CEM
Center for Empirical Macroeconomics

Working Paper No. 77

Asset Pricing - Constrained by Past Consumption Decisions

by

Lars Grüne and Willi Semmler

University of Bielefeld
Department of Economics
Center for Empirical Macroeconomics
P.O. Box 100 131
33501 Bielefeld, Germany
(http://www.wiwi.uni-bielefeld.de/∼cem)

Asset Pricing - Constrained by Past Consumption Decisions *

Lars Grüne
Mathematisches Institut
Fakultät für Mathematik und Physik
Universität Bayreuth
95440 Bayreuth, Germany
lars.gruene@uni-bayreuth.de

Willi Semmler
Center for Empirical Macroeconomics
Bielefeld and
New School University, New York
wsemmler@wiwi.uni-bielefeld.de

August 26, 2004

Abstract: The attempt to match asset price characteristics such as the risk-free interest rate, equity premium and the Sharpe ratio with data for models with instantaneous consumption decisions and time separable preferences has not been very successful. Many recent versions of asset pricing models have, in order to match those financial characteristics better with the data, employed habit formation where past consumption acts as a constraint on current consumption. In those models, surplus consumption, consumption over and above past consumption, improves welfare, yet habit formation gives rise to an additional state variable. By studying such a model we also allow for adjustment costs of investment. The asset price characteristics that one obtains from those models may depend on the solution techniques employed. In this paper a stochastic version of a dynamic programming method with adaptive grid scheme is applied to compute the above mentioned asset price characteristics where past consumption decisions are treated as an additional state variable. Since, as shown in Grüne and Semmler (2004), our method produces only negligible errors it is suitable to be used as solution technique for such models with more complicated decision structure. Using our solution methods shows that there are still remaining puzzles for the consumption based asset pricing model.

JEL Classification: C60, C61, C63, D90, G12

Keywords: stochastic growth models, habit formation, stochastic dynamic programming, adaptive grid, asset pricing

*We want to thank Martin Lettau and Buz Brock for helpful communications.

1 Introduction

Intertemporal asset pricing models with time separable preferences, such as power utility or log utility, have been shown to have difficulties to match financial market characteristics such as risk-free interest rate, equity premium and the Sharpe-ratio, a measure of the risk-return trade off. In those models the risk-free interest rate turns out to be too high (and too smooth) and the mean equity premium and Sharpe-ratio too low as compared to what one finds in time series data.

The conjecture has been that the solution methods to solve the optimal consumption path in feedback form and to use the growth of marginal utility of consumption as discount factor for pricing assets have been insufficient. Thus, one needs to be concerned with the accuracy of the solution method to solve the model. One conjecture in the literature was thus that the solution of stochastic growth models through linearizations for cases with more complicated decision structure may not be appropriate. Recently, global solution techniques to the Hamilton-Jacobi-Bellman equation have been developed that can address this concern. A global solution method that is useful in this context is stochastic dynamic programming with discretization of the state space and adaptive gridding strategy which generates quite accurate solutions.[1] A full discussion of the literature on this and other methods is given in sect. 2.

Another concern has been that asset pricing models have often used models with exogenous dividend stream[2] and the difficulties to match stylized financial statistics may have come from the fact that consumption is not endogenized. There is a tradition of asset pricing models that is based on the stochastic growth model with production originating in Brock and Mirman (1972) and Brock (1979, 1982) which endogenizes consumption. The Brock approach extends the asset pricing strategy beyond endowment economies to economies that have endogenous state variables including capital stocks that are used in production. Authors, building on this tradition,[3] have argued that it is crucial how consumption is endogenized. In stochastic growth models the randomness occurs to the production function of firms and consumption and dividends are derived endogenously. Yet, models with production have turned out to be even less successful. Given a production shock, consumption can be smoothed through savings and thus asset market features are even harder to match.[4]

Recent development of asset pricing studies has therefore turned to extensions of intertemporal models conjecturing that the difficulties to match real and financial time series characteristics may be related to the simple structure of the basic model. In order to match better asset price characteristics of the model to the data, economic research has extended the baseline stochastic growth model to include different utility functions,

[1] For deterministic versions, see Grüne (1997), Santos and Vigo–Aguiar (1998), and Grüne and Semmler (2004).

[2] Those models originate in Lucas (1978) and Breeden (1979) for example.

[3] See Rouwenhorst (1995, Akdeniz and Dechert (1997), Jerman (1998), Boldrin, Christiano and Fisher (2001), Lettau and Uhlig (1999) and Hansen and Sargent (2002), the latter in a linear-quadratic economy. The Brock model has also been used to evaluate the effect of corporate income tax on asset prices, see McGrattan and Prescott (2001).

[4] For a recent account of the gap between models and facts, see Boldrin, Christiano and Fisher (2001), Cochrane (2001, ch. 21), Lettau, Gong and Semmler (2001) and Semmler (2003, chs. 9-10).

in particular habit formation, adjustment costs of investment, idiosyncratic technology shocks to firms or the effect of leverage on firm value.[5] In this paper we will focus on an intertemporal decision model with habit formation and adjustment costs of investment.

Since, as aforementioned, time separable preferences fail to match financial market characteristics an enormous effort has been invested into models with time non-separable preferences, such as habit formation models, which allow for adjacent complementarity in consumption. Past consumption enters here as a constraint, defined as external habit persistence where the aggregate level of consumption serves as a benchmark level, or internal habit persistence where a household's own past consumption is viewed as a benchmark over and above welfare is considered to be increasing. If one chooses internal habit persistence, given by past consumption, as benchmark, it is then in general time varying.

There is a long tradition in economic theory where it is assumed that habits are formed through past consumption.[6] Habit persistence is nowadays used to understand a wide range of issues in growth theory (Carrol et al. 1997, 2000, Alvarez-Cuadrado et al. 2004) macroeconomics (Fuhrer, 2002), and business cycle theory (Boldrin et al, 2001). In all of those models of habit persistence high level of consumption in the past depresses current welfare and high current consumption depresses future welfare. This can be written as ratios of current over past consumption (Abel 1990, 1999) or in difference form as $(1-\alpha)c_t + \alpha(c_t - c_{t-1})$ with c_t current, c_{t-1} past consumption and α a respective weight. This form of habit formation will be chosen in this paper.

This type of habit specification gives rise to time non-separable preferences where risk aversion and intertemporal elasticity substitution are separated and a time variation of risk aversion will arise. If we define surplus consumption as $s_t = \frac{c_t - X_t}{c_t}$ with X_t, the habit, and γ, the risk aversion parameter, then the time variation of risk-aversion is $\frac{\gamma}{s_t}$: the risk aversion falls with rising surplus consumption and the reverse holds for falling surplus consumption. A high volatility of the surplus consumption will lead to a high volatility of the growth of marginal utility and thus to a high volatility of the stochastic discount factor.

Habit persistence in asset pricing has been introduced by Constantinides (1990) in order to account for high equity premia. Asset pricing models along this line have been further explored by Campbell and Cochrane (1999), Jerman (1998), and Boldrin et al. (2001). Yet, asset pricing introducing habit persistence in stochastic models with production may just produce smoother consumption. But with income different from consumption, for example due to shocks, habit formation amplifies investment and demand for capital goods. Yet, Boldrin et al. (2001) have argued if there is, however, perfectly elastic supply of capital there is no effect on the volatility of the return on equity. As the literature has demonstrated (Jerman 1998, and Boldrin et al. 2001) one also needs adjustment costs of investment to minimize the elasticity of the supply of capital. It seems to be both habit persistence and adjustment costs for investment which are needed to generate higher equity premia. By choosing such a model we will not, following Jerman (1998), allow for elastic labor supply, but rather employ a model with fixed labor supply, since the latter,

[5] For further detailed studies of those extensions see, for example, Campbell and Cochrane (1999), Jerman (1998), Boldrin, Christiano and Fisher (2001) and Cochrane (2001, ch. 21).

[6] See the description in Marshall (1920), Veblen (1899) and Duesenberry (1949). For a first use of habit persistence in a dynamic decision model see Ryder and Heal (1973).

as shown in Lettau and Uhlig (2000), provides the most favorable case for matching the model with the financial market characteristics.

Since accuracy of the solution method is an intricate issue for models with more complicated decision structure, we first have to have sufficient confidence in the accuracy of the stochastic dynamic programming method that we will use. In our method we do not use fixed grids, but adaptive space discretization. In the method applied in our paper efficient and reliable local error estimation is undertaken and used as a basis for a local refinement of the grid in order to deal with regions of steep slopes or other non-smooth properties of the value function (such as non-differentiability). This procedure allows for a global dynamic analysis of deterministic as well as stochastic intertemporal decision problems.

In Grüne and Semmler (2004) a stochastic dynamic programming algorithm with flexible grid size has been tested for the most basic stochastic growth model as based on Brock and Mirman (1972) and Brock (1979, 1982). This model can analytically be solved for the sequence of optimal consumption in feedback form. Asset prices, the risk-free interest rate, the equity premium and the Sharpe-ratio, can, once the model is solved analytically for the sequence of optimal consumption, easily be solved numerically and those solutions can be compared to the numerical solutions obtained from our numerical procedure. As has been shown in Grüne and Semmler (2004) the errors, as compared to the analytical solutions, are negligibly small. Thus, the method we employ here can confidently be applied to extensions of the basic model with more complicated decision structure.

The paper is organized as follows. Section 2 discusses related literature. Section 3 presents the stochastic dynamic programming algorithm. Section 4 introduces our model of asset pricing with habit persistence and adjustment costs of investment and the measures of the financial characteristics we want to study. Section 5 reports the numerical results of our study which are evaluated in section 6. Section 7 concludes the paper.

2 Related Literature on Solution Methods

In the literature on solving asset pricing models one can find a vast amount of different approaches most of them using linear approximations.[7] The most promising approaches are those ones that are employing the dynamic programming approach since it is closely related to the Hamilton-Jacobi-Bellman equation for asset pricing. Many of the recent versions of dynamic programming use state–of–the art mathematical and numerical techniques for making this approach more efficient. Here we apply an adaptive gridding algorithm that works for very general Hamilton-Jacobi-Bellman equations, see Section 3 for details. In the present section we briefly review similar approaches and highlight similarities and differences to our approach.

One of the fundamental difficulties with the dynamic programming approach is that the computational load grows exponentially with the dimension of the problem, a phenomenon known as the "curse of dimensionality" (see Rust (1996) for a comprehensive account on complexity issues). In our case, for computing asset pricing in the context of stochastic growth models, starting with Brock and Mirman (1972) as suggested in the literature,

[7] For an extensive survey of those techniques, see Taylor and Uhlig (1990).

the problem to be solved is two dimensional, hence this is not a crucial aspect here. Nevertheless, for the sake of completeness we want to mention approaches like randomly distributed grid points (Rust (1997)) or so called low discrepancy grids (Rust (1996), Reiter (1999)) which are able to break the curse of dimensionality. In principle also Monte–Carlo techniques like in Keane and Wolpin (1994) allow for breaking the curse of dimensionality, but as Rust (1997) points out, the specific algorithm in Keane and Wolpin (1994) uses an interpolation technique which again is subject to exponential growth of the numerical cost in the space dimension.

For low dimensional problems the goal of the numerical strategy is not to avoid the curse of dimensionality but rather to reduce the computational cost for a problem of fixed dimension. For this purpose, two main approaches can be found in the literature, namely higher order approximations and adaptive gridding techniques; the latter will be used in our numerical approach.

The idea of high order approximations lies in exploiting the smoothness of the optimal value function: if the optimal value function turns out to be sufficiently smooth, then methods using approximations by smooth functions, like Chebyshev polynomials (Rust (1996), Judd (1996), Jermann (1998)), Splines (Daniel (1976), Johnson et al. (1993), Trick and Zin (1993, 1997)) or piecewise high–order approximations (Falcone and Ferretti (1998)) can be very efficient. Smoothness is also the basis of other high–order strategies, like in finite difference approximations (Candler (2001)), Gaussian Quadrature discretization (Tauchen and Hussey (1991), Burnside (2001)) and in perturbation techniques (Judd (1996)). Yet, the last should also work if the value function is only piecewise smooth.[8]

Some of these methods (like Spline and piecewise high order approximation) use a (fixed) grid discretization of the state space similar to our approach. The combination of adaptive grids with higher order approximation is currently under investigation and it will be interesting to see whether adaptive discretization ideas based on our local error estimation technique work equally well with these approximation techniques.

Concerning discretization techniques it should be noted that from the complexity point of view it turns out to be optimal to solve the dynamic programming problem on successively finer grids, using a one–way multigrid strategy (Chow and Tsitsiklis (1991), see also Rust (1996)). In fact, our adaptive gridding algorithm is similar to this approach since the approximation on the previous grid Γ_i is always used as the initial value for the computation on the next finer adaptive grid Γ_{i+1}. This also explains the large reduction in computation time observed for our approach compared to the computation on *one* fixed equidistant grid.

Let us now turn to the methodology employed here, i.e., adaptive gridding techniques. Perhaps closest to our approach are the techniques discussed in Munos and Moore (2002). Here a number of heuristic techniques are compared which lead to local and global error indicators which can in turn be used for an adaptive grid generation. Some of the indicators discussed in this paper bear some similarity with our residual based estimator, though rigorous estimates as given in our paper below are not given there. In any case, the authors report that these techniques are unsatisfactory and argue for a completely

[8]For an early survey of those methods, see Taylor and Uhlig (1990) where one can find a comparative numerical study of several methods.

different approach which measures the influence of local errors in certain regions on the global error by analyzing the information flow on the Markov chain related to the discretization of the (deterministic) problem at hand. The reason for this lies in the fact that the model problem treated by Munos and Moore (2002) has a discontinuous optimal value function, which often happens in technical problems with boundary conditions. In fact, also our adaptive scheme performs rather poorly in presence of discontinuities but since our economic problems do always have continuous optimal value functions, Munos' and Moore's conclusions do not apply here. A roughly similar technique is the endogenous oversampling used by Marcet (1994). This is again a heuristic method, which, however, does not lead to adaptive grids but rather selects suitable parts of the state space where the optimally controlled trajectories stay with high probability.

Probably the adaptive approaches with the most solid mathematical background are presented in the papers of Trick and Zin (1993, 1997).[9] In these papers an alternative approach for the solution of the fully discrete problem is developed using advanced linear programming techniques which are capable of solving huge linear programs with many unknowns and constraints. In Trick and Zin (1993) an adaptive selection of constraints in the linear program is used based on estimating the impact of the missing constraint, a method which is closely related to the chosen solution method but only loosely connected to our adaptive gridding approach. The later paper (Trick and Zin (1997)), however, presents an idea which is very similar to our approach. Due to the structure of their solution they can ensure that the numerical approximation is greater than or equal to the true optimal value function. On the other hand, the induced suboptimal optimal control strategy always produces a value which is lower than the optimal value. Thus, comparing these values for each test point in space one can compute an interval in which the true value must lie, which produces a mathematically concise error estimate that can be used as a refinement criterion. While this approach is certainly a good way to measure errors, which could in particular be less conservative than our measure for an upper bound, we strongly believe that it is less efficient for an adaptive gridding scheme, because (i) the estimated error measured by this procedure is not a local quantity (since it depends on the numerical along the whole suboptimal trajectory), which means that regions may be refined although the real error is large elsewhere, and (ii) compared to our approach it is expensive to evaluate, because for any test point one has to compute the whole suboptimal trajectory, while our residual based error estimate needs only one step of this trajectory.

Let us comment on the idea of a posteriori error estimation. In fact, the idea to evaluate residuals can also be found in the papers of Judd (1996) and Judd and Guu (1997), using, however, not the dynamic programming operator but the associated Euler equation. In these references the resulting residual was used to estimate the quality of the approximating solution, but to our knowledge it has not been used to control adaptive gridding strategies, and we are not aware of any estimates such as ours which is a crucial property for an efficient and reliable adaptive gridding scheme, particularly needed to solve stochastic problems in asset pricing models.

Summarizing our discussion, there are a number of adaptive strategies around which are

[9] As mentioned above, this approach also uses splines, i.e., a smooth approximation, but the ideas developed in these papers do also work for linear splines which do not require smoothness of the approximated optimal value function.

all reported to show good results, however, they are either heuristic[10] and better suited for other classes of problems than asset pricing models or they have nice theoretical features but are practically inconvenient because their implementation is numerically much more expensive than our approach.

3 Stochastic Dynamic Programming

Next we describe the stochastic dynamic programming algorithm that we use to solve the asset pricing characteristics of the intertemporal decision model we want to study. Our approach is characterized by using a combined value function and policy iteration and permitting grid refinements due to error estimates.

We consider the discrete stochastic dynamic programming equation

$$V(x) = \max_{c \in C} E\{u(x,c,\varepsilon) + \beta(x,\varepsilon)V(\varphi(x,c,\varepsilon))\}. \tag{3.1}$$

Here $x \in \Omega \subset \mathbb{R}^2$, $C \subset \mathbb{R}$, Ω and C are compact sets and ε is a random variable with values in $\mathbb{R}$. The mappings $\varphi : \Omega \times C \times \mathbb{R} \to \mathbb{R}^2$ and $g : \Omega \times C \times \mathbb{R} \to \mathbb{R}$ are supposed to be continuous and Lipschitz continuous in x. Furthermore, we assume that either $\varphi(x,c,z) \in \Omega$ almost surely for all $x \in \Omega$ and all $c \in C$, or that suitable boundary values $V(x)$ for $x \notin \Omega$ are specified, such that the right hand side of (3.1) is well defined for all $x \in \Omega$. The value $\beta(x,\varepsilon)$ is the (possibly state and ε dependent) discount factor which we assume to be Lipschitz and we assume that there exists $\beta_0 \in (0,1)$ such that $\beta(x,\varepsilon) \in (0,\beta_0)$ holds for all $x \in \Omega$. We can relax this condition if no maximization takes place, in this case it suffices that all trajectories end up in a region where $\beta(x,\varepsilon) \in (0,\beta_0)$ holds. This is the situation for the asset price problem, cf. the discussion in Cochrane (2001:27).

Associated to (3.1) we define the dynamic programming operator

$$T : C(\Omega,\mathbb{R}) \to C(\Omega,\mathbb{R})$$

given by

$$T(W)(x) := \max_{c \in C} E\{u(x,c,\varepsilon) + \beta(x,\varepsilon)W(\varphi(x,c,\varepsilon))\}. \tag{3.2}$$

The solution V of (3.1) is then the unique fixed point of (3.2), i.e.,

$$T(V) = V. \tag{3.3}$$

For the numerical solution of (3.3) we use a discretization method that goes back to Falcone (1987) and in Santos and Vigo–Aguiar (1998) in the deterministic case. Here we use unstructured rectangular grids: We assume that $\Omega \subset \mathbb{R}^n$ is a rectangular and consider a grid Γ covering Ω with rectangular elements Q_l and nodes x_j and the space of continuous and piecewise multilinear functions

$$\mathcal{W}_\Gamma := \{W \in C(\Omega,\mathbb{R}) \,|\, W(x+\alpha e_j) \text{ is linear in } \alpha \text{ on each } Q_l \text{ for each } j = 1,\ldots,n\}$$

[10] In order to avoid misunderstandings: We do not claim that heuristic methods cannot perform well; in fact they can show very good results. Our main concern about these methods is that one can never be sure about the quality of the final solution of a heuristic method.

where the e_j, $j = 1, \ldots, n$ denote the standard basis vectors of the $\mathbb{R}^n$, see Grüne (2003) for details of the grid construction. With $\pi_\Gamma : C(\Omega, \mathbb{R}) \to \mathcal{W}_\Gamma$ we denote the projection of an arbitrary continuous function to $\mathcal{W}_\Gamma$, i.e.,

$$\pi_\Gamma(W)(x_j) = W(x_j) \text{ for all nodes } x_j \text{ of the grid } \Gamma.$$

Note that our approach easily carries over to higher order approximations, the use of multilinear approximations is mainly motivated by its ease of implementation, especially for adaptively refined grids.[11] Also, the approach can easily be extended to higher dimensions.

We now define the discrete dynamic programming operator by

$$T_\Gamma : C(\Omega, R) \to \mathcal{W}_\Gamma, \quad T_\Gamma = \pi_\Gamma \circ T \tag{3.4}$$

with T from (3.2). Then the discrete fixed point equation

$$T_\Gamma(V_\Gamma) = V_\Gamma. \tag{3.5}$$

has a unique solution $V_\Gamma \in \mathcal{W}_\Gamma$ which converges to V if the size of the elements Q_l tends to zero. The convergence is linear if V is Lipschitz on Ω, see Falcone (1987), and quadratic if V is $C2$, see Santos and Vigo–Aguiar (1998).

For the solution of (3.5) as well as for the computation of $\eta(x)$ we need to evaluate the operator T_Γ. More precisely, we need to evaluate

$$\max_{c \in C} E\{u(x_j, c, \varepsilon) + \beta(x_j, \varepsilon) W(\varphi(x_j, c, \varepsilon))\}.$$

for all nodes x_j of Γ.

This first includes the numerical evaluation of the expectation E. If ε is a finite random variable then this is straightforward, if ε is a continuous random variable then the corresponding integral

$$\int (u(x, c, \varepsilon) + \beta(x, \varepsilon) V(\varphi(x, c, \varepsilon))) f(\varepsilon) d\varepsilon$$

has to be computed, where f is the probability density of ε. In our implementation we approximated this integral by a trapezoidal rule with 10 equidistant intervals.

The second difficulty in the numerical evaluation of T lies in the maximization over c. In our implementation we used a recursive discrete approximation of the feasible values in the set C, i.e., of those values $c \in C$ with $c - \alpha X_t > 0$. The maximum is approximated by comparing finitely many values in C, then a neighborhood of this candidate is refined to obtain a new approximate maximum, a procedure which is repeated recusrively for several times. It can be shown that for unimodal functions this procedure indeed converges to the maximum and even though for our functions this property cannot be shown rigorously this procedure shows very good results in practice.

For the solution of the fixed point equation (3.5) we use the Gauss–Seidel type value space iteration where we subsequently compute $V_{i+1} = S_\Gamma(V_i)$ with S_Γ being a Gauss–Seidel type iteration operator (including the maximization over c) obtained from T_Γ. This iteration

[11] The combination of adaptive grids and higher order approximations is currently under investigation.

is coupled with a policy space iteration: Once a prescribed percentage of the maximizing u–values in the nodes remains constant from one iteration to another we fix all control values and compute the associated value function by solving a linear system of equations using a modified Gauss-Seidel-method, (CGS or BICGSTAB method, in our examples the CGS method turned out to show more reliable convergence behavior). After convergence of this method we continue with the value space iteration using S_Γ until the control values again converge, switch to the linear solver and so on. This combined policy–value space iteration turns out to be much more efficient (often more than 90 percent faster) than the plain Gauss–Seidel value space iteration using S_Γ[12] The details of the adaptive gridding strategy based on error estimates are presented in the appendix.

4 The Stochastic Decision Problem in Asset Pricing

Our stochastic decision problem arising from the stochastic growth model in the Brock tradition which we want to solve and for which we want to compute certain financial measures is as follows. Before we introduce the baseline stochastic growth model that we want to apply our solution technique too, see sect. 5, we outline an asset pricing model in a very generic form. The problem we are concerned with is to solve an optimal control, c_t, for the dynamic decision problem

$$V(k,z) = \max_{c_t} E\left(\sum_{t=0}^{\infty} \beta^i u(c_t, X_t)\right) \tag{4.1}$$

with habit X_t, subject to the dynamics

$$\begin{aligned} k_{t+1} &= \varphi_1(k_t, z_t, c_t, \varepsilon_t) \\ z_{t+1} &= \varphi_2(k_t, z_t, c_t, \varepsilon_t) \\ X_{t+1} &= c_t \end{aligned}$$

using the constraints $c_t \geq 0$ and $k_t \geq 0$ and the initial value $k_0 = k, z_0 = z, X_0 = X$. Here $(k_t, z_t, X_t) \in \mathbb{R}^3$ is the state and ε_t are i.i.d. random variables. We abbreviate $x_t = (k_t, z_t, X_t)$ and $\varphi(x, c, \varepsilon) = (\varphi_1(k, z, c, \varepsilon), \varphi_2(k, z, c, \varepsilon), c_t)$, i.e.,

$$x_{t+1} = \varphi(x_t, c_t, \varepsilon_t). \tag{4.2}$$

This optimal decision problem should allow the computation of c in feedback form, i.e. $c_t = c(x_t)$ for some suitable map $c : \mathbb{R}^2 \to \mathbb{R}$. Based on this c we compute the stochastic discount factor[13]

$$m(x_t) = \beta \frac{u'(c(x_{t+1}))}{u'(c(x_t))} \tag{4.3}$$

(note that m depends on ε_t and the derivative u' is taken with respect to c_t), which serves as an ingredient for the next step, which consists of solving the asset pricing problem

$$p(x) = E\left(\sum_{t=1}^{\infty} \prod_{s=1}^{t} m(x_s) d(x_t)\right), \tag{4.4}$$

[12]The latter in turn is considerably faster than the Banach iteration $V_{i+1} = T_\Gamma(V_i)$.
[13]The following financial measures are introduced and studied in detail in Cochrane (2001).

where $d(x_t)$ denotes the dividend at x_t and $x_0 = x$ and the dynamics are given by

$$x_{t+1} = \varphi(x_t, c(x_t), \varepsilon_t)$$

with c from above.

Finally, we use these values to compute the Sharpe ratio, which represents the ratio of the equity premium to the standard deviation of the equity return. Hereby R^f is the risk-free interest rate.

$$S = \left| \frac{E(R(x)) - R^f(x)}{\sigma(R(x))} \right| = \frac{-R^f(x)\mathrm{cov}\Big(m(x), R(x)\Big)}{\sigma(R(x))}. \tag{4.5}$$

The upper bound of the Sharpe-ratio is

$$S_B = \frac{\sigma(m(x))}{E(m(x))}. \tag{4.6}$$

Here

$$R^f(x) = \frac{1}{E(m(x))} \tag{4.7}$$

is the risk-free interest rate and

$$R(x_t) = \frac{d(x_{t+1}) + p(x_{t+1})}{p(x_t)} \tag{4.8}$$

is the gross return.

Note that the equality $E(m(x)R(x)) = 1$ holds, which can serve as a indicator for the accuracy of our numerical solution.

We solve the asset pricing problem in the following three steps: (i) We compute the optimal value function V of the underlying optimal control problem, and compute c from V, (ii) we compute the prices $p(x)$ from c and m, and (iii) we compute the risk-free interest rate, the equitiy premium and the Sharpe ratio S (and its bound S_B) from c, m and p.

For our baseline stochastic growth model introduced below, which we solve numerically, both c and p are actually available analytically. This allows us to test each single step of our algorithm by replacing the numerically computed c in (ii) and (iii) and/or p in (iii) by their exact values.

For each of the steps we do now sketch our technique for the numerical computation using the algorithm described above in Section 3.

Step (i):

For the solution of the optimal control problem we use a dynamic programming algorithm with adaptive grid. In order to solve (4.1) we solve the equivalent dynamic programming equation

$$V(x) = \max_c E\left(u(c) + \beta V(\varphi(x, c, \varepsilon))\right) =: T(V)(x) \tag{4.9}$$

For solving this equation we choose a computational domain $\Omega \subset \mathbb{R}^2$ and approximate V on a rectangular grid Γ covering Ω, using multilinear interpolation between the grid

nodes. For continuous and piecewise multilinear function V_Γ on Γ, we solve the fixed point equation

$$V_\Gamma(x_i) = T(V_\Gamma)(x_i) \tag{4.10}$$

for all nodes x_i of Γ.

As mentioned in sect. 3 in order to solve (4.10) we use a mixed value-policy iteration method: We first iterate (4.10) (using a Gauss–Seidel like acceleration method) until we observe convergence of the maximizing control values c. Then we fix the values and solve the resulting system of linear equations (using the CGS method). Iteratively, we continue this process until convergence.

As also aforementioned, here in each step of the iteration we have to perform a maximization over c and an evaluation of the expectation. Knowing the density function $p(\varepsilon)$ of the underlying random variable ε, the latter problem lies in evaluating the integral

$$\int V(\varphi(t,x,c,\varepsilon))p(\varepsilon)d\varepsilon,$$

which is efficiently accomplished by a numerical quadrature rule.

The maximization turns out to be a more severe numerical problem. In our implementation we have used a simple and straightforward method by discretizing the set C of possible values of c and maximizing by comparing the finitely many discrete values.[14]

Another crucial point in solving (4.10) is the choice of an appropriate grid. Here we make use of an adaptive gridding strategy as outlined in the appendix. After the solution V_Γ is computed, we evaluate the error estimates

$$\eta(x) = |V_\Gamma(x) - T(V_\Gamma)(x)|.$$

This value gives an upper and lower bound for the real global error. Thus, we evaluate η in a number of test points in each grid element and refine those elements carrying a large error estimate. This way, we can iteratively construct a grid which is adjusted to the problem.

Once V_Γ is computed with sufficient accuracy we can obtain the optimal control value $c(x)$ in each point by choosing $c(x)$ such that (4.9) is maximized, i.e., such that

$$E\left(u(c(x)) + \beta V_\Gamma(\varphi(x,c(x),\varepsilon))\right) = \max_c E\left(u(c) + \beta V_\Gamma(\varphi(x,c,\varepsilon))\right)$$

holds. Once c is known, m of equ. (4.2) can be computed from this value.

Step (ii):

For computing $p(x)$ we follow the same approach as in Step (i), except that here $c(x)$ is known in advance and hence no maximization needs to be done.

For the computation of p we first solve the dynamic programming equation

$$\tilde{p}(x) = E(d(x) + m(x)p(\varphi(x,c(x),\varepsilon)))$$

[14]However, it turns out that for higher accuracy this method is not so efficient and is suggested to be replaced by a more efficient method in future research.

which is simply a system of linear equations which we solve using the CGS method. This yields a numerical approximation of the function

$$\tilde{p}(x) = E\left(\sum_{t=0}^{\infty}\prod_{s=1}^{t} m(x_s)d(x_t)\right)$$

(with the convention $\prod_{s=1}^{0} m(x_s) = 1$), from which we obtain p by

$$p(x) = \tilde{p}(x) - d(x).$$

In our numerical computations for the computation of p we have always used the same grid Γ as in the previous computation of V in Step (i). The reason is that it did not seem justified to use a finer grid here, because the accuracy of the entering values c from Step (i) is limited by the resolution of Γ, anyway. However, it might nevertheless be that using a different grid (generated e.g. by an additional adaptation routine) in Step (ii) could increase the numerical accuracy.

Step (iii):

The last step is in principle straightforward, because we do now have all the necessary ingredients to compute the risk-free interest rate, the equity premium and Sharpe ratio S. However, since all the numerical values entering these computations are subject to numerical errors we have to be concerned with the numerical stability of the respective magnitudes. While the bound $S_B = \sigma(m(x))/E(m(x))$ for the Sharpe ratio turns out to be numerically nice, the first formula for the Sharpe ratio, wher the numerator represents the equity premium, as the spread between the expected equity return and the risk-free interest rate,

$$\left|\frac{E(R(x)) - R^f(x)}{\sigma(R(x))}\right|. \tag{4.11}$$

This turns out to be considerably less precise than the second formula

$$\frac{-R^f(x)\mathrm{cov}\Big(m(x), R(x)\Big)}{\sigma(R(x))}. \tag{4.12}$$

Since the denominator is the same in both formulas, the difference can only be caused by the different numerators. A further investigation reveals that the numerator of the first formula can be rewritten as

$$R^f(x)(1 - E(m(x))E(R(x)))$$

while that of the second formula reads

$$R^f(x)(E(m(x)R(x)) - E(m(x))E(R(x))).$$

Note that in both formulas we have to subtract values which have approximately the same values, which considerably amplifies the numerical errors. As mentioned above, we know that $E(m(x)R(x)) = 1$, which shows that these formulas are theoretically equivalent. Yet the second formula is more accurate.[15]

[15]The higher accuracy of the second formula can be explained as follows: Assume that we have a small

5 Numerical results

We have applied our numerical scheme to the model given by

$$\begin{aligned} k_{t+1} = \varphi_1(k_t, z_t, c_t, \varepsilon_t) &= k_t + \frac{k_t}{1-\varphi}\left[\left(\frac{I_t}{k_t}\right)^{1-\varphi} - 1\right] \\ \ln z_t = \varphi_2(k_t, z_t, c_t, \varepsilon_t) &= \rho \ln z_t + \varepsilon_t, \end{aligned}$$

where $I_t = z_t A k_t^\alpha - c_t$, where in our numerical computations we used the variable $y_t = \ln z_t$ instead of z_t as the second variable.

The utility function is given by

$$u(c_t, X_t) = \frac{(c_t - bX_t)^{1-\gamma} - 1}{1-\gamma}$$

for $\gamma \neq 1$ and by

$$u(c_t, X_t) = \ln(c_t - bX_t)$$

for $\gamma = 1$. Since we are working with internal habit, in our case, we have $X_t = C_{t-1}$.

For our numerical experiments we employed the values

$$A = 5, \;\; \alpha = 0.34, \;\; \rho = 0.9, \;\; \beta = 0.95, \;\; b = 0.5$$

and ε_t was chosen as a Gaussian distributed random variable with standard deviation $\sigma = 0.008$, which we restricted to the interval $[-0.032, 0.032]$. With this choice of parameters it is easily seen that the interval $[-0.32, 0.32]$ is invariant for the second variable y_t.

Motivated by our 2d studies (Grüne and Semmler (2004)) we would like to solve our problem for k_t in the interval $[0.1, 10]$. However, the habit persistence implies that for a given habit X_t only those value c_t are admissible for which $c_t - bX_t > 0$ holds, which defines a constraint from below on c_t depending on the habit X_t. On the other hand, the condition $I_t \geq 0$ defines a constraint from above on c_t depending on k_t and $y_t = \ln z_t$. As a consequence, there exist states $x_t = (k_t, y_t, X_t)$ for which the set of admissible control

---

systematic additive numerical error in $R(x)$, e.g., $R_{num}(x) \approx R(x) + \delta$. Such errors are likely to be caused by the interpolation process on the grid. Then, using $R^f(x) = 1/E(m(x))$, in the first formula we obtain

$$\begin{aligned} R^f(x)(1 - E(m(x))E(R_{num}(x))) &\approx R^f(x)(1 - E(m(x))E(R(x) + \delta)) \\ &\approx R^f(x)(1 - E(m(x))E(R(x))) - \delta, \end{aligned}$$

while in the second formula we obtain

$$\begin{aligned} & R^f(x)(E(m(x)R_{num}(x)) - E(m(x))E(R_{num}(x))) \\ &\approx R^f(x)(E(m(x)(R(x) + \delta)) - E(m(x))E(R(x) + \delta)) \\ &\approx R^f(x)(E(m(x)R(x)) + E(m(x))\delta - E(m(x))E(R(x)) - E(m(x))\delta) \\ &\approx R^f(x)(E(m(x)(R(x))) - E(m(x))E(R(x))), \end{aligned}$$

i.e., systematic additive errors cancel out in the second formula.

values c_t is empty, i.e., for which the problem is not feasible. On the one hand, we want to exclude these points from our computation, on the other hand we want to have a computational domain which is of a simple shape. A solution for this problem is given by a coordinate transformation which transforms a suitable set $\widetilde{\Omega}$ of feasible points to the set $\Omega = [0.1, 10] \times [-0.32, 0.32] \times [0, 7]$ on which we perform our computation. The coordinate transformation $\Psi : \mathbb{R}^3 \to \mathbb{R}^3$ we use for this purpose shifts the k–component of a point $x = (k, y, X)$ in such a way that the non–feasible points are mapped to a point $\tilde{x}_t \notin \Omega$. It is given by

$$\Psi(k, y, X) := (k - s(y, X), y, X)$$

with

$$s(y, X) = \left(\frac{s_0 + bX}{\exp(y)A} \right)^{\frac{1}{\alpha}} - 0.1$$

where $s_0 = 0.1^\alpha \exp(-0.32)A$ is chosen such that for $y = -0.32$ and $X = 0$ the coordinate change is the identity. This map is built in such a way that for all points $x_t \in \widetilde{\Omega} = \Psi^{-1}(\Omega)$ a value c_t with $c_t - bX_t \geq s_0$ is admissible. Note that forward invariance of $\widetilde{\Omega}$ is not automatically guaranteed by this construction and indeed there are parameter combinations for which this property does not hold.

This coordinate transformation allows us to set up our dynamic programming algorithm on a set of feasible points without having to deal with a complicated domain $\widetilde{\Omega}$, because numerically we can now work on the simple set Ω using the transformed dynamics

$$x_{t+1} = \Psi \circ \varphi(\Psi^{-1}(x_t), c_t, \varepsilon_t)$$

instead of (4.2).

In addition to the parameters specified above, for our numerical experiments we have used the following sets of parameters:

$$\begin{array}{lllll} \text{(a)} & \varphi = 0, & \gamma = 1, & b = 0 \\ \text{(b)} & \varphi = 0, & \gamma = 1, & b = 0.5 \\ \text{(c)} & \varphi = 0, & \gamma = 3, & b = 0.5 \\ \text{(d)} & \varphi = 0.8, & \gamma = 1, & b = 0.5 \\ \text{(e)} & \varphi = 0.8, & \gamma = 3, & b = 0.5 \end{array}$$

Note that (a) corresponds to the setting from Grüne and Semmler (2004).

Our first set of numerical results shows the behavior of the k_t–component of the optimal trajectories of the optimal control problem (4.1) in Figure 5.1 (a)–(e), the stochastic discount factor along these trajectories in Figure 5.2(a)–(e) and the consumption in Figure 5.3(a)–(e). For all trajectories we used the initial value $(k_0, y_0, X_0) = (2, 0, 4)$ which is near the point around which the optimal trajectory oscillates. Furthermore, for all trajectories we have used the same sequence of the random variables ε_t.

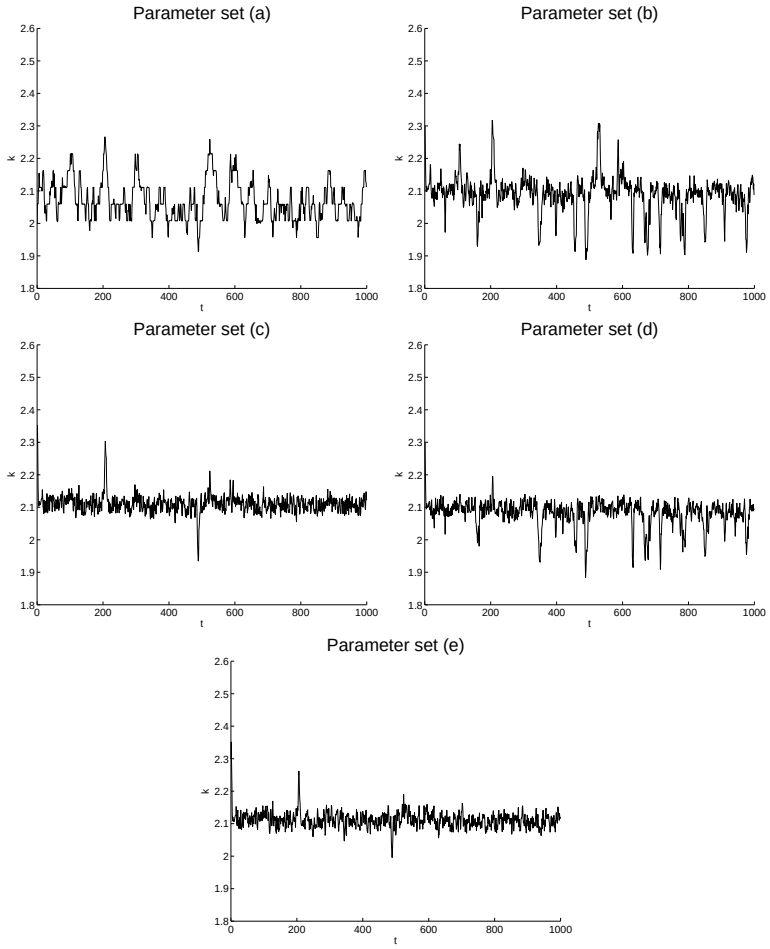

Figure 5.1: k–component of optimal trajectories

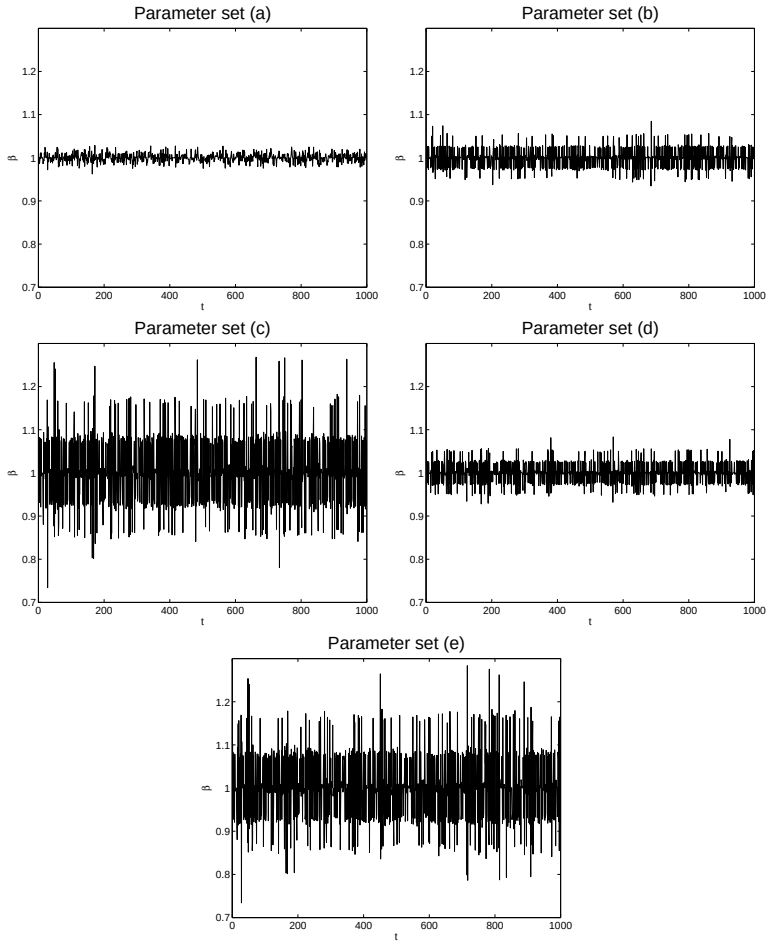

Figure 5.2: Stochastic discount factor β along optimal trajectories

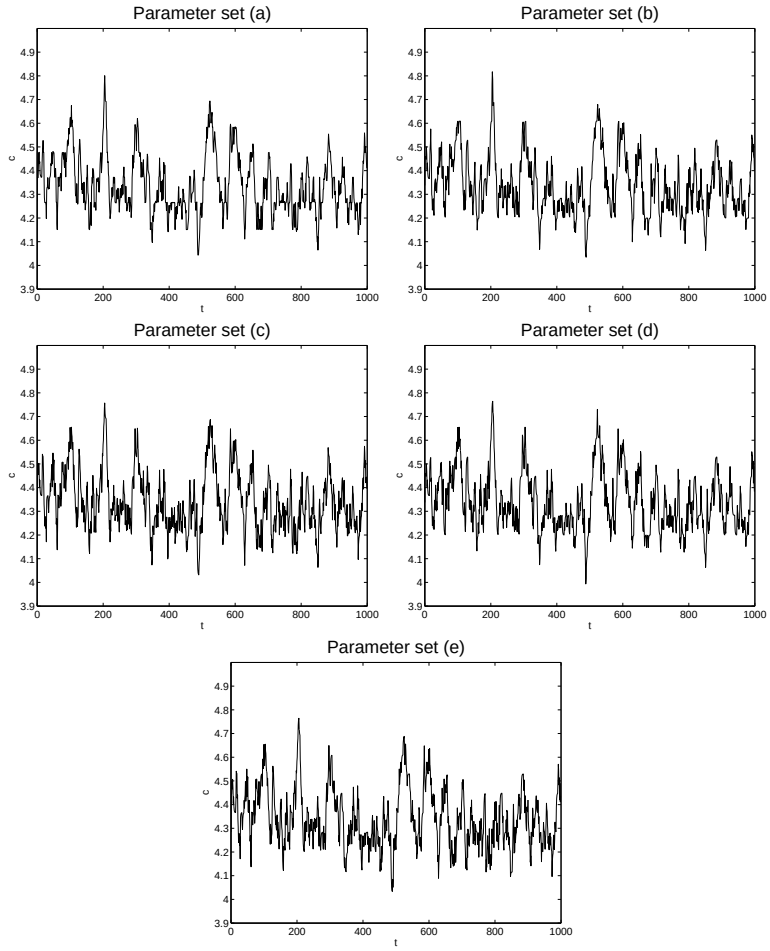

Figure 5.3: Consumption c along optimal trajectories

The following Table 5.1 shows several characteristic values obtained from our algorithm. The values were obtained by averaging the values along the optimal trajectories described above.

parameter set	Sharpe ratio	equity premium	risk free interest rate
(a)	0.0085	0.00008	1.051
(b)	0.0153	0.00021	1.049
(c)	0.0540	0.00227	1.084
(d)	0.0201	0.00031	1.060
(e)	0.0572	0.00329	1.085

Table 5.1: Numerically computed values

As can be observed from the figures 5.2 (c) and (e) the stochastic discount factor is most volatile for the combination of a high γ and habit persistence, whereas habit persistence by itself increases the stochastic discount factor only moderately. Moreover, as the figures 5.3 (a)-(b) show, the consumption path itself is only very little affected by habit persistence and adjustment costs of capital.[16] From the table 5.1 we can observe that the Sharpe ratio and equity premium increase strongly with habit persistence and adjustment costs, though not sufficiently to match empirical facts, but the risk free interest rate is still much too high.

6 Interpretation of the Results

It is interesting to compare the numerical results that we have obtained, by using stochastic dynamic programming, to previous quantitative studies undertaken for habit formation, but using other solution techniques. We in particular will restrict ourselves to a comparison with the results obtained by Boldrin et al. (2001) and Jerman (1998).

Whereas Boldrin et al. use a model with log utility for internal habit, but endogenous labor supply in the household's preferences, Jerman studies the asset price implication of a stochastic growth model, also with internal habit formation but, as in our model, labor effort is not a choice variable. All three papers Boldrin et al. (2001), Jerman (1998) and our variant use adjustment costs of investment in the model with habit formation. Both previous studies claim that habit formation models with adjustment costs can match the financial characteristics of the data. Yet, both studies have chosen parameters that appear to be conducive to results which replicate better the financial characteristics such as risk free rate, equity premium and the Sharpe ratio.

In comparison to their parameter choice we have chosen parameters that have commonly been used for stochastic growth models[17] and that seem to describe the first and second moments of the data well. Table 5.2 reports the parameters and the results.

[16] Note that our result on habit persistence is a result that Lettau and Uhlig (2000) have also predicted.
[17] See Santos and Vigo-Aguiar (1998).

Both, the study by Boldrin et al. (2001) and Jerman (1998) have chosen a parameter, $\varphi = 4.05$, in the adjustment costs of investment, a very high value which is at the very upper bound found in the data.[18] Since the parameter φ smoothes the fluctuation of the capital stock and makes the supply of capital very inelastic, we have rather worked with a $\varphi = 0.8$ in order to avoid such strong volatility of returns generated by high φ. Moreover, both papers use a higher parameter for past consumption, b, than we have chosen. Both papers have also selected a higher standard deviation of the technology shock. Boldrin et al. take $\sigma = 0.018$, and Jerman takes a $\sigma = 0.01$, whereas we use $\sigma = 0.008$ which has been employed in many models.[19] Those parameters increase the volatility of the stochastic discount factor, a crucial ingredient to raise the equity premium and the Sharpe ratio.

Boldrin et al.[a)]	Jerman[b)]	Grüne and Semmler	US Data[c)] (1954-1990)
b= 0.73-0.9 φ=4.15 σ= 0.018 ρ= 0.9 β= 0.999 γ= 1	b= 0.83 φ=4.05 σ= 0.01 ρ= 0.99 β= 0.99 γ= 5	b= -0.5 φ= 0.8 σ= 0.008 ρ= 0.9 β= 0.95 γ= 1-3	
$R^f = 1.2$	$R^f = 0.81$	$R^f = 20.4 - 34.0$	$R^f = 0.8$
$E(R) - R^f$ =6.63	$E(R) - R^f$ =6.2	$E(R) - R^f$ =1.32	$E(R) - R^f$ =6.18
SR= 0.36	SR= 0.33	SR=0.11	SR=0.35

a) Boldrin et al.(2001) use a model with endogenous labor supply, log utility for habit formation and adjustment costs, σ quarterly, return data and Sharpe ratio are, in percentage terms annualized.

b) Jerman (1998) uses a model with exogenous labor supply, habit formation with coefficient of RRA of 5, and adjustment costs, σ quarterly, return data and Sharpe ratio are annualized.

c) The return dara and the Sharpe ratio are annualized.

d) The following financial characteristics of the data are reported in Jerman (1998), annualized returns and Sharpe ratio in percentage terms.

Table 5.2: Habit formation models

Jerman, in addition, takes a very high parameter of relative risk aversion, a $\gamma = 5$, which also increases the volatility of the discount factor and increases the equity premium when used for the pricing of assets. Jerman also takes a much higher persistence parameter for the technology shocks, a $\rho = 0.99$, from which one knows that it will make the stochastic discount factor more volatile too. All in all, both studies have chosen parameters which are known to bias the results toward the empirically found financial characteristics.

We also want to remark that both papers do not provide any accuracy test for their procedure that they have chosen to solve the intertemporal decision problem. Boldrin et

[18]See for example, Kim (2002) for a summary of the empirical results reported on φ in empirical studies.
[19]This value of σ has also been used by Santos and Vigo-Aguiar (1998).

al. use the Lagrangian multiplier from the corresponding planner's problem to solve for asset prices with no accuracy test for the procedure. Jerman uses a log-linear approach to solve the model and an accuracy test of this procedure is also not provided in the paper. We also want to note that there is a crucial constraint in habit formation models, namely that the surplus consumption has to remain non-negative when the optimal solution, C_t, is computed.[20] As we have shown in section 5 this constraint has to be treated properly in the numerical solution method.

We also want to note, since the risk aversion, $\frac{\gamma}{s_t}$ for power utility, rises with the consumption surplus ratio, given by $s_t = \frac{c_t - X_t}{c_t}$, that the habit persistence model predicts a rising risk aversion (rising Sharpe ratio) in reversions and falling risk aversion (falling Sharpe ratio in booms, for details see Cochrane (2001, p. 471). Thus, the risk aversion and the Sharpe ratio move over time. A state dependent risk aversion and Sharpe ratio are also observable in our computations, for details see Grüne and Semmler (2004). What we have depicted in table 3.1 and 5.2 are averages along the optimal trajectories in the neighborhood of the steady state.

A further remark is needed on the high risk free rate, in table 5.1 computed as $1 + R^f$ and for quarterly data and in table 5.2, for annualized in percent terms. As table 5.2 shows the risk free rate is highest for the high γ, and with thus the annual risk free rate in table 5.2 can go up to value above 30 percent. (In the basic model with $\gamma = 1$ and not habit persistence and adjustment cost, f is about 20 percent.) This is much too high as compared to empirical data where the annual risk free rate is usually reported to be about 1 to 2 percent (depending on time periods considered). Yet, as well known in the literature[21] the equity premium and Sharpe ratio are favorably be impacted by a higher γ produces also a higher risk free rate. this is the reason, why we are successful, to increase the equity premium and Sharpe ratio in our model but, the risk free rate moves in the wrong direction, namely it rises too. In our context the high risk free rate is also computed by a low β (we have chosen $\beta = 0.95$, as compared to a $\beta = 0.99$.. in the study by Boldrin et al. and Jerman). As also well known in the literature a higher β will reduce again the risk free rate.[22]

Overall, one is, inclined to conclude that previous studies because of, first, the specific parameter choice and, second, lacking accuracy tests of the solution procedure have not satisfactorily solved the dynamics of asset prices and the equity premium puzzle. As can be observed from table 5.2 our results show that even if habit formation is jointly used with adjustment costs of investment there are still puzzles remaining for the consumption-based asset pricing models. Finally, we want to note that in our study we have chosen a model variant with no endogenous labor supply, which, as Lettau and Uhlig (2000) show, is the most favorable model for asset pricing in a production economy, since including labor supply as a choice variable, would even reduce the equity premium and the Sharpe ratio.

---

[20] Boldrin et al. (2001:154) just make a general statement "that $C_t \leq bC_{t-1}$...[is] never observed in the Monte Carlo simulations..."

[21] See Cochrane (2001, ch. 21) and Campbell, Lo and MacKinley (1997, ch. 8.2).

[22] See, for example, Cochrane (2001, ch. 21) chosen this. We have taken low value of β in order to make our study comparable to previous studies solving the stochastic growth model with dynamic programming. A too high risk free interest rate can always be reduced by a higher β, see Cochrane (2001, ch. 21) and Campbell, Lo and MacKinley (1997, ch. 82.).

7 Conclusion

Extensive research effort has recently been devoted to study the asset price characteristics, such as the risk-free interest rate, the equity premium and the Sharpe ratio, arising from the stochastic growth model of the Brock type. The failure of the basic model to match the empirical characteristics of asset prices and returns has given rise to numerous attempts to extend the basic model by allowing for different preferences and technology shocks, adjustment costs of investment, the effect of leverage on asset prices and heterogenous households and firms.[23]

The aim of this paper was two-fold. First, we wanted to study the financial characteristics of a model with the most basic and promising extensions. We have chosen a model with more complex decision structure, a model with habit persistence, and augmented it, along the line of Boldrin et al (2001) and Jerman (1998), with adjustment costs of investment. Second, we intended to apply and explore a solution method, a stochastic dynamic programming algorithm, that provides rather accurate global solutions. We apply this numerical procedure to an extended version of the basic stochastic growth model.

The algorithm, we apply here, has been tested for a basic stochastic growth model, where asset prices and the Sharpe ratio can analytically be computed and the algorithm tested. Our computations for the basic model, see Grüne and Semmler (2004) show that the optimal consumption, the value function and the Sharpe ratio can be computed with small absolute errors. Overall our accuracy test is very encouraging and our method thus can safely be applied to extended versions of the stochastic growth model.

In this paper we have employed the most promising extensions of the basic model stochastic growth model, namely habit persistence and adjustment costs of investment. By doing so we, however, employ not extreme, but rather realistic parameter values and solve for the asset price characteristics. Our results, based on an algorithm with reliable accuracy test, shows that, even if habit persistence is jointly used with adjustment costs of investment, there are still puzzles remaining for consumption based asset pricing models.

[23] A model with heterogenous firms in the context of a Brock type stochastic growth model can be found in Akdeniz and Dechert (1997) who are able to match, to some extent, the equity premium by building on idiosynchratic stochastic shocks to firms.

Appendix: Adaptive Gridding Strategy

The basic idea of our adaptive gridding algorithm lies in evaluating the residual of the operator T applied to V_Γ, and as described in sect. 3 as made precise in the following definition. Here for any subset $B \subset \Omega$ and any function $W \in C(\Omega, \mathbb{R})$ we use

$$\|W\|_{\infty,B} := \max_{x \in B} |W|.$$

(i) We define the *a posteriori error estimate* η as a continuous function $\eta \in C(\Omega, \mathbb{R})$ by

$$\eta(x) := |T(V_\Gamma)(x) - V_\Gamma(x)|.$$

(ii) For any element Q_l of the grid Γ we define the *elementwise error estimate*

$$\eta_l := \|\eta\|_{\infty,Q_l}$$

(iii) We define the *global error estimate* $\eta_{\max}$ by

$$\eta_{\max} := \max_l \eta_l = \|\eta\|_\infty.$$

It is shown in Grüne (2003), that for this error estimate the inequalities

$$\frac{\eta_{\max}}{1+\beta_0} \le \|V - V_\Gamma\|_\infty \le \frac{\eta_{\max}}{1-\beta_0}$$

holds. These inequalities show that the error estimate is *reliable* and *efficient* in the sense of numerical error estimator theory, which is extensively used in the numerical solution of partial differential equations. Furthermore, $\eta(x)$ is continuous and one can show that a similar upper bound holds for the error in the derivative of V and V_Γ.

If the size of a grid element tends to zero then also the corresponding error estimate tends to zero, even quadratically in the element size if V_Γ satisfies a suitable "discrete $C2$" condition, i.e., a boundedness condition on the second difference quotient.

This observation shows that refining elements carrying large error estimates is a strategy that will eventually reduce the element error and consequently the global error, and thus forms the basis of the adaptive grid generation method which we will describe in the next section.

Clearly, in general the values $\eta_l = \max_{x \in Q_l} \eta(x)$ can not be evaluated exactly since the maximization has to be performed over infinitely many points $x \in Q_l$. Instead, we approximate η_l by

$$\tilde{\eta}_l = \max_{x_T \in X_T(Q_l)} \eta(x_T),$$

where $X_T(Q_l)$ is a set of test points. In our numerical experiments we have used the test points indicated in Figure 7.4.

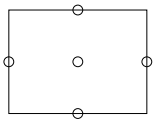

Figure 7.4: Test points $X_T(Q_l)$ for a 2d element Q_l

The adaptive grid itself was implemented on a tree data structure in the programming language C. The adaptive refinement follows the standard practice in numerical schemes and works as follows:

(0) Choose an initial grid Γ_0, set $i = 0$, fix a refinement threshold $\theta \in (0, 1)$

(1) Compute V_{Γ_i} and the (approximated) error estimates $\tilde{\eta}_l$ and $\tilde{\eta}_{\max}$. If a desired accuracy or a maximally allowed number of nodes is reached, then stop

(2) Refine all elements Q_l with $\tilde{\eta}_l \geq \theta\tilde{\eta}_{\max}$, denote the new grid by Γ_{i+1}

(3) Set $i := i + 1$ and go to (1)

Here for the solution of V_{Γ_i} for $i \geq 1$ we use the previous solution $V_{\Gamma_{i-1}}$ as the initial value for the iteration described in Section 3, which turns out to be very efficient.

During the adaptation routine it might happen that the error estimate causes refinements in regions which later turn out to be very regular. It is therefore advisable to include a coarsening mechanism in the above iteration. This mechanism can, e.g., be controlled by comparing the approximation V_{Γ_i} with its projection $\pi_{\widetilde{\Gamma}_i} V_{\Gamma_i}$ onto the grid $\widetilde{\Gamma}_i$ which is obtained from Γ_i by coarsening each element once. Using a specified coarsening tolerance $tol \geq 0$ one can add the following step after Step (2).

(2a) Coarsen all elements Q_l with $\tilde{\eta}_l < \theta\tilde{\eta}_{\max}$ and $\|V_{\Gamma_i} - \pi_{\widetilde{\Gamma}_i} V_{\Gamma_i}\|_{\infty,Q_l} \leq tol$.

This procedure also allows to start from rather fine initial grids Γ_0, which have the advantage of yielding a good approximation $\tilde{\eta}_l$ of η_l. Unnecessarily fine elements in the initial grids will this way be coarsened afterwards.

In addition, it might be desirable to add additional refinements in order to avoid large differences in size between adjacent elements, e.g., to avoid degeneracies. Such regularization steps could be included as a step (2b) after the error based refinement and coarsening has been performed. In our implementation such a criterion was used; there the difference in refinement levels between two adjacent elements was restricted to at most one. Note that the values in the hanging nodes (these are the nodes appearing at the interface between two elements of different refinement level) have to be determined by interpolation in order to ensure continuity of V_Γ.

In addition, our algorithm allows for the anisotropic refinement of elements: consider an element Q of Γ (we drop the indices for notational convenience) and let $X_{new,i}$ be the set of potential new nodes which would be added to Γ if the element Q_l was refined in coordinate direction e_i, cf. Figure 7.5.

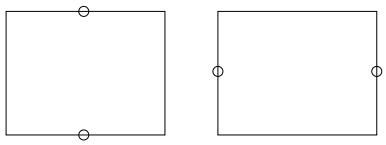

Figure 7.5: Potential new nodes $X_{new,1}$ (left) and $X_{new,2}$ (right) for a 2d element Q

Define the error estimate in these nodes for each coordinate direction e_i by $\eta_{dir,i} := \max_{x \in X_{new,i}} \eta(x)$ and define the overall error measured in these potential new nodes by $\eta_{dir} := \max_{i=1,\dots,n} \eta_{dir,i}$. Note that $\eta_{dir} \leq \eta_l$ always holds. If we include all the points in $X_{new} := \bigcup_{i=1,\dots,n} X_{new,i}$ in our set of test points $X_T(Q)$ (which is reasonable because in order to compute $\eta_{dir,i}$ we have to evaluate $\eta(x)$ for $x \in X_{new}$, anyway) then we can also ensure $\eta_{dir} \leq \tilde{\eta}_l$.

Now we refine the element only in those directions for which the corresponding test points yield large values, i.e., if the error estimate $\eta_{dir,1}$ is large we refine in x–direction and if the error estimate $\eta_{dir,2}$ is large we refine in y–directions (and, of course, we refine in both directions if all test points have large error estimates).

Anisotropic refinement can considerably increase the efficiency of the adaptive gridding strategy, in particular if the solution V has certain anisotropic properties, e.g., if V is linear or almost linear in one coordinate direction. Note that this is the case in our example and the anisotropic refinement is clearly visible in Figure 3.1. On the other hand, a very anisotropic grid Γ can cause degeneracy of the function V_Γ like, e.g., large Lipschitz constants or large (discrete) curvature even if V is regular, which might slow down the convergence. However, according to our numerical experience the positive effects of anisotropic grids are usually predominant.

References

[1] Abel, A. (1990), Asset prices under habit formation and catching up with the Joneses, in: American Economic Review 40(2), 38-42.

[2] Abel, A. (1999), Risk premia and term premia in general equilibrium, in: Journal of Monetary Economics, 43(1): 3-33.

[3] Akdeniz, L. and W.D. Dechert (1997), Do CAPM results hold in a dynamic economy? Journal of Economic Dynamics and Control 21: 981-1003.

[4] Alvarez-Cuadradoo F., G. Monteiro and S.J. Turnovsky (2004), Habit Formation, Catching Up with the Joneses, and Economic Growth, in: Journal of Economic Growth, 9, 47-80.

[5] Breeden, D.T. (1979), An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7: 231-262.

[6] Boldrin, M., L.J. Christiano and J.D.M. Fisher (2001), Habit persistence, asset returns and the business cycle. American Economic Review, vol. 91, 1: 149-166.

[7] Brock, W. (1979) An integration of stochastic growth theory and theory of finance, part I: the growth model, in: J. Green and J. Schenkman (eds.), New York, Academic Press: 165-190.

[8] Brock, W. (1982) Asset pricing in a production economy, in: The Economies of Information and Uncertainty, ed. by J.J. McCall, Chicgo, University pf Chicago Press: 165-192.

[9] Brock, W. and L. Mirman (1972), Optimal economic growth and uncertainty: the discounted case, Journal of Economic Theory 4: 479-513.

[10] Burnside, C. (2001), Discrete state–space methods for the study of dynamic economies, in: Marimon, R. and A. Scott, eds., Computational Methods for the Study of Dynamic Economies, Oxford University Press, 95–113.

[11] Campbell, J.Y. and J.H. Cochrane (1999), Explaining the poor performance of consumption-based asset pricing models, Working paper, Harvard University.

[12] Caroll, C:D., J. Overland and D.N. Weil (1997), Comparison utility in a growth model, in: Journal of Economic Growth, 2: 339-367.

[13] Carroll, C.D., J. Overland and D.N. Weil (2000), Saving and growth with habit formation, in: The American Economic Review, June 2000.

[14] Chow, C.-S. and J.N. Tsitsiklis (1991), An optimal one–way multigrid algorithm for discrete–time stochastic control. IEEE Trans. Autom. Control 36:898–914.

[15] Cochrane, J. (2001), Asset pricing, Princeton University Press, Princeton.

[16] Constantinides, G.M. (1990), Habit formation: a resolution of the equity premium puzzle, in: Journal of POlitical Economy 98: 519-543.

[17] Daniel, J.W. (1976), Splines and efficiency in dynamic programming, J. Math. Anal. Appl. 54:402–407.

[18] Duesenberry, J.S. (1949), Income, saving, and the theory of consumer behavior. Cambridge, MA: Harvard University Press.

[19] Falcone, M. (1987), A numerical approach to the infinite horizon problem of deterministic control theory, Appl. Math. Optim., 15: 1–13

[20] Falcone, M. and R. Ferretti (1998), Convergence analysis for a class of high-order semi-Lagrangian advection schemes. SIAM J. Numer. Anal. 35:909–940.

[21] Fuhrer, J.C. (2000), Habit formation in consumption and its implications for monetary-policy models, in: American Economic Review 90:367-390.

[22] Grüne, L. (1997), An adaptive grid scheme for the discrete Hamilton-Jacobi-Bellman equation, Numer. Math., 75: 1288-1314.

[23] L. Grüne and W. Semmler (2003), Using dynamic programming with adaptive grid scheme for optimal control problems in economics. Working Paper No. 38, Center for Empirical Macroeconomics, University of Bielefeld. Forthcoming, Journal of Economic Dynamics and Control.

[24] Grüne, L. and W. Semmler (2004), Solving asset pricing models with stochastic dynamic programming, CEM working paper no. 54, Bielefeld University.

[25] Hansen, L.P. and T. Sargent (2002) Robust control, book manuscript, Stanford University.

[26] Jerman, U.J. (1998), Asset pricing in production economies, Journal of Monetary Economies 41: 257-275.

[27] Johnson, S.A., J.R. Stedinger, C.A. Shoemaker, Y. Li, J.A. Tejada–Guibert (1993), Numerical solution of continuous–state dynamic programs using linear and spline interpolation, Oper. Research 41:484–500.

[28] Judd, K.L. (1996), Approximation, perturbation, and projection methods in economic analysis, Chapter 12 in: Amman, H.M., D.A. Kendrick and J. Rust, eds., Handbook of Computational Economics, Elsevier, pp. 511–585.

[29] Judd, K.L. and S.-M. Guu (1997), Asymptotic methods for aggregate growth models. Journal of Economic Dynamics & Control 21: 1025-1042.

[30] Keane, M.P. and K.I. Wolpin (1994), The Solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence, The Review of Economics & Statistics, 76:648–672.

[31] Kim, J. (2002), Indeterminacy and investment adjustment costs: an analytical result. Dept. of Economics, Unviersity of Virginia, mimeo.

[32] Lettau, M. and H. Uhlig (1999), Volatility bounds and preferences: an analytical approach, revised from CEPR Discussion Paper No. 1678.

[33] Lettau, M. and H. Uhlig (2000), Can habit formation be reconciled with business cycle facts? Review of Economic Dynamics 3: 79-99.

[34] Lettau, M. G. Gong and W. Semmler (2001), Statistical estimation and moment evaluation of a stochastic growth model with asset market restrictions, Journal of Economic Organization and Behavior, Vol. 44: 85-103.

[35] Lucas, R. Jr. (1978), Asset prices in an exchange economy. Econometrica 46: 1429-1446.

[36] Marcet, A. (1994) Simulation analysis of stochastic dynamic models: applications to theory and estimation, in: C.A. Sims, ed., Advances in Econometrics, Sixth World Congress of the Econometric Society, Cambridge University Press, pp. 81–118.

[37] Marshall, A. (1920), Principles of economics: an introductory volume. 8th. ed. London: Macmillan.

[38] McGrattan, E.R. and E.C. Prescott (2001) Taxes, regulation and asset prices. Working paper, Federal Reserve Bank of Minneapolis.

[39] Munos, R. and A. Moore (2002), Variable resolution discretization in optimal control, Machine Learning 49:291–323.

[40] Reiter, M. (1999), Solving higher–dimensional continuous–time stochastic control problems by value function regression, J. Econ. Dyn. Control, 23:1329–1353.

[41] Rouwenhorst, K.G. (1995), Asset pricing implications of equilibrium business cycle models, in: T. Cooley: Frontiers of Business Cycle Research. Princeton, Princeton University Press: 295-330.

[42] Rust, J. (1996), Numerical dynamic programming in economics, in: Amman, H.M., D.A. Kendrick and J. Rust, eds., Handbook of Computational Economics, Elsevier, pp. 620–729.

[43] Rust. J. (1997), Using randomization to break the curse of dimensionality, Econometrica 65:478–516.

[44] Ryder, Jr.H. and G. Heal (1973), Optimal growth with intertemporally dependent preferences, in: Review of Economic Studies, 40: 1-31.

[45] Santos, M.S. and J. Vigo–Aguiar (1998), Analysis of a numerical dynamic programming algorithm applied to economic models, Econometrica 66: 409-426.

[46] Semmler, W. (2003), Asset prices, booms and recessions, Springer Publishing House, Heidelberg and New York.

[47] Tauchen, G. and R. Hussey (1991), Quadrature–based methods for obtaining approximate solutions to nonlinear asset–price models, Econometrica, 59:371–396.

[48] Taylor J.B. and H. Uhlig (1990), Solving non-linear stochastic growth models: a comparison of alternative solution methods. Journal of Business and Economic Studies, 8: 1-18.

[49] Trick, M.A. and S.E. Zin (1993), A linear programming approach to solving stochastic dynamic programs, Working Paper, Carnegie–Mellon University.

[50] Trick, M.A. and S.E. Zin (1997), Spline approximations to value functions: a linear programming approach, Macroeconomic Dynamics, 1:255-277.

[51] Veblen, T.B. (1899), The theory of the leisure class: an economic study of institutions. New York: Modern Library.