency corresponding to that of the observed data. The correct time aggregation is crucial and non-trivial; the appropriate formulas provided by Calvet and Czellar (2015) are given in Appendix 4.A.1. We denote the model-implied processes that are time-aggregated to the base frequency (b) of the observed data as $\boldsymbol{y}_M^{(b)}(\boldsymbol{\xi}^M)$ and $\boldsymbol{y}_P^{(b)}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)$.

The indirect inference estimation strategy requires to generate numerically finite realizations $\{\boldsymbol{y}_{M,s}^{(b)}(\boldsymbol{\xi}^M)\}_{s=1}^S$ and $\{\boldsymbol{y}_{P,s}^{(b)}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)\}_{s=1}^S$ given $\boldsymbol{\xi}^M$ and $\boldsymbol{\xi}^P$. Section 2.2.4 describes how to simulate such LRR model-implied data. The frequency of the simulated time series is determined by the decision interval of the LRR investor, which BY assume to be one month; subsequent aggregation allows to transform the simulated data to a lower frequency, e.g. to match the frequency of the empirical data. Corresponding to the simulated, model-implied processes there are the observed vector processes $\boldsymbol{w}_M^{(b)} \equiv \{\boldsymbol{w}_{M,t}^{(b)}, t \geq 1\}$, where $\boldsymbol{w}_{M,t}^{(b)} = \left(g_t^{(b)}, g_{d,t}^{(b)}\right)'$, and $\boldsymbol{w}_P^{(b)} \equiv \{\boldsymbol{w}_{P,t}^{(b)}, t \geq 1\}$, where $\boldsymbol{w}_{P,t}^{(b)} = \left(r_{m,t}^{(b)}, r_{f,t}^{(b)}, z_{m,t}^{(b)}\right)'$. Of these processes, which are also assumed to be stationary and ergodic, we observe finite realizations $\{\boldsymbol{w}_{M,t}^{(b)}\}_{t=1}^T$ and $\{\boldsymbol{w}_{P,t}^{(b)}\}_{t=1}^T$. Indirect inference estimation is based on the assumption that there exists a unique set of parameters $\boldsymbol{\xi}_0^M \in \boldsymbol{\Theta}^M$ and $\boldsymbol{\xi}_0^P \in \boldsymbol{\Theta}^P$ such that the realizations of $\boldsymbol{w}_M^{(b)}$ and $\boldsymbol{w}_P^{(b)}$ on the one hand, and the realizations of $\boldsymbol{y}_M^{(b)}(\boldsymbol{\xi}_0^M)$ and $\boldsymbol{y}_P^{(b)}(\boldsymbol{\xi}_0^M, \boldsymbol{\xi}_0^P)$ on the other hand, are drawn from the same distribution.

The philosophy of indirect inference estimation and the inherently recursive LRR model structure suggests to perform the estimation of the macro parameters $\boldsymbol{\xi}^M$ and the estimation of the preference parameters $\boldsymbol{\xi}^P$ in two consecutive steps. The reasons are twofold. First, the separate indirect inference estimation of the macro parameters benefits from a simpler data simulation, because the solution for the endogenous model parameters is only required to simulate $\{\boldsymbol{y}_{P,s}^{(b)}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)\}_{s=1}^S$. It is not needed to obtain $\{\boldsymbol{y}_{M,s}^{(b)}(\boldsymbol{\xi}^M)\}_{s=1}^S$. Second, and more importantly, the auxiliary models to be

employed in each step can be tailored such that the diverse properties of the LRR model can be accounted for. Consider the macro dynamics in Equations (2.1)-(2.4), which only depend on $\boldsymbol{\xi}^M$, and in which the presence of two latent processes poses a challenge for the auxiliary model. It should be tractable, but also capture the intricate time series properties induced by these latent processes. The estimation of the investor preference parameters imposes different requirements on the auxiliary model. For that purpose, the LRR model-implied distributional properties of the market portfolio return and the risk-free rate should be reflected by the auxiliary model. Entangling the information about these diverse aspects—time series dynamics, asset pricing relations, and preferences—does not seem prudent: Monte Carlo experiments revealed that the joint estimation of all LRR model parameters yields unstable results. The advantage of a two-step indirect inference strategy that separates the estimation of $\boldsymbol{\xi}^M$ and $\boldsymbol{\xi}^P$ is that we can use specialized and customized auxiliary models in each step that are only required to capture the properties of $\boldsymbol{y}_{M,s}^{(b)}$ or $\boldsymbol{y}_{P,s}^{(b)}$ but not both.

4.2.2 First step: macro parameter estimation

The first indirect inference estimation step thus only deals with the estimation of the macro parameters $\boldsymbol{\xi}^M$. For that purpose, we must specify an auxiliary model that captures the properties of the LRR model-implied macro process $\boldsymbol{y}_M^{(b)}(\boldsymbol{\xi}^M)$. Let us collect the first-step auxiliary model parameters in the vector $\boldsymbol{\theta}^M \in \boldsymbol{\Xi}^M \subset \mathcal{R}^{k_M}$, where k_M is at least as large as the number of macro parameters, and presume that auxiliary parameter estimates $\hat{\boldsymbol{\theta}}^M$ can be obtained by maximizing the criterion function $Q_T^M(\{\boldsymbol{w}_{M,t}^{(b)}\}_{t=1}^T, \boldsymbol{\theta}^M)$.

The challenge for the first-step auxiliary model is to account for the predictable growth component x_t, which induces small but very persistent serial correlations in

the growth series. These deviations from i.i.d. growth let the asset pricing implications of the LRR model unfold. A parsimonious way to capture the autocorrelation structure of a persistent process is the HAR specification proposed by Corsi (2009). It is used in the realized volatility literature to capture the long-memory properties of squared and absolute returns by accounting for different sampling frequencies in an autoregressive model. To set up the first-step auxiliary model, we therefore use the following HAR specification for log consumption and dividend growth observed at the base frequency:[2]

$$\begin{bmatrix} g_t^{(b)} \\ g_{d,t}^{(b)} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \sum_{\iota=1}^{\tau} \mathbf{\Phi}_\iota L^\iota \begin{bmatrix} g_t^{(b)} \\ g_{d,t}^{(b)} \end{bmatrix} + \mathbf{\Phi}_{\tau+1} \begin{bmatrix} g_{t-1}^{(f(h_1))} \\ g_{d,t-1}^{(f(h_1))} \end{bmatrix} + \mathbf{\Phi}_{\tau+2} \begin{bmatrix} g_{t-1}^{(f(h_2))} \\ g_{d,t-1}^{(f(h_2))} \end{bmatrix} + \begin{bmatrix} \zeta_{1,t} \\ \zeta_{2,t} \end{bmatrix}, \quad (4.1)$$

where $\mathbf{\Phi}_\iota$ are parameter matrices and $\boldsymbol{\zeta}_t = (\zeta_{1,t}, \zeta_{2,t})'$ are orthogonal Gaussian white noise innovations. In an empirical application, the base frequency (b) could be quarterly (as in Hasseltoft, 2012) or annual (as in Constantinides and Ghosh, 2011). $f(h_1)$ and $f(h_2)$ denote lower frequencies that result from a time aggregation of the base frequency data over h_i periods. With a quarterly base frequency, we would use $h_1 = 4$ and $h_2 = 12$ to obtain annual and triannual data. The time aggregation of consumption and dividend growth is based on the formulas given in Appendix 4.A.1. Compared with a standard vector-autoregressive process, the HAR specification can account for the long-run impact of shocks to consumption and dividend growth in a parsimonious way, as the large required number of lagged growth rates gets replaced by few aggregates. The auxiliary parameters that result from the HAR specification are collected in the vector

$$\boldsymbol{\theta}^{HAR} = (c_1, c_2, \mathrm{vec}(\mathbf{\Phi}_1)', \ldots, \mathrm{vec}(\mathbf{\Phi}_{\tau+2})', \mathrm{vec}(\mathbf{\Sigma}_{\boldsymbol{\zeta}})')', \quad (4.2)$$

[2] We are grateful to George Tauchen for suggesting the use of the HAR specification as an auxiliary model.

where $\boldsymbol{\Sigma}_{\boldsymbol{\zeta}}$ is the covariance matrix of $\boldsymbol{\zeta}_t$. The flexibility of the first-step auxiliary model is enhanced by extending the auxiliary parameter vector to include the means and standard deviations of the two growth processes and their time-aggregates,

$$\boldsymbol{g}_t = \left(g_t^{(b)}, g_{d,t}^{(b)}, g_t^{(f(h_1))}, g_{d,t}^{(f(h_1))}, g_t^{(f(h_2))}, g_{d,t}^{(f(h_2))} \right)', \tag{4.3}$$

which we collect in the vectors $\boldsymbol{\mu_g}$ and $\boldsymbol{\sigma_g}$. The complete vector of first-step auxiliary parameters is then given by $\boldsymbol{\theta}^M = \left(\boldsymbol{\theta}^{HAR\prime}, \boldsymbol{\mu_g}', \boldsymbol{\sigma_g}' \right)'$. OLS regressions yield the estimates of $\boldsymbol{\theta}^{HAR}$, and sample moments are used to estimate $\boldsymbol{\mu_g}$ and $\boldsymbol{\sigma_g}$. Assuming for the auxiliary model that $\boldsymbol{w}_M^{(b)}$ is a Gaussian process (a natural assumption as the innovations in Equations (2.1)-(2.4) are i.i.d. $\mathcal{N}(0,1)$), the elements of $\hat{\boldsymbol{\theta}}_T^M$ can be interpreted as pseudo-maximum likelihood estimates, and the criterion Q_T^M as a pseudo-likelihood function.

The number of auxiliary parameters exceeds the number of macro parameters, so that we use the following first-step indirect inference estimator of $\boldsymbol{\xi}^M$:

$$\hat{\boldsymbol{\xi}}_T^M = \underset{\boldsymbol{\xi}^M \in \boldsymbol{\Theta}^M}{\operatorname{argmin}} \quad \Delta^M(\boldsymbol{\xi}^M)' \, \boldsymbol{W}_T^M \, \Delta^M(\boldsymbol{\xi}^M), \tag{4.4}$$

where $\Delta^M(\boldsymbol{\xi}^M) = \hat{\boldsymbol{\theta}}_T^M - \tilde{\boldsymbol{\theta}}_S^M(\boldsymbol{\xi}^M)$. $\tilde{\boldsymbol{\theta}}_S^M(\boldsymbol{\xi}^M)$ denotes the estimate of $\boldsymbol{\theta}^M$ that is obtained when the auxiliary parameters are estimated on simulated LRR model-implied data of sample size S, where S is chosen as a fixed multiple H of T. $\boldsymbol{W}_T^M$ is a symmetric and positive definite weighting matrix, $\boldsymbol{W}_T^M \underset{p}{\to} \boldsymbol{W}^M$, a non-stochastic positive definite matrix. The weighting matrix $\boldsymbol{W}_T^M$ can be used to enforce precise matches of elements of $\hat{\boldsymbol{\theta}}_T^M$ and $\tilde{\boldsymbol{\theta}}_S^M$.

Under the assumptions stated by Gourieroux et al. (1993), the first-step indirect inference estimator in Equation (4.4) is a consistent estimator of $\boldsymbol{\xi}_0^M$. In addition to

stationarity and ergodicity of the data generating processes, we have to assume that the criterion function $Q_T^M(\{\boldsymbol{w}_{M,t}^{(b)}\}_{t=1}^T, \boldsymbol{\theta}^M)$ converges uniformly and almost surely to a non-stochastic limit function $Q_\infty^M(F_0, \boldsymbol{\xi}_0^M, \boldsymbol{\theta}^M)$, where F_0 denotes the true distribution function of the fundamental innovations in Equations (2.1)-(2.4). Moreover, we have to assume that the limit function is continuous in $\boldsymbol{\theta}^M$ and has $\boldsymbol{\theta}_0^M$ as the unique maximum. Defining

$$b(F, \boldsymbol{\xi}^M) = \underset{\boldsymbol{\theta}^M \in \boldsymbol{\Xi}^M}{\operatorname{argmax}} \; Q_\infty^M(F, \boldsymbol{\xi}^M, \boldsymbol{\theta}^M), \tag{4.5}$$

we have $\boldsymbol{\theta}_0^M = b(F_0, \boldsymbol{\xi}^M)$. Consistency requires that the binding function

$$b(F_0, \cdot) : \boldsymbol{\xi}^M \to b(F_0, \boldsymbol{\xi}^M) \tag{4.6}$$

is injective and that $\frac{\partial b(F_0, \boldsymbol{\xi}_0^M)}{\partial \boldsymbol{\xi}^{M\prime}}$ is of full column rank.

While the rank condition is fulfilled, as can be assessed by simulation, the injectivity condition cannot be formally checked since the binding function is not available in closed form. Connections between auxiliary and structural parameters are obvious, though. The autoregressive parameter matrices $\boldsymbol{\Phi}$ should provide information about the persistence parameter ρ and the leverage ratio on expected consumption growth ϕ. The parameters c_1, c_2, and $\boldsymbol{\mu_g}$ are linked to the unconditional expected values of log consumption and dividend growth, μ_c and μ_d, while the second moments in $\boldsymbol{\Sigma_\zeta}$ and $\boldsymbol{\sigma_g}$ should contribute to the identification of the variance-scaling parameters φ_e and φ_d and the parameters of the stochastic volatility process. To assess the feasibility of the estimation approach and to provide simulation-based evidence on the injectivity of the binding function, we conduct a Monte Carlo study and check whether the indirect inference strategy can reliably recover the true structural parameters $\boldsymbol{\xi}_0^M$ when a large sample size is available (see Section 4.3.2).

Under the regularity conditions and assumptions stated by Smith (1993) and Gourieroux et al. (1993), the first-step indirect inference estimator $\hat{\boldsymbol{\xi}}_T^M$ in Equation (4.4) is asymptotically normal. As an alternative to using the large sample formulas, we rely on bootstrap-based inference, which we describe in Section 4.2.5.

4.2.3 Second step: preference parameter estimation

The second estimation step focuses on the preference parameters $\boldsymbol{\xi}^P$, taking $\hat{\boldsymbol{\xi}}_T^M$ as given, and uses an auxiliary model that aims to capture the asset pricing implications of the LRR model. The second-step auxiliary parameters are collected in the vector $\boldsymbol{\theta}^P \in \boldsymbol{\Xi}^P \subset \mathcal{R}^{k_P}$, where $k_P > 3$ (we use an over-identified auxiliary model), and estimates $\hat{\boldsymbol{\theta}}^P$ can be obtained by maximizing the criterion function $Q_T^P(\{\boldsymbol{w}_{P,t}^{(b)}\}_{t=1}^T, \boldsymbol{\theta}^P)$.

The LRR model-implied equations for the risk-free rate and the market equity premium (see Equations (2.20) and (A-16)) guide our selection of the second-step auxiliary parameters. The mean of the log risk-free rate $\mathbb{E}(r_f) = \mu_{r_f}$ should convey information about the subjective time preference δ, the propensity for intertemporal substitution ψ, and also precautionary savings due to risk aversion γ. The equity premium $\mu_{r_m^e} = \mathbb{E}(r_m - r_f)$—albeit a function of all three preference parameters—should primarily reflect relative risk aversion. To disentangle risk aversion from intertemporal substitution, we exploit that the contemporaneous relationship between the log P/D ratio and the log risk-free rate implied by the LRR model is predominantly determined by the IES but largely unaffected by the RRA coefficient, which should promote the identification of ψ.[3]

[3] Section 4.A.3 in the Appendix shows that the analytical expression of the covariance between r_f and z_m is dominated by ψ.

Moreover, Equation (2.13) implies that $\mathbb{E}(z_m) = \mu_{z_m}$ depends on all preference parameters, while the standard deviation of z_m (σ_{z_m}) only depends on γ and ψ, so using μ_{z_m} and σ_{z_m} as auxiliary parameters provides separate information about risk aversion and time preference. Including the standard deviations of the market excess return ($\sigma_{r_m^e}$) and the log risk-free rate (σ_{r_f}) among the set of auxiliary parameters lends further flexibility to the second-step auxiliary model. The complete vector of auxiliary model parameters then reads:

$$\boldsymbol{\theta}^P = \left(\beta, \alpha, \mu_{r_m^e}, \mu_{r_f}, \mu_{z_m}, \sigma_{r_m^e}, \sigma_{r_f}, \sigma_{z_m}\right)', \tag{4.7}$$

where β and α are the parameters of an orthogonal projection of z_m on r_f and a constant. The second-step auxiliary model parameters are estimated by sample moments and a linear regression of z_m on r_f. Specifying the auxiliary model such that $\boldsymbol{y}_P^{(b)}(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)$ is a Gaussian process, the sample moments and OLS estimates of the auxiliary parameters can then be conceived of as pseudo-maximum likelihood estimates.

The second-step indirect inference estimator is then given by:

$$\hat{\boldsymbol{\xi}}_T^P = \underset{\boldsymbol{\xi}^P \in \boldsymbol{\Theta}^P}{\operatorname{argmin}} \quad \Delta^P(\hat{\boldsymbol{\xi}}_T^M, \boldsymbol{\xi}^P)' \, \boldsymbol{W}_T^P \, \Delta^P(\hat{\boldsymbol{\xi}}_T^M, \boldsymbol{\xi}^P), \tag{4.8}$$

where $\Delta^P(\hat{\boldsymbol{\xi}}_T^M, \boldsymbol{\xi}^P) = \hat{\boldsymbol{\theta}}_T^P - \tilde{\boldsymbol{\theta}}_S^P(\hat{\boldsymbol{\xi}}_T^M, \boldsymbol{\xi}^P)$. $\hat{\boldsymbol{\theta}}_T^P$ denotes the estimate of the auxiliary model parameters $\boldsymbol{\theta}^P$ based on empirical data with T time series observations. $\tilde{\boldsymbol{\theta}}_S^P(\hat{\boldsymbol{\xi}}_T^M, \boldsymbol{\xi}^P)$ are the corresponding estimates obtained when the auxiliary parameters are estimated on simulated LRR model-implied data with sample size $S = HT$. This simulation takes $\hat{\boldsymbol{\xi}}^M$ as given and leaves it unchanged during optimization. $\boldsymbol{W}_T^P$ is a

symmetric and positive definite weighting matrix that can depend on the observed sample. $\boldsymbol{W}_T^P \underset{p}{\to} \boldsymbol{W}^P$, a non-stochastic positive definite matrix.

Generating LRR model-implied data during the second estimation step entails solving for the endogenous parameters in Equations (2.8) through (2.13). As pointed out in Section 2.2.3, the solution may not exist, which would cause the estimation to break down if the optimization algorithm probes inadmissible parameter combinations. Unfortunately, constrained indirect inference estimation as proposed by Calzorari, Fiorentini, and Sentana (2004) cannot be employed, because the constraint would not have to be imposed on the auxiliary model parameters but on the structural model parameters. Moreover, it is impossible to formulate explicit constraints that would ensure that only eligible (structural) parameter combinations are used. Our solution is to use a large penalty (we use 10^3) that is added to the value of the objective function whenever the optimization algorithm tries structural parameter values that would imply an unsolvable model.

Under the assumptions stated by Gourieroux et al. (1993), and using the consistent first-step estimator $\hat{\boldsymbol{\xi}}_T^M$ instead of $\boldsymbol{\xi}_T^M$ when generating LRR model-implied data, the second-step indirect inference estimator in Equation (4.8) is a consistent estimator of $\boldsymbol{\xi}_0^P$. The second-step binding function cannot be expressed in closed form, so an analytical check of the injectivity condition is not possible. To assess the feasibility of the second estimation step, we therefore extend the Monte Carlo study and check whether it is possible to reliably recover the true structural parameters $\boldsymbol{\xi}_0^P$ when a large sample is available.

Inference about the second-step estimator $\hat{\boldsymbol{\xi}}_T^P$ cannot rely on the standard asymptotic theory of indirect inference estimation. Section 4.2.5 explains how to obtain bootstrap inference on the parameters instead.

4.2.4 An alternative representation

The two-step indirect inference approach presented in the previous sections is equivalent to obtaining $\hat{\boldsymbol{\xi}}_T^M$ and $\hat{\boldsymbol{\xi}}_T^P$ as the solution of the following system of equations:

$$\begin{bmatrix} \frac{\partial \Delta^M(\hat{\boldsymbol{\xi}}^M)'}{\partial \boldsymbol{\xi}^M} & \mathbf{0} \\ \mathbf{0} & \frac{\partial \Delta^P(\hat{\boldsymbol{\xi}}^M, \hat{\boldsymbol{\xi}}^P)'}{\partial \boldsymbol{\xi}^P} \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_T^M & \mathbf{0} \\ \mathbf{0} & \boldsymbol{W}_T^P \end{bmatrix} \begin{bmatrix} \Delta^M(\hat{\boldsymbol{\xi}}^M) \\ \Delta^P(\hat{\boldsymbol{\xi}}^M, \hat{\boldsymbol{\xi}}^P) \end{bmatrix} = \mathbf{0}. \tag{4.9}$$

Note that this is not equivalent to stacking both auxiliary parameter vectors into one,

$$\Delta(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) = \begin{bmatrix} \Delta^M(\boldsymbol{\xi}^M) \\ \Delta^P(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P) \end{bmatrix}, \tag{4.10}$$

and using the indirect inference estimator

$$\begin{bmatrix} \hat{\boldsymbol{\xi}}_T^M \\ \hat{\boldsymbol{\xi}}_T^P \end{bmatrix} = \underset{\boldsymbol{\xi}^M, \boldsymbol{\xi}^P}{\operatorname{argmin}} \quad \Delta(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)' \, \boldsymbol{W}_T \, \Delta(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P), \tag{4.11}$$

which implies the first-order conditions

$$\begin{bmatrix} \frac{\partial \Delta^M(\hat{\boldsymbol{\xi}}^M)'}{\partial \boldsymbol{\xi}^M} & \frac{\partial \Delta^P(\hat{\boldsymbol{\xi}}^M, \hat{\boldsymbol{\xi}}^P)'}{\partial \boldsymbol{\xi}^M} \\ \mathbf{0} & \frac{\partial \Delta^P(\hat{\boldsymbol{\xi}}^M, \hat{\boldsymbol{\xi}}^P)'}{\partial \boldsymbol{\xi}^P} \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_T^M & \boldsymbol{W}_T^{12} \\ \boldsymbol{W}_T^{21} & \boldsymbol{W}_T^P \end{bmatrix} \begin{bmatrix} \Delta^M(\hat{\boldsymbol{\xi}}^M) \\ \Delta^P(\hat{\boldsymbol{\xi}}^M, \hat{\boldsymbol{\xi}}^P) \end{bmatrix} = \mathbf{0}. \tag{4.12}$$

Both Equation (4.9) and Equation (4.12) set linear combinations of auxiliary parameter matches to zero. Yet, while the weights of the linear combinations in Equation (4.12) lead to an inevitable interference of the auxiliary parameter matches $\Delta^P(\boldsymbol{\xi}^M, \boldsymbol{\xi}^P)$ with the estimation of the macro parameters $\boldsymbol{\xi}^M$, even if we use a block-diagonal weighting matrix such that $\boldsymbol{W}_T^{21} = \boldsymbol{W}_T^{12} = \mathbf{0}$, the weights in Equation (4.9) prevent the second-step auxiliary model from interfering with the estimation of the

macro parameters. The two-step indirect inference approach thus takes into account the findings presented in Chapter 3, which imply that the entanglement of macro and financial moment matches in a one-step GMM or SMM estimation of the LRR model should be avoided as it yields unreliable parameter estimates. Our experiences with alternative one-step indirect inference estimation strategies lead to the same conclusion.

4.2.5 Bootstrap inference

The two-step indirect inference approach implies that standard theory of asymptotic inference is not applicable. However, the LRR model structure permits the use of a bootstrap simulation to obtain parameter standard errors and to construct confidence intervals. The procedure can be characterized as a parametric residual bootstrap that works as follows.

After performing the two-step estimation on the empirical data, which yields the estimates $\hat{\boldsymbol{\xi}}^M$ and $\hat{\boldsymbol{\xi}}^P$, we independently draw $4 \times (T^* + L)$ standard normally distributed random variables to obtain realizations of the i.i.d. innovations $\{\eta_t\}_{t=1}^{T^*+L}$, $\{e_t\}_{t=1}^{T^*+L}$, $\{u_t\}_{t=1}^{T^*+L}$, and $\{w_t\}_{t=1}^{T^*+L}$ in Equations (2.1)–(2.4). The appropriate time series length T^* is determined by the number of observations and sampling frequency of the empirical data, as well as the assumed decision frequency of the investor. For example, the data used for our empirical application comprise T=271 quarterly observations. We assume a monthly decision frequency, such that T^*=813. The simulated innovations are used to generate time series of length T^*+L of LRR model-implied macro and financial variables, as described in Section 2.2.4. For that purpose, $\hat{\boldsymbol{\xi}}^M$ and $\hat{\boldsymbol{\xi}}^P$ serve as "true" parameters. The first L observations are discarded to mitigate the effect of the choice of starting values. We use $L = 100$ as a default. If the empirical data frequency is lower than the decision frequency, the

simulated time series are time-aggregated, using the formulas in Appendix 4.A.1, to match the empirical data frequency.

The two indirect inference estimation steps are then performed on the bootstrap sample. Data simulation and estimation are repeated R independent times, with new i.i.d. draws of standard normally distributed innovations, simulation of the LRR model variables, and two-step estimation performed on the simulated samples. The resulting sets of estimates $\{\hat{\boldsymbol{\xi}}^M_{(r)}\}_{r=1}^R$ and $\{\hat{\boldsymbol{\xi}}^P_{(r)}\}_{r=1}^R$ are used to compute parameter standard errors and to construct confidence intervals. The latter are obtained by the percentile method, which amounts to using the appropriate quantiles of the bootstrap distribution as upper and lower bounds (cf. Efron and Tibshirani, 1993). The bootstrap simulation is computationally intensive, such that fast and reliable auxiliary model estimation is even more important.

To assess its validity, we have to check the conditions under which the bootstrap is consistent, meaning that the bootstrap estimator of the distribution function (cdf) of the statistic of interest (here: one of the parameter estimates in $\hat{\boldsymbol{\xi}}^M$ or $\hat{\boldsymbol{\xi}}^P$) is uniformly close to the statistic's asymptotic cdf for large T. The formal definition and the conditions for consistency of the bootstrap are stated by Horowitz (2001).[4] Briefly, consistency requires that the cdf of the probability distribution from which the data are sampled and its bootstrap estimator are uniformly close to each other when T is large, and that suitable continuity conditions regarding the asymptotic cdf of the statistic of interest hold.

While the conditions for consistency cannot be formally checked in the present application, we argue that the proposed procedure is not subject to those issues that are known to provoke a failure of the bootstrap. As Horowitz (2001) notes, failures of the bootstrap are associated with heavy-tailed or dependent data, or true param-

[4] See Horowitz's (2001) Definition 2.1 and Theorem 2.1, originally formulated by Beran and Ducharme (1991).

eters that lie on the boundary of the parameter space. However, the i.i.d. draws of innovations from the standard normal distribution along with economically plausible LRR model parameters preclude generating heavy-tailed data. Moreover, the parametric residual bootstrap avoids drawing directly from the macro and financial data series, which may exhibit considerable serial dependence. Provided that the parameter estimates are consistent, the bootstrap estimate should therefore constitute a good approximation of the true cdf of the data for large T. Violations of the continuity assumption regarding the asymptotic cdfs of the parameter estimates are also not indicated. In particular, the aforementioned intricate parameter space should not affect the validity of the bootstrap. We do have to assume, however, that the LRR model is solvable in the neighborhood of the true parameters; in other words, we have to rule out that the true parameters lie on the boundaries of the admissible parameter space. It should also be noted that in the present application the bootstrap does not provide asymptotic refinement, as the statistics of interest—the elements of $\hat{\boldsymbol{\xi}}^M$ and $\hat{\boldsymbol{\xi}}^P$—are not pivotal.

4.3 Monte Carlo study

4.3.1 Design

The Monte Carlo study is designed to check the feasibility of the two-step indirect inference estimation strategy and to assess the estimation precision that can be expected when using empirically available sample sizes. For that purpose we generate 400 independent LRR model-implied data series of g, g_d, r_m, r_f, and z_m using as true parameter values the calibration by BY reported in Table 2.1, and perform the two-step indirect inference estimation on the simulated data. The calibrated values correspond to a monthly decision frequency. We assume that data and decision

frequency are identical, such that time aggregation is not required. The lengths of the simulated data series are T=275, 1k, and 100k, respectively.

As mentioned previously, an analytical validation of the assumptions for consistency, in particular of the injectivity of the binding functions, is not possible. The T=100k study should provide a substitute check whether the estimation strategy is viable such that it can recover the true parameters when using a large sample. Assuming a monthly sampling frequency, T=1k represents a large but not implausible sample size for an empirical application that relies on monthly data, while T=275 corresponds to the number of observations currently available at a quarterly frequency.

In the simulated economy, growth expectations are very persistent, ρ=0.979, which is pivotal for the asset pricing implications of the LRR model. On the other hand, the predictable growth component x_t is small, as a result of scaling consumption volatility σ_t by φ_e=0.044. Consumption growth expectations are leveraged into dividend growth expectations by ϕ=3. The expected values of consumption and dividend growth are identical, μ_c=μ_d=0.0015. However, dividend growth volatility is considerably larger than the volatility of consumption growth as $\varphi_d = 4.5$. Moreover, while the stochastic variance process is highly persistent, it is not very volatile due to the fact that σ_w is small. The LRR investor has positive time preferences as δ is close to but smaller than one. The risk aversion parameter γ=10 lies at the upper bound of economic plausibility.[5] The intertemporal elasticity of substitution is larger than one (ψ=1.5), which is a crucial factor for the ability of the LRR model to resolve the equity premium and risk-free rate puzzle, as shown in Section 2.3.

The estimates of ρ, μ_c, and μ_d are restricted to values between 0 and 1 by means of a logit transform and the estimates for σ, ϕ, φ_e, and φ_d are restricted to positive

[5] The canonical reference is Mehra and Prescott (1985), who consider a range for γ between 1 and 10 to be plausible.

values by an exponential transform of the unrestricted parameters. We use H=10 for T=100k and T=1k, following Smith (1993) who recommends using $S = 10T$ as the lower bound for which the inflation of the variance covariance matrix induced by simulation error becomes sufficiently small. For the T=275 study we use H=100, as initial estimations indicated that the stability of the numerical optimization benefits from a larger simulated sample size. To ensure robust, yet fast optimization of the indirect inference objective functions, we use the Nelder-Mead (1965) algorithm. To provide a safeguard against false convergence close to favorably chosen starting values, optimizations are started from initial values distant from the known true parameters.[6]

4.3.2 Monte Carlo results: macro parameters

As an initial feasibility check we tried to estimate all macro parameters $\boldsymbol{\xi}^M$ using the first-step auxiliary model described in Section 4.2.2 based on simulated samples with T=100k. These experiments revealed that the subset of the macro parameters $\boldsymbol{\xi}^{M_*} = (\mu_c, \mu_d, \rho, \varphi_e, \sigma, \phi, \varphi_d)'$ could be reliably recovered by maximizing the first-step objective function (4.4) but not the SV parameters ν_1 and σ_w, for which we obtain vastly different estimates $\hat{\nu}_1$ and $\hat{\sigma}_w$ when using different initial values. This result raises the concern that the first-step auxiliary model may be unable to identify ν_1 and σ_w. Extending the auxiliary model in various directions does not alleviate the

[6] This is a safety measure to prevent reporting overly optimistic results, but it makes the optimization more difficult. As a result, the optimization could not be successfully accomplished for some replications, in particular for small T. The optimization algorithm either exceeded the maximum number of iterations, or converged to implausible values (more than ten times larger than the true value in absolute terms). We consider these cases as failed estimation attempts and exclude them in the tables and plots that summarize the simulation study results. In the second estimation step, an estimation is also classified as failed if the LRR model is not solvable at the parameter values to which the optimization converges. In an empirical study, such problematic data could receive special treatment, by increasing the maximum number of iterations, or by using alternative optimization algorithms. Such an expensive handling is not tenable in a Monte Carlo study.

problem. Including a heterogeneous autoregressive conditional heteroscedasticity model, as discussed in Appendix 4.A.4, does not allow to identify ν_1 and σ_w either. Figure 4.1 suggests a possible explanation. It shows that in BY's calibrated LRR economy the volatility of volatility is indeed very small, which suggests that the signal-to-noise ratio may be too low to estimate ν_1 and σ_w.

These findings suggest an alternative estimation strategy, in which the conditional variance σ_t^2 is predicted by its unconditional expectation, $\mathbb{E}(\sigma_t^2) = \sigma^2$. Estimating σ^2 within the first-step indirect inference estimation procedure entails replacing σ_t^2 by σ^2 when generating LRR model-implied data. While ν_1 and σ_w are not estimated in the first step, the true data-generating process still exhibits stochastic volatility: we do not change the model, but deliver an alternative estimate of σ_t^2. The Monte Carlo study investigates the consequences for the quality of the other parameter estimates.

In each replication we therefore estimate the reduced set of macro parameters $\boldsymbol{\xi}^{M_*}$ by minimizing the indirect inference objective function in Equation (4.4). The auxiliary parameter vector $\boldsymbol{\theta}^M$ is constructed as described in Section 4.2.2, and with the following customization. In the HAR specification in Equation (4.1) we account for consumption and dividend growth on the annual and the triannual level by choosing $h_1 = 12$ and $h_2 = 36$. The first few monthly lags should be particularly informative for the estimation of the persistence parameter ρ, so we set $\tau = 6$. Initial estimations indicated that a precise match of the means and standard deviations of consumption and dividend growth can enhance the precision of the estimates of μ_c and μ_d and that of the variance-scaling parameters φ_e and φ_d, which prove difficult to estimate. This match is accomplished by using a diagonal weighting matrix $\boldsymbol{W}_T^M$ with values of 1 on the main diagonal, except for the entries that correspond to the first two elements of $\boldsymbol{\mu_g}$ and $\boldsymbol{\sigma_g}$, which receive a large weight (10^4).

As a benchmark, we also perform a GMM estimation that relies on moment matches inspired by the studies of Hasseltoft (2012) and Constantinides and Ghosh (2011). For that purpose we exploit that the population moments of log consumption and dividend growth implied by the LRR model can be expressed as functions of the parameter vector $\boldsymbol{\xi}^M$. The GMM strategy is based on exact identification using the seven moments given in Appendix 4.A.2.

Table 4.1 reports the medians and root mean squared errors (RMSEs) of the first-step indirect inference estimates (Panels A and B) and the GMM estimates (Panel C). Figure 4.2 illustrates the indirect inference results using kernel estimates. In addition, Appendix 4.A.5 provides a comparison of the two estimation approaches regarding the precision of model-implied moment matches. The T=100k results show that the proposed indirect inference estimation strategy is feasible and works well. Biases and the RMSEs shrink, there are no estimation failures, and the bell-shaped kernel estimates center closely around the true parameter values. Using σ^2 instead of σ_t^2 when simulating LRR model-implied data does not affect the quality of the other parameter estimates. Panel B of Table 4.1 shows the results assuming that ν_1 and σ_w are known. These results do not differ qualitatively from those in Panel A, which reports the results when σ_t^2 is predicted by σ^2. This conclusion holds for all simulated sample sizes. The estimation precision is different across macro parameters. Not surprisingly, the estimates of the parameter φ_e, which scales the variance of the latent expected growth component x_t, and ϕ, the parameter that leverages the effect of x_t on expected dividend growth, are less precise. However, compared with the GMM results reported in Panel C, the indirect inference RMSEs are much smaller. A considerably smaller RMSE is also obtained for the persistence parameter ρ. Figure 4.3 shows that the distribution of the indirect inference estimate $\hat{\rho}$ is much more closely centered around the true value than the GMM counterpart.

Precise estimation becomes more difficult using smaller sample sizes, as indicated by the increase in the RMSE and the wider distribution of the estimates around the true parameters. Efficiency varies across parameters in a similar way as in the large sample. As could be expected from the 100k results, the critical parameters φ_e and ϕ prove most difficult to estimate precisely. However, we do assert that the optimization of the indirect inference objective function yields reliable results in that the algorithm converges to the same minimum, independent of the starting values.[7]

We conclude that the indirect inference strategy is reliable. Using the currently available sample sizes one should not expect a high estimation precision for some of the structural parameters, though. We believe that the simulation study draws a realistic picture of the estimation precision that can be expected in an empirical study.

4.3.3 Monte Carlo results: preference parameters

Preference parameter estimates $\hat{\boldsymbol{\xi}}^P$ are obtained by minimizing the objective function in Equation (4.8) using the second-step auxiliary parameter vector $\boldsymbol{\theta}^P$ in Equation (4.7) with $\boldsymbol{W}_T^P = \boldsymbol{I}_8$.[8] To evaluate the performance of the second estimation step independently of the precision of the first-step input, we first perform the estimation of $\boldsymbol{\xi}^P$ assuming that all macro parameters $\boldsymbol{\xi}^M$ are known. Panel B in Table 4.2 reports median, RMSE, and 95% confidence bounds of the resulting preference parameter estimates. The T=100k study again serves as a check of the validity of the estimation strategy, which is corroborated by shrinking RMSEs, tight

[7] The GMM estimation strategy does not provide such robustness. Varying the starting values yields different results for smaller samples. Hence we refrain from reporting the GMM results for the smaller sample sizes.

[8] Starting values for the optimization are found by an initial grid search to mimic the recommended procedure in an empirical application. Again we purposefully avoid starting from the known true values to prevent the danger of false convergence to a point conveniently near the true parameters.

confidence bounds around the true parameters, and the absence of estimation failures.

It is a noteworthy result that the preference parameters can be efficiently estimated also for the smaller sample sizes. Although the second-step auxiliary model is simple and easy to estimate, and despite the more complicated data simulation procedure that requires a model solution, the indirect inference strategy delivers precise preference parameter estimates.

Moreover, Panel A of Table 4.2 shows that predicting the conditional volatility σ_t^2 by its unconditional expectation σ^2 does not impair the estimation of $\boldsymbol{\xi}^P$. If the interest lies in estimating the preference parameters, it therefore suffices to focus on estimating the unconditional volatility σ^2. This conclusion is based on BY's calibrated model economy but it should also extend beyond it. We are using unconditional moments of the equity premium and the risk-free rate to estimate the investor's subjective time preference, risk aversion, and IES. It is plausible that the knowledge of the dynamics of conditional volatility does not substantially improve the precision of the preference parameter estimation.

To assess the efficiency that can be expected when $\boldsymbol{\xi}^M$ is unknown, we also estimate $\boldsymbol{\xi}^P$ based on the first-step estimates of $\boldsymbol{\xi}^{M_*}$. The results are reported in Panel C of Table 4.2. The T=100k results corroborate our conjecture that the two-step estimation strategy is able to recover the true parameters as RMSEs decrease and confidence bounds narrow, while Figure 4.4 shows that the bell-shaped kernel estimates center closely around the true values. Compared to the case in which the macro parameters are known, the (asymptotic) efficiency is inevitably reduced. For the smaller sample sizes, the subjective discount factor can still be estimated accurately, whereas the RRA and IES estimates become less precise. Table 4.2 shows that the RMSEs are influenced by some large estimates that produce the

right-skewed kernel estimates for $\hat{\gamma}$ and $\hat{\psi}$ depicted in Figure 4.4. The mass of the distributions remains centered around the true values, though.

Overall, the Monte Carlo study shows that the second-step auxiliary model is suitable for estimating the preference parameters. Yet, the results also emphasize the importance of precise macro parameter estimates as an input for the second estimation step.

4.4 Data

The empirical application of the two-step estimation strategy is based on quarterly U.S. data from 1947Q2 to 2014Q4. The construction of the data base follows closely Beeler and Campbell (2012). Consumption growth is computed from real personal consumption per capita of non-durable goods and services obtained from the Bureau of Economic Analysis. The market portfolio return, dividend growth, and the price-dividend ratio are calculated for the CRSP value-weighted market portfolio. Conversions into real terms are performed using the consumer price index obtained from the Bureau of Labor Statistics. For the calculation of the risk-free rate proxy, we use the three-month nominal T-bill yield from the CRSP database. Following Beeler and Campbell (2012), we approximate the ex-ante risk-free rate by using a forecast for the ex-post real rate, where the predictors are the quarterly T-bill yield and the average of quarterly log inflation across the past year. Figure 4.5 shows time series plots of the data.

Dividend payments occur irregularly, such that the quarterly dividend growth series depicted in Panel (b) of Figure 4.5 is quite erratic.[9] The time series exhibits

[9] Dividend growth is less volatile at the annual frequency, but in that case the number of observations is small. For example, Constantinides and Ghosh (2011) base their econometric analyses of the LRR model on 79 annual observations.

a strong, negative first-order autocorrelation that cannot be accounted for by the dividend growth process in Equation (2.3). We deal with this problem by following Hasseltoft (2012) in taking the average of the current period's log dividend growth and that of the previous three quarters to obtain the smoothed dividend growth series in Panel (d) of Figure 4.5. Descriptive statistics of the variables used in the empirical application are provided in Table 4.3.

4.5 Empirical Results

To apply the two-step indirect inference estimation strategy to these data we follow BY and assume a monthly decision frequency. Time aggregation of the simulated monthly data to the quarterly frequency of the empirical data is performed as described in Section 4.2.1 and the auxiliary models are set up as described in Sections 4.2.2 and 4.2.3. As supported by the results of the Monte Carlo study, we replace σ_t^2 by $\sigma^2 = \mathbb{E}(\sigma_t^2)$ when generating LRR model-implied data. We include annual and triannual aggregates in the HAR model in Equation (4.1) by setting h_1=4 and h_2=12, and we use S=100k to mitigate simulation inaccuracy. Apart from that, the specification of the auxiliary models is the same as in the simulation study. Table 4.4 reports the parameter point estimates along with the bounds of the 95% bootstrap confidence intervals.[10]

Table 4.4 shows that the estimates of the macro parameters are consistent with the LRR paradigm in that they corroborate the existence of a small persistent growth component. The lower bound of the 95% confidence interval for $\hat{\varphi}_e$ is distinctly greater than zero ($\varphi_e = 0$ would imply i.i.d. growth processes), and the 95% confidence interval for the difference $\hat{\rho} - \hat{\varphi}_e$ does not include zero ($\varphi_e = \rho$ would

[10] The selection criteria for successful bootstrap replications that are included in the calculation of the confidence bounds are the same as for the simulation study.

imply an AR(1) consumption growth process). The estimate $\hat{\rho} = 0.991$ indicates a strong persistence of growth expectations. With an estimated base volatility of $\hat{\varphi}_e \cdot \hat{\sigma} \cdot \sqrt{12}$=0.053%, the growth component is indeed small when compared to the estimated base volatility of consumption growth innovations, $\hat{\sigma} \cdot \sqrt{12}$=0.83%, and when compared to the estimated base volatility of dividend growth innovations, $\hat{\varphi}_d \cdot \hat{\sigma} \cdot \sqrt{12}$=2.54%.

Moreover, the estimate $\hat{\phi} = 5.14$ indicates that the effect of expected consumption growth on dividend growth is leveraged, as conjectured. The estimates $\hat{\mu}_c$ and $\hat{\mu}_d$ imply plausible mean growth rates of 2.0% p.a. and 2.3% p.a. for consumption and dividends, respectively. We assert that these estimates are robust in that we obtain the same values and the same minimum of the first-step indirect inference objective function for very different starting values. The first-stage estimation problem is well-defined and we are confident that the reported estimates represent the global minimum of the objective function. The same result holds for the estimates from each bootstrap replication.

The estimation precision reflected in the bootstrap confidence intervals and its variation across parameters corresponds to the Monte Carlo results. While the 95% confidence bands contain plausible parameter values, one may consider the intervals to be rather wide. However, we believe that they provide a realistic view on the estimation precision, given the small sample size and the intricate properties of the estimated stochastic processes. Shephard and Harvey (1990) note that it is very difficult to distinguish between a purely i.i.d. process and one that incorporates a small persistent component. Bansal, Gallant, and Tauchen (2007a, henceforth BGT), who estimate the LRR model by EMM, discuss identification issues that entail the necessity to calibrate several time series parameters. In the light of these results, it is quite remarkable that some econometric studies have reported very precise estimates of all

LRR macro parameters. These papers propose LRR model extensions that contain even more structural parameters, yet they employ estimation methods that are less sophisticated than the EMM estimation strategy applied by BGT. Identification of such nonlinear models is difficult to ascertain, and convergence to a local optimum of the objective function can easily be overlooked. The asymptotic inference that is often applied requires that the neighborhood of the estimates is well-defined, which may well be the case at a local optimum, and which consequently yields favorably small parameter standard errors.

The estimation precision for the macro parameters is unlikely to be improved by exploiting the LRR model's asset pricing implications. Consumption and dividend growth are exogenous processes that are independent of investor preferences. Accordingly, only the model-implied asset pricing relations should help to estimate the investor preference parameters, the first-step auxiliary model is not useful for that purpose. The second-step auxiliary model must accomplish the difficult task of disentangling risk aversion from intertemporal substitution and it does not seem prudent to burden it with the additional task of identifying parameters of an intricate vector stochastic process. Asset pricing relationships inevitably interfere with the estimation of the macro parameters in any one-step estimation strategy, and there is evidence that such an entanglement hampers the econometric analysis of the LRR model.

The second-step estimate for the subjective discount factor results in positive time preferences ($\hat{\delta} = 0.99998$), and the estimate of the RRA coefficient implies reasonable risk preferences ($\hat{\gamma} = 11.8$). These point estimates are comparable to the calibrated values in Table 2.1 and the estimation precision corresponds to what could be expected from the Monte Carlo study. The 95% confidence interval is narrow for $\hat{\delta}$ and wide for $\hat{\gamma}$. As can be seen from Table 4.5, the estimates are also comparable to

those obtained by BGT, who report a narrower confidence band for the risk aversion coefficient. However, BGT resort to calibrating the third preference parameter, the intertemporal elasticity of substitution, because of identification problems. In particular, they report that the EMM objective function is flat in ψ, and therefore, instead of estimating the IES, they calibrate $\psi = 2$, which is a crucial choice. As noted by BY, the ability of the LRR model to account for the large equity premium and relatively small risk-free rate hinges on an IES larger than 1.

Table 4.4 shows that our IES point estimate is smaller than 1 ($\hat{\psi} = 0.29$), although the 95% confidence interval includes values larger than 1 as well. Our IES estimate is comparable to that reported by Calvet and Czellar (2015) who also estimate all three preference parameters of the LRR model (see Table 4.5).[11]

While there is an ongoing debate about whether a plausible IES should be smaller or larger than 1, empirical estimates tend to be quite small (cf. the results by Yogo (2006) reported in Table 4.5). As noted by Beeler and Campbell (2012), an estimate of the IES can be obtained from the slope of a regression of log consumption growth on the log risk-free rate and a constant. Using our empirical data, the OLS estimate amounts to $\hat{\psi}_{OLS} = 0.23$, which is comparable to the indirect inference estimate but considerably smaller than the calibrated IES. To provide evidence that the OLS approach yields a reasonable IES estimate, we run the regression on simulated LRR model data, for which we use BY's calibration as true parameter values. Based on a sample size of T=100k, we obtain $\hat{\psi}_{OLS} = 1.446$ on a monthly level, and $\hat{\psi}_{OLS} = 1.443$ for quarterly aggregates. Both estimates are close to the true parameter value $\psi = 1.5$.

[11] We note that some of the estimation results reported by Calvet and Czellar (2015) are not unanimously favorable for the LRR model. Besides some implausible macro parameter estimates like negative expected dividend growth, Table 4.5 shows that their estimates imply negative time preferences and a very high risk aversion.

While the point estimates of the macro parameters support the LRR paradigm, as we find that consumption growth indeed features a small, highly persistent component, the evidence regarding the asset pricing implications of the LRR model is less conclusive. Even though our estimated confidence band for $\hat{\psi}$ includes also values greater than 1, our point estimate is rather small. With $\psi < 1$, the LRR model no longer produces the desired asset pricing implications. As Table 4.5 shows, some results reported in previous literature are more favorable for the LRR model paradigm in that the reported IES are greater than 1. However, some of those values are conveniently calibrated, whereas others result from one-step GMM or SMM estimation attempts that should be considered with caution, as argued previously. It is unlikely that the identification problems addressed by BGT can be resolved by replacing efficient moment matches by ad hoc choices.

In line with the results obtained in the simulation study, our estimates have rather wide confidence bounds. The low estimation precision is likewise reflected by the implications of the estimates regarding the essential moments characterizing the LRR model variables. Table 4.6 compares the means and standard deviations of the empirical data to their counterparts implied by the point estimates reported in Table 4.4. To illustrate the precision of the moments, the related LRR model-implied distributions of the means and standard deviations resulting from the bootstrap distributions are also included. While the moments of the growth rates g and g_d are matched precisely, certain moments of the financial data differ notably from the LRR model implications entailed by the point estimates. This discrepancy must be attributed to the two-step estimation, in which the first-step parameter estimates for the macro dynamics pre-determine key features of the financial variables due to the LRR model structure. The wide range of quantiles of the LRR model-implied moments fits into the general picture in that any econometric analysis must be

based on a relatively small number of time series observations, which inevitably limits estimation precision. The 95% confidence interval for the RRA estimate $\hat{\gamma}$ ranges from 2.2 to 110.3, the confidence interval for $\hat{\psi}$ encompasses values between 0.22 and 1.20. Table 4.5 shows that the EMM approach by BGT enhances the estimation precision for the RRA coefficient, however, the advantage comes at the cost of having to calibrate the IES. We conclude that both our Monte Carlo study and our empirical application draw a realistic picture of the efficiency that can be attained when estimating the parameters of the LRR model based on the currently available data.

4.6 Conclusion

Asset pricing with long-run consumption risk has become an important paradigm in financial economics, but the estimation of the parameters of the LRR model is challenging due to its intricate macroeconomic growth processes and asset pricing properties. LRR model-implied data can be simulated, so that provided an appropriate auxiliary model is available, indirect inference estimation presents itself as an obvious econometric strategy. However, the attempt to simultaneously estimate the parameters that govern the model's consumption and dividend growth processes and those that describe investor preferences entails problems. BGT resort to calibrating several model parameters, among them the all-important intertemporal elasticity of substitution, as a result of identification problems.

This chapter proposes a two-step indirect inference estimation strategy that employs two separate, customized auxiliary models. It exploits the recursive nature of the LRR model, in which dividend and consumption growth processes determine the model-implied asset pricing relations but not vice versa. The first-step auxil-

iary model therefore focuses on estimating the parameters that describe the time series properties of the observable and latent macroeconomic growth processes. The second-step auxiliary model is designed to identify the three dimensions of investor preferences: subjective time preference, propensity for intertemporal substitution, and risk aversion, taking the first-step estimates as given. A bootstrap procedure is used to assess the estimation precision.

The discussion provided by BGT indicates that identification issues should be a major concern for any econometric analysis of the LRR model. Formal checks of the conditions for consistency are unavailable and Monte Carlo studies that explore the validity and efficiency of the various estimation strategies are scarce. However, such an analysis is an important reality check. Some recent contributions that rely on GMM or SMM report remarkably precise estimates of the complete set of LRR model parameters. Given the unresolved identification problems, this is a counterintuitive result, in particular since BGT employ the efficient method of moments that should be superior to any GMM/SMM estimation strategy.

Our Monte Carlo study ascertains that the two-step indirect inference approach yields reliable results and it documents the efficiency that can be expected using empirically available sample sizes. Moreover, it shows that using an auxiliary model that captures the LRR model's asset pricing implications, the investor preference parameters can be efficiently estimated, provided that accurate estimates of the macro parameters are available. The parameters of the stochastic volatility process prove difficult to estimate, and instead of relying on estimates of weakly identified SV parameters, we propose to estimate the conditional volatility σ_t^2 by its unconditional expected value. The simulation study shows that concentrating out stochastic volatility in this way does not hamper the estimation of the other model parameters.

The empirical application contributes to literature that investigates whether measurement or specification of consumption growth is responsible for the apparent empirical failure of the consumption-based asset pricing paradigm (prominent examples are Parker and Julliard, 2005; Yogo, 2006; Savov, 2011). In that vein, the LRR approach assumes that consumption growth is not an i.i.d. process, but that there exists a small persistent growth component that matters for long-horizon investors. Calibrations show that when accounting for such a predictable growth component, the data generated from a suitably parametrized LRR model can replicate some key properties of the data. Our econometric analysis investigates what model parameter values are compatible with the empirical data, recognizing limits of identification and the information content of a small sample.

In our empirical application, we do find support of the LRR paradigm, in particular there is evidence for the existence of a small persistent growth component, a plausible and precisely estimated subjective time preference parameter, and a reasonable point estimate of the risk aversion coefficient. A point estimate of the intertemporal elasticity of substitution below unity is a less favorable result, though. The IES is usually calibrated to values greater than 1, as the ability of the LRR model to explain the prominent asset pricing puzzles requires that the substitution effect dominates the income effect, which in turn requires a large IES. The estimation precision is in line with the Monte Carlo study results. The available data series are relatively short, which entails wide confidence bounds. The confidence interval for the IES does include values greater than 1, so the LRR paradigm can still be considered as compatible with the data. The evidence in favor of the LRR approach is, however, not as conclusive as implicated by some previous studies.

Our Monte Carlo study shows that when high-quality macro parameter input is available, the preference parameters can be efficiently estimated, even for smaller

samples. It is rather the first estimation step, in particular the estimation of the parameters of the latent growth process, for which it would be desirable to enhance estimation precision, which should in turn increase the efficiency of the preference parameter estimation. We thus conclude that efforts to improve the accuracy of the preference parameter estimates—which are, from an economic point of view, the most interesting ones—should focus on increasing the estimation precision of the macroeconomic parameters of the LRR model.

4.A Appendix

4.A.1 Time aggregation of LRR processes

The formulas for the time aggregation of the LRR model variables over h periods provided by Calvet and Czellar (2015) are as follows:

$$g_t^{(f(h))} = \ln \frac{\sum_{i=(t-1)h+1}^{th} \exp\left[\sum_{j=(t-1)h+1}^{i} g_j^{(b)}\right]}{1 + \sum_{i=(t-2)h+2}^{(t-1)h} \exp\left[-\sum_{j=i}^{(t-1)h} g_j^{(b)}\right]}, \tag{4.13}$$

$$g_{d,t}^{(f(h))} = \ln \frac{\sum_{i=(t-1)h+1}^{th} \exp\left[\sum_{j=(t-1)h+1}^{i} g_{d,j}^{(b)}\right]}{1 + \sum_{i=(t-2)h+2}^{(t-1)h} \exp\left[-\sum_{j=i}^{(t-1)h} g_{d,j}^{(b)}\right]}, \tag{4.14}$$

$$z_{m,t}^{(f(h))} = z_{m,th}^{(b)} + \sum_{i=(t-1)h+1}^{th} g_{d,i}^{(b)} - \ln\left[\sum_{i=(t-1)h+1}^{th} \exp\left(\sum_{j=(t-1)h+1}^{i} g_{d,j}^{(b)}\right)\right], \tag{4.15}$$

$$r_{m,t}^{(f(h))} = \sum_{i=(t-1)h+1}^{th} r_{m,i}^{(b)}, \tag{4.16}$$

$$r_{f,t}^{(f(h))} = \sum_{i=(t-1)h+1}^{th} r_{f,i}^{(b)}. \tag{4.17}$$

4.A.2 Theoretical moments of log consumption and dividend growth

The LRR model implies the following theoretical moments, which are matched with their empirical counterparts to obtain GMM estimates of $\boldsymbol{\xi}^{M_*}$ as an alternative to the first-step indirect inference estimation:

$$\mathbb{E}(g_t) = \mu_c, \tag{4.18}$$

$$\mathbb{E}(g_{d,t}) = \mu_d, \tag{4.19}$$

$$\mathbb{E}(g_t^2) = \mu_c^2 + \frac{\varphi_e^2\sigma^2}{1-\rho^2} + \sigma^2, \tag{4.20}$$

$$\mathbb{E}(g_{d,t}^2) = \mu_d^2 + \phi^2\frac{\varphi_e^2\sigma^2}{1-\rho^2} + \varphi_d^2\sigma^2, \tag{4.21}$$

$$\mathbb{E}(g_{d,t}g_t) = \mu_c\mu_d + \phi\frac{\varphi_e^2\sigma^2}{1-\rho^2}, \tag{4.22}$$

$$\mathbb{E}(g_{t+1}g_t) = \mu_c^2 + \rho\frac{\varphi_e^2\sigma^2}{1-\rho^2}, \tag{4.23}$$

$$\mathbb{E}(g_{t+2}g_t) = \mu_c^2 + \rho^2\frac{\varphi_e^2\sigma^2}{1-\rho^2}. \tag{4.24}$$

4.A.3 Identification of the IES in the second-step auxiliary model

The expression for the log risk-free rate in Equation (2.20) can be written as:

$$r_{f,t} = A_{0,f} + A_{1,f}x_t + A_{2,f}\sigma_t^2, \tag{4.25}$$

where $A_{0,f}$ collects all terms of the right-hand side of Equation (2.20) that do not depend on either of the two state variables, $A_{1,f}$ collects all terms related to x_t, and $A_{2,f}$ collects all terms of Equation (2.20) that depend on σ_t^2. It can be shown that $A_{0,f}$ and $A_{2,f}$ depend on all three preference parameters, while $A_{1,f}$ depends only on ψ:

$$A_{1,f} = \left[1-\theta+\frac{\theta}{\psi}-(1-\theta)A_1(1-\kappa_1\rho)\right] = \left[1-\theta+\frac{\theta}{\psi}-(1-\theta)(1-\frac{1}{\psi})\right] = \frac{1}{\psi}. \tag{4.26}$$

Using the expression for $z_{m,t}$ in Equation (2.13), the contemporaneous covariance of $z_{m,t}$ and $r_{f,t}$ is given by:

$$\text{Cov}(z_{m,t}, r_{f,t}) = A_{1,m}A_{1,f}\text{Var}(x_t) + A_{2,m}A_{2,f}\text{Var}(\sigma_t^2), \tag{4.27}$$

where $A_{1,m}$ is given in Equation (2.18), and $A_{2,m}$ is given in Equation (2.19). Equation (2.18) shows that of the three preference parameters only ψ affects $A_{1,m}$.

For economically plausible parameter values, such as the BY calibration, the expression for $\text{Var}(x_t)$ is several orders of magnitude larger than $\text{Var}(\sigma_t^2)$. Hence, the covariance of $z_{m,t}$ and $r_{f,t}$ is dominated by the term $A_{1,m}A_{1,f}\text{Var}(x_t)$, which only depends on ψ but not on δ and γ. The influence of the subjective discount factor and the RRA coefficient on the covariance of $z_{m,t}$ and $r_{f,t}$ is negligible. The identification of the IES is thus facilitated by the slope parameter of a contemporaneous regression of $z_{m,t}$ on $r_{f,t}$, which is thus included in the second-step auxiliary parameter vector.

4.A.4 A HARCH approach for SV estimation: Discussion

It is obvious that the persistence ν_1 and the volatility σ_w of the stochastic variance process are not well represented by the HAR model specified in Section 4.2.2. The properties of the SV process could be better accounted for by a separate autoregressive model that captures those features of the SV process. We try to estimate the SV parameters together with the remainder of the macro parameters by extending the macro auxiliary model by an ARCH-type model that captures the autoregressive pattern in the squared residuals $\boldsymbol{\zeta}_t^2$ of the HAR model in Equation (4.1).

In the spirit of the HAR model, we construct a heterogeneous autoregressive conditional heteroscedasticity (HARCH) model, in which the squared residuals of the HAR regression are modeled as functions of their own lags and their aggregates in order to capture the persistence in the squared residuals in a parsimonious way:

$$\begin{bmatrix} \zeta_{1,t}^2 \\ \zeta_{2,t}^2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \sum_{\iota=1}^{\tau} \boldsymbol{\Psi}_\iota L^\iota \begin{bmatrix} \zeta_{1,t}^{2(b)} \\ \zeta_{1,t}^{2(b)} \end{bmatrix} + \boldsymbol{\Psi}_{\tau+1} \begin{bmatrix} \zeta_{1,t-1}^{2(f(h_1))} \\ \zeta_{1,t-1}^{2(f(h_1))} \end{bmatrix} + \boldsymbol{\Psi}_{\tau+2} \begin{bmatrix} \zeta_{1,t-1}^{2(f(h_2))} \\ \zeta_{1,t-1}^{2(f(h_2))} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}. \tag{4.28}$$

For the identification of the volatility of volatility parameter σ_w, we complement the parameters of the HARCH model by the covariance matrix $\boldsymbol{\Sigma}_\epsilon$ of the HARCH model residuals $\boldsymbol{\epsilon}_t = (\epsilon_{1,t}, \epsilon_{2,t})'$. These additional parameters are then, jointly with the means and standard deviations of the aggregates, added to the auxiliary model parameter vector for the macro parameter estimation. The full macro auxiliary model parameter vector thus reads:

$$\begin{aligned}\boldsymbol{\theta}^M = (&c_1, c_2, \text{vec}(\boldsymbol{\Phi}_1)', \ldots, \text{vec}(\boldsymbol{\Phi}_{\tau+2})', \text{vec}(\boldsymbol{\Sigma}_{\boldsymbol{\zeta}})', \boldsymbol{\mu_g}', \boldsymbol{\sigma_g}', \\ &a_1, a_2, \text{vec}(\boldsymbol{\Psi}_1)', \ldots, \text{vec}(\boldsymbol{\Psi}_{\tau+2})', \text{vec}(\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})', \boldsymbol{\mu}_{\boldsymbol{\zeta}^2}', \boldsymbol{\sigma}_{\boldsymbol{\zeta}^2}')'. \end{aligned} \tag{4.29}$$

We use a large sample size, T=10k, to find out whether this extended auxiliary model can identify all macro parameters $\boldsymbol{\xi}^M$=$(\mu_c, \mu_d, \rho, \varphi_e, \nu_1, \sigma_w, \sigma, \phi, \varphi_d)'$. However, we observe a strong starting value-dependence in the parameter estimates for ν_1 and, in particular, for σ_w. Including these parameters in the estimation process, using the extended auxiliary model, even renders impossible a reliable estimation of the other parameters. This leads us to the conclusion that the additional information that is supposed to identify the SV parameters rather introduces noise into the auxiliary model, which not only leads to starting value-dependent results for the newly added parameters ν_1 and σ_w but also for the other parameters.

It is intuitively clear that the coefficients of the lagged squared residuals should contain information about ν_1 and that the covariance matrix of the resulting residuals should contain information about σ_w. However, our results suggest that the information in the squared residuals is dominated by the unobserved shocks to consumption and dividend growth η_{t+1} and u_{t+1}, which are much larger than σ_t. Put differently, the signal-to-noise ratio in the conditional volatility is too low and therefore we cannot extract sufficient information about the SV parameters from the

HAR model residuals. We used several variations of the auxiliary model, excluding the aggregates, excluding their means and standard deviations, however, this did not change the results.

Therefore, we rely on the approach of estimating the time-varying σ_t^2 by its unconditional mean $\mathbb{E}(\sigma_t^2) = \sigma^2$. The results in Tables 4.1 and 4.2 show that this replacement hardly changes the estimation results for the remainder of the LRR model parameters, both in the macro parameter estimation and in the preference parameter estimation.

4.A.5 Moment matches for GMM and indirect inference estimation

We use GMM estimation as a benchmark for our HAR model approach, because moment-based estimation is a standard method applied in other empirical studies, e.g. by Constantinides and Ghosh (2011), Bansal et al. (2012b), and, with simulated moments, by Hasseltoft (2012). The difference in the quality of the parameter estimates is assessed and illustrated in Table 4.1 and Figure 4.3. In addition, we compare the stylized facts of the data to the model-implied stylized facts.

To conduct this comparison for the Monte Carlo study, we compare the true moment vector implied by the true parameter values (as calibrated by Bansal and Yaron, 2004) to the distributions of moments implied by the 400 estimated parameter sets obtained from the GMM approach and the indirect inference estimation in the Monte Carlo study. Panel A of Table 4.7 shows the comparison between the true moment vector and the distribution of the moments (2.5%, 50%, and 97.5% quantiles) implied by the GMM estimates, whereas in Panel B the true moment vector is compared to the distribution of moments according to the HAR-based

indirect inference estimation. As a goodness-of-fit measure, we additionally compute the RMSE between the estimated moments and the true moment values.

The results clearly reflect the properties of the different estimation approaches. In the GMM approach, 7 moments ($\mathbb{E}(g)$, $\mathbb{E}(g_d)$, $\mathbb{E}(g^2)$, $\mathbb{E}(g_d^2)$, $\mathbb{E}(g \cdot g_d)$, $\mathbb{E}(g_t \cdot g_{t-1})$, $\mathbb{E}(g_t \cdot g_{t-2})$) are exactly matched to estimate 7 parameters. Given that the sample size used is T=10k, we should expect the data moments to be reasonably close to the theoretical values. Therefore, the estimates from this approach should imply values very close to the moment values resulting from the true parameters for the 7 moments used in the GMM estimation. As a consequence, it is not surprising that for the first 5 moments in Table 4.7—which are among the 7 moments used in the GMM approach—the GMM estimation results imply a closer moment match (both in terms of median and extreme quantiles) than the indirect inference estimation.

By contrast, the GMM estimation procedure does not account well for the slow-moving long-run risk component. The autocorrelation structure of consumption and dividend growth is much better accounted for in the indirect inference estimation, in which the auxiliary model is designed to capture the persistence of the latent growth process. This is reflected in the closer match between the auto-moments of consumption and dividend growth implied by the true parameters and the HAR model, whereas the values of the auto-moments in the GMM approach are too small with wide upper and lower bounds. To obtain precise estimates of the parameters that determine the defining feature of the LRR model, the indirect inference approach based on a HAR auxiliary model is thus more suitable than the standard GMM benchmark.

Comparing the RMSEs, we find that the HAR model performs relatively similar to the GMM estimation regarding the first and second moments, where the GMM approach yields slightly smaller RMSEs. However, for the autocorrelation structure,

in particular for dividend growth, the moments implied by the indirect inference estimates have a clearly better fit compared to the GMM approach. This is mainly due to the difference in estimation quality of ρ and ϕ, which is pronounced, as can be seen from Table 4.1.

Table 4.1: Monte Carlo results: first-step estimates.
The table reports the medians (in italics) and the RMSE (normal font) of the first-step macro parameter estimates obtained in the Monte Carlo study. The last column contains the number of successfully estimated replications $\tilde{R}$.

	μ_c	μ_d	ρ	φ_e	σ	ϕ	φ_d	$\tilde{R}$
true values	0.0015	0.0015	0.979	0.044	0.0078	3	4.5	
Panel A: Indirect inference, σ_t^2 predicted by σ^2								
T=275	*0.0016*	*0.0023*	*0.946*	*0.0745*	*0.0075*	*3.38*	*4.57*	245
	0.0013	0.0034	0.265	0.1198	0.0013	3.07	1.95	
T=1k	*0.0015*	*0.0018*	*0.965*	*0.0555*	*0.0080*	*2.97*	*4.36*	348
	0.0006	0.0017	0.218	0.0642	0.0005	1.85	0.24	
T=100k	*0.0015*	*0.0015*	*0.980*	*0.0430*	*0.0078*	*2.95*	*4.50*	400
	0.0001	0.0002	0.004	0.0046	0.0000	0.17	0.02	
Panel B: Indirect inference, ν_1 and σ_w known								
T=275	*0.0016*	*0.0023*	*0.938*	*0.0736*	*0.0076*	*3.45*	*4.54*	252
	0.0012	0.0034	0.328	0.1176	0.0015	3.69	1.06	
T=1k	*0.0015*	*0.0017*	*0.967*	*0.0496*	*0.0079*	*3.00*	*4.33*	347
	0.0006	0.0016	0.274	0.0696	0.0005	2.48	0.27	
T=100k	*0.0015*	*0.0015*	*0.981*	*0.0429*	*0.0078*	*2.94*	*4.50*	400
	0.0001	0.0002	0.004	0.0046	0.0000	0.17	0.02	
Panel C: GMM								
T=100k	*0.0015*	*0.0015*	*0.958*	*0.0620*	*0.0078*	*2.83*	*4.51*	397
	0.0001	0.0002	0.110	0.0559	0.0000	1.62	0.04	

Table 4.2: Monte Carlo results: second-step estimates.
The table reports medians (in italics) and RMSE (normal font) along with the 95% confidence bounds (in brackets), of the second-step indirect inference parameter estimates obtained in the Monte Carlo study. $\tilde{R}$ denotes the number of successfully estimated replications.

	Panel A		**Panel B**		**Panel C**	
	$\boldsymbol{\xi}^{M_*}$ known		$\boldsymbol{\xi}^{M}$ known		$\boldsymbol{\xi}^{M_*}$ estimated	
δ=0.998						
T=275	*0.9979*	0.0005	*0.9979*	0.0005	*0.9973*	0.0017
	[0.9969	0.9987]	[0.9969	0.9987]	[0.9938	1.0004]
T=1k	*0.9980*	0.0002	*0.9980*	0.0002	*0.9975*	0.0015
	[0.9975	0.9984]	[0.9976	0.9985]	[0.9945	0.9998]
T=100k	*0.9980*	0.0000	*0.9980*	0.0000	*0.9980*	0.0002
	[0.9979	0.9980]	[0.9980	0.9980]	[0.9977	0.9984]
γ=10						
T=275	*9.7*	2.0	*9.5*	1.9	*13.4*	19.8
	[5.6	13.7]	[5.6	12.9]	[3.5	72.9]
T=1k	*10.5*	1.2	*10.2*	1.0	*12.8*	14.5
	[8.3	12.6]	[8.2	11.9]	[5.7	55.4]
T=100k	*10.3*	0.3	*10.0*	0.1	*10.0*	1.1
	[10.1	10.6]	[9.8	10.2]	[8.3	12.5]
ψ=1.5						
T=275	*1.50*	0.02	*1.50*	0.02	*2.18*	3.71
	[1.46	1.55]	[1.46	1.55]	[0.42	11.36]
T=1k	*1.51*	0.02	*1.51*	0.02	*2.12*	2.68
	[1.48	1.54]	[1.48	1.54]	[0.86	10.65]
T=100k	*1.51*	0.01	*1.51*	0.01	*1.48*	0.16
	[1.51	1.51]	[1.51	1.51]	[1.20	1.85]
$\tilde{R}$						
T=275	400		400		169	
T=1k	400		400		284	
T=100k	400		400		400	

Table 4.3: Data descriptives.
The table reports descriptives of the empirical data. The data are sampled at a quarterly frequency and range from 1947Q2 to 2014Q4. AC(1) is the first-order autocorrelation. A four-quarter moving average of the raw log dividend growth time series is used to obtain g_d.

	mean	**std. dev.**	**AC(1)**
log consumption growth g	0.0048	0.0051	0.3116
log dividend growth g_d	0.0066	0.0247	0.4443
log market return r_m	0.0176	0.0825	0.0840
log risk-free rate r_f	0.0017	0.0045	0.9138
log price-dividend ratio z_m	3.4979	0.4217	0.9804

Table 4.4: Estimation results from the empirical application.
The table reports two-step indirect inference estimates obtained from the empirical data along with upper and lower bootstrap 95% confidence bounds obtained by the percentile method.

	μ_c	μ_d	ρ	φ_e	σ	ϕ	φ_d	δ	γ	ψ
estimate	0.0017	0.0019	0.991	0.0643	0.0024	5.14	3.06	0.99998	11.8	0.29
lower b.	0.0011	0.0000	0.757	0.0220	0.0002	2.72	1.68	0.98399	2.2	0.22
upper b.	0.0033	0.0088	1.000	0.2687	0.0029	8.96	16.29	1.00036	110.3	1.20

Table 4.5: Comparison of preference parameter estimates.
The table reports point estimates of the preference parameters and the bounds of the 95% confidence intervals (in brackets). The two-step indirect inference results are compared with the results reported in the studies by Yogo (2006), Bansal, Gallant, and Tauchen (2007a) (BGT), Constantinides and Ghosh (2011) (CG), Hasseltoft (2012), Bansal, Kiku, and Yaron (2012b) (BKY), Calvet and Czellar (2015) (CC), and with the calibrated values chosen by Bansal and Yaron (2004) (BY). Calibrated parameters are indicated by (c). The confidence bounds for the other studies are computed using the reported standard errors. Sample size and data frequency are given in the last column.

	$\hat{\delta}$	$\hat{\gamma}$	$\hat{\psi}$	T/Freq.
Two-step ind. inference	0.99998	11.8	0.29	271/Q
	[0.98399 1.00036]	[2.2 110.3]	[0.22 1.20]	
Yogo (2006)	0.9000	191.4	0.024	204/Q
	[0.7922 1.0078]	[93.7 289.2]	[0.006 0.042]	
BGT (2007)	0.9996	7.1	2	73/Y
	[0.9989 1.0002]	[-0.3 14.6]	(c)	
CG (2011)	0.968	9.3	1.41	79/Y
	[0.8563 1.0797]	[-0.1 18.8]	[-4.35 7.17]	
Hasseltoft (2012)	0.9992	6.8	2.51	223/Q
	(c)	[3.6 9.9]	[1.06 3.96]	
BKY (2012)	0.9989	7.4	2.05	80/Y
	[0.9969 1.0009]	[4.4 10.5]	[0.40 3.70]	
CC (2015)	1.0081	27.1	0.20	247/Q
	[1.0034 1.0129]	n.a.	[0.04 0.36]	
BY (2004)	0.9980	10	1.5	70/Y
	(c)	(c)	(c)	

Table 4.6: Implications of the empirical parameter estimates and their distributions.
The table reports means and standard deviations of the observable LRR model variables g, g_d, z_m, r_m, and r_f implied by the point estimates obtained in the empirical application and the corresponding bootstrap distribution. The first column contains the means and standard deviations of the empirical data for comparison. All quantities computed relate to a quarterly frequency. Since the time-aggregation of the moments of interest is non-trivial, the parameter estimates are used to simulate LRR model-implied data for 10^6 months that are subsequently aggregated to the quarterly level before computing estimates of the respective moments.

	data	**model**	**model-implied quantiles**				
			0.005	0.025	0.5	0.975	0.995
$\mathbb{E}(g)$	0.0048	0.0050	0.0003	0.0043	0.0071	0.0186	0.0266
$\mathbb{E}(g_d)$	0.0066	0.0056	0.0000	0.0000	0.0097	0.0436	0.0810
$\mathbb{E}(z_m)$	3.4979	3.6305	2.2770	2.8461	3.3585	3.6370	5.3209
$\mathbb{E}(r_m)$	0.0176	0.0322	0.0283	0.0363	0.0451	0.0927	0.1184
$\mathbb{E}(r_f)$	0.0017	0.0262	0.0243	0.0296	0.0395	0.0890	0.1156
$\sigma(g)$	0.0051	0.0049	0.0008	0.0036	0.0056	0.0101	0.0145
$\sigma(g_d)$	0.0247	0.0208	0.0159	0.0175	0.0234	0.0382	0.0553
$\sigma(z_m)$	0.4217	0.1153	0.0105	0.0202	0.1061	0.2303	1.2759
$\sigma(r_m)$	0.0825	0.0315	0.0164	0.0177	0.0283	0.0473	0.1146
$\sigma(r_f)$	0.0045	0.0120	0.0000	0.0022	0.0142	0.0308	0.0456

Table 4.7: Comparison between macro moments for GMM and HAR estimates. The table compares the moments implied by the true parameters to the distribution of the moments implied by the simulation study results for T=10k. Panel A contains the moments for the GMM results, Panel B uses the results from the indirect inference estimation based on the HAR model. The first column holds the moment values implied by the true parameters. The subsequent columns hold the 2.5%, the 50%, and the 97.5% quantiles of the same moments obtained from the parameter estimates, respectively. The last column holds the RMSE for all moments. All values are scaled by a factor of 10^3.

	true moment	**quantiles**			**RMSE**
		2.5%	50%	97.5%	
Panel A: GMM					
$\mathbb{E}(g)$	1.5000	1.1831	1.5016	1.8224	0.1670
$\mathbb{E}(g_d)$	1.5000	0.3731	1.5781	2.6515	0.5893
$\mathbb{E}(g^2)$	0.0659	0.0625	0.0661	0.0702	0.0020
$\mathbb{E}(g_d^2)$	1.2598	1.1830	1.2611	1.3370	0.0381
$\mathbb{E}(g \cdot g_d)$	0.0108	0.0046	0.0109	0.0176	0.0033
$\mathbb{E}(g_t \cdot g_{t-12})$	0.0044	0.0017	0.0038	0.0060	0.0014
$\mathbb{E}(g_t \cdot g_{t-24})$	0.0040	0.0016	0.0032	0.0060	0.0014
$\mathbb{E}(g_t \cdot g_{t-36})$	0.0036	0.0016	0.0029	0.0060	0.0014
$\mathbb{E}(g_t \cdot g_{t-48})$	0.0033	0.0015	0.0028	0.0060	0.0014
$\mathbb{E}(g_t \cdot g_{t-60})$	0.0030	0.0015	0.0028	0.0060	0.0014
$\mathbb{E}(g_{d,t} \cdot g_{d,t-12})$	0.0220	0.0009	0.0116	0.0667	0.0425
$\mathbb{E}(g_{d,t} \cdot g_{d,t-24})$	0.0176	0.0007	0.0076	0.0667	0.0415
$\mathbb{E}(g_{d,t} \cdot g_{d,t-36})$	0.0141	0.0005	0.0059	0.0667	0.0405
$\mathbb{E}(g_{d,t} \cdot g_{d,t-48})$	0.0115	0.0005	0.0051	0.0667	0.0396
$\mathbb{E}(g_{d,t} \cdot g_{d,t-60})$	0.0094	0.0004	0.0045	0.0667	0.0388
Panel B: Indirect inference with HAR					
$\mathbb{E}(g)$	1.5000	1.2037	1.5296	1.8640	0.1751
$\mathbb{E}(g_d)$	1.5000	0.4378	1.6900	2.8173	0.6183
$\mathbb{E}(g^2)$	0.0659	0.0626	0.0663	0.0705	0.0020
$\mathbb{E}(g_d^2)$	1.2598	1.1794	1.2581	1.3337	0.0382
$\mathbb{E}(g \cdot g_d)$	0.0108	0.0074	0.0114	0.0175	0.0076
$\mathbb{E}(g_t \cdot g_{t-12})$	0.0044	0.0022	0.0047	0.0064	0.0010
$\mathbb{E}(g_t \cdot g_{t-24})$	0.0040	0.0020	0.0042	0.0060	0.0010
$\mathbb{E}(g_t \cdot g_{t-36})$	0.0036	0.0019	0.0038	0.0058	0.0010
$\mathbb{E}(g_t \cdot g_{t-48})$	0.0033	0.0018	0.0035	0.0056	0.0010
$\mathbb{E}(g_t \cdot g_{t-60})$	0.0030	0.0017	0.0033	0.0055	0.0010
$\mathbb{E}(g_{d,t} \cdot g_{d,t-12})$	0.0220	0.0042	0.0224	0.0326	0.0067
$\mathbb{E}(g_{d,t} \cdot g_{d,t-24})$	0.0176	0.0017	0.0179	0.0274	0.0061
$\mathbb{E}(g_{d,t} \cdot g_{d,t-36})$	0.0141	0.0014	0.0150	0.0247	0.0056
$\mathbb{E}(g_{d,t} \cdot g_{d,t-48})$	0.0115	0.0012	0.0126	0.0224	0.0053
$\mathbb{E}(g_{d,t} \cdot g_{d,t-60})$	0.0094	0.0011	0.0105	0.0207	0.0051

Figure 4.1: Consumption growth and stochastic volatility.
The figure displays simulated data of length T=1k for log consumption growth g_t and stochastic volatility σ_t using the parameter values given in Table 2.1. The figure also depicts the unconditional volatility $\sigma = \sqrt{\mathbb{E}(\sigma_t^2)}$.

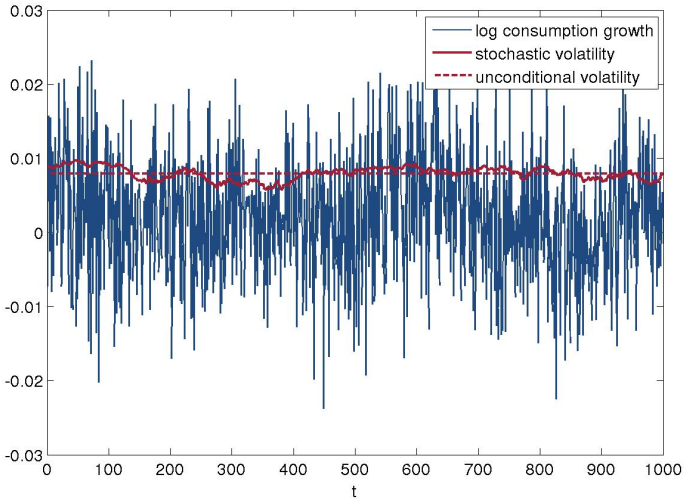


Figure 4.2: Monte Carlo results: distribution of first-step indirect inference estimates.
The figure shows kernel estimates across different simulated sample sizes. The beta kernel proposed by Chen (1999) is used with the bandwidth selector by Silverman (1986) adjusted for variable kernels. Vertical lines indicate the positions of the true parameters.

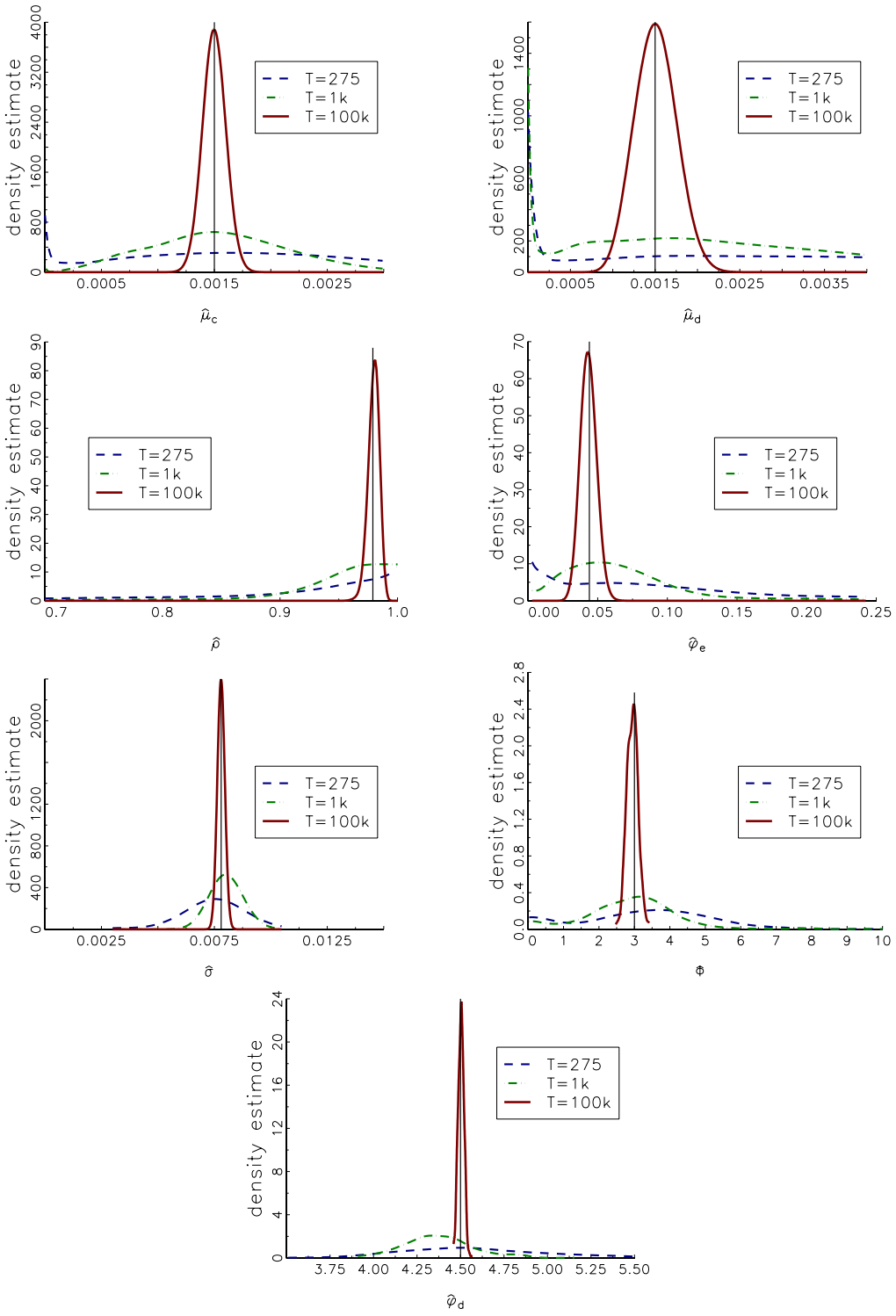


Figure 4.3: Monte Carlo results: asymptotic efficiency of indirect inference vs. GMM.

The figure shows kernel estimates of the LRR parameter estimate $\hat{\rho}$ implied by the indirect inference estimation strategy and GMM. The sample size is T =100k. The beta kernel proposed by Chen (1999) is used with the bandwidth selector by Silverman (1986) adjusted for variable kernels. The vertical line indicates the position of the true parameter.

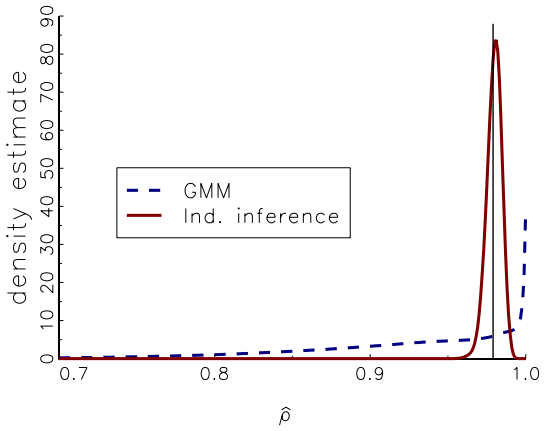


Figure 4.4: Monte Carlo results: distribution of second-step indirect inference estimates.
The figure displays kernel estimates for $\hat{\delta}$, $\hat{\gamma}$, and $\hat{\psi}$ obtained in the second estimation step. The first-step indirect inference estimates of $\boldsymbol{\xi}^{M_*}$ are taken as given. Vertical lines indicate the positions of the true parameters.

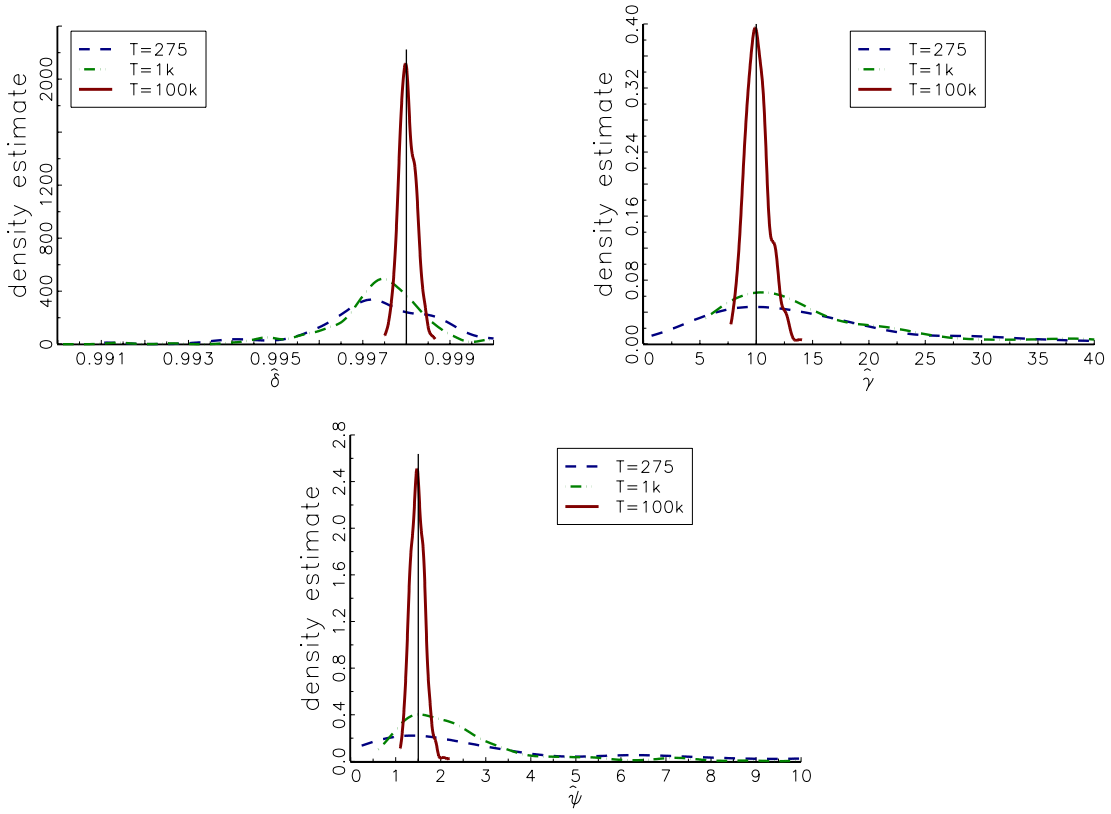
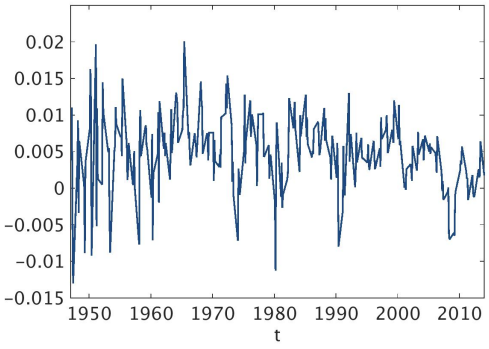
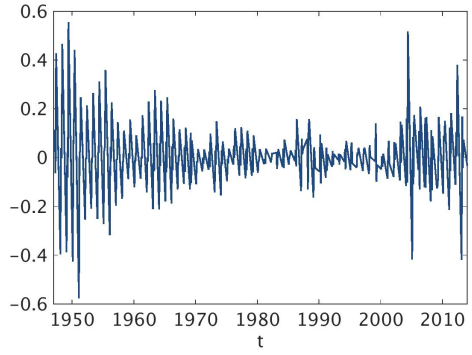


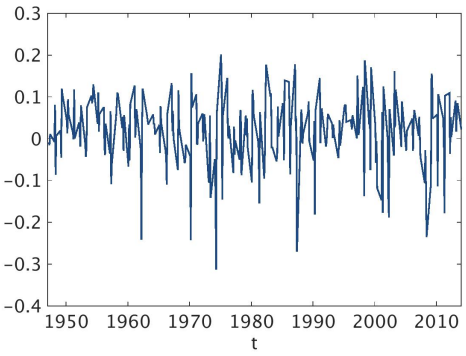
Figure 4.5: Empirical data series.
The figure displays the time series used in the empirical application. The sample period is 1947Q2 to 2014Q4.



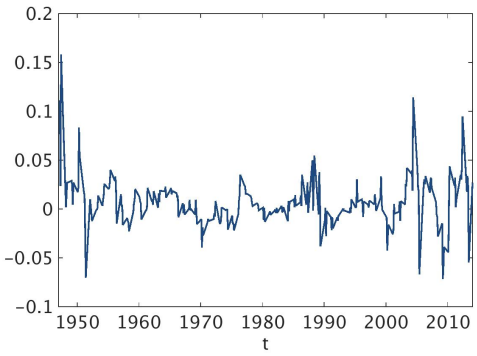
(a) log consumption growth g_t



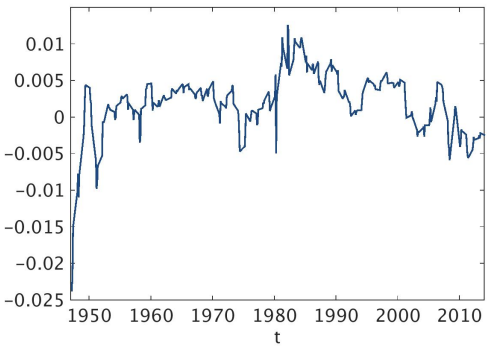
(b) log dividend growth (raw)



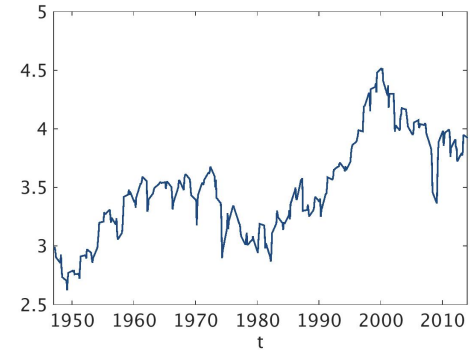
(c) log market return $r_{m,t}$



(d) log dividend growth moving avg. $g_{d,t}$



(e) log risk-free rate $r_{f,t}$



(f) log price-dividend ratio $z_{m,t}$

Chapter 5

Filtering-based maximum likelihood estimation of the long-run risk model

5.1 Introduction

Previous attempts to estimate the long-run risk model revealed serious methodological issues and low estimation precision of the existing econometric approaches, as discussed in Chapters 3 and 4. The estimation of the stochastic volatility parameters of the model has proven to be difficult or even impossible. The imprecise estimates of the representative investor's preference parameters have been attributed to the lack of efficiency in the parameter estimates for the macro sub-model. The three-step estimation strategy suggested in this chapter increases efficiency of the macro parameter estimates by resorting to asymptotically efficient maximum likelihood estimation, a novel approach among the existing econometric analyses of the LRR model. Despite the presence of latent variables, ML estimation is possible through

the application of filtering methods. An extensive Monte Carlo study demonstrates the ability of the estimation approach to identify all structural parameters of the LRR model, and it illustrates the estimation precision that can be expected in an empirical application with the presently available sample size. Using monthly U.S. data from 1947 to 2014, the LRR model is finally estimated for empirical data.

As mentioned in Chapter 2, Bansal and Yaron calibrated the LRR model in their seminal paper to demonstrate that it can reproduce the stylized facts of post-war U.S. data, and thereby provides a solution for the notorious asset pricing puzzles (cf. Mehra and Prescott, 1985; Weil, 1989). Since then, multiple studies have endeavored to come up with an estimation strategy that can identify some or all LRR model parameters. Bansal et al. (2007a) provide the first estimation results for a cointegrated LRR model variant using the efficient method of moments. However, key parameters like the intertemporal elasticity of substitution are calibrated. As detailed in Chapter 2, a number of other studies apply moment-based approaches, using either analytical or simulated moments (cf. Constantinides and Ghosh, 2011; Bansal et al., 2012b; Hasseltoft, 2012), while Calvet and Czellar (2015) suggest an indirect inference estimation approach. Schorfheide, Song, and Yaron (2014) conduct a Bayesian estimation of an extended LRR model that also uses filtering methods. They estimate the model at a monthly frequency, but several parameters are set to fixed values and their choices of prior distributions are somewhat narrow for important parameters, such as the investor's risk aversion.

The above estimation approaches employ a wide range of econometric techniques, yet all attempt to estimate the model in one step. However, the previous chapters have shown that the recursive model structure, consisting of an independent macroeconomic basis and a representative investor, whose choices are influenced by the macroeconomy but not vice versa, should be accounted for also in the estimation

process. A one-step estimation entangles macro data and preference parameter estimates, adapting the macroeconomic dynamics such that they best fit the financial data and the investor's preferences, instead of keeping the macroeconomy independent of the representative agent's decision making process, as implied by the LRR model.

The three-step estimation approach follows the idea to adhere to the model structure also in the estimation process. In a first step, the parameters that determine the dynamics of the macroeconomy are estimated by a maximum likelihood approach. The latent variables required for the ML estimation are obtained by applying a Kalman filter. For that purpose, the LRR macro dynamics are cast in state-space form. Due to nonlinearities in the innovations, the SV parameters cannot be identified by a Kalman filter-based approach, which constitutes an optimal *linear* filtering method. However, this does not affect the estimation of the remainder of the parameters, since the use of the Kalman filter does not imply any assumptions about the stochastic volatility parameters. In the second step, the SV parameters are estimated by using a particle filter-based maximum likelihood approach. Unlike the Kalman filter, which is limited to forecasting and updating the conditional mean only, the particle filter yields an estimate of the conditional distribution of the filtered process at each point in time. This feature allows to estimate the parameters of the stochastic variance process, which scales the conditional volatility of the macroeconomy. In the third step, the preference parameters are estimated using the indirect inference approach suggested in Chapter 4, as a closed-form likelihood function is not available for the asset pricing model of the LRR model. This approach has been shown to perform well, given high-quality macro parameter estimates. Overall, using the Kalman filter within a maximum likelihood approach yields more precise estimates for the macro parameters than previous comparable moment-based ap-

proaches. In particular, the use of the particle filter in the second-step estimation results in excellent estimates for the SV parameters, which have been found to be difficult to identify in previous studies. For larger samples, the estimation precision of the preference parameters is also good, while for small samples, the estimates of the preference parameters have rather wide confidence bounds. In the empirical application, the point estimate of the subjective discount factor is close to but below 1, the IES estimate is greater than 1, and the estimate for the risk aversion lies well above 10—the upper bound of plausible values, according to Mehra and Prescott (1985).

The remainder of the chapter is structured as follows. Section 5.2 presents the LRR macro sub-model in state-space form. In Section 5.3, the three-step estimation approach is explained in detail. Section 5.4 provides evidence of the validity and feasibility of the estimation strategy by means of a Monte Carlo study. Section 5.5 describes the data used in the empirical application, the results of which are presented in Section 5.6. Section 5.7 concludes.

5.2 LRR macro model in state-space form

The LRR model structure is described in detail in Chapter 2. To apply a maximum likelihood estimation to the macro sub-model, it is convenient to cast the dynamics presented in Section 2.2.1 in a state-space representation. The equations of the financial variables given in Section 2.2.2 remain unchanged.

The LRR macro dynamics comprise two observable growth processes, the conditional means and variances of which are driven by two latent processes. Since the estimation strategy involves the use of the Kalman filter, it is convenient to cast the macro sub-model of BY's LRR model in state-space form. The observation

vector $\mathbf{y}_{t+1}$ contains log consumption growth g_{t+1} and log dividend growth $g_{d,t+1}$, de-meaned by their respective unconditional means μ_c and μ_d, and it is driven by the state vector $\boldsymbol{\alpha}_t$ from the previous period and a contemporaneous vector of innovations $\mathbf{u}_{t+1}$

$$\mathbf{y}_{t+1} = \mathbf{H}\boldsymbol{\alpha}_t + \mathbf{u}_{t+1}, \tag{5.1}$$

$$\begin{bmatrix} g_{t+1} - \mu_c \\ g_{d,t+1} - \mu_d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \phi & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \sigma_t^2 - \sigma^2 \end{bmatrix} + \begin{bmatrix} \sigma_t \eta_{t+1} \\ \varphi_d \sigma_t u_{t+1} \end{bmatrix}, \tag{5.2}$$

where η_{t+1} and u_{t+1} are i.i.d. $\sim N(0,1)$, respectively, and where

$$\mathbf{u}_{t+1|t} \sim N(\mathbf{0}, \mathbf{R}_{t+1}) \quad \text{with} \quad \mathbf{R}_{t+1} = \begin{bmatrix} \sigma_t^2 & 0 \\ 0 & \varphi_d^2 \sigma_t^2 \end{bmatrix}. \tag{5.3}$$

The latent state vector $\boldsymbol{\alpha}_t$ comprises the small, latent growth component, which constitutes the predictable fraction of consumption and dividend growth, and the unobserved stochastic variance process σ_t^2, de-meaned by its unconditional mean σ^2. The state vector follows an autoregressive process with a vector of contemporaneous shocks $\mathbf{v}_t$

$$\boldsymbol{\alpha}_t = \mathbf{F}\boldsymbol{\alpha}_{t-1} + \mathbf{v}_t, \tag{5.4}$$

$$\begin{bmatrix} x_t \\ \sigma_t^2 - \sigma^2 \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ 0 & \nu_1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \sigma_{t-1} - \sigma^2 \end{bmatrix} + \begin{bmatrix} \varphi_e \sigma_{t-1} e_t \\ \sigma_w w_t \end{bmatrix}, \tag{5.5}$$

where e_t and w_t are both i.i.d. $\sim N(0,1)$, and where

$$\mathbf{v}_{t|t-1} \sim N(\mathbf{0}, \mathbf{Q}_t) \quad \text{with} \quad \mathbf{Q}_t = \begin{bmatrix} \varphi_e^2 \sigma_{t-1}^2 & 0 \\ 0 & \sigma_w^2 \end{bmatrix}. \tag{5.6}$$

This representation permits a compact notation of the Kalman filter equations for the estimation of the macro parameters in the following section.

5.3 Econometric methodology

5.3.1 Kalman filtering within a maximum likelihood framework

The recursive model structure is reflected in the three-step estimation strategy. The first estimation step focuses on the macroeconomic basis of the LRR model. Estimating the macro sub-model with a maximum likelihood approach is supposed to overcome the previously encountered lack of efficiency in the macro parameter estimates. However, the LRR macro dynamics are driven by two latent variables, which precludes the application of standard ML estimation. To estimate the macro parameters $\boldsymbol{\xi}^M$, a technique to extract the latent variables from the observables is required. The Kalman filter (cf. Kalman, 1960), which provides a minimum mean squared error estimator for linear Gaussian systems, constitutes such a technique. The macro sub-model, as presented in Section 5.2, is a Gaussian system and therefore in principle permits the application of the Kalman filter. However, the system is not entirely linear in the parameters due to the multiplicative term in the innovations, where the stochastic volatility is multiplied with the i.i.d. normal shocks and their volatility-scaling parameters. Using the Kalman filter is nevertheless possible for this particular model, as derived in Appendix 5.A. The use of the Kalman filter inevitably entails a loss of information related to the nonlinearity in the LRR model, of which the repercussions for parameter estimation will be discussed in more detail in Section 5.3.2.

To allow for a parsimonious notation of the Kalman filter equations, Section 5.2 casts the LRR macro dynamics in state-space representation, consisting of the state equation that describes the dynamics of the latent variables and the observation equation that specifies the relationship between the state and the observed vari-

ables. The Kalman filter algorithm performs sequential linear projections of the state variables that are subsequently updated using the observations of the following period. The prediction step of the Kalman filter for the LRR macro model is performed by computing

$$\hat{\boldsymbol{\alpha}}_{t|t} = \mathbf{F}\hat{\boldsymbol{\alpha}}_{t-1|t} \tag{5.7}$$

$$\mathbf{P}_{t|t} = \mathbf{F}\mathbf{P}_{t-1|t}\mathbf{F}' + \mathbf{Q}_t \tag{5.8}$$

$$\tilde{\mathbf{y}}_{t+1} = \mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t} = \mathbf{y}_{t+1} - \mathbf{H}\hat{\boldsymbol{\alpha}}_{t|t} \tag{5.9}$$

$$\mathbf{f}_{t+1|t} = \mathbf{H}\,\mathbf{P}_{t|t}\mathbf{H}' + \mathbf{R}_{t+1},$$

where $\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t}$ is the mean-zero forecast error committed in the prediction step, and $\mathbf{f}_{t+1|t}$ is its variance.[1] The subsequent updating step is given by

$$\hat{\boldsymbol{\alpha}}_{t|t+1} = \hat{\boldsymbol{\alpha}}_{t|t} + \mathbf{K}_{t+1}\left(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t}\right) \tag{5.10}$$

$$\mathbf{K}_{t+1} = \mathbf{P}_{t|t}\mathbf{H}'\,\mathbf{f}_{t+1|t}^{-1} \tag{5.11}$$

$$\mathbf{P}_{t|t+1} = \mathbf{P}_{t|t} - \mathbf{K}_{t+1}\mathbf{H}\,\mathbf{P}_{t|t}. \tag{5.12}$$

The Kalman filter is initialized with the unconditional mean of the state vector, $\hat{\boldsymbol{\alpha}}_0 = (0,0)'$, and the unconditional covariance of the state, $\mathbf{P}_0 = \text{diag}\left(\frac{\varphi_e^2 \cdot \sigma^2}{1-\rho^2}, \frac{\sigma_w^2}{1-\nu_1^2}\right)$. The joint density of the observations $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T$ is given by

$$p(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T) = p\left(\mathbf{y}_1\right) \prod_{t=2}^{T} p\left(\mathbf{y}_t | \mathcal{Y}_{t-1}\right), \tag{5.13}$$

[1] Due to the specific time structure of the LRR model, in which the state enters the observation equation with a lag of one period, and not contemporaneously, the timing of the usual Kalman filter equations is also shifted by one period.

where $\mathcal{Y}_t = \left(\mathbf{y}_t', \mathbf{y}_{t-1}', \ldots, \mathbf{y}_1'\right)$, which implies the conditional density of the forecast errors $\tilde{\mathbf{y}}_{t+1} = \mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t}$ as

$$p(\tilde{\mathbf{y}}_2, \ldots, \tilde{\mathbf{y}}_T | \tilde{\mathbf{y}}_1) = \prod_{t=2}^{T} p\left(\tilde{\mathbf{y}}_t | \mathcal{Y}_{t-1}\right). \tag{5.14}$$

Consequently, the conditional log likelihood function is obtained as

$$\begin{aligned} \ln \mathcal{L}\left(\boldsymbol{\xi}^M\right) = &-(T-1)\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T-1} \ln\left|\mathbf{f}_{t+1|t}\right| \\ &- \frac{1}{2}\sum_{t=1}^{T-1} (\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})' \mathbf{f}_{t+1|t}^{-1} (\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t}). \end{aligned} \tag{5.15}$$

The parameter estimates are obtained by maximizing the log likelihood function with respect to the parameters. Compared to the computationally intensive simulation-based estimation methods that are frequently applied to the LRR model, this is a rather elegant and both computationally and econometrically efficient way of obtaining parameter estimates of the macro sub-model.

5.3.2 Non-linearity and the Kalman filter

A successful application of the Kalman filter requires the observability of the system to obtain reliable estimates of all latent states. As described by Southall, Buxton, and Marchant (1998), the dynamic system is observable if a finite number k of observations between two points in time t_1 and t_k suffices to recover the initial condition of the state at t_0. Formally, observability is established if the row rank of the observability matrix $\mathbf{O} = (\mathbf{H}', \mathbf{F}'\mathbf{H}')'$ corresponds to the number of states in the system (cf. e.g. Tangirala, 2014). However, due to the structure of the LRR model, $rk(\mathbf{O}) = 1$, and thus the system is not observable. The latent stochastic variance process is an unobservable state, as the observation equation does not permit any

conclusions regarding the second state. Since the matrix $(\mathbf{F}' - \lambda\mathbf{I}, \mathbf{H}')'$ is of full rank for $|\lambda| \geq 1$, the system is asymptotically observable and hence the necessary condition for a consistent estimation of the unobserved state is fulfilled (cf. Sontag, 2013).

The result of the Kalman filtering process described in Section 5.3.1 is independent of the parameter values for the SV parameters ν_1 and σ_w, as both the forecast error and the Kalman gain are unaffected by those parameters. The forecast error vector $\tilde{\mathbf{y}}_{t+1}$ is given by

$$\tilde{\mathbf{y}}_{t+1} = \begin{bmatrix} g_{t+1} - \mu_c - \hat{x}_{t|t} \\ g_{d,t+1} - \mu_d - \phi\hat{x}_{t|t} \end{bmatrix}, \tag{5.16}$$

where $\hat{x}_{t|t}$ denotes the Kalman-filtered value for x_t given $\mathcal{Y}_t$, which reveals that the forecast error $\tilde{\mathbf{y}}_{t+1}$ does not depend on the SV parameters. The Kalman gain $\mathbf{K}_{t+1}$ corresponds to the product of the diagonal matrix $\mathbf{P}_{t|t}$, $\mathbf{H}'$, and $\mathbf{f}_{t+1|t}^{-1}$. Since $\mathbf{H}'$ has a zero lower row, the result of $\mathbf{P}_{t|t}\mathbf{H}'$, and thereby the Kalman gain, also has a zero lower row by construction, irrespective of the nature of $\mathbf{f}_{t+1|t}^{-1}$. Consequently, the only non-zero entry of $\mathbf{K}_{t+1}\mathbf{H}\,\mathbf{P}_{t|t}$ in Equation (5.12) is the top-left element.

In the updating step, there is no information to be gained from the new observation $\mathbf{y}_{t+1}$ regarding the second state: updating the filtered value of the state vector from $\hat{\boldsymbol{\alpha}}_{t|t}$ to $\hat{\boldsymbol{\alpha}}_{t|t+1}$ according to Equation (5.10) does not involve any change in the second state, and the update of the covariance matrix of the state from $\mathbf{P}_{t|t}$ to $\mathbf{P}_{t|t+1}$ according to Equation (5.12) only affects the top-left element of the covariance matrix. As a consequence, the Kalman filter's optimal forecast for the second state inevitably is the unconditional mean of zero. If the Kalman filter is not initialized with the unconditional mean of the state vector, the filtered series for $\sigma_t^2 - \sigma^2$ converges nevertheless to its unconditional mean of zero over time since $|\nu_1| < 1$.

As the structure of the matrix $\mathbf{H}$ eliminates the impact of the parameters ν_1 and σ_w in the Kalman filter recursion, the ML estimation does not allow for the identification of the stochastic volatility parameters ν_1 and σ_w that are notoriously difficult to estimate.

5.3.3 Maximum likelihood estimation of the SV parameters

A maximum likelihood estimation of the parameters that define the persistence and the volatility of the stochastic variance process requires a filter that models the entire conditional distribution of the state and the observation vectors at every point in time, since the parameters in question affect the conditional variance of the observation but not the conditional mean. Sequential Monte Carlo methods, so-called particle filters, introduced to economics and econometrics research by the contributions of Kim et al. (1998) and Pitt and Shephard (1999), provide such a series of conditional distributions. Arulampalam, Maskell, Gordon, and Clapp (2002) review different versions of the generic particle filter, among them the Sampling Importance Resampling (SIR) filter, which will be used in this study.

Similar to the Kalman filter, the generic particle filter can also be considered as a repeated forecasting and updating procedure. Unlike the Kalman filter, particle filters do not operate on the conditional mean only, but estimate a point-mass representation of the conditional distribution. The forecasting step of the particle filter consists of a draw of N particles from the conditional distribution $p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}, \mathcal{Y}_t)$. Since the latent state follows an autoregressive process of order 1, this is equivalent to drawing from $p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1})$, which is straightforward as the density is known. In the subsequent resampling step—the analogon to the updating step in the Kalman

filter—the current observation $\mathbf{y}_{t+1}$ must be incorporated to obtain a particle representation of $p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$,

$$p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1}) \approx \sum_{j=1}^{N} \tilde{w}_t^j \delta(\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_t^j), \tag{5.17}$$

where $\delta(\cdot)$ denotes the Kronecker delta, j is the particle index and $\tilde{w}_t^j$ stands for the weight of particle j at time t, which is yet to be determined.

The updating step is more involved, as $p(\boldsymbol{\alpha}|\mathcal{Y}_{t+1})$ is unknown. Thus, it is impossible to sample directly from $p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$. However, the idea of importance sampling allows for a solution. If a function $\pi(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1}) \propto p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$ can be evaluated and if it is possible to draw from an importance density $q(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$, we can determine the non-normalized weight $w_t^j \propto \frac{p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})}{q(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})}$ for each particle $j = 1, \ldots, N$ as

$$w_t^j \propto \frac{\pi(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})}{q(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})}. \tag{5.18}$$

As filtering is a recursive process, it is convenient to rewrite the weight w_t^j as a function of w_{t-1}^j:

$$w_t^j \propto w_{t-1}^j \frac{p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1})}{q(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}, \mathcal{Y}_{t+1})}, \tag{5.19}$$

which is derived in Appendix 5.A.2. Choosing $p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1})$ for $q(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}, \mathcal{Y}_{t+1})$ in the importance density, we arrive at

$$w_t^j \propto w_{t-1}^j \, p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t). \tag{5.20}$$

Equation (5.20) allows to determine the weights up to proportionality. The exact weights $\tilde{w}_t^j$ are obtained by normalizing w_t^j by $\sum_{k=1}^{N} w_t^k$ such that $\sum_{j=1}^{N} \tilde{w}_t^j = 1$.

The resulting weights are used to resample the particles, where the probability to draw particle j at time t corresponds to $\tilde{w}_t^j$. This is achieved by drawing uniform random numbers $\tilde{u}^j$ and choosing the corresponding particle i, such that

$$\sum_{k=1}^{i-1} \tilde{w}_t^j < \tilde{u}^j \leq \sum_{k=1}^{i} \tilde{w}_t^j. \tag{5.21}$$

The resulting sample of N particles is a point-mass representation of $p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$.

The SIR particle filter is implemented following the pseudo-code of Arulampalam et al. (2002). The detailed scheme is given in Appendix 5.A.3. A stratified sampling approach, as suggested by Flury and Shephard (2011), further improves the quality of the point-mass representation of the density $p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$ without increasing the number of particles. For that purpose, the draws u^j from the uniform distribution $U(0,1)$ are evenly redistributed over the interval $[0;1]$ by using the transformation

$$\tilde{u}^j = \frac{u^j}{N} + \frac{j-1}{N}. \tag{5.22}$$

Since the importance weights are incorporated in the selection of particles after resampling, the weights for all particles are reset to $w_{t-1}^j = \tilde{w}_{t-1}^j = \frac{1}{N}$ after moving on to the next period t. Therefore, the weights at time t are immediately obtained from

$$w_t^j = p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t^j). \tag{5.23}$$

The likelihood function is given by

$$\mathcal{L}(\boldsymbol{\xi}^M) = \prod_{t=1}^{T} p(\mathbf{y}_t) = \prod_{t=1}^{T} \left(\int p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t) p(\boldsymbol{\alpha}_t) \mathrm{d}\boldsymbol{\alpha}_t \right), \tag{5.24}$$

which can be approximated by the particle filter estimator

$$\widehat{\mathcal{L}}(\boldsymbol{\xi}^M) = \prod_{t=1}^{T-1} \left(\sum_{j=1}^{N} p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t^j) p(\boldsymbol{\alpha}_t^j) \right) = \prod_{t=1}^{T-1} \left(\frac{1}{N} \sum_{j=1}^{N} w_t^j \right), \tag{5.25}$$

and the corresponding particle filter estimator of the log likelihood

$$\widehat{\ln \mathcal{L}}(\boldsymbol{\xi}^M) = \sum_{t=1}^{T-1} \ln \left(\frac{1}{N} \sum_{j=1}^{N} w_t^j \right). \tag{5.26}$$

5.3.4 A three-step estimation approach

As extensively discussed in Chapters 3 and 4, the estimation of the macro parameters $\boldsymbol{\xi}_M$ and the preference parameters $\boldsymbol{\xi}_P$ should be disentangled to avoid an interference of the equilibrium asset pricing implications with the estimation of the macro parameters, which is inevitable in a one-step estimation. Such an interference would be at odds with the LRR model structure and, what is more, using the time series dynamics of g_t and $g_{d,t}$ should be more appropriate for the estimation of the parameters that drive the macro dynamics. Moreover, the information on the financial market equilibrium is required for the identification of the preference parameters, in particular for the intricate disentanglement of the risk aversion γ and the intertemporal elasticity of substitution ψ.

Using the Kalman filter-based log likelihood function in Equation (5.15), only a subset of the macro parameters $\boldsymbol{\xi}^M$, $\boldsymbol{\xi}^{M_*} = (\mu_c, \mu_d, \rho, \varphi_e, \sigma, \phi, \varphi_d)'$, can be estimated, due to the non-observability of the system, as discussed in Section 5.3.2. Both the Kalman-filtered series and the likelihood function in Equation (5.15) are independent of the values of ν_1 and σ_w if the Kalman filter is initialized with the unconditional means. Thus, ν_1 and σ_w can be set to arbitrary values in the estimation process since all information regarding the values of these parameters is eliminated in the

filtering process. This will be verified in the Monte Carlo study in the following section.

To estimate the SV parameters, a filtering technique based on conditional means is not sufficient, as ν_1 and σ_w only affect the conditional variance of the observation vector $\mathbf{y}_t$. The particle filter instead provides an estimate of the conditional density at each point in time t and thereby allows for the identification of the SV parameters. The second-step estimation of ν_1 and σ_w is hence performed by maximizing the log likelihood function estimate in Equation (5.26), in which the macro parameter estimates from the first-step estimation are taken as given.

The challenge in estimating the preference parameters $\boldsymbol{\xi}^P$ lies in the joint identification and disentanglement of γ and ψ. The indirect inference estimation described in Section 4.2.3 yields reliable and precise results given the availability of good macro parameter estimates. It consists in matching auxiliary model parameters estimated from simulated financial variables r_m, r_f, and z_m with their empirical counterparts, where the key auxiliary model parameters for the disentanglement of γ and ψ result from a regression of $z_{m,t}$ on $r_{f,t}$ and a constant. Beyond that, the auxiliary model is complemented by means and standard deviations of the aforementioned observable financial variables of the LRR model.[2] This approach will be used for the third-step estimation of $\boldsymbol{\xi}^P$, taking the estimates of $\boldsymbol{\xi}^M$ obtained from the first and second estimation step as given.

[2] The auxiliary model for the indirect inference estimation of the preference parameters $\boldsymbol{\xi}^P$ is described in detail in Section 4.2.3.

5.4 Monte Carlo study

5.4.1 Monte Carlo setup

To illustrate the performance and the validity of the three-step estimation approach, LRR model-implied data series of various sample sizes are simulated using BY's calibrated parameter values given in Table 2.1. The samples comprise different numbers of observations T=650, T=1k, 5k, and 10k, ranging from the currently available sample size to a very large sample to illustrate the asymptotic properties of the estimation strategy. For each sample size, a set of 400 independent data series of the observable variables g, g_d, r_m, r_f, and z_m is generated. Applying the estimation approach described in Section 5.3 to the simulated data yields the approximate distribution of the parameter estimates for the respective sample size. The calibrated parameter values—and thus also the simulated data—correspond to a monthly frequency of the representative investor's investment decision. The smallest sample of T=650 months, corresponding to 54 years of empirical data, can serve to gauge the estimation quality of an empirical application: with monthly consumption data being available from 1959, the results for this sample size can be considered as a lower bound for the precision that is to be expected for an estimation on empirical data, assuming that the model describes the true data generating processes well.

When simulating LRR model-implied data, the starting values of the autoregressive processes are set to their unconditional means. A burn-in sample of 100 observations is dropped to avoid any starting value dependence of the simulated series. It is necessary to ensure the non-negativity of the observations for σ_t^2, which is not necessarily the case (cf. Equation (5.4)). This is achieved by setting any negative values to zero. To simulate data for the financial variables r_m, r_f, and z_m, the endogenous means $\bar{z}$ and $\bar{z}_m$ are solved for numerically.

For the estimation, the parameter estimates for μ_c, μ_d, ρ, and ν_1 are restricted to lie between 0 and 1, the parameter estimates of φ_e, ϕ, φ_d, σ_w are restricted to positive values. When filtering the unobserved stochastic variance by the Kalman filter, a constraint to non-negative values is not required since the filtered value for σ_t^2 corresponds to the unconditional mean σ^2 for all t. For the particle filter, however, a similar constraint as in the data simulation process is applied: if the filtered value for the stochastic variance is negative, it is set to a small positive value of $1e^{-20}$. It cannot be set to zero, however, since this would impede the evaluation of the density $p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t)$ due to a conditional variance of zero for both elements in $\mathbf{y}_{t+1}$. In the application of the particle filter, the number of particles used corresponds to N=100k throughout the simulation study.

In the third-step indirect inference estimation, the simulated sample size corresponds to $10\,T$ for all T. Model solvability is ensured by a penalty term of size 10^3 in the third estimation step, which is added to the objective function value in the case of an unsolvable model. For the first two estimation steps, a model solution is not required, since the financial variables that depend on the endogenous parameters are not involved in the estimation.

5.4.2 Monte Carlo results: macro parameter estimates

In the first estimation step, the macro parameters $\boldsymbol{\xi}^{M_*}$ are estimated for all simulated samples. For various reasons, it is challenging to estimate the parameters that determine the macro dynamics. In the BY calibration, g and in particular g_d are rather volatile; therefore, the unconditional means of consumption and dividend growth μ_c and μ_d cannot be precisely determined when the sample size is small. The parameter ρ sets the persistence of the latent, autoregressive process x_t, which constitutes the predictable fraction of consumption and dividend growth; it

is calibrated close to 1 ($\rho = 0.979$), such that the simulated data for x_t are on the verge of non-stationarity, which exacerbates the estimation. Since the variation in dividend growth is typically larger than the variation in consumption growth, the parameter ϕ that leverages the impact of the latent growth component x_t on dividend growth $g_{d,t+1}$ is set to $\phi = 3$. For the same reason, also the volatility-scaling parameters φ_e and φ_d are vastly different in size. While $\varphi_e = 0.044$, the volatility of the dividend growth process is inflated by $\varphi_d = 4.5$. Since both parameters co-occur multiplicatively with the lagged stochastic volatility, their values are often challenging to estimate. In the first-step estimation, the latent process x_t is approximated by the Kalman filtered series, which complicates the estimation of the parameters that determine x_t and its impact on the observable variables g_{t+1} and $g_{d,t+1}$.

The estimation results for the macro parameters $\boldsymbol{\xi}^{M_*}$ are reported in Table 5.1. Panel A displays the medians and root mean squared errors (RMSE) of $\hat{\boldsymbol{\xi}}^{M_*}$ resulting from the Kalman filter-based ML estimation with $\nu_1 = 0$ and $\sigma_w = 0$, while Panel B displays medians and RMSE for $\hat{\boldsymbol{\xi}}^{M_*}$ resulting from the same estimation approach, except that ν_1 and σ_w are set to their true values. As indicated in Section 5.3.4, the results are by construction numerically identical. The last column in both panels displays the number of successful estimations out of 400 simulated data sets.[3] Overall, the medians are remarkably close to the true values even for a sample size of T=650, and increasingly so for T=1k, 5k and 10k. For μ_c ρ, σ, and φ_d the estimation precision is good for all sample sizes, for μ_d, φ_e, and ϕ precise estimates can only be obtained for the larger samples. The distributions of the parameter estimates can be analyzed in more detail by means of the kernel density plots in Figure 5.1. In the case of ϕ, the rather large RMSE for the small samples are due to

[3] Estimations are deemed unsuccessful if the Nelder and Mead (1965) algorithm does not converge, or if one of the parameter estimates deviates by more than a factor of 10 from the true value.

only weakly bell-shaped density estimates with additional outliers on the right-hand side, whereas for a sample size of T=5k or T=10k the kernel density is clearly more concentrated, and more closely centered around the true value. For φ_e the kernel densities for T=650 and T=1k are somewhat right-skewed, but also become increasingly precise and symmetric for larger T. The lack of precision of μ_d is reflected in an almost flat kernel density for the smallest sample, whereas for larger samples the distribution becomes more bell-shaped and symmetric. The graph shows that the estimation approach works in principle, but the high volatility in the dividend growth series severely exacerbates the estimation for small samples.

Comparing the estimation results in Table 5.1 to results previously obtained from moment-based approaches or indirect inference estimation reveals considerable efficiency gains. Comparing medians and RMSE for T=1k to the results from Chapter 4 shows that the ML estimation outperforms the indirect inference approach for all parameters except ϕ. The medians of the ML estimates are closer to the true values, and the deviations are smaller on average. The differences in RMSE are particularly pronounced for ρ (reduction by 66%), φ_e (38%), and φ_d (38%). Moreover, the number of successful estimations is clearly higher for the ML approach (397 vs. 348 out of 400), such that the sample selection effect should rather work in favor of the indirect inference results.

Instead of matching selected moments or auxiliary model parameters, which represent isolated properties of the joint distribution of the observed series for g and g_d, the Kalman filter-based maximum likelihood approach uses the entire distribution and allows to take the latent growth process x_t into account explicitly. Consequently, the efficiency gains are particularly high for those parameters that determine the conditional distribution of x_t. Overall, maximizing the log likelihood function from

Equation (5.15) clearly provides more efficient results than matching a collection of selected properties of the observed series, in particular for small samples.

5.4.3 Monte Carlo results: SV parameter estimates

In a second step, estimates for the SV parameters can be obtained by maximizing the log likelihood function estimate in Equation (5.26) using the generic particle filter according to the implementation scheme in Appendix 5.A.3. The role of the stochastic variance σ_t^2 in the LRR model is to introduce fluctuating economic uncertainty, which allows to model higher risk premia in times of high economic uncertainty and lower risk premia in more moderate periods. The parameter ν_1 determines the persistence of σ_t^2, while σ_w scales its volatility. Similar to ρ, also ν_1 is calibrated to a value close to 1 ($\nu_1 = 0.987$), which implies that not only the latent growth process x_t, but also the latent stochastic variance process σ_t^2 is close to non-stationarity, which exacerbates the estimation of the autoregressive parameter ν_1. The calibrated value for the conditional volatility σ_w is very small (σ_w=2.3e-06), which implies a rather slow-moving pattern of σ_t^2, given the high persistence implied by the choice of ν_1. The values chosen by BY imply that the economy tends to remain in its present state of volatility, meaning that the LRR model can emulate prolonged crisis periods as well as lasting quiet periods of the economy.

Both SV parameters have proven rather difficult to estimate in previous estimation attempts. As shown in an extensive moment sensitivity study in Chapter 3, few available moments of observable series are related to ν_1 and σ_w at all. The first and second (auto-) moments of the observable macro variables do not depend on the SV parameters by construction. Their fourth moments, used by Constantinides and Ghosh (2011), only exhibit a very low level of sensitivity to ν_1 and σ_w. Finally, also the moments of the observable financial variables are rather insensitive to the

values of the SV parameters, whereas they are highly sensitive to the preference parameters. Own attempts to develop an estimation strategy for the SV parameters based on moment matching failed, not only due to a lack of precision but also due a lack of identification, in particular to identify whether ν_1 is close to 1 or close to 0.

Since the particle filter-based likelihood estimate is based on a set of 10^5 particles, the maximization of the objective function is rather computationally intensive, in particular within the framework of a simulation study. Despite the high number of particles, the estimate of the likelihood function is not perfectly smooth, which prevents even powerful optimization algorithms from moving freely on the objective function surface. This issue is resolved by evaluating the objective function along two consecutive grids of parameter values. Initially, a coarse-meshed grid is used to locate the area of the objective function's maximum. Subsequently, around the associated parameter values a finer grid is set up, for which the second, more precise maximum and the related parameter values are determined. Finally, the Nelder and Mead (1965) algorithm is used to find the local maximum. This technique has been thoroughly scrutinized. Panel A of Figure 5.2 shows an example of the objective function surface for a coarse grid, and Panel B depicts the surface of the same objective function for the subsequent finer grid. The graphs illustrate that the necessity to apply a grid-based method is unrelated to identification issues, as the objective function is well-behaved with a distinct area of the global optimum. The different steps in the optimization are rather required to find the local optimum on a non-smooth surface. This approach ensures the validity and consistency of the parameter estimates as far as possible with feasible computational effort.

For the estimation of the SV parameters, the estimate of the unconditional mean σ^2 is taken as given from the first-step estimation, as well as the remainder of the macro parameters in $\boldsymbol{\xi}^{M_*}$. The log likelihood function estimate in Equation (5.26) is

maximized with respect to the parameters ν_1 and σ_w. To begin with, the validity of the particle filter-based likelihood approach is ascertained by conducting the second-step estimation assuming the true values of the macro parameters $\boldsymbol{\xi}^{M_*}$ are known. If the estimation strategy works, the estimates should be closely centered around the true values for all sample sizes, and the estimation precision should increase in T. After the method is established in this manner, the estimation is conducted based on the estimated macro parameters $\hat{\boldsymbol{\xi}}^{M_*}$. Proceeding in this way also allows to assess the importance of the different sources of estimation uncertainty, in particular of the quality of the underlying macro parameter values.

The results of the second-step estimation are reported in Table 5.2. Panel A displays the medians and RMSE of the estimated SV parameters based on the true macro parameter values for $\boldsymbol{\xi}^{M_*}$, while Panel B displays the medians and RMSE of the parameter estimates relying on the macro parameter estimates $\hat{\boldsymbol{\xi}}^{M_*}$ from the first-step estimation. In each panel, the last column labeled by R contains the number of successful estimations out of 400 simulated data sets. Even for the smallest sample of size $T = 650$, the medians are very close to the true values. Also the estimation precision in terms of RMSE is remarkably high. The convergence of the medians toward the true values, as well as the decrease in RMSE with increasing sample size, provides evidence for the validity of the particle filter-based ML estimation and serves as a simulation-based consistency check. Notably, the difference between using the true and the estimated macro parameters is rather small. While the medians based on $\hat{\boldsymbol{\xi}}^{M_*}$ are slightly further away from the true values for all T, the RMSE are very similar, in particular for the larger samples.

The kernel density plots in Figure 5.3 confirm that the distribution of the estimates is indeed very similar between Panel A (true macro parameter values) and Panel B (estimated macro parameters). A comparison of Panels A and B of the ker-

nel density plots shows that the quality of the macro parameters only plays a minor role for the estimation precision of $\hat{\nu}_1$ and $\hat{\sigma}_w$ compared to the effect of the sample size. Overall, the quality of the estimates is excellent, relative to the precision of the macro parameter estimates, and taking into account the difficulty of obtaining any reliable estimates of the parameters ν_1 and σ_w at all.

5.4.4 Monte Carlo results: preference parameter estimates

To provide a full set of parameter estimates for the LRR model, the indirect inference estimation strategy described in detail in Section 4.2.3 is employed for the estimation of the preference parameters $\boldsymbol{\xi}^P$. The main challenge in the estimation of the representative investor's preference parameters is to disentangle risk aversion γ and the intertemporal elasticity of substitution ψ, while the estimation of the subjective discount factor δ is usually feasible (cf. Chapter 4). The calibration assumes the investor to be risk averse with γ=10, and to prefer consumption in the present month over consumption in the following month by a discount factor of δ=0.998. The elasticity of substitution is calibrated to ψ=1.5. Hence, the substitution effect is assumed to dominate the wealth effect, such that the investor is supposed to be responsive to interest rate changes and to reduce the consumption smoothing behavior if the interest rate increases.

Table 5.3 contains the medians and RMSE of the estimates obtained from the third-step indirect inference estimation of $\boldsymbol{\xi}^P$. Panel A shows the results of an estimation that assumes all macro parameters $\boldsymbol{\xi}^M$ to be known, which serves as a proof of concept for the estimation strategy. Panel B displays the results based on the estimated macro parameters $\hat{\boldsymbol{\xi}}^{M_*}$ resulting from the first-step estimation and $\hat{\nu}_1$ and $\hat{\sigma}_w$ from the second-step estimation. The median converges toward the true

parameter values and the RMSE decreases with increasing sample size, which can be interpreted as simulation-based evidence of consistency. The kernel density plots in Panel A of Figure 5.4 illustrate that the estimation results are very precise when the true macro parameter values are known, while for the estimation based on $\hat{\boldsymbol{\xi}}^{M_*}$ the variation in the estimates is larger when the sample size is small, as depicted in Panel B. This applies in particular to $\hat{\gamma}$ and $\hat{\psi}$, for which the estimation quality suffers from large outliers. For larger samples, the number of outliers is greatly reduced, which implies considerably lower values for the RMSE.

The preference parameter estimates reported in Section 4.3.3 are produced by a two-step indirect inference estimation approach, in which the macro parameter estimates are less efficient, as discussed in Section 5.4.2. Furthermore, the SV parameters could not be estimated and were therefore set to zero. Both studies use, among others, the sample size T=1k, for which the results can be compared. The median values of all parameter estimates are closer to the true values for the three-step estimation approach: the median value for $\hat{\psi}$ is 1.85 (vs. 2.12 for the two-step estimation), the median of $\hat{\gamma}$ is 12.0 (vs. 12.8), and for $\hat{\delta}$ the median is 0.9976 (vs. 0.9975). The most pronounced improvement lies in the higher precision of the risk aversion estimate in terms of RMSE (8.0 vs. 14.5); the RMSE of $\hat{\psi}$, however, is somewhat increased by the three-step estimation approach (3.58 vs. 2.68). The results for $\hat{\delta}$ are almost the same, irrespective of the underlying values for $\boldsymbol{\xi}^M$, both regarding the medians (0.9976 vs. 0.9975) and the RMSE (0.0018 vs. 0.0015). Overall, the improved precision of the macro parameter estimates in the three-step estimation, and the accomplishment of being able to obtain reliable estimates of the SV parameters as well, has a beneficial effect on the results for the preference parameter estimates.

Moreover, the results for the two-step estimation are computed from a much smaller—and probably favorable—subset of estimation results, since the approach yielded only 284 (out of 400) successfully estimated sets of preference parameters, whereas in the three-step approach 373 out of 400 estimations could be successfully completed. This difference is partly due to the higher number of failed estimations when applying the macro parameter estimation strategy from Section 4.2.2, where 52 macro parameter estimations fail in the first estimation step only, in contrast to a failure of 3 out of 400 estimations in the first- and second-step estimation of the present estimation strategy combined.

Overall, the third-step estimation results constitute progress in the estimation of the LRR preference parameters since the number of successful LRR model estimations is greatly increased. Furthermore, the preference parameter estimation is based on estimates of all remaining model parameters $\boldsymbol{\xi}^M$, including the SV parameters. Finally, the estimation results show that the three-step estimation works well and that it is able to identify all LRR model parameters with good point estimates and reasonable precision given the available sample size.

5.5 Data

In the literature, the LRR model is typically estimated on quarterly data (cf. Hasseltoft, 2012; Calvet and Czellar, 2015) or annual data (cf. Bansal et al., 2007a; Constantinides and Ghosh, 2011; Bansal et al., 2012b). The BY calibration is, however, based on a monthly frequency, implying a monthly decision interval of the representative investor, which is typically considered as the most plausible choice in the literature and also supported by an empirical result of Bansal et al. (2012b). The latter derive analytical time aggregation formulas for the moments matched in

their GMM estimation to be able to estimate the model at a monthly frequency using annual data. Constantinides and Ghosh (2011) implicitly assume an annual decision frequency as they estimate the LRR model by GMM estimation on annual data without aggregation. The remainder of the studies mentioned above estimate the monthly dynamics from lower-frequency data by applying simulation-based estimation techniques that allow for a time-aggregation of the simulated processes. Such a time aggregation is not feasible for the log likelihood functions in Equations (5.15) and (5.26). Monthly U.S. data from February 1959 to December 2014—i.e. a total of 671 observations—are used to estimate the model for a realistic decision frequency of the representative investor. This choice also mitigates the problem of estimating a large number of parameters of a complex structural model from a very limited set of data. The annual data sets available, often preferred over quarterly data because they are considered to be most reliable and not subject to notorious problems like seasonality in dividend payments, only comprise about 80 observations, which seems scarce for the estimation of 12 structural parameters.

Consumption data are obtained from the Bureau of Economic Analysis. The standard choice, real personal consumption in non-durable goods and services, is only available on a monthly basis from 1999, which would imply a very short data series. Monthly data of nominal personal consumption expenditures are used instead to compute log consumption growth, which is converted into real terms by using the Consumer Price Index (CPI) data from the Bureau of Labor Statistics. The monthly series for dividend growth, market portfolio return, and the log price-dividend ratio are obtained from the CRSP value-weighted market portfolio, again the CPI is used for conversion into real terms. The well-known seasonal pattern in dividend growth is corrected for by a 12-month trailing average, since the LRR model cannot account for the strong negative autocorrelation in raw dividend growth data by construction (cf.

Equation (5.1)). The risk-free rate is approximated on the basis of the one-month nominal T-bill yield obtained from CRSP. As suggested by Beeler and Campbell (2012), the ex ante risk-free rate is obtained from a predictive regression of the monthly ex-post real log yield on the nominal monthly log yield, the monthly log inflation rate averaged over the past year, and a constant.

Since the first- and second-step estimations are based on the macro data for g and g_d only, whereas the third-step estimation is solely based on financial market data for r_m, r_f, z_m but not on the macro series, it is possible to use time series of different lengths for the consecutive estimation steps. As longer time series are available for the financial market data series than for consumption growth, the sample for the third-step estimation is extended back to February 1947, which yields an additional 144 observations, such that the third-step estimation is based on 815 observations. The limiting factor are the seasonally adjusted CPI data, which are available starting in 1947.

Descriptive statistics of the data are provided in Table 5.4. The mean of monthly consumption growth corresponds to an annual growth rate of 2.8%, dividend growth amounts to approximately 2.0% p.a. on average. The dividend growth series is substantially more volatile than consumption growth; both series have a small negative autocorrelation. The average return of the market portfolio aggregates to 7.0% p.a., while the average annual risk-free rate only equals 0.4%. The market return is volatile with a rather low autocorrelation, whereas the log price-dividend ratio and the risk-free rate are highly persistent with first-order autocorrelations of 0.99 and 0.97, respectively. Figure 5.5 illustrates the time series in detail.

5.6 Empirical application

By using the monthly data described in Section 5.5, the model parameters can be estimated at a monthly frequency, which allows for an immediate comparison of the point estimates with the BY calibration. Using the data described in Section 5.5 and applying the estimation strategy outlined in Section 5.3 yields the parameter estimates shown in Table 5.5.

The point estimates for the unconditional means of consumption and dividend growth $\hat{\mu}_c$=0.0023 and $\hat{\mu}_d$=0.0018 are very close to the means of the data series for g and g_d, which is a plausible result. These results amount to annual growth rates of consumption and dividends of 2.8% and 2.2%, respectively. The data provide evidence that indeed consumption and dividend growth are not i.i.d., but that there is a small predictable growth component. In line with the fundamental idea of the LRR model, the latent growth component is indeed estimated to be highly persistent with an autoregressive parameter estimate of $\hat{\rho}$=0.944. Furthermore, also the latent stochastic variance process, which determines the conditional variance of the macro processes, is estimated to be comparatively persistent with $\hat{\nu}_1$=0.877. Accordingly, the economic uncertainty in one period largely determines the uncertainty of the next period, albeit to a clearly lesser extent than in BY's calibration. The fluctuation in economic uncertainty is estimated to be notably higher than in BY's calibration: the volatility parameter of the stochastic variance process is estimated as $\hat{\sigma}_w$=6.3e-06, which is more than twice the calibrated value. The average level of economic uncertainty is, in contrast, somewhat lower than in the calibration, as the constant parameter of the stochastic variance process is estimated as $\hat{\sigma}$=0.0057. Overall, we can conclude that for the sample period the estimated stochastic volatility is moderate. In particular, the half-life of a high-volatility period, i.e. a crisis, is considerably shorter than in the calibration, since the persistence is lower and the volatility is

higher. However, since the fluctuation in volatility is sizeable, high-volatility periods also occur more frequently. The volatility-scaling parameters of consumption and dividend growth are both lower than in the BY calibration, with estimates of $\hat{\varphi}_e$=0.029 and $\hat{\varphi}_d$=1.97, respectively. The leverage parameter that scales the impact of the small predictable growth component on dividend growth is rather high with an estimate of $\hat{\phi}$=8.8, which implies that the estimation identifies a considerably larger persistent component in dividend growth than in consumption growth.

The estimates for the preference parameters exhibit a phenomenon frequently encountered in the literature: the risk aversion parameter estimate $\hat{\gamma}$=54.1 is very high, exceeding by far the value of 10, the upper bound for plausible values stated by Mehra and Prescott (1985). Also, the estimate of the intertemporal elasticity of substitution $\hat{\psi}$=2.31 is rather large. The subjective discount factor is estimated as $\hat{\delta}$=0.9815, which implies a plausible extent of preference for present instead of future consumption. As already indicated in the Monte Carlo study, the estimation precision for the preference parameter estimates that can be expected for the present sample size is rather limited, even for data generated by the LRR model. This applies all the more to empirical data.

The estimation precision of the empirical estimation is assessed by means of a parametric bootstrap. The bootstrap is conducted as follows. 400 data sets are simulated, using as true parameter values the point estimates obtained from the three-step estimation. In three consecutive steps, those 400 sets of simulated data are then used for parameter estimation, proceeding precisely in the same fashion as in the estimation on the empirical data. Empirical 2.5% and 97.5% quantiles of the resulting distribution of parameter estimates are finally used to estimate the lower and upper bounds of the confidence intervals, respectively.

The 95% confidence intervals for the empirical parameter estimates are also given in Table 5.5. The estimation precision is comparatively high for $\hat{\mu}_c$, $\hat{\rho}$, $\hat{\nu}_1$, $\hat{\sigma}$, $\hat{\varphi}_d$, and $\hat{\delta}$, whereas the confidence intervals are rather wide for $\hat{\mu}_d$, $\hat{\phi}$, $\hat{\gamma}$, and $\hat{\psi}$. This result is in line with the findings in Section 5.4. As observed in the Monte Carlo study, the precision of the preference parameter estimates that can be expected for a sample size comparable to the available set of empirical data is rather limited. The confidence interval of the subjective discount factor estimate $\hat{\delta}$ lies entirely below 1, plausibly indicating that we can reject that the investor favors consumption in the future over consumption in the present on a 5% significance level. Furthermore, the confidence interval for $\hat{\gamma}$ indicates that a low or moderate risk aversion of the investor can be rejected, since values for γ between 1 and 10 can be rejected on a 5% significance level. Finally, the confidence interval for $\hat{\psi}$ is wide, such that it accommodates both for values of the IES below and above 1.

Table 5.6 illustrates the low estimation precision from a different angle. The table contains the means and standard deviations of g, g_d, z_m, r_m, and r_f implied by the point estimates from the empirical estimation and compares them to their data counterparts. Quantiles of the means and standard deviations implied by the estimation are obtained from the bootstrap distribution of the point estimates. The macro moments are matched rather closely, while the features of the asset pricing model cannot be matched precisely because of limitations imposed on the financial variables by the estimation of the macro parameters, which are due to the model structure. Considering those implications resulting from the model-implied distributions of the point estimates, the properties of the empirical data on the financial market—such as the high equity premium—cannot be reproduced by the LRR model.

The high risk aversion parameter estimate can be explained by the low overall estimated volatility in the macro model and by the estimates of the persistence

parameters ρ and ν_1, which are notably lower than the calibrated values by BY. Therefore, the macroeconomic risk in the estimated model is considerably less severe than in the calibration; however, a high level of risk is required to achieve high risk premia with a moderate risk aversion. Due to the three-step estimation procedure, which does not allow to adapt the macro parameter values in a way that is convenient to explain the asset pricing properties of the model, the only way to account for high risk premia in the presence of low or moderate risk is to adjust the preference parameter estimates accordingly.

5.7 Conclusion

This study introduces a novel three-step strategy for the estimation of Bansal and Yarons's (2004) LRR model that is able to reliably identify all structural parameters, including precise estimates for the SV parameters that proved difficult to estimate in previous studies. The method used in the first step relies on a Kalman filter-based maximum likelihood estimation to obtain the estimates of the parameters that determine the dynamics of consumption and dividend growth, as well as their latent persistent growth component. In a second step, the application of a particle filter within a maximum likelihood approach allows for the estimation of the persistence parameter and the volatility of the stochastic variance process. Finally, in a third step, the preference parameters are estimated by indirect inference. The estimation strategy thus adheres to the recursive model structure, which consists of an independent set of macroeconomic processes that influence the financial variables and the decisions of the representative investor but not vice versa. A Monte Carlo study shows that the use of maximum likelihood for the estimation of the macro model parameters in the first two steps indeed enhances the precision of the resulting pa-

rameter estimates, and also the quality of the preference parameter estimates in the final estimation step.

Applying the estimation strategy to monthly U.S. data provides some support for the idea of long-run risk in the macroeconomy by identifying a persistent latent growth component in consumption and dividend growth and a persistent stochastic variance. However, the estimates of the autoregressive parameters in the growth expectations and the fluctuating macroeconomic uncertainty are not as close to 1 as in the BY calibration, implying a considerably less severe degree of macroeconomic risk. As a consequence, the high observed equity premium leads to a rather large estimate for the relative risk aversion. Moreover, the estimation yields a plausible subjective discount factor estimate close to but below 1, and an IES estimate greater than 1. Bansal and Yaron (2004) calibrate an IES value larger than 1, as this choice typically allows for a high equity premium and a low risk-free rate at the same time. Even though an IES value larger than 1 should thus have favorable implications for the long-run risk paradigm, the empirical results do not yield a close match between the properties of the empirical data and the model-implied features. In particular, the parameter estimates implied by the empirical data do not permit to replicate the features of the observed financial data series, due to the restrictions imposed by the LRR model. Thus, for the present monthly data set, the LRR model cannot serve to explain the equity premium puzzle and the risk-free rate puzzle at the same time.

5.A Appendix

5.A.1 Kalman filter derivation

State forecast:

$$\hat{\boldsymbol{\alpha}}_{t|t} = \hat{\mathbb{E}}\left(\boldsymbol{\alpha}_t|\mathcal{Y}_t\right) = \hat{\mathbb{E}}\left(\mathbf{F}\boldsymbol{\alpha}_{t-1}\right) = \mathbf{F}\hat{\boldsymbol{\alpha}}_{t-1|t}$$
$$\text{where} \quad \mathcal{Y}_t = \left(\mathbf{y}_t', \mathbf{y}_{t-1}', \ldots, \mathbf{y}_1'\right)$$

Mean squared forecast error:

$$\begin{aligned}
\mathbf{P}_{t|t} &= \mathbb{E}\left[(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})'\right] \\
&= \mathbb{E}\left[\left(\mathbf{F}\boldsymbol{\alpha}_{t-1} + \mathbf{v}_t - \mathbf{F}\hat{\boldsymbol{\alpha}}_{t-1|t}\right)\left(\mathbf{F}\boldsymbol{\alpha}_{t-1} + \mathbf{v}_t - \mathbf{F}\hat{\boldsymbol{\alpha}}_{t-1|t}\right)'\right] \\
&= \mathbf{F}\,\mathbb{E}\left[\left(\boldsymbol{\alpha}_{t-1} + \hat{\boldsymbol{\alpha}}_{t-1|t}\right)\left(\boldsymbol{\alpha}_{t-1} - \hat{\boldsymbol{\alpha}}_{t-1|t}\right)'\right]\mathbf{F}' + \mathbb{E}\left[\mathbf{v}_t\mathbf{v}_t'\right] \\
&= \mathbf{F}\,\mathbf{P}_{t-1|t}\mathbf{F}' + \mathbf{Q}_t
\end{aligned}$$

Observation forecast:

$$\begin{aligned}
\hat{\mathbf{y}}_{t+1|t} &= \hat{\mathbb{E}}\left(\mathbf{y}_{t+1}|\mathcal{Y}_t\right) = \hat{\mathbb{E}}\left(\mathbf{H}\boldsymbol{\alpha}_t + \mathbf{u}_{t+1}|\mathcal{Y}_t\right) \\
&= \mathbf{H}\,\hat{\mathbb{E}}\left(\boldsymbol{\alpha}_t|\mathcal{Y}_t\right) = \mathbf{H}\hat{\boldsymbol{\alpha}}_{t|t}
\end{aligned}$$

Observation forecast error:

$$\begin{aligned}
\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t} &= \mathbf{y}_{t+1} - \mathbf{H}\hat{\boldsymbol{\alpha}}_{t|t} \\
&= \mathbf{H}\boldsymbol{\alpha}_t + \mathbf{u}_{t+1} - \mathbf{H}\hat{\boldsymbol{\alpha}}_{t|t} \\
&= \mathbf{H}\left(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t}\right) + \mathbf{u}_{t+1} \\
\mathbf{f}_{t+1|t} &= \mathbb{E}\left[(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})'\right] \\
&= \mathbb{E}\left[\mathbf{H}(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})'\mathbf{H}'\right] + \mathbb{E}\left[\mathbf{u}_{t+1}\mathbf{u}_{t+1}'\right] \\
&= \mathbf{H}\,\mathbf{P}_{t|t}\mathbf{H}' + \mathbf{R}_{t+1}
\end{aligned}$$

Since $\boldsymbol{\alpha}_t|\mathcal{Y}_t$ and $\tilde{\mathbf{y}}_{t+1} = (\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})$ are jointly Gaussian, $\boldsymbol{\alpha}_t|\mathcal{Y}_t, \tilde{\mathbf{y}}_{t+1} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with:

$$\begin{aligned}
\boldsymbol{\mu} &= \mathbb{E}(\boldsymbol{\alpha}_t|\mathcal{Y}_t) + \mathrm{Cov}(\boldsymbol{\alpha}_t, \tilde{\mathbf{y}}_{t+1}|\mathcal{Y}_t)\mathrm{Var}(\tilde{\mathbf{y}}_{t+1}|\mathcal{Y}_t)^{-1}\left(\tilde{\mathbf{y}}_{t+1} - \mathbb{E}(\tilde{\mathbf{y}}_{t+1})\right) \\
\boldsymbol{\Sigma} &= \mathrm{Var}(\boldsymbol{\alpha}_t|\mathcal{Y}_t) - \mathrm{Cov}(\boldsymbol{\alpha}_t, \tilde{\mathbf{y}}_{t+1}|\mathcal{Y}_t)\mathrm{Var}(\tilde{\mathbf{y}}_{t+1}|\mathcal{Y}_t)^{-1}\mathrm{Cov}(\tilde{\mathbf{y}}_{t+1}, \boldsymbol{\alpha}_t|\mathcal{Y}_t).
\end{aligned}$$

Upon arrival of time $t+1$ information, the estimate of the unknown state can thus be updated as:

$$\begin{aligned}
\hat{\boldsymbol{\alpha}}_{t|t+1} &= \hat{\mathbb{E}}\left(\boldsymbol{\alpha}_t|\mathcal{Y}_t, \mathbf{y}_{t+1}\right) \\
&= \hat{\mathbb{E}}\left(\boldsymbol{\alpha}_t|\mathcal{Y}_t\right) + \mathrm{Cov}(\boldsymbol{\alpha}_t, \tilde{\mathbf{y}}_{t+1}|\mathcal{Y}_t)\mathrm{Var}(\tilde{\mathbf{y}}_{t+1}|\mathcal{Y}_t)^{-1}\left(\tilde{\mathbf{y}}_{t+1} - \mathbb{E}(\tilde{\mathbf{y}}_{t+1})\right).
\end{aligned}$$

Since the forecast error has mean zero and variance $\mathbf{f}_{t+1|t-1}$, we have:

$$\begin{aligned}
\hat{\boldsymbol{\alpha}}_{t|t+1} &= \hat{\boldsymbol{\alpha}}_{t|t} + \mathbb{E}\left[(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})'\right]\mathbf{f}_{t+1|t}^{-1}\left(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t}\right) \\
&= \hat{\boldsymbol{\alpha}}_{t|t} + \mathbb{E}\left[(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})(\mathbf{H}\left(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t}\right) + \mathbf{u}_{t+1})'\right]\mathbf{f}_{t+1|t}^{-1}\left(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t}\right) \\
&= \hat{\boldsymbol{\alpha}}_{t|t} + \mathbf{P}_{t|t}\mathbf{H}'\mathbf{f}_{t+1|t}^{-1}\left(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t}\right) \\
\mathbf{P}_{t|t+1} &= \mathbb{E}\left[(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t+1})(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t+1})'\right] \\
&= \mathbb{E}\left[(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})'\right] - \mathbb{E}\left[(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})'\right]\mathrm{Var}(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})^{-1} \\
&\quad \mathbb{E}\left[(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})(\boldsymbol{\alpha}_t - \hat{\boldsymbol{\alpha}}_{t|t})'\right] \\
&= \mathbf{P}_{t|t} - \mathbf{P}_{t|t}\mathbf{H}'\,\mathbf{f}_{t+1|t}^{-1}\,\mathbf{H}\,\mathbf{P}_{t|t}.
\end{aligned}$$

5.A.2 Particle filter derivation

Following Arulampalam et al. (2002), the weight $w_t^j \propto \frac{p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})}{q(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})}$ for each particle $j = 1, \ldots, N$ is determined up to proportionality by

$$w_t^j \propto \frac{\pi(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})}{q(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})}, \tag{5.27}$$

where $\pi(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1}) \propto p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$. In a first step, a function $\pi(\cdot)$ must be derived that is proportional to $p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$, and that can be evaluated:

$$p(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1}) = \frac{p(\boldsymbol{\alpha}_t, \mathcal{Y}_{t+1})}{p(\mathcal{Y}_{t+1})} = \frac{p(\mathcal{Y}_{t+1}|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t)}{p(\mathbf{y}_{t+1}, \mathcal{Y}_t)}. \tag{5.28}$$

With conditional independence of $\mathbf{y}_{t+1}$ and $\mathcal{Y}_t$ given $\boldsymbol{\alpha}_t$:

$$= \frac{p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t)p(\mathcal{Y}_t|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t)}{p(\mathbf{y}_{t+1}|\mathcal{Y}_t)p(\mathcal{Y}_t)} \tag{5.29}$$

$$= \frac{p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t|\mathcal{Y}_t)}{p(\mathbf{y}_{t+1}|\mathcal{Y}_t)} \tag{5.30}$$

$$= \frac{p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1})p(\boldsymbol{\alpha}_{t-1}|\mathcal{Y}_t)}{p(\mathbf{y}_{t+1}|\mathcal{Y}_t)} \tag{5.31}$$

$$\propto p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1})p(\boldsymbol{\alpha}_{t-1}|\mathcal{Y}_t). \tag{5.32}$$

Furthermore, choosing the importance density $q(\boldsymbol{\alpha}_t|\mathcal{Y}_{t+1})$ as the product $q(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}, \mathcal{Y}_{t+1})q(\boldsymbol{\alpha}_{t-1}|\mathcal{Y}_t)$ and inserting (5.32) into (5.27) yields:

$$w_t^j \propto \frac{p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1})p(\boldsymbol{\alpha}_{t-1}|\mathcal{Y}_t)}{q(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}, \mathcal{Y}_{t+1})q(\boldsymbol{\alpha}_{t-1}|\mathcal{Y}_t)} \tag{5.33}$$

$$\propto w_{t-1}^j \frac{p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1})}{q(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}, \mathcal{Y}_{t+1})}. \tag{5.34}$$

5.A.3 Particle filter implementation

The implementation of the particle filter is equivalent to the Sampling Importance Resampling (SIR) algorithm described by Arulampalam et al. (2002), which is also used by Flury and Shephard (2011), where it is referred to as a generic particle filter.

1. Set the initial values for the states to the unconditional expected values, $\boldsymbol{\alpha}_{-1} = (0,0)'$, and draw N initial particles from $p(\boldsymbol{\alpha}_0|\mathcal{Y}_0, \boldsymbol{\xi}^M)$. Set the initial value of the log likelihood function estimate $\widehat{\ln \mathcal{L}}_0$ to 0. Set $t = 1$.

2. Draw $\boldsymbol{\alpha}_t^j$ for $j = 1, \ldots, N$ from the conditional distribution $p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}^j, \mathcal{Y}_t, \boldsymbol{\xi}^M)$

3. Compute the weights $w_t^j = p(\mathbf{y}_{t+1}|\boldsymbol{\alpha}_t^j)$ and normalize $\tilde{w}_t^j = \frac{w_t^j}{\sum_{k=1}^N w_t^k}$ for $j = 1, \ldots, N$.

4. Add the log likelihood contribution for t: $\widehat{\ln \mathcal{L}}_t = \widehat{\ln \mathcal{L}}_{t-1} + \ln\left(\frac{1}{N}\sum_{j=1}^N w_t^j\right)$

5. Draw N uniform random numbers $u^j \sim U(0,1)$ and ensure an even distribution over the interval $[0,1]$ by transforming to $\tilde{u}^j = \frac{u^j}{N} + \frac{j-1}{N}$ for $j = 1, \ldots, N$.

6. To resample the particles, for every $j = 1, \ldots, N$, select particle i^j that fulfills the inequality $\sum_{k=1}^{i^j-1} \tilde{w}_t^k < \tilde{u}^j \leq \sum_{k=1}^{i^j} \tilde{w}_t^k$.

7. To obtain the filtered series, record the filtered state values $\hat{\boldsymbol{\alpha}}_t = \frac{1}{N}\sum_{j=1}^N \boldsymbol{\alpha}_t^j$.

8. Set $t = t + 1$ and go back to 2. Repeat until $t = T - 1$.

Due to the temporal structure of the model, the estimate of the likelihood function is given by $\ln \mathcal{L}_{T-1}$. Computing $\ln\left(\frac{1}{N}\sum_{j=1}^N w_T^j\right)$ would require an observation for $\mathbf{y}_{T+1}$. Therefore, the state vector can only be filtered up to $T - 1$.

Table 5.1: Comparison of macro parameter estimates.
The table reports medians *(in italics)* and RMSE (in normal font) of the macro parameter estimates $\boldsymbol{\xi}^{M_*}$ obtained from the Monte Carlo study using Kalman filter-based likelihood estimation. For the results reported in Panel A ν_1 and σ_w are both set to 0 in the Kalman filter, while the results in Panel B are based on ν_1=0.987 and σ_w=$2.3e^{-6}$, which corresponds to the true values. The last column contains the number of successful estimations R out of 400 for each sample size.

	μ_c	μ_d	ρ	φ_e	σ	ϕ	φ_d	R
Panel A: no SV								
T=650	*0.0015*	*0.0015*	*0.969*	*0.0489*	*0.0078*	*3.05*	*4.49*	389
	0.0006	0.0018	0.103	0.0442	0.0005	2.72	0.18	
T=1k	*0.0015*	*0.0016*	*0.973*	*0.0472*	*0.0078*	*2.95*	*4.49*	397
	0.0005	0.0016	0.075	0.0395	0.0004	2.44	0.15	
T=5k	*0.0015*	*0.0016*	*0.978*	*0.0453*	*0.0078*	*3.01*	*4.49*	400
	0.0002	0.0008	0.010	0.0089	0.0002	0.48	0.07	
T=10k	*0.0015*	*0.0016*	*0.978*	*0.0446*	*0.0078*	*3.00*	*4.50*	400
	0.0002	0.0006	0.005	0.0055	0.0001	0.33	0.05	
Panel B: true SV								
T=650	*0.0015*	*0.0015*	*0.969*	*0.0489*	*0.0078*	*3.05*	*4.49*	389
	0.0006	0.0018	0.103	0.0442	0.0005	2.72	0.18	
T=1k	*0.0015*	*0.0016*	*0.973*	*0.0472*	*0.0078*	*2.95*	*4.49*	397
	0.0005	0.0016	0.075	0.0395	0.0004	2.44	0.15	
T=5k	*0.0015*	*0.0016*	*0.978*	*0.0453*	*0.0078*	*3.01*	*4.49*	400
	0.0002	0.0008	0.010	0.0089	0.0002	0.48	0.07	
T=10k	*0.0015*	*0.0016*	*0.978*	*0.0446*	*0.0078*	*3.00*	*4.50*	400
	0.0002	0.0006	0.005	0.0055	0.0001	0.33	0.05	

Table 5.2: SV parameter estimates.
The table reports medians *(in italics)* and RMSE (in normal font) of the stochastic volatility parameter estimates $\hat{\nu}_1$ and $\hat{\sigma}_w$ obtained from the Monte Carlo study using particle filter-based likelihood estimation. The results reported in Panel A are based on the true macro parameters $\boldsymbol{\xi}^{M_*}$, while the results reported in Panel B are based on the macro parameter estimates $\hat{\boldsymbol{\xi}}^{M_*}$ from the first estimation step. The columns labeled by R contain the numbers of successful estimations out of 400 for each sample size.

	Panel A true macro			**Panel B estimated macro**		
	ν_1	σ_w	R	ν_1	σ_w	R
T=650	*0.985*	*2.53e-06*	388	*0.976*	*2.87e-06*	388
	0.066	2.27e-06		0.092	2.64e-06	
T=1k	*0.985*	*2.56e-06*	396	*0.979*	*2.65e-06*	397
	0.062	1.50e-06		0.060	1.69e-06	
T=5k	*0.987*	*2.32e-06*	400	*0.986*	*2.35e-06*	400
	0.005	4.29e-07		0.006	4.31e-07	
T=10k	*0.987*	*2.31e-06*	400	*0.987*	*2.31e-06*	400
	0.004	2.99e-07		0.003	2.86e-07	

Table 5.3: Preference parameter estimates.
The table reports medians *(in italics)* and RMSE (in normal font) of the preference parameter estimates $\boldsymbol{\xi}^P$ obtained from the Monte Carlo study using indirect inference estimation. The results reported in Panel A are based on the true macro parameters $\boldsymbol{\xi}^M$, while the results reported in Panel B are based on the macro parameter estimates $\hat{\boldsymbol{\xi}}^M$ obtained in the first and second estimation step. The columns labeled by R contain the numbers of successful estimations out of 400 for each sample size.

	Panel A true macro				**Panel B estimated macro**			
	δ	γ	ψ	R	δ	γ	ψ	R
T=650	*0.9979*	*9.6*	*1.50*	400	*0.9975*	*12.3*	*1.86*	336
	0.0003	1.2	0.02		0.0436	9.1	5.61	
T=1k	*0.9980*	*10.1*	*1.50*	400	*0.9976*	*12.0*	*1.85*	373
	0.0002	1.0	0.01		0.0018	8.0	3.58	
T=5k	*0.9980*	*9.9*	*1.50*	400	*0.9979*	*10.8*	*1.59*	398
	0.0001	0.5	0.01		0.0006	3.5	0.79	
T=10k	*0.9980*	*9.9*	*1.50*	400	*0.9979*	*10.2*	*1.53*	400
	0.0001	0.3	0.00		0.0004	2.2	0.31	

Table 5.4: Descriptive statistics.
The table reports means, standard deviations, and the first-order autocorrelation of the monthly data used in the empirical application. The three-step estimation strategy allows for different sample periods for macro and financial data. The time span used for each variable is given in the last column.

	mean	**std. dev.**	**AC(1)**	**time span**
log consumption growth g_t	0.0023	0.0057	-0.1281	[1959/02 – 2014/12]
log dividend growth $g_{d,t}$	0.0017	0.0120	-0.0070	[1959/02 – 2014/12]
log market return $r_{m,t}$	0.0058	0.0431	0.0852	[1947/02 – 2014/12]
log risk-free rate $r_{f,t}$	0.0003	0.0015	0.9696	[1947/02 – 2014/12]
log price-dividend ratio $z_{m,t}$	3.4973	0.4197	0.9943	[1947/02 – 2014/12]

Table 5.5: Empirical application results.
The table reports the point estimates with 95% confidence bounds for the empirical application. The first row contains the parameter estimates, while the second and third row comprise the lower and upper bound of the 95% confidence interval, respectively. The confidence bounds are obtained as 2.5% and 97.5% quantiles of the empirical distribution resulting from a parametric bootstrap.

μ_c	μ_d	ρ	φ_e	ν_1	σ_w	σ	ϕ	φ_d	δ	γ	ψ
0.0023	0.0018	0.944	0.0293	0.877	6.3e-06	0.0057	8.80	1.97	0.9815	54.1	2.31
0.0019	0.0000	0.800	0.0059	0.738	3.4e-06	0.0052	3.44	1.80	0.9798	17.2	0.40
0.0029	0.0038	0.977	0.0902	0.945	7.7e-06	0.0061	53.53	2.13	0.9854	73.5	17.73

Table 5.6: Implications of the empirical parameter estimates and their distributions.
The table reports means and standard deviations of the observable LRR model variables g, g_d, z_m, r_m, and r_f implied by the point estimates obtained in the empirical application and the corresponding bootstrap distribution. The first column contains the means and standard deviations of the empirical data for comparison. All quantities computed relate to a monthly frequency.

	data	**model**	**model-implied quantiles**				
			0.005	0.025	0.5	0.975	0.995
$\mathbb{E}(g)$	0.0023	0.0023	0.0018	0.0019	0.0024	0.0029	0.0030
$\mathbb{E}(g_d)$	0.0017	0.0018	0.0000	0.0000	0.0019	0.0038	0.0048
$\mathbb{E}(z_m)$	3.4973	3.9763	3.9503	3.9575	3.9864	4.0529	4.0710
$\mathbb{E}(r_m)$	0.0058	0.0204	0.0186	0.0188	0.0202	0.0219	0.0222
$\mathbb{E}(r_f)$	0.0003	0.0183	0.0178	0.0181	0.0186	0.0207	0.0210
$\sigma(g)$	0.0057	0.0057	0.0051	0.0053	0.0057	0.0062	0.0063
$\sigma(g_d)$	0.0120	0.0120	0.0107	0.0111	0.0120	0.0130	0.0132
$\sigma(z_m)$	0.4197	0.0574	0.0102	0.0173	0.0479	0.0954	0.1126
$\sigma(r_m)$	0.0431	0.0217	0.0133	0.0144	0.0200	0.0261	0.0283
$\sigma(r_f)$	0.0015	0.0002	0.0000	0.0000	0.0002	0.0014	0.0019

Figure 5.1: Monte Carlo results: distribution of first-step maximum likelihood estimates.
The figure displays kernel estimates of the macro parameters $\boldsymbol{\xi}^{M_*}$ across different simulated sample sizes. The beta kernel proposed by Chen (1999) is used with the bandwidth selector by Silverman (1986) adjusted for variable kernels. Vertical lines indicate the positions of the true parameters.

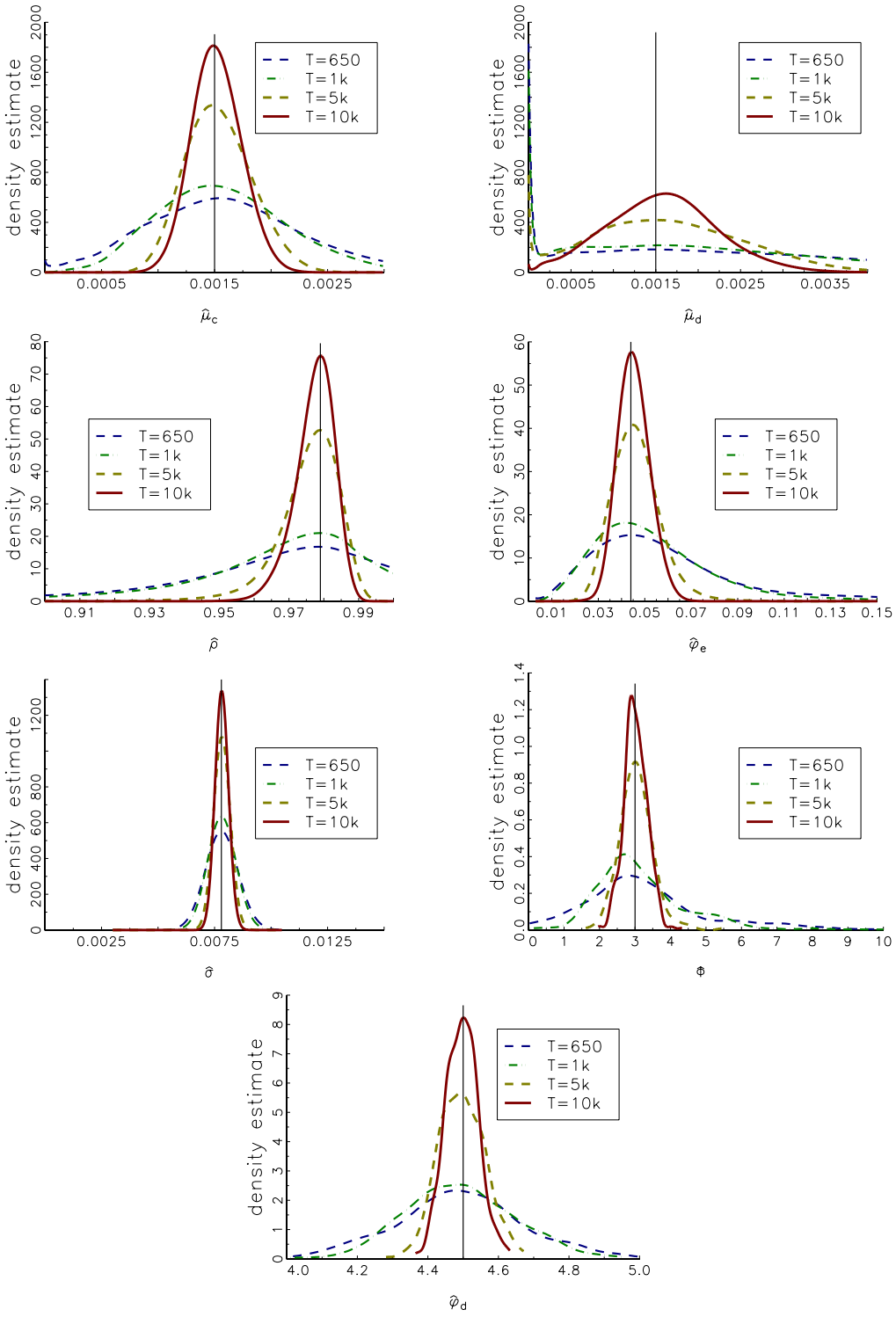


Figure 5.2: Simulation: second-step objective function.
The figure displays the objective function of the second-step estimation for a simulated sample of size T=1k based on the true macro parameter values $\boldsymbol{\xi}^{M_*}$. Panel A illustrates the objective function surface from two angles for the coarse-meshed grid; Panel B shows the same objective function for the finer grid.

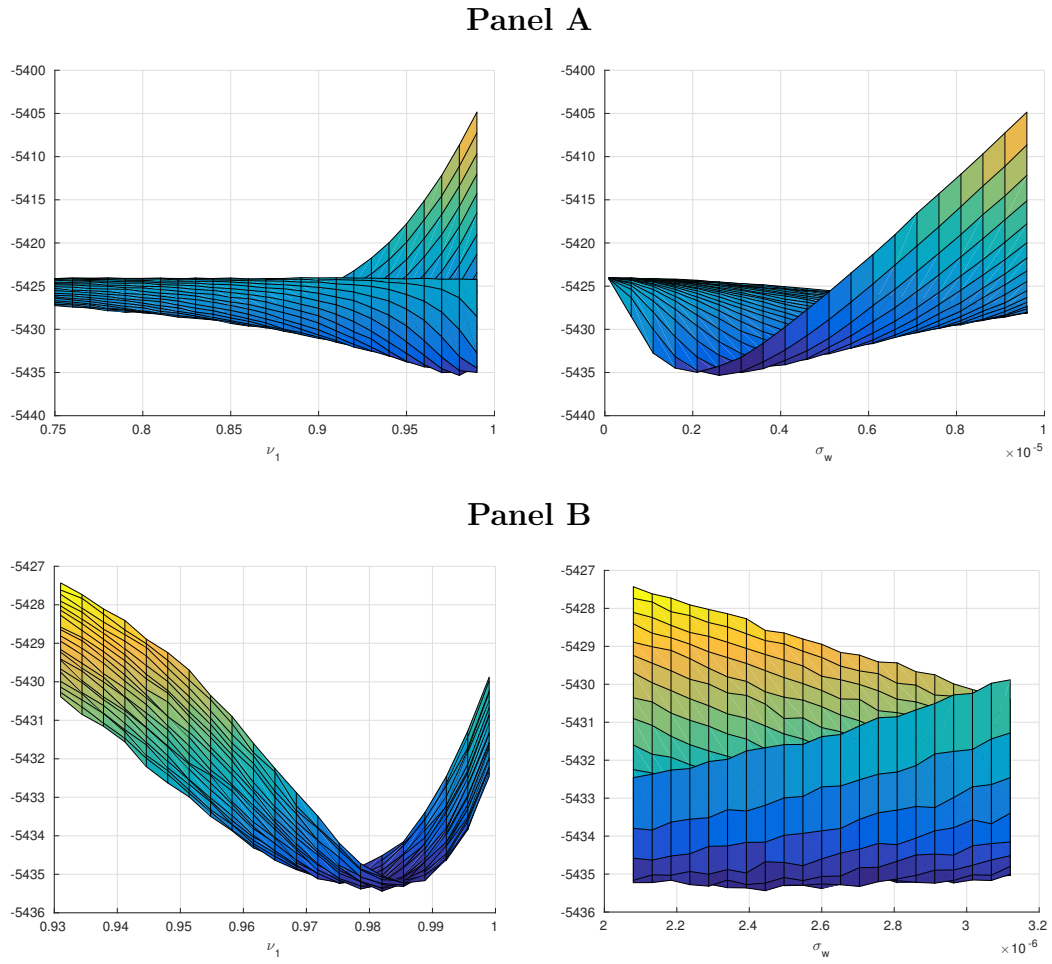
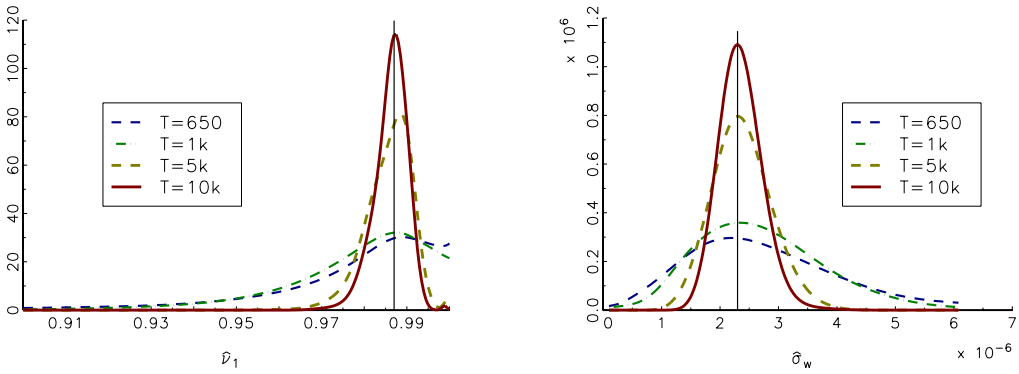


Figure 5.3: Monte Carlo results: distribution of second-step maximum likelihood estimates.
The figure displays kernel estimates of the SV parameters ν_1 and σ_w across different simulated sample sizes. The graphs in Panel A are based on the true values of the macro parameters $\boldsymbol{\xi}^{M_*}$, whereas those in Panel B are based on the first-step macro parameter estimates. The beta kernel proposed by Chen (1999) is used with the bandwidth selector by Silverman (1986) adjusted for variable kernels. Vertical lines indicate the positions of the true parameters.

Panel A



Panel B

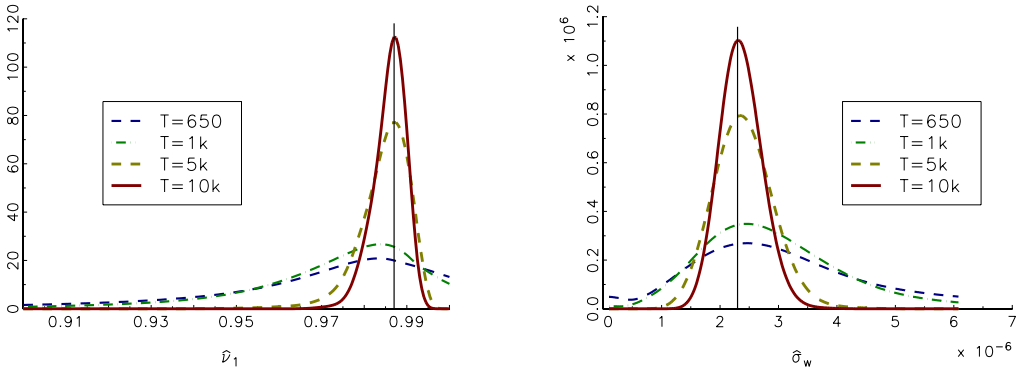


Figure 5.4: Monte Carlo results: distribution of third-step maximum likelihood estimates.
The figure displays kernel estimates of the preference parameters δ, γ, and ψ across different simulated sample sizes. The parameter estimates illustrated in Panel A are based on the true macro parameter values, while the kernel estimates in Panel B are based on the first-step macro parameter estimates and the second-step SV parameter estimates. Vertical lines indicate the positions of the true parameters.

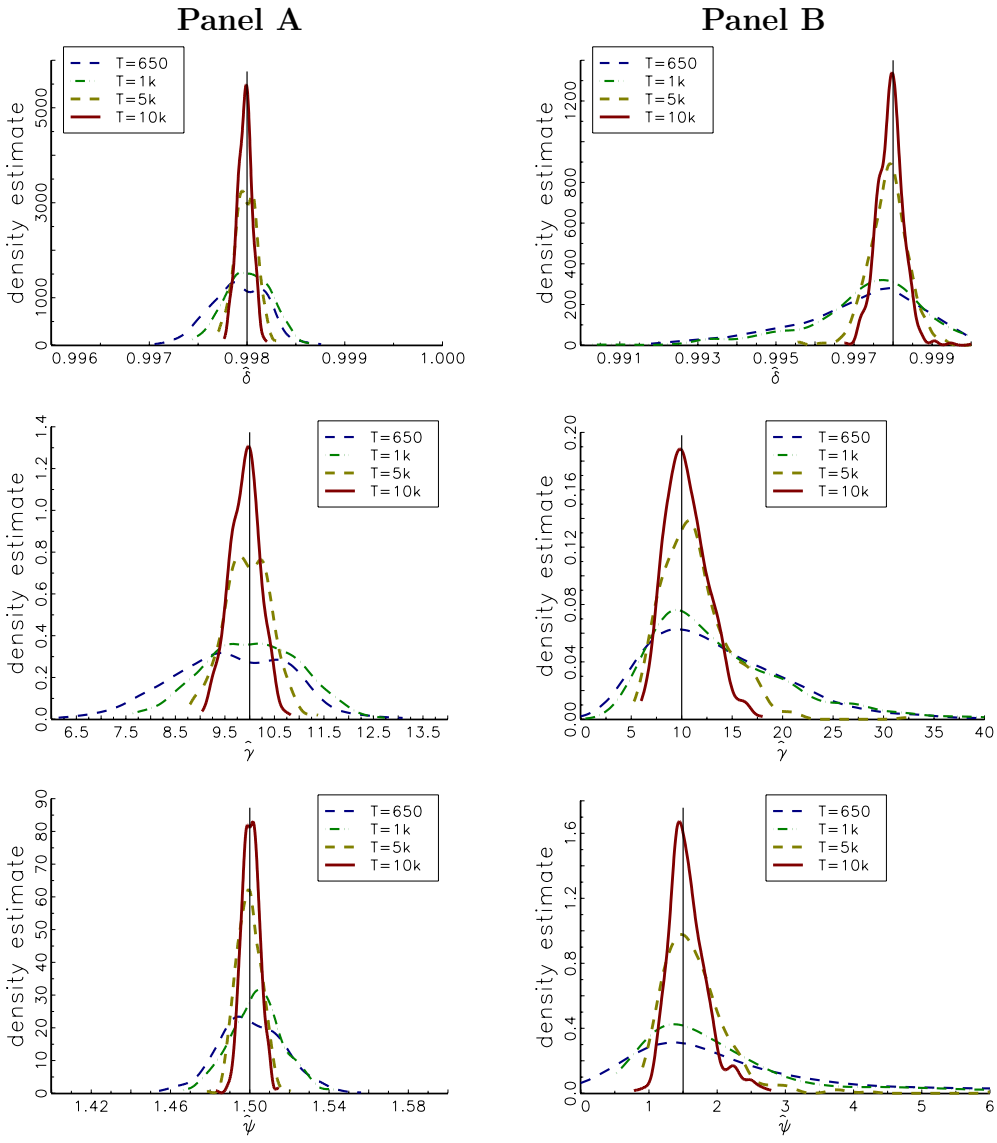
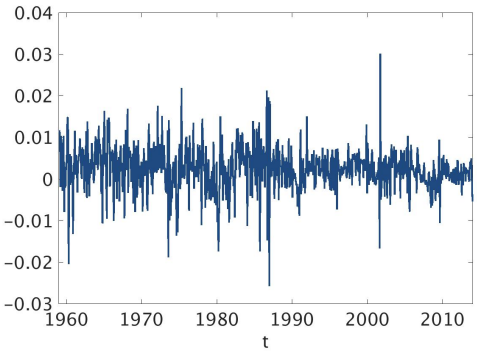
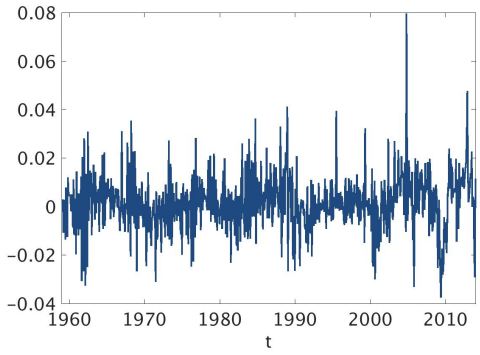


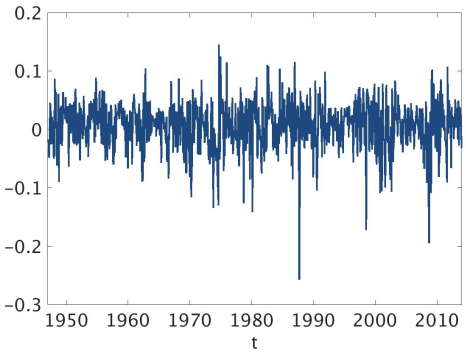
Figure 5.5: Empirical data series.
The figure illustrates the time series used in the empirical application. The sample period spans February 1959 to December 2014 for g and g_d and February 1947 to December 2014 for r_m, r_f, and z_m.



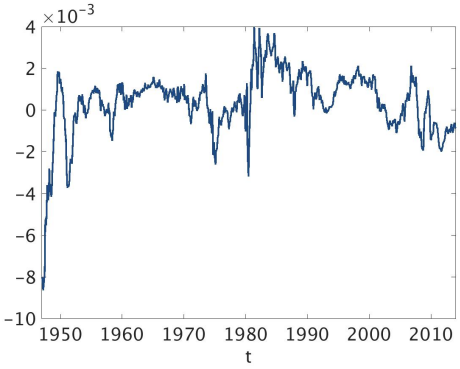
(a) log consumption growth g



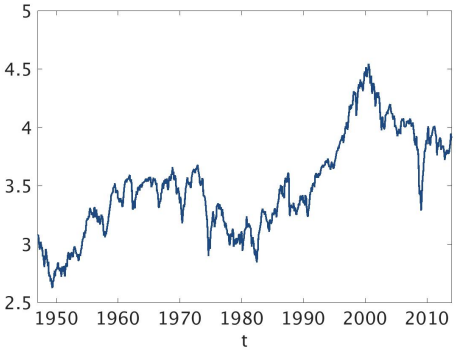
(b) log dividend growth g_d



(c) log market return r_m



(d) log risk-free rate r_f



(e) log price-dividend ratio z_m

Chapter 6

Conclusion

This thesis analyzes the impact of long-run risk on the decisions of a representative investor. It provides a methodological and empirical evaluation of the idea that non-diversifiable fluctuations in long-run expectations of consumption growth and economic uncertainty explain the considerable risk premia observed in financial markets in the past 70 years. A theoretical framework—the long-run risk asset pricing model suggested by Bansal and Yaron (2004)—that accounts for those sources of risk is subjected to a broad spectrum of econometric methods. Due to the intricate model structure and the presence of persistent latent variables and endogenous parameters, the econometric analysis of the LRR model is highly demanding. This work seeks to overcome identification issues detected in the previous literature and to provide a realistic picture of the estimation quality that can be expected when estimating a complex structural asset pricing model based on the limited macro-finance data that are currently available.

The defining concept behind all estimation strategies is their consistency with the recursive model structure, meaning that the model is always estimated in multiple steps. In line with the model, the parameters that determine the shape of the

macro dynamics are estimated independently of the representative investor's preference parameters. For all estimation strategies, identification of the parameters is ascertained by means of Monte Carlo studies, which demonstrate the validity of each method.

The two-step GMM/SMM estimation strategy introduced in Chapter 3 exploits the availability of analytical moments for the macro sub-model. A GMM estimation is applied in the first step, while SMM is used for the estimation of the preference parameters in the second step. For the latter task, SMM is ideally suited, since the necessity to solve for the endogenous parameters calls for simulation-based methods. The Monte Carlo study shows that long series of auto-moments of consumption and dividend growth are required to elicit the persistence of the small, latent growth process, the driving feature of long-run consumption risk. Furthermore, a precise estimation of the preference parameters is found to be possible, given high-quality macro parameter estimates. A moment sensitivity analysis helps to carefully select moment matches that provide meaningful information on the parameters in question. This analysis reveals that none of the considered moments, not even the fourth moments of the growth processes, measurably respond to the persistence and the volatility of the latent stochastic variance process. This finding substantiates the serious identification issues related to moment-based estimation of the stochastic volatility parameters. As a solution, we approximate the time-varying stochastic variance in the second-step estimation by the unconditional variance estimate from the first estimation step. This approach does not impair the estimation quality of the remaining parameters, as documented in the Monte Carlo study. The empirical results lend some support to the long-run risk paradigm on the one hand, as the important intertemporal elasticity of substitution is estimated to be larger than 1. On the other hand, the risk aversion parameter estimate is still very large, despite

accounting for long-run risk, implying that introducing this additional source of risk does not suffice to reconcile the large equity premium with plausible preference parameters within a consumption-based asset pricing model. Moreover, estimating the LRR model at a quarterly frequency, which is inevitable for this estimation strategy when using quarterly data, complicates tracing the supposedly high persistence of the long-run risk component at the monthly frequency. A persistence parameter very close to 1, however, is key for the model's ability to produce high risk premia without inflating the risk aversion parameter. The precision of the preference parameter estimates is low, due to imprecise macro parameter estimates owing to the limited amount of data. We conclude that the quality of the macro estimates must be improved to obtain precise estimates for the preference parameters.

The indirect inference estimation strategy presented in Chapter 4 breaks the link between model and data frequency, as the method is entirely simulation-based and thus permits arbitrary aggregation of the simulated data to the frequency of the empirical data. Most importantly, this feature allows for a reasonable decision interval of the representative investor, which is typically assumed to be one month. Taking account of the findings of Chapter 3, indirect inference is also conducted in two consecutive steps, thereby observing the LRR model structure. In each step, the auxiliary model used for the estimation is tailored to the key characteristics of the relevant variables. In particular, the slow-moving long-run risk component can be detected in a more parsimonious way compared to the two-step GMM/SMM approach, which requires a large number of auto-moment matches. Instead, the persistence in growth expectations is captured by means of a heterogeneous autoregressive model designed for long-memory data, which uses past aggregates of the growth rates. The lack of identification of the stochastic volatility parameters encountered in the GMM/SMM estimation cannot be resolved, regardless of the broader range of possi-

bilities afforded by indirect inference estimation. Again, the time-varying stochastic variance must be approximated by its unconditional expected value. Provided precise macro parameter estimates are available, the investor's preference parameters can be estimated accurately. Estimating the parameters at a monthly frequency from quarterly empirical data reveals that there is indeed evidence for the existence of a small predictable component in consumption and dividend growth. Another favorable result is the comparatively low estimate of the risk aversion parameter, leaving room for hope that the LRR paradigm can ultimately explain the equity premium puzzle without large risk aversion of the investor. However, the estimate of the intertemporal elasticity of substitution is below 1, which rules out large model-implied equity premia. Furthermore, the precision of the empirical results is rather limited. All things considered, the implications of this empirical test of the LRR model are somewhat inconclusive and warrant further investigation.

Given the results of the two-step GMM/SMM estimation and the indirect inference method, two major methodological issues remain unresolved. In the first place, the persistence and the volatility of the stochastic variance process, which represents the fluctuating economic uncertainty, cannot be estimated using the aforementioned approaches. Furthermore, the precision of the macro estimates is low for small samples, which in turn hampers a precise estimation of the preference parameters. Despite the presence of latent variables, filtering methods allow for a maximum likelihood estimation strategy, which is introduced in Chapter 5. This approach can indeed identify all structural parameters, including the stochastic volatility parameters. By taking into account the full distribution of the macro variables instead of isolated moments only, an efficiency gain is realized for the macro parameter estimates. The method disentangles the estimation of the macro parameters and that of the preference parameters and thus adheres to the LRR model structure. In a

first step, a Kalman filter-based maximum likelihood estimation is conducted that identifies the macro parameters except for the stochastic volatility parameters, to which it is invariant. In contrast to the Kalman filter, which only provides updated projections of the conditional mean, a particle filter provides estimates of the full conditional distribution of the latent variance process. Thus, applying a particle filter within a maximum likelihood estimation in a second step allows to estimate the persistence and the volatility of the stochastic variance. Finally, in the third step, the preference parameters are estimated using the indirect inference approach proven and tested in Chapter 4. Applying the estimation strategy to empirical data again provides support for the LRR paradigm by detecting a persistent component in consumption and dividend growth. The intertemporal elasticity of substitution is estimated to be larger than 1, as postulated by the architects of the long-run risk concept. However, the risk aversion estimate is large and the estimation precision is limited.

This comprehensive econometric analysis of the LRR model shows that the estimation of its structural parameters is highly demanding and that seemingly straightforward estimation approaches are subject to serious identification issues. A reliable econometric test of the LRR concept should observe the model structure and beware of turning into a mere goodness-of-fit exercise that introduces links in the estimation procedure where there are none intended in the model. The insights collected from the application of different estimation strategies lead to the conclusion that the currently available sample sizes prevent a precise estimation of the LRR model parameters. However, the estimation strategies presented in this thesis provide guidance for the estimation of complex structural (asset pricing) models and draw a realistic picture of what we can expect in terms of precision when estimating such models from small macro-finance data sets.

Appendix

A.1 Linear approximations

To model the dependence between log returns and the log price-dividend ratio, Bansal and Yaron (2004) resort to a linear approximation suggested by Campbell and Shiller (1988). The linear relationship between the log return h_t and the log dividend-price ratio δ_t, as suggested by Campbell and Shiller (1988), can be derived as follows:

$$\begin{aligned} h_t &= \ln\left(P_{t+1} + D_t\right) - \ln(P_t) \\ &= \ln\left(P_t + D_{t-1}\right) + \Delta \ln\left(P_{t+1} + D_t\right) - \ln(P_t). \end{aligned}$$

We use a first-order Taylor series expansion for $\Delta \ln\left(P_{t+1} + D_t\right)$ at $P_{t+1} = P_t$ and $D_t = D_{t-1}$:

$$\begin{aligned} \Delta \ln\left(P_{t+1} + D_t\right) = \ln(P_{t+1} + D_t) - \ln(P_t + D_{t-1}) &= \ln\left(\frac{P_{t+1} + D_t}{P_t + D_{t-1}}\right) \\ &\approx \ln(1) + \frac{1}{P_t + D_{t-1}}\left[P_{t+1} + D_t - P_t - D_{t-1}\right] \\ &= \frac{P_{t+1} - P_t}{P_t + D_{t-1}} + \frac{D_t - D_{t-1}}{P_t + D_{t-1}}. \end{aligned}$$

Assuming the price is a constant fraction ρ of the price including the dividends, $P_t \approx \rho\,(P_t + D_{t-1})$, and hence, $D_{t-1} \approx (1-\rho)\,(P_t + D_{t-1})$, we can approximate:

$$\begin{aligned}\Delta \ln\left(P_{t+1} + D_t\right) &\approx \rho \frac{P_{t+1} - P_t}{P_t} + (1-\rho)\frac{D_t - D_{t-1}}{D_{t-1}} \\ &\approx \rho \Delta \ln(P_{t+1}) + (1-\rho)\Delta \ln(D_t).\end{aligned}$$

Inserting these results into the expression for h_t yields:

$$\begin{aligned}h_t &\approx \ln\left(P_t + D_{t-1}\right) + \rho \Delta \ln(P_{t+1}) + (1-\rho)\Delta \ln(D_t) - \ln(P_t) \\ &= \ln\left(P_t + D_{t-1}\right) + \rho\left(p_{t+1} - p_t\right) + (1-\rho)\left(d_t - d_{t+1}\right) - p_t \\ &= \ln\left(P_t + D_{t-1}\right) + \rho p_{t+1} + (1-\rho)d_t - (1-\rho)(d_{t-1} - p_t) - 2p_t \\ &= \ln\left(\frac{P_t + D_{t-1}}{P_t}\right) - (1-\rho)\delta_t + \rho p_{t+1} + (1-\rho)d_t - p_t \\ &\approx -\ln(\rho) - (1-\rho)\delta_t + \rho p_{t+1} + (1-\rho)d_t - p_t \\ &= k + \rho p_{t+1} + (1-\rho)d_t - p_t.\end{aligned}$$

Note that Campbell and Shiller (1988) model a log dividend-price ratio δ_t, whereas the LRR model refers to the log price-dividend ratio $z_{m,t}$. Translating this result into the notation used by Bansal and Yaron (2004) yields:

$$\begin{aligned}r_{m,t} &= -\ln(\rho) - (1-\rho)(d_{t-1} - p_t) + \rho p_{t+1} + (1-\rho)d_t - p_t \\ &= -\ln(\rho) + (1-\rho)(p_t - d_{t-1}) + \rho(p_{t+1} - d_t) - (p_t - d_{t-1}) + d_t - d_{t-1} \\ &= \kappa_{0,m} + \kappa_{1,m} z_{m,t} - z_{m,t-1} + g_{d,t},\end{aligned}$$

where $\kappa_{0,m}$ and $\kappa_{1,m}$ are given by:

$$\kappa_{0,m} = -\ln(\rho) + (1-\rho)z_{m,t-1}$$
$$\kappa_{1,m} = \rho.$$

We can rewrite $\kappa_{0,m}$ and $\kappa_{1,m}$ as follows:

$$\begin{aligned}\kappa_{1,m} &\approx \frac{P_t}{P_t + D_{t-1}} = \frac{1}{\frac{P_t + D_{t-1}}{P_t}} = \frac{1}{1 + \frac{1}{\exp(z_{m,t})}} \\ &= \frac{\exp(z_{m,t})}{1 + \exp(z_{m,t})}.\end{aligned}$$

Because ρ and thus $\kappa_{1,m}$ should be a constant ratio, we use a time average to obtain a constant value:

$$\kappa_{1,m} \approx \frac{\exp(\bar{z}_m)}{1 + \exp(\bar{z}_m)}.$$

For $\kappa_{0,m}$ to be a constant, we also use a time average to obtain a constant value:

$$\begin{aligned}\kappa_{0,m} &\approx -\ln(\kappa_{1,m}) + (1-\kappa_{1,m})\bar{z}_m = \ln\left(\frac{1+\exp(\bar{z}_m)}{\exp(\bar{z}_m}\right) + \bar{z}_m - \kappa_{1,m}\bar{z}_m \\ &= \ln\left(1 + \exp(\bar{z}_m)\right) - \kappa_{1,m}\bar{z}_m.\end{aligned}$$

A.2 Return on the aggregate wealth portfolio

To find the expressions for the coefficients A_0, A_1, and A_2 in Equation (2.12), we use the basic asset pricing equation with the SDF from Equation (2.7):

$$\mathbb{E}_t\left[\delta^\theta G_{t+1}^{-\frac{\theta}{\psi}} R_{a,t+1}^{-(1-\theta)} R_{i,t+1}\right] = 1.$$

Taking the logarithm of Equation (2.7) yields:

$$m_{t+1} = \ln(M_{t+1}) = \theta\ln(\delta) - \frac{\theta}{\psi} g_{t+1} + (\theta-1) r_{a,t+1},$$
$$\text{where} \quad g_{t+1} = \ln(G_{t+1}) \quad \text{and} \quad r_{a,t+1} = \ln(R_{a,t+1}).$$

It follows that

$$\begin{aligned} 1 &= \mathbb{E}_t\left[\exp(\ln(M_{t+1}) + r_{i,t+1})\right] \\ &= \mathbb{E}_t\left[\exp(\theta\ln(\delta) - \frac{\theta}{\psi} g_{t+1} + (\theta-1) r_{a,t+1} + r_{i,t+1})\right]. \end{aligned}$$

The model must price any return, so the Euler equation also holds for $r_{i,t+1} = r_{a,t+1}$:

$$\begin{aligned} 1 &= \mathbb{E}_t\left[\exp\left(\theta\ln(\delta) - \frac{\theta}{\psi} g_{t+1} + \theta r_{a,t+1}\right)\right] \\ &= \exp\left(\mathbb{E}_t\left[m_{t+1} + r_{a,t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[m_{t+1} + r_{a,t+1}\right]\right) \\ 0 &= \mathbb{E}_t\left[m_{t+1} + r_{a,t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[m_{t+1} + r_{a,t+1}\right]. \end{aligned}$$

Inserting the linear approximation for $r_{a,t+1}$, we obtain:

$$\begin{aligned}0 = {} & \theta \ln \delta - \frac{\theta}{\psi}\mathbb{E}_t(g_{t+1}) \\ & + \theta\left[\kappa_0 + \kappa_1 A_0 + \kappa_1 A_1 \mathbb{E}_t(x_{t+1}) + \kappa_1 A_2 \mathbb{E}_t(\sigma^2_{t+1}) - A_0 - A_1 x_t - A_2\sigma^2_t + \mathbb{E}_t(g_{t+1})\right] \\ & + \frac{1}{2}\left[\left(\theta - \frac{\theta}{\psi}\right)^2 \mathrm{Var}_t(g_{t+1}) + \theta^2\left(\kappa_1^2 A_1^2 \mathrm{Var}_t(x_{t+1}) + \kappa_1^2 A_2^2 \mathrm{Var}_t(\sigma^2_{t+1})\right)\right]\end{aligned}$$

because $\mathrm{Cov}_t(g_{t+1}, x_{t+1}) = 0$, $\mathrm{Cov}_t(g_{t+1}, \sigma^2_{t+1}) = 0$ and $\mathrm{Cov}_t(x_{t+1}, \sigma^2_{t+1}) = 0$.

It follows that:

$$\begin{aligned}0 = {} & \theta \ln \delta - \frac{\theta}{\psi}(\mu_c + x_t) + \theta\left[\kappa_0 + \kappa_1 A_0 + \kappa_1 A_1 \rho x_t + \kappa_1 A_2(\sigma^2 + \nu_1(\sigma^2_t - \sigma^2)) - A_0\right. \\ & \left. - A_1 x_t - A_2\sigma^2_t + \mu_c + x_t\right] + \frac{1}{2}\left[\left(\theta - \frac{\theta}{\psi}\right)^2 \sigma^2_t + \theta^2\left(\kappa_1^2 A_1^2 \varphi_e^2 \sigma^2_t + \kappa_1^2 A_2^2 \sigma^2_w\right)\right].\end{aligned} \tag{A-1}$$

Equation (A-1) must hold for all values of x_t, which means that all terms involving x_t must cancel out:

$$\begin{aligned}-\frac{\theta}{\psi}x_t + \theta\kappa_1 A_1 \rho x_t - \theta A_1 x_t + \theta x_t &\stackrel{!}{=} 0 \\ -\frac{\theta}{\psi}x_t + \theta\left[\kappa_1 A_1 \rho x_t - A_1 x_t + x_t\right] &= 0.\end{aligned} \tag{A-2}$$

Equation (A-1) also has to hold for all values of σ^2_t:

$$\begin{aligned}\theta\kappa_1 A_2 \nu_1 \sigma^2_t - \theta A_2 \sigma^2_t + \frac{1}{2}\left[\left(\theta - \frac{\theta}{\psi}\right)^2 \sigma^2_t + \theta^2 A_1^2 \kappa_1^2 \varphi_e^2 \sigma^2_t\right] &\stackrel{!}{=} 0 \\ \left[\theta\left(\kappa_1 \nu_1 A_2 - A_2\right) + \frac{1}{2}\left(\theta - \frac{\theta}{\psi}\right)^2 + \frac{1}{2}(\theta A_1 \kappa_1 \varphi_e)^2\right]\sigma^2_t &= 0.\end{aligned} \tag{A-3}$$

Equation (A-2) leads to the expression for the parameter A_1:

$$\begin{aligned}-\frac{\theta}{\psi} + \theta\left[\kappa_1 A_1 \rho - A_1 + 1\right] &= 0 \\ A_1 = \frac{(\frac{\theta}{\psi} - \theta)}{\theta\kappa_1\rho - \theta} &= \frac{1 - \frac{1}{\psi}}{1 - \kappa_1\rho}.\end{aligned} \tag{A-4}$$

Equation (A-3) leads to the expression for the parameter A_2:

$$\left[\theta\left(\kappa_1\nu_1 - 1\right)A_2 + \frac{1}{2}\left(\theta - \frac{\theta}{\psi}\right)^2 + \frac{1}{2}(\theta A_1\kappa_1\varphi_e)^2\right]\sigma_t^2 = 0$$
$$-\frac{1}{2}\left[\left(\theta - \frac{\theta}{\psi}\right)^2 + (\theta A_1\kappa_1\varphi_e)^2\right] = \theta(\kappa_1\nu_1 - 1)A_2$$

$$A_2 = \frac{1}{2}\frac{\left(\theta - \frac{\theta}{\psi}\right)^2 + (\theta A_1\kappa_1\varphi_e)^2}{\theta[1 - \kappa_1\nu_1]}. \tag{A-5}$$

The constant A_0 can be obtained by setting the sum of all x_t and σ_t^2 terms in Equation (A-1) to zero:

$$\begin{aligned}
0 &= \theta\ln\delta - \frac{\theta}{\psi}\mu_c + \theta\left[\kappa_0 + \kappa_1 A_0 + \kappa_1 A_2(1-\nu_1)\sigma^2 - A_0 + \mu_c\right] + \frac{1}{2}\theta^2\left(\kappa_1^2 A_2^2\sigma_w^2\right)\\
0 &= \ln\delta + \left(1 - \frac{1}{\psi}\right)\mu_c + \kappa_0 + (\kappa_1 - 1)A_0 + \kappa_1(1-\nu_1)\sigma^2 A_2 + \frac{1}{2}\theta(\kappa_1 A_2\sigma_w)^2\\
A_0 &= \frac{1}{1-\kappa_1}\left[\ln\delta + \left(1 - \frac{1}{\psi}\right)\mu_c + \kappa_0 + \kappa_1 A_2\sigma^2(1-\nu_1) + \frac{\theta}{2}(\kappa_1 A_2\sigma_w)^2\right].
\end{aligned} \tag{A-6}$$

A.3 Representation of the market return

According to Equation (2.9), combined with Equations (2.1) and (2.13), $r_{m,t+1}$ is given by:

$$\begin{aligned}
r_{m,t+1} = {} & \kappa_{0,m} + \kappa_{1,m}\left[A_{0,m} + A_{1,m}x_{t+1} + A_{2,m}\sigma_{t+1}^2\right]\\
& - \left[A_{0,m} + A_{1,m}x_t + A_{2,m}\sigma_t^2\right] + \mu_d + \phi x_t + \varphi_d\sigma_t u_{t+1}.
\end{aligned}$$

Applying the basic pricing equation to $r_{m,t}$, we can derive the expressions for $A_{0,m}$, $A_{1,m}$, and $A_{2,m}$:

$$1 = \mathbb{E}_t\left[\exp(m_{t+1} + r_{m,t+1})\right]$$

$$1 = \exp\left(\mathbb{E}_t\left[m_{t+1} + r_{m,t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[m_{t+1} + r_{m,t+1}\right]\right)$$

$$\begin{aligned} 0 = \theta \ln\delta - \frac{\theta}{\psi}(\mu_c + x_t) + (\theta - 1)\Big[\kappa_0 + \kappa_1 A_0 + \kappa_1 A_1 \rho x_t + \kappa_1 A_2(\sigma^2 + \nu_1(\sigma_t^2 - \sigma^2)) \\ - A_0 - A_1 x_t - A_2\sigma_t^2 + \mu_c + x_t\Big] + \kappa_{0,m} + \kappa_{1,m}A_{0,m} \\ + \kappa_{1,m}A_{1,m}\rho x_t + \kappa_{1,m}A_{2,m}(\sigma^2 + \nu_1(\sigma_t^2 - \sigma^2)) - A_{0,m} \\ - A_{1,m}x_t - A_{2,m}\sigma_t^2 + \mu_d + \phi x_t + \frac{1}{2}\mathrm{Var}_t(m_{t+1} + r_{m,t+1}). \end{aligned} \tag{A-7}$$

Derive the expression for $\mathrm{Var}_t(m_{t+1} + r_{m,t+1})$:

$$\begin{aligned} \mathrm{Var}_t(m_{t+1} + r_{m,t+1}) &= \mathrm{Var}_t\left[\theta\ln\delta - \frac{\theta}{\psi}g_{t+1} + (\theta-1)r_{a,t+1} + r_{m,t+1}\right] \\ &= \mathrm{Var}_t\Bigg[-\frac{\theta}{\psi}g_{t+1} + (\theta-1)\Big[\kappa_0 + \kappa_1(A_0 + A_1x_{t+1} + A_2\sigma_{t+1}^2) \\ &\quad - A_0 - A_1x_t - A_2\sigma_t^2 + g_{t+1}\Big] + \kappa_{0,m} + \kappa_{1,m}A_{0,m} \\ &\quad + \kappa_{1,m}A_{1,m}x_{t+1} + \kappa_{1,m}A_{2,m}\sigma_{t+1}^2 - A_{0,m} - A_{1,m}x_t - A_{2,m}\sigma_t^2 \\ &\quad + \mu_d + \phi x_t + \varphi_d\sigma_t u_{t+1}\Bigg] \\ &= \mathrm{Var}_t\Bigg[\left(\theta - 1 - \frac{\theta}{\psi}\right)g_{t+1} + (\theta-1)\Big[\kappa_1 A_1 x_{t+1} \\ &\quad + \kappa_1 A_2\sigma_{t+1}^2\Big] + \kappa_{1,m}A_{1,m}x_{t+1} + \kappa_{1,m}A_{2,m}\sigma_{t+1}^2 + \varphi_d\sigma_t u_{t+1}\Bigg] \\ &= \mathrm{Var}_t\Bigg[\left(\theta - 1 - \frac{\theta}{\psi}\right)g_{t+1} + \left((\theta-1)\kappa_1 A_1 + \kappa_{1,m}A_{1,m}\right)x_{t+1} \\ &\quad + \left((\theta-1)\kappa_1 A_2 + \kappa_{1,m}A_{2,m}\right)\sigma_{t+1}^2 + \varphi_d\sigma_t u_{t+1}\Bigg]. \end{aligned}$$

Finally:

$$\begin{aligned} \mathrm{Var}_t(m_{t+1} + r_{m,t+1}) = \left(\theta - 1 - \frac{\theta}{\psi}\right)^2\sigma_t^2 + \left(\left[(\theta-1)\kappa_1 A_1 + \kappa_{1,m}A_{1,m}\right]^2\varphi_e^2 + \varphi_d^2\right)\sigma_t^2 \\ + \left[(\theta-1)\kappa_1 A_2 + \kappa_{1,m}A_{2,m}\right]^2\sigma_w^2, \end{aligned} \tag{A-8}$$

because $\text{Cov}_t(g_{t+1}, x_{t+1}) = 0$, $\text{Cov}_t(g_{t+1}, \sigma^2_{t+1}) = 0$ and $\text{Cov}_t(x_{t+1}, \sigma^2_{t+1}) = 0$.

To derive the coefficient $A_{1,m}$, we insert Equation (A-8) into Equation (A-7) and collect all terms that involve x_t. They are set to zero, because the Euler equation must hold for all values of the state variables:

$$\begin{aligned}
-\frac{\theta}{\psi}x_t + (\theta-1)\left[\kappa_1 A_1 \rho x_t - A_1 x_t + x_t\right] + \kappa_{1,m} A_{1,m} \rho x_t - A_{1,m} x_t + \phi x_t &\stackrel{!}{=} 0 \\
-\frac{\theta}{\psi} + (\theta-1)\left[A_1(\kappa_1\rho - 1) + 1\right] + A_{1,m}(\kappa_{1,m}\rho - 1) + \phi &= 0 \\
-\frac{\theta}{\psi} + (\theta-1)\left[\left(\frac{1}{\psi} - 1\right) + 1\right] + A_{1,m}(\kappa_{1,m}\rho - 1) + \phi &= 0
\end{aligned}$$

$$A_{1,m} = \frac{-\frac{\theta}{\psi} + (\theta-1)\frac{1}{\psi} + \phi}{1 - \kappa_{1,m}\rho} = \frac{\phi - \frac{1}{\psi}}{1 - \kappa_{1,m}\rho}. \tag{A-9}$$

To derive the coefficient $A_{2,m}$, we collect all terms involving σ^2_t and set them to zero, because the Euler equation must hold for all values of the state variables:

$$\begin{aligned}
&(\theta-1)(\kappa_1 A_2 \nu_1 - A_2) + \kappa_{1,m} A_{2,m} \nu_1 - A_{2,m} \\
&+ \frac{1}{2}\left[\left(\theta - 1 - \frac{\phi}{\psi}\right)^2 + \left(\kappa_{1,m} A_{1,m} \varphi_e - (1-\theta)\kappa_1 A_1 \varphi_e\right)^2 + \varphi_d^2\right] \stackrel{!}{=} 0,
\end{aligned}$$

with $(\theta - 1 - \frac{\theta}{\psi}) = \lambda_{m,\eta}$, $(\kappa_{1,m} A_{1,m} \varphi_e) = \beta_{m,e}$, and $((1-\theta)\kappa_1 A_1 \varphi_e) = \lambda_{m,e}$:

$$(1-\theta)(\kappa_1\nu_1 - 1)A_2 - \frac{1}{2}\left[\lambda^2_{m,\eta} + (\beta_{m,e} - \lambda_{m,e})^2 + \varphi_d^2\right] = A_{2,m}(\kappa_{1,m}\nu_1 - 1)$$

$$A_{2,m} = \frac{(1-\theta)(1-\kappa_1\nu_1)A_2 + \frac{1}{2}[\lambda^2_{m,\eta} + (\beta_{m,e} - \lambda_{m,e})^2 + \varphi_d^2]}{(1 - \kappa_{1,m}\nu_1)}. \tag{A-10}$$

To derive $A_{0,m}$, we set the sum of all terms involving x_t and σ_t^2 in Equation (A-7) to zero:

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

$$\begin{aligned}
A_{0,m} = \frac{1}{(1-\kappa_{1,m})}\Bigg[&\theta \ln \delta - \frac{\theta}{\psi}\mu_c + (\theta - 1)\Bigg[\kappa_0 + \kappa_1 A_0 + \kappa_1 A_2(1-\nu_1)\sigma^2 \\
&- A_0 + \mu_c\Bigg] + \kappa_{0,m} + \kappa_{1,m}A_{2,m}\sigma^2(1-\nu_1) + \mu_d \\
&+ \frac{1}{2}\left[(\theta-1)\kappa_1 A_2 + \kappa_{1,m}A_{2,m}\right]^2\sigma_w^2\Bigg].
\end{aligned} \tag{A-11}$$

A.4 Representation of the risk-free rate

The formula for the risk-free rate can be derived by substituting $r_{f,t}$ for $r_{i,t+1}$ into the basic pricing equation:

$$\begin{aligned}
1 &= \mathbb{E}_t\left[\exp(\theta \ln(\delta) - \frac{\theta}{\psi}g_{t+1} + (\theta-1)r_{a,t+1} + r_{f,t})\right] \\
1 &= \exp\left(\mathbb{E}_t\left[\theta \ln(\delta) - \frac{\theta}{\psi}g_{t+1} + (\theta-1)r_{a,t+1} + r_{f,t}\right] + \frac{1}{2}\mathrm{Var}_t\left[-\frac{\theta}{\psi}g_{t+1} + (\theta-1)r_{a,t+1}\right]\right) \\
0 &= \theta \ln(\delta) - \frac{\theta}{\psi}\mathbb{E}_t(g_{t+1}) + (\theta-1)\mathbb{E}_t(r_{a,t+1}) + r_{f,t} + \frac{1}{2}\mathrm{Var}_t\left[\frac{\theta}{\psi}g_{t+1} + (1-\theta)r_{a,t+1}\right].
\end{aligned}$$

The risk-free rate is thus given by:

$$r_{f,t} = -\theta \ln(\delta) + \frac{\theta}{\psi}\mathbb{E}_t(g_{t+1}) + (1-\theta)\mathbb{E}_t(r_{a,t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1}).$$

In addition, $\mathbb{E}_t(r_{a,t+1})$ can be obtained from the definition of $r_{a,t+1}$:

$$\begin{aligned} r_{a,t+1} &= \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{t+1} \\ &= \kappa_0 + \kappa_1 \left[A_0 + A_1 x_{t+1} + A_2 \sigma^2_{t+1}\right] - \left[A_0 + A_1 x_t + A_2 \sigma^2_t\right] + g_{t+1} \\ \mathbb{E}_t(r_{a,t+1}) &= \kappa_0 + \kappa_1 \left[A_0 + A_1 \rho x_t + A_2(\sigma^2 + \nu_1(\sigma^2_t - \sigma^2))\right] - A_0 - A_1 x_t - A_2 \sigma^2_t + \mu_c + x_t. \end{aligned}$$

$\mathrm{Var}_t\left(m_{t+1}\right)$ is computed as follows:

$$\begin{aligned} \mathrm{Var}_t\left(m_{t+1}\right) &= \mathrm{Var}_t\left[\frac{\theta}{\psi} g_{t+1} + (1-\theta) r_{a,t+1}\right] \\ &= \mathrm{Var}_t\left[\frac{\theta}{\psi} g_{t+1} + (1-\theta)(\kappa_1 A_1 x_{t+1} + \kappa_1 A_2 \sigma^2_{t+1} + g_{t+1})\right] \\ &= \mathrm{Var}_t\left[\left(\frac{\theta}{\psi} + 1 - \theta\right) g_{t+1} + (1-\theta)\kappa_1 A_1 x_{t+1} + (1-\theta)\kappa_1 A_2 \sigma^2_{t+1}\right]. \end{aligned}$$

With $\mathrm{Cov}_t(g_{t+1}, x_{t+1}) = 0$:

$$\begin{aligned} &= \left(\frac{\theta}{\psi} + 1 - \theta\right)^2 \mathrm{Var}_t(g_{t+1}) + (1-\theta)^2(\kappa_1 A_1)^2 \mathrm{Var}_t(x_{t+1}) \\ &\quad + (1-\theta)^2(\kappa_1 A_2)^2 \mathrm{Var}_t(\sigma^2_{t+1}) \\ &\text{with } (-\tfrac{\theta}{\psi} + \theta - 1) = \lambda_{m,\eta},\ ((1-\theta)\kappa_1 A_1 \varphi_e) = \lambda_{m,e}, \text{ and } (1-\theta)\kappa_1 A_2 = \lambda_{m,w}: \\ &= \lambda^2_{m,\eta}\sigma^2_t + \lambda^2_{m,e}\sigma^2_t + \lambda^2_{m,w}\sigma^2_w. \end{aligned}$$

A.5 Risk premia

The gross risk-free rate is given by:

$$R_{f,t+1} = \frac{1}{\mathbb{E}_t(M_{t+1})}.$$

The log return on the risk-free asset is given by:

$$\begin{aligned}
\ln(R_{f,t+1}) &= -\ln\left[\mathbb{E}_t(M_{t+1})\right] \\
r_{f,t+1} &= -\ln\left[\mathbb{E}_t\left(\exp(m_{t+1})\right)\right] \\
&= -\ln\left[\exp\left(\mathbb{E}_t(m_{t+1}) + \frac{1}{2}\mathrm{Var}_t(m_{t+1})\right)\right] \\
&= -\mathbb{E}_t(m_{t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1}). \qquad \text{(A-12)}
\end{aligned}$$

To derive the risk premium on the aggregate wealth portfolio, the following Euler equation can be used to obtain the relation between $\mathbb{E}_t(m_{t+1})$ and $\mathbb{E}_t(r_{a,t+1})$:

$$\begin{aligned}
\mathbb{E}_t\left[M_{t+1}R_{a,t+1}\right] &= 1 \\
\mathbb{E}_t\left[\exp\left(m_{t+1} + r_{a,t+1}\right)\right] &= 1 \\
\exp\left(\mathbb{E}_t(m_{t+1} + r_{a,t+1}) + \frac{1}{2}\mathrm{Var}_t(m_{t+1} + r_{a,t+1})\right) &= 1 \\
\mathbb{E}_t(m_{t+1}) + \mathbb{E}_t(r_{a,t+1}) + \frac{1}{2}\mathrm{Var}_t(m_{t+1} + r_{a,t+1}) &= 0.
\end{aligned}$$

Finally:

$$\mathbb{E}_t(m_{t+1}) = -\mathbb{E}_t(r_{a,t+1}) - \frac{1}{2}\left[\mathrm{Var}_t(m_{t+1}) + \mathrm{Var}_t(r_{a,t+1}) + 2\ \mathrm{Cov}_t(m_{t+1}, r_{a,t+1})\right]. \qquad \text{(A-13)}$$

In a next step, Equation (A-12) can be combined with Equation (A-13) to determine the risk premium on the aggregate wealth portfolio:

$$\begin{aligned}
\mathbb{E}_t\left[r_{a,t+1} - r_{f,t+1}\right] &= \mathbb{E}_t\left[r_{a,t+1} + \mathbb{E}_t(m_{t+1}) + \frac{1}{2}\mathrm{Var}_t(m_{t+1})\right] \\
&= \mathbb{E}_t\Big[r_{a,t+1} - \mathbb{E}_t(r_{a,t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1}) - \frac{1}{2}\mathrm{Var}_t(r_{a,t+1}) \\
&\quad - \mathrm{Cov}_t(m_{t+1}, r_{a,t+1}) + \frac{1}{2}\mathrm{Var}_t(m_{t+1})\Big] \\
&= -\mathrm{Cov}_t\left[m_{t+1}, r_{a,t+1}\right] - \frac{1}{2}\mathrm{Var}_t(r_{a,t+1}) \\
&= -\mathrm{Cov}_t\left[m_{t+1} - \mathbb{E}_t(m_{t+1}), r_{a,t+1} - \mathbb{E}_t(r_{a,t+1})\right] - \frac{1}{2}\mathrm{Var}_t(r_{a,t+1}).
\end{aligned}$$

To write the risk premium in detail, the expressions for $m_{t+1} - \mathbb{E}_t(m_{t+1})$ and $r_{a,t+1} - \mathbb{E}_t(r_{a,t+1})$ must be derived explicitly:

$$\begin{aligned}
r_{a,t+1} &= \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{t+1} \\
&= \kappa_0 + \kappa_1\left[A_0 + A_1 x_{t+1} + A_2\sigma^2_{t+1}\right] - A_0 - A_1 x_t - A_2\sigma^2_t + g_{t+1} \\
\mathbb{E}_t(r_{a,t+1}) &= \kappa_0 + \kappa_1\left[A_0 + A_1\rho x_t + A_2(\sigma^2 + \nu_1(\sigma^2_t - \sigma^2))\right] - A_0 - A_1 x_t - A_2\sigma^2_t \\
&\quad + \mu_c + x_t \\
r_{a,t+1} - \mathbb{E}_t(r_{a,t+1}) &= \kappa_1 A_1\left[x_{t+1} - \rho x_t\right] + \kappa_1 A_2\left[\sigma^2_{t+1} - \sigma^2 - \nu_1(\sigma^2_t - \sigma^2)\right] + \left[g_{t+1} - \mu_c - x_t\right] \\
&= \kappa_1 A_1\varphi_e\sigma_t e_{t+1} + \kappa_1 A_2\sigma_w w_{t+1} + \sigma_t\eta_{t+1}
\end{aligned}$$

$$
\begin{aligned}
m_{t+1} &= \theta \ln \delta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1} \\
\mathbb{E}_t(m_{t+1}) &= \theta \ln \delta - \frac{\theta}{\psi} \left[\mu_c + x_t\right] + (\theta - 1) \mathbb{E}_t(r_{a,t+1}) \\
m_{t+1} - \mathbb{E}_t(m_{t+1}) &= -\frac{\theta}{\psi} \left[\sigma_t \eta_{t+1}\right] + (\theta - 1) \left[\sigma_t \eta_{t+1} + \kappa_1 A_1 \varphi_e \sigma_t e_{t+1} + \kappa_1 A_2 \sigma_w w_{t+1}\right] \\
&= \left[\theta - 1 - \frac{\theta}{\psi}\right] \sigma_t \eta_{t+1} - (1-\theta) \kappa_1 A_1 \varphi_e \sigma_t e_{t+1} - (1-\theta) \kappa_1 A_2 \sigma_w w_{t+1} \\
&= \lambda_{m,\eta} \sigma_t \eta_{t+1} - \lambda_{m,e} \sigma_t e_{t+1} - \lambda_{m,w} \sigma_w w_{t+1}. \qquad \text{(A-14)}
\end{aligned}
$$

The risk premium on the aggregate wealth portfolio is given by:

$$
\begin{aligned}
\mathbb{E}_t \left[r_{a,t+1} - r_{f,t+1}\right] = & - \mathrm{Cov}_t \left[m_{t+1} - \mathbb{E}_t(m_{t+1}), r_{a,t+1} - \mathbb{E}_t(r_{a,t+1})\right] - \frac{1}{2} \mathrm{Var}_t(r_{a,t+1}) \\
= & - \mathbb{E}_t \Big[\left(\lambda_{m,\eta} \sigma_t \eta_{t+1} - \lambda_{m,e} \sigma_t e_{t+1} - \lambda_{m,w} \sigma_w w_{t+1}\right) \\
& \qquad \left(\sigma_t \eta_{t+1} + \kappa_1 A_1 \varphi_e \sigma_t e_{t+1} + \kappa_1 A_2 \sigma_w w_{t+1}\right) \Big] \\
& - \frac{1}{2} \Bigg(\mathbb{E}_t \left[\sigma_t^2 \eta_{t+1}^2\right] + \mathbb{E}_t \left[(\kappa_1 A_1 \varphi_e)^2 \sigma_t^2 e_{t+1}^2\right] + \mathbb{E}_t \left[\kappa_1^2 A_2^2 \sigma_w^2 w_{t+1}^2\right] \Bigg) \\
= & - \lambda_{m,\eta} \sigma_t^2 + \lambda_{m,e} (\kappa_1 A_1 \varphi_e) \sigma_t^2 + \kappa_1 A_2 \lambda_{m,w} \sigma_w^2 \\
& - \frac{1}{2} \Bigg((1 + (\kappa_1 A_1 \varphi_e)^2) \sigma_t^2 + (\kappa_1 A_2)^2 \sigma_w^2 \Bigg).
\end{aligned}
$$

To derive the risk premium on the market portfolio, the following Euler equation is used to obtain the relation between $\mathbb{E}_t(m_{t+1})$ and $\mathbb{E}_t(r_{m,t+1})$:

$$
\begin{aligned}
\mathbb{E}_t \left[M_{t+1} R_{m,t+1}\right] &= 1 \\
\mathbb{E}_t \left[\exp\left(m_{t+1} + r_{m,t+1}\right)\right] &= 1 \\
\exp\left(\mathbb{E}_t(m_{t+1} + r_{m,t+1}) + \frac{1}{2} \mathrm{Var}_t(m_{t+1} + r_{m,t+1})\right) &= 1 \\
\mathbb{E}_t(m_{t+1}) + \mathbb{E}_t(r_{m,t+1}) + \frac{1}{2} \mathrm{Var}_t(m_{t+1} + r_{m,t+1}) &= 0.
\end{aligned}
$$

Finally:

$$\mathbb{E}_t(m_{t+1}) = -E_t(r_{m,t+1}) - \frac{1}{2}\left[\mathrm{Var}_t(m_{t+1}) + \mathrm{Var}_t(r_{m,t+1}) + 2\ \mathrm{Cov}(m_{t+1}, r_{m,t+1})\right]. \tag{A-15}$$

Equations (A-12) and (A-15) can then be used to determine the risk premium:

$$\begin{aligned}
\mathbb{E}_t\left[r_{m,t+1} - r_{f,t+1}\right] &= \mathbb{E}_t\left[r_{m,t+1} + \mathbb{E}_t(m_{t+1}) + \frac{1}{2}\mathrm{Var}_t(m_{t+1})\right] \\
&= \mathbb{E}_t\Big[r_{m,t+1} - \mathbb{E}_t(r_{m,t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1}) - \frac{1}{2}\mathrm{Var}_t(r_{m,t+1}) \\
&\quad - \mathrm{Cov}_t(m_{t+1}, r_{m,t+1}) + \frac{1}{2}\mathrm{Var}_t(m_{t+1})\Big] \\
&= -\mathrm{Cov}_t\left[m_{t+1}, r_{m,t+1}\right] - \frac{1}{2}\mathrm{Var}_t(r_{m,t+1}) \\
&= -\mathrm{Cov}_t\left[m_{t+1} - \mathbb{E}_t(m_{t+1}), r_{m,t+1} - \mathbb{E}_t(r_{m,t+1})\right] - \frac{1}{2}\mathrm{Var}_t(r_{m,t+1}).
\end{aligned}$$

To write the risk premium in detail, first derive the expression for $r_{m,t+1} - \mathbb{E}_t(r_{m,t+1})$:

$$\begin{aligned}
r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1} \\
&= \kappa_{0,m} + \kappa_{1,m}\left[A_{0,m} + A_{1,m} x_{t+1} + A_{2,m}\sigma^2_{t+1}\right] - A_{0,m} - A_{1,m} x_t - A_{2,m}\sigma^2_t \\
&\quad + \mu_d + \phi x_t + \varphi_d \sigma_t u_{t+1} \\
\mathbb{E}_t(r_{m,t+1}) &= \kappa_{0,m} + \kappa_{1,m} A_{0,m} + \kappa_{1,m} A_{1,m}\rho x_t + \kappa_{1,m} A_{2,m}(\sigma^2 + \nu_1(\sigma^2_t - \sigma^2)) \\
&\quad - A_{0,m} - A_{1,m} x_t - A_{2,m}\sigma^2_t + \mu_d + \phi x_t \\
r_{m,t+1} - \mathbb{E}_t(r_{m,t+1}) &= \kappa_{1,m} A_{1,m}\left[x_{t+1} - \rho x_t\right] + \kappa_{1,m} A_{2,m}(\sigma^2_{t+1} - \sigma^2 - \nu_1(\sigma^2_t - \sigma^2)) + \varphi_d \sigma_t u_{t+1} \\
&= \kappa_{1,m} A_{1,m}\varphi_e \sigma_t e_{t+1} + \kappa_{1,m} A_{2,m}\sigma_w w_{t+1} + \varphi_d \sigma_t u_{t+1}.
\end{aligned}$$

The risk premium on the market portfolio is given by:

$$\begin{aligned}
\mathbb{E}_t\left[r_{m,t+1} - r_{f,t+1}\right] = & -\operatorname{Cov}_t\left[m_{t+1} - \mathbb{E}_t(m_{t+1}), r_{m,t+1} - \mathbb{E}_t(r_{m,t+1})\right] - \frac{1}{2}\operatorname{Var}_t(r_{m,t+1}) \\
= & -\mathbb{E}_t\Big[(\lambda_{m,\eta}\sigma_t\eta_{t+1} - \lambda_{m,e}\sigma_t e_{t+1} - \lambda_{m,w}\sigma_w w_{t+1}) \\
& \qquad (\varphi_d\sigma_t u_{t+1} + \kappa_{1,m}A_{1,m}\varphi_e\sigma_t e_{t+1} + \kappa_{1,m}A_{2,m}\sigma_w w_{t+1})\Big] \\
& -\frac{1}{2}\Big[\mathbb{E}_t\left(\varphi_d^2\sigma_t^2 u_{t+1}^2\right) + \mathbb{E}_t\left((\kappa_{1,m}A_{1,m}\varphi_e)^2\sigma_t^2 e_{t+1}^2\right) \\
& \qquad + \mathbb{E}_t\left(\kappa_{1,m}^2 A_{2,m}^2\sigma_w^2 w_{t+1}^2\right)\Big] \\
= & \lambda_{m,e}\kappa_{1,m}A_{1,m}\varphi_e\sigma_t^2 + \lambda_{m,w}\kappa_{1,m}A_{2,m}\sigma_w^2 \\
& -\frac{1}{2}\left[\varphi_d^2\sigma_t^2 + (\kappa_{1,m}A_{1,m}\varphi_e)^2\sigma_t^2 + (\kappa_{1,m}A_{2,m})^2\sigma_w^2\right]. \qquad \text{(A-16)}
\end{aligned}$$

Bibliography

Aldrich, E. M., Gallant, A. R., 2011. Habit, long-run risks, prospect? A statistical inquiry. Journal of Financial Econometrics 9 (4), 589–618.

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.

Arulampalam, M. S., Maskell, S., Gordon, N., Clapp, T., 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Transactions on Signal Processing 50 (2), 174–188.

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

Bansal, R., Kiku, D., Yaron, A., 2007b. A note on the economics and statistics of predictability: A long-run risks perspective.

Bansal, R., Kiku, D., Yaron, A., 2012a. An empirical evaluation of the long-run risks model for asset prices. Critical Finance Review 1 (1), 183–221.

Bansal, R., Kiku, D., Yaron, A., 2012b. Risks for the long run: Estimation with time aggregation. Working Paper 18305, National Bureau of Economic Research.

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. Journal of Finance 59 (4), 1481–1509.

Barro, R. J., 2006. Rare disasters and asset markets in the twentieth century. The Quarterly Journal of Economics, 823–866.

Barro, R. J., 2009. Rare disasters, asset prices, and welfare costs. The American Economic Review 99 (1), 243–264.

Beeler, J., Campbell, J. Y., 2012. The long-run risks model and aggregate asset prices: An empirical assessment. Critical Finance Review 1 (1), 141–182.

Beran, R. J., Ducharme, G. R., 1991. Asymptotic Theory for Bootstrap Methods in Statistics. Centre de Recherches Mathematiques.

Breeden, D., Litzenberger, R. H., Jia, T., 2014. Consumption-based asset pricing: research and applications. Annual Review of Financial Economics 7.

Breeden, D. T., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7 (3), 265–296.

Bui, M. P., 2007. Long-run risks and long-run predictability: A comment. Unpublished Paper, Harvard University.

Calvet, L. E., Czellar, V., 2015. Through the looking glass: Indirect inference via simple equilibria. Journal of Econometrics 185 (2), 343–358.

Calzorari, G., Fiorentini, G., Sentana, E., 2004. Constrained indirect estimation. Review of Economic Studies 71 (4), 945–973.

Campbell, J. Y., 1993. Intertemporal asset pricing without consumption data. The American Economic Review, 487–512.

Campbell, J. Y., 2003. Consumption-based asset pricing. Handbook of the Economics of Finance 1, 803–887.

Campbell, J. Y., Cochrane, J. H., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. The Journal of Political Economy 107 (2), 205–251.

Campbell, J. Y., Cochrane, J. H., 2000. Explaining the poor performance of consumption-based asset pricing models. The Journal of Finance 55 (6), 2863–2878.

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

Chen, S. X., 1999. Beta kernel estimators for density functions. Computational Statistics & Data Analysis 31 (2), 131–145.

Cochrane, J., 2005. Asset Pricing. Princeton University Press, Princeton.

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. Review of Asset Pricing Studies 1 (1), 96–136.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7 (2), 174–196.

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

Duffie, D., Singleton, K. J., 1993. Simulated moments estimation of Markov models of asset prices. Econometrica 61 (4), 929–952.

Efron, B., Tibshirani, R. J., 1993. An introduction to the bootstrap. Chapman & Hall.

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.

Flury, T., Shephard, N., 2011. Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models. Econometric Theory 27 (05), 933–956.

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

Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8 (S1), S85–S118.

Grammig, J., Küchlin, E.-M., 2016a. Give me strong moments and time: Combining GMM and SMM to estimate the long-run risk asset pricing model, Working Paper.

Grammig, J., Küchlin, E.-M., 2016b. Estimating the long-run risk asset pricing model with a two-step indirect inference approach, Working Paper.

Hall, A., 2005. Generalized Method of Moments. Advanced Texts in Econometrics. Oxford University Press.

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.

Horowitz, J. L., 2001. The Bootstrap. In: Handbook of Econometrics. Vol. 5. Elsevier, pp. 3159–3228.

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

Jagannathan, R., Wang, Y., 2007. Lazy investors, discretionary consumption, and the cross-section of stock returns. The Journal of Finance 62 (4), 1623–1661.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82 (1), 35–45.

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

Küchlin, E.-M., 2016. Filtering methods for the estimation of the long-run risk asset pricing model, Working Paper.

Lettau, M., Ludvigson, S., 2001. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. Journal of Political Economy 109 (6), 1238–1287.

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. The Review of Economics and Statistics 47 (1), 13–37.

Mehra, R., Prescott, E. C., 1985. The equity premium – A puzzle. Journal of Monetary Economics 15, 145–161.

Merton, R. C., 1973. An intertemporal capital asset pricing model. Econometrica 41 (5), 867–887.

Mossin, J., 1966. Equilibrium in a capital asset market. Econometrica 34 (4), 768–783.

Nelder, J. A., Mead, R., 1965. A simplex method for function minimization. Computer Journal 7 (4), 308–313.

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

Pitt, M. K., Shephard, N., 1999. Filtering via simulation: Auxiliary particle filters. Journal of the American Statistical Association 94 (446), 590–599.

Rietz, T. A., 1988. The equity risk premium: A solution. Journal of Monetary Economics 22 (1), 117–131.

Rubinstein, M., 1976. The valuation of uncertain income streams and the pricing of options. The Bell Journal of Economics 7 (2), 407–425.

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.

Savov, A., 2011. Asset pricing with garbage. Journal of Finance 66, 177–201.

Schorfheide, F., Song, D., Yaron, A., 2014. Identifying long-run risks: A Bayesian mixed-frequency approach. Working paper, National Bureau of Economic Research.

Sharpe, W. F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. The Journal of Finance 19 (3), 425–442.

Shephard, N. G., Harvey, A., 1990. On the probability of estimating a deterministic component in the local level model. Journal of Time Series Analysis 11, 339–347.

Silverman, B. W., 1986. Density estimation for statistics and data analysis. Chapman & Hall.

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

Smith, A. A., 1993. Estimating nonlinear time-series models using simulated vector autoregressions. Journal of Applied Econometrics 8 (S1), 63–84.

Sontag, E. D., 2013. Mathematical control theory: Deterministic finite dimensional systems. Vol. 6. Springer Science & Business Media.

Southall, B., Buxton, B. F., Marchant, J. A., 1998. Controllability and observability: Tools for Kalman filter design. In: British Machine Vision Conference. pp. 1–10.

Tangirala, A. K., 2014. Principles of System Identification: Theory and Practice. CRC Press.

Weil, P., 1989. The equity premium puzzle and the risk-free rate puzzle. Journal of Monetary Economics 24 (3), 401–421.

Yogo, M., 2006. A consumption-based explanation of expected stock returns. Journal of Finance 61 (2), 539–580.