Journal of Economic Dynamics & Control 36 (2012) 468–500

Contents lists available at SciVerse ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Small noise methods for risk-sensitive/robust economies

Evan W. Anderson[a,*], Lars Peter Hansen[b], Thomas J. Sargent[c]

[a] Northern Illinois University, Zulauf 515, DeKalb, IL 60115, United States
[b] University of Chicago and NBER, United States
[c] New York University and NBER, United States

ARTICLE INFO

Article history:
Received 1 September 2009
Accepted 10 June 2011
Available online 6 December 2011

JEL classification:
C63
D81
E20
G12

Keywords:
Computational economics
Recursive utility
Perturbation methods
Stochastic growth model

ABSTRACT

We provide small noise expansions for the value function and decision rule for the recursive risk-sensitive preferences specified by Hansen and Sargent (1995), Hansen et al. (1999), and Tallarini (2000). We use the expansions (1) to provide a fast method for approximating solutions of dynamic stochastic problems and (2) to quantify the effects on decisions of uncertainty and concerns about robustness to misspecification.

Small noise expansions justify standard measures of risk aversion and precaution in models of decision-making under uncertainty (see Pratt, 1964). In a static or a two-period decision problem, risk aversion measures are computed as amounts that compensate a consumer for accepting risk. These measures become arbitrarily accurate when the variance of the risk approaches zero. Similarly, measures of precautionary saving are derived by approximating the response of savings to increased risk, where again arbitrary accuracy is achieved when the variance becomes sufficiently small. One aim of the present paper is to compute such measures for a class of discounted, infinite-horizon control problems. Another aim is to seek measures of aversion to bearing model uncertainty.

We study small noise expansions for discrete-time infinite horizon control problems with risk sensitivity or equivalently with a concern about robustness to model misspecification. We follow Epstein and Zin (1989) and model the preferences of the decision-maker recursively.[1] Hansen and Sargent (1995) showed that for linear–quadratic, Gaussian control problems, a recursive formulation of risk sensitivity preserves the tractability of risk-sensitive control theory and for infinite-horizon control problems delivers decision rules with time-invariant risk adjustments. The solution of the risk-sensitive control problem of Hansen and Sargent (1995) is identical to the solution of a particular type of robust control problem. That equivalence allows us to interpret our expansions as providing a way to quantify the effects of concerns about robustness to model misspecification.

* Corresponding author.
E-mail addresses: ewanderson@niu.edu (E.W. Anderson), l-hansen@uchicago.edu (L.P. Hansen), thomas.sargent@nyu.edu (T.J. Sargent).

[1] In the control theory literature, infinite-horizon, risk-sensitive problems have been studied mostly in models without discounting (e.g. see Whittle, 1990; Glover and Doyle, 1988; Runolfsson, 1994).

0165-1889/$ - see front matter
doi:10.1016/j.jedc.2011.11.007

Our small noise expansions are closely related to expansions in the control-theory literature [e.g. see Fleming and Souganidis, 1986; James, 1992; James et al., 1994; Fleming and Yang, 1994, and especially Campi and James, 1996] but differ from expansions that are typically used in economics. Judd and Guu (1993, 1997) and Judd (1996, 1998) explore small noise expansions in the shock standard deviation in conjunction with expansions in the state variables about a deterministic steady state. Because we expand with respect to the shock variance only, we can approximate around a deterministic trajectory. Some recent work in economics has contributed expansions in other parameters. For example, Kogan and Uppal (2003) consider an expansion in a risk aversion preference parameter. As we will see, our first-order expansion can be re-interpreted as a joint expansion in the shock variance and a parameter governing either risk sensitivity or robustness.

Several recent papers have considered alternative perturbation methods for risk-sensitive (robust) problems. Chen and Zadrozny (2003) obtain a fourth-order expansion for problems in which preferences are quadratic. Following Kogan and Uppal (2003), Trojani and Vanini (2002) obtain an expansion in an analog of a risk-aversion parameter. Hansen et al. (2007, 2008) also extend Kogan and Uppal (2003) by expanding in the reciprocal of the intertemporal elasticity of substitution.

We focus on using small noise expansions for two different purposes:

1. To provide a fast method for approximating solutions of dynamic stochastic problems.
2. To quantify the effects on decision of uncertainty and concerns about robustness to misspecification on optimal decisions.

In the stochastic growth model, we show that a small noise expansion around a deterministic trajectory (which we refer to as the path method) typically provides a significantly more accurate solution than a small noise expansion around a steady state (which we refer to as the steady state method). We also provide examples of how small noise expansions can be used to measure the long-run effects of robustness on capital accumulation, consumption, and asset returns.

The remainder of this paper is organized as follows. In Section 1 we describe the risk-sensitive control problem that is the focus of this paper. In Section 2 we give a heuristic derivation of the Campi–James small noise value function expansion appropriately modified for the infinite-horizon, recursive formulation of risk-sensitive preferences of Hansen and Sargent (1995). In Section 3, we provide a first-order small noise approximation for the decision rule. In Section 4, we describe an algorithm for computing expansions and unconditional expectations. In Section 5, we describe expansions for asset prices. In Section 6, we provide two simple examples where the expansion can be computed analytically. In Section 7, we discuss the relationship between the path expansions described in this paper with the typical steady state expansion used in economics and in Section 8 compare their accuracy for a stochastic growth model. In Section 9 we investigate the robustness interpretation of risk-sensitive preferences. In Section 10 we study the implications of second-order expansions in a stochastic growth model. In Section 11 we extend the expansions to the Kreps–Porteus CES specification. Our conclusions are in Section 12.

1. A risk-sensitive control problem

A decision-maker has an infinite horizon and an increasing sequence of sigma algebras (information sets) $\{\mathcal{F}_t : t=0,1,\ldots\}$. Let i_t denote a time t control vector such as investment; i_t must be adapted to the conditioning information set $\mathcal{F}_t$. There is a sequence of i.i.d. shocks $\{w_t : t=0,1,\ldots\}$ in which w_t is normally distributed with mean zero and covariance I. We take this process to be adapted to $\{\mathcal{F}_t\}$, with w_{t+1} independent of $\mathcal{F}_t$. Let the state vector be denoted x_t where the process $\{x_t : t=0,1,\ldots\}$ is also adapted to $\{\mathcal{F}_t\}$. The state vector evolves as

$$x_{t+1}=A(x_t,i_t)+\sqrt{\epsilon}\Lambda(x_t)w_{t+1} \tag{1}$$

with the initial value x_0 known (x_0 is in $\mathcal{F}_0$). The parameter ϵ enables "small noise" expansions of the value function and the decision rule.

As in Hansen and Sargent (1995), Hansen et al. (1999), Tallarini (2000), and Anderson et al. (2003), we model preferences of the decision-maker using the recursion:

$$U_t=u(x_t,i_t)+\mathcal{R}_t(\beta U_{t+1})$$

where

$$\mathcal{R}_t(\beta U_{t+1})\equiv-\frac{1}{\sigma}\log E[\exp(-\sigma\beta U_{t+1})|\mathcal{F}_t]$$

The function $\mathcal{R}_t$ makes an additional risk adjustment to the continuation value function and is the vehicle by which we introduce risk sensitivity. As emphasized by Hansen and Sargent (1995), the log−exp specification of the recursion provides a bridge between risk-sensitive control theory and a more general recursive utility specification of Epstein and Zin (1989). The degree of risk sensitivity is quantified by σ. When $\sigma=0$, we have the usual Von Neumann–Morgenstern form of state additivity: $\mathcal{R}_t(\beta U_{t+1})\equiv\beta E(U_{t+1}|\mathcal{F}_t)$. Values of σ greater than zero increase aversion to risk *vis a vis* the Von Neumann–Morgenstern specification.

The control problem that is the focus of this paper is to maximize the time zero utility index by a control process adapted to $\mathcal{F}_t$. Depending on specifications of the constraints and criterion function, we shall impose additional restrictions on the set of feasible controls. By construction, the control problem is Markovian with a time invariant solution.

We assume when $\epsilon=0$ there is a unique control, $\{i_t\}_{t=0}^{\infty}$ which obtains the maximum achievable lifetime utility. The analysis in subsequent sections requires this assumption, though it is possible to generalize our approach to remove this requirement and there are interesting problems for which the generalization is needed. For example, the assumption would be violated in a portfolio choice problem in which all assets had the same return when $\epsilon=0$ but different returns when ϵ is positive.

1.1. Markov formulation

We can characterize the solution of the control problem with a functional equation for the value function. Let W^ϵ denote a value function indexed by ϵ, a measure of noise. The Bellman equation is

$$W^\epsilon(x)=\max_i[u(x,i)+\mathcal{T}^\epsilon(\beta W^\epsilon)(x,i)] \tag{2}$$

where $\mathcal{T}^\epsilon$ is the Markov counterpart to the risk-sensitivity operator $\mathcal{R}_t$:

$$\mathcal{T}^\epsilon(\beta W)(x,i)\equiv-\frac{1}{\sigma}\log E(\exp[-\beta\sigma W(y)]|x,i) \tag{3}$$

with $y=A(x,i)+\sqrt{\epsilon}\Lambda(x)w$. Here w is a normally distributed random vector with mean zero and covariance matrix I. The expectation operator in Eq. (3) denotes integration with respect to w. The first-order condition for the control vector i:

$$\left.\frac{\partial u(x,i)}{\partial i}\right|_{i=i^\epsilon}+\beta\frac{\partial A(x,i)'}{\partial i}\left.\frac{E(\exp[-\sigma\beta W^\epsilon(y)]\frac{\partial W^\epsilon(y)}{\partial y}|x,i)}{E(\exp[-\sigma\beta W^\epsilon(y)]|x,i)}\right|_{i=i^\epsilon}=0 \tag{4}$$

where i^ϵ denotes the optimal decision indexed by the amount of uncertainty in the economy.

1.2. An example: the stochastic growth model

Later sections of this paper will study the numerical accuracy of expansions in a stochastic growth model in which agents have power reward functions and preferences are defined recursively as

$$U_t=\frac{C_t^{1-\gamma}}{1-\gamma}-\frac{1}{\sigma}\log E[\exp(-\sigma\beta U_{t+1})|\mathcal{F}_t]$$

where C_t is consumption at time t and γ is a parameter. When $\gamma=1$ we interpret the reward function as logarithmic.

We assume there is an underlying exogenous state vector a_t that evolves as

$$a_{t+1}=\Omega_0+\Omega_a a_t+\sqrt{\epsilon}\Omega_v w_{t+1} \tag{5}$$

where conditioned on information available at time t, w_{t+1} is a vector of independent standard normal random variables. We assume w_s is independent of w_v for any s and $v\neq s$. Values of a_t and w_t are observed at time t. Here Ω_0,Ω_a, and Ω_v are constant matrices.

We let there be a single capital stock K that evolves according to

$$K_t=(1-\delta)K_{t-1}+I_t \tag{6}$$

where δ is the depreciation rate of capital and I_t is investment. Output at time t is

$$Y_t=\exp(Pa_t)K_{t-1}^\alpha \tag{7}$$

where P is a constant row vector, the scalar $\exp(Pa_t)$ is productivity, and α is a scalar parameter that satisfies $0<\alpha\leq 1$. Aggregate resource constraints at time t require that consumption and investment do not exceed output:

$$C_t+I_t\leq Y_t \tag{8}$$

It is convenient to write consumption as

$$C_t=(1-\delta)K_{t-1}+\exp(Pa_t)K_{t-1}^\alpha-K_t$$

by substituting for investment and output from Eqs. (6) and (7). Because agents are never satiated, Eq. (8) holds with equality.

This risk-sensitive stochastic growth model is a special case of the general formulation presented earlier with

$$x_t=\begin{bmatrix}\log K_{t-1}\\ a_t\end{bmatrix},\quad i_t=\log K_t,\quad A(x,i)=\begin{bmatrix}i\\ \Omega_0+\Omega_a\Gamma_a x\end{bmatrix},\quad \Lambda(x)=\begin{bmatrix}0\\ \Omega_v\end{bmatrix}$$

and

$$u(x,i) = \frac{[(1-\delta)\exp(\Gamma_k x)+\exp(P\Gamma_a x+\alpha\Gamma_k x)-\exp(i)]^{1-\gamma}}{1-\gamma}$$

where Γ_k is a selector vector that, when multiplied by x, selects the first element of x (the log capital stock) and Γ_a is a selector matrix that, when multiplied by x, selects the rest of the elements of x (the exogenous state variables). There are many other possible ways of mapping the stochastic growth model into the general framework and Section 6 will discuss an alternative method.

2. First-order expansion for the value function

In this section, we construct a first-order small noise expansion for the value function that is the recursive counterpart to the expansion of Campi and James (1996). We modify their first-order small-noise expansion of a value function to capture our recursive formulation of a risk-sensitive control problem. This expansion allows us to interpret how risk sensitivity alters the control problem. As in Campi and James, we are interested in a value function expansion of the form:

$$W^{\epsilon}(x) = W^o(x)+\epsilon[W_g(x) \;\; W_n(x)]\begin{bmatrix}\sigma \\ 1\end{bmatrix}+o(\epsilon) \tag{9}$$

as ϵ tends to zero.

The baseline value function W^o is the one associated with the corresponding deterministic ($\epsilon=0$) control problem. The "derivative" W_n measures the contribution of the noise to the decision problem in the absence of risk sensitivity and the derivative σW_g measures the incremental contribution of risk sensitivity or concerns about robustness. Our derivation is admittedly heuristic and lacks the rigor offered by Campi and James (1996), who provide sufficient conditions for their approximation to be uniform on compact sets.

2.1. An approximate risk adjustment of the value function

As a step toward our ultimate goal, this section considers a problem in which there is uncertainty between this period and next but none from next period onward. Between this period and next, the state evolves as described in Eq. (1), but at future dates $\epsilon=0$. Since there is no uncertainty after next period, the deterministic value function evaluated at next period's state gives the agent's continuation value. Therefore, the continuation value from next period onward is

$$W^o(z+\sqrt{\epsilon}Hw)$$

where $z=A(x,i)$, $H=\Lambda(x)$, and w is a random vector. To solve the problem, we need to compute

$$-\frac{1}{\sigma}\log\int \exp[-\beta\sigma W^o(z+\sqrt{\epsilon}Hw)]f(w)\,dw \tag{10}$$

where $f(w)$ is the density for a normal random vector with mean zero and covariance matrix I.

Our first step is to obtain a small ϵ approximation to the risk adjustment of $W^o(z+\sqrt{\epsilon}Hw)$. Assume that the value function W^o for the deterministic control problem is twice continuously differentiable. Santos (1994) provides sufficient conditions for deterministic value functions to be twice differentiable. He shows that if the return function is of class C2 and strongly concave, and if the optimal solution always lies in the interior, then the value function is of class C2. Under this presumed smoothness

$$W^o(z+\sqrt{\epsilon}Hw) \approx W^o(z)+\sqrt{\epsilon}\frac{\partial W^o(z)'}{\partial z}Hw+\frac{\epsilon}{2}w'H'\frac{\partial^2 W^o(z)}{\partial z\,\partial z'}Hw \tag{11}$$

Initially, suppose that the quadratic expansion for the value function is exact. That allows us to use a complete-the-square argument that Jacobson (1973) applied to risk-sensitive control problems with quadratic value functions. The complete-the-square argument implies that Eq. (10) can be written as

$$-\frac{1}{\sigma}\log\int\left[\exp\left(-\sigma\beta\left[W^o(z)+\sqrt{\epsilon}\frac{\partial W^o(z)'}{\partial z}Hw+\frac{\epsilon}{2}w'H'\frac{\partial^2 W^o(z)}{\partial z\partial z'}Hw\right]\right)f(w)\,dw\right]$$
$$=\beta\left[W^o(z)-\frac{\beta\epsilon\sigma}{2}\frac{\partial W^o(z)'}{\partial z}H\Gamma(\epsilon)H'\frac{\partial W^o(z)}{\partial z}-\frac{1}{\sigma}\log\det[\Gamma(\epsilon)]\right]$$

where f is the multivariate standard normal density and covariance matrix I, and

$$\Gamma(\epsilon)\equiv\left[I+\epsilon\beta\sigma H'\frac{\partial^2 W^o(z)}{\partial z\,\partial z'}H\right]^{-1}$$

For small ϵ, we see that

$$-\frac{1}{\sigma}\log\int\left[\exp\left(-\beta\sigma\left[W^o(z)+\sqrt{\epsilon}\frac{\partial W^o(z)'}{\partial z}Hw+\frac{\epsilon}{2}w'H'\frac{\partial^2 W^o(z)}{\partial z\,\partial z'}Hw\right]\right)f(w)\,dw\right]$$
$$=\beta\left[W^o(z)-\frac{\epsilon\beta\sigma}{2}\frac{\partial W^o(z)'}{\partial z}HH'\frac{\partial W^o(z)}{\partial z}+\frac{\epsilon}{2}\mathrm{tr}\left[HH'\frac{\partial^2 W^o(z)}{\partial z\,\partial z'}\right]\right]+o(\epsilon)$$

where we have used the fact that

$$-\frac{1}{\sigma}\log\det[\Gamma(\epsilon)]=\frac{\epsilon\beta}{2}\mathrm{tr}\left[HH'\frac{\partial^2 W^o(z)}{\partial z\,\partial z'}\right]+o(\epsilon)$$

Although our presentation of this small ϵ approximation took as given a quadratic value function, so long as the value function is twice continuously differentiable, this approximation extends directly to control problems for which the value function is not necessarily quadratic, as in Eq. (11). Therefore, we have

$$-\frac{1}{\sigma}\log\int\exp[-\beta\sigma W^o(z+\sqrt{\epsilon}Hw)]f(w)\,dw=\beta\left[W^o(z)-\frac{\epsilon\beta\sigma}{2}\frac{\partial W^o(z)'}{\partial z}HH'\frac{\partial W^o(z)}{\partial z}+\frac{\epsilon}{2}\mathrm{tr}\left[HH'\frac{\partial^2 W^o(z)}{\partial z\,\partial z'}\right]\right]+o(\epsilon) \tag{12}$$

In what follows, the term

$$\beta\left[-\frac{\beta\sigma}{2}\frac{\partial W^o(z)'}{\partial z}HH'\frac{\partial W^o(z)}{\partial z}+\frac{1}{2}\mathrm{tr}\left[HH'\frac{\partial^2 W^o(z)}{\partial z\partial z'}\right]\right]$$

will be used to infer a derivative with respect to ϵ. As we will see, the second component gives rise to the familiar second derivative contribution used in local characterizations of risk aversion. The incremental contribution contributed by risk sensitivity or concerns about robustness is reflected in the first term, whose magnitude is scaled by σ. The first component is unambiguously negative. Whether the second one is negative depends on the concavity of the value function.

2.2. An approximate risk adjustment for the small noise value function

In the preceding section, we assumed that there is no uncertainty from next period onward. We now relax that assumption and assume that the state evolves according to Eq. (1) at all dates. One implication of this is that the deterministic value function no longer gives the continuation value from next period onward. To solve the agent's problem, now we need to compute

$$-\frac{1}{\sigma}\log\int\exp\left[-\beta\sigma W^\epsilon(z+\sqrt{\epsilon}Hw)\right]f(w)\,dw \tag{13}$$

where W^ϵ is the value function that solves the Bellman Eq. (2). In this section, we seek an approximation to Eq. (13) when W^ϵ is represented by its first-order approximation. Write

$$W^\epsilon(z)=W^o(z)+\epsilon h(z)+o(\epsilon) \tag{14}$$

where h is a correction term presumed to be continuous in the state vector. Then

$$W^\epsilon[z+\sqrt{\epsilon}Hw]=W^o[z+\sqrt{\epsilon}Hw]+\epsilon(h[z+\sqrt{\epsilon}Hw]-h(z))+\epsilon h(z)+o(\epsilon) \tag{15}$$

We obtain a formula for the approximate risk adjustment to the function on the left side of Eq. (15) by deducing the approximate risk adjustment to the right side of Eq. (15). The term $\epsilon(h[z+\sqrt{\epsilon}Hw]-h(z))$ on the right side of Eq. (15) makes only a higher order (in ϵ) contribution to the risk adjustment and can be ignored. Moreover,

$$-\frac{1}{\sigma}\log\int\exp[-\beta\sigma W^o(z+\sqrt{\epsilon}Hw)+\epsilon h(z)]f(w)\,dw=-\frac{1}{\sigma}\log\int\exp[-\beta\sigma W^o(z+\sqrt{\epsilon}Hw)]f(w)\,dw+\epsilon h(z) \tag{16}$$

Therefore, substituting from Eqs. (12), (14), and (16), we obtain

$$-\frac{1}{\sigma}\log\int\exp[-\beta\sigma W^\epsilon(z+\sqrt{\epsilon}Hw)]f(w)\,dw=\beta W^o(z)+\beta\left[\epsilon h(z)-\frac{\epsilon\beta\sigma}{2}\frac{\partial W^o(z)'}{\partial z}HH'\frac{\partial W^o(z)}{\partial z}+\frac{\epsilon}{2}\mathrm{tr}\left[HH'\frac{\partial^2 W^o(z)}{\partial z\,\partial z'}\right]\right]+o(\epsilon) \tag{17}$$

Hence, the assumption that βW^ϵ is not known perfectly means that its first-order correction term alters the approximate risk adjustment by the same amount as the correction.

2.3. A recursion for the small noise correction

We can now deduce a recursion for the correction term h in Eq. (14) that comes directly from Bellman equation (2). It is straightforward to deduce once we observe that we can ignore the small ϵ contribution to the control law because the optimal control law satisfies Eq. (4), making the impact of maximization be of second order. It suffices to look at the recursion:

$$W^{\epsilon}(x) = u[i^{o}(x),x] + \mathcal{T}^{\epsilon}(\beta W^{\epsilon})[x,i^{o}(x)]$$

where $i^{o}(x)$ is the optimal control law for the deterministic control problem. Substituting the adjustment terms from Eq. (14) into the left side and from Eq. (17) into the right side, it follows that:

$$h(x) = \beta h[A^{o}(x)] - \frac{\beta^2 \sigma}{2} \frac{\partial W^{o}(y)'}{\partial y} \Lambda(x)\Lambda(x)' \frac{\partial W^{o}(y)}{\partial y}\bigg|_{y=A^{o}(x)} + \frac{\beta}{2} \mathrm{tr}\left[\Lambda(x)\Lambda(x)' \frac{\partial^2 W^{o}(y)}{\partial y \partial y'}\right]\bigg|_{y=A^{o}(x)}$$

where $A^{o}(x) = A[x,i^{o}(x)]$ and where we have used the fact that $H = \Lambda(x)$ and $z = A^{o}(x)$.

Solving this recursion forward, we obtain

$$h(x) = [W_g(x) \;\; W_n(x)] \begin{bmatrix} \sigma \\ 1 \end{bmatrix}$$

where

$$W_g(x) = -\frac{\beta}{2} \sum_{j=1}^{\infty} \left[\beta^j \frac{\partial W^{o}(y)'}{\partial y} \Lambda[A^{o,j-1}(x)]\Lambda[A^{o,j-1}(x)]' \frac{\partial W^{o}(y)}{\partial y}\bigg|_{y=A^{o,j}(x)} \right] \tag{18}$$

$$W_n(x) = \frac{1}{2} \sum_{j=1}^{\infty} \left[\beta^j \mathrm{tr}\left(\Lambda[A^{o,j-1}(x)]\Lambda[A^{o,j-1}(x)]' \frac{\partial^2 W^{o}(y)}{\partial y \partial y'}\bigg|_{y=A^{o,j}(x)} \right] \right) \tag{19}$$

and where $A^{o,j}$ denotes the jth iterate of the function A^{o}. Here $A^{o,o}$ is the identity function so that $A^{o,o}(x) = x$ for any x.

Since the term W_g is multiplied by the risk-sensitivity parameter σ, its magnitude gives us one way to assess how risk sensitivity alters the value function. A standard result from the risk-sensitive control literature is that this term injects a concern about robustness to misspecification of the transition law. The parameter $\theta = 1/\sigma$ penalizes perturbations to the baseline model and measures the decision maker's concerns about misspecification. Larger values of θ translate into smaller concerns about misspecification. The term W_g is necessarily negative semidefinite. The term W_n is the familiar second derivative contribution that represents a dynamic version of a local measure of risk aversion. Both terms apply to discounted infinite horizon control problems.

In summary, our approximation to the value function is given by Eq. (9) with W_g and W_n being given by Eqs. (18) and (19).

3. First-order expansion for the decision-rule

In this section, we construct a first-order small noise decision-rule expansion. Although a decision-rule expansion does not come into play when deducing the first-order value function expansion, it is important for characterizing how the shock variances influence actions, for example, consumption and investment in models of precautionary saving. The decision-rule expansion is also a crucial ingredient of higher order value function expansions. The decision-rule expansion can be obtained by computing the appropriate derivatives of the first-order conditions (4) for optimality, or alternatively the envelope condition.

We begin by defining alternative versions of the risk-sensitivity and noise corrections

$$\hat{W}_g(x,y) = -\frac{\beta^2}{2} \left[\frac{\partial W^{o}(y)'}{\partial y} \Lambda(x)\Lambda(x)' \frac{\partial W^{o}(y)}{\partial y} \right] + \beta W_g(y) \tag{20}$$

$$\hat{W}_n(x,y) = \frac{\beta}{2} \left[\mathrm{tr}\left(\Lambda(x)\Lambda(x)' \frac{\partial^2 W^{o}(y)}{\partial y \, \partial y'} \right] \right) + \beta W_n(y) \tag{21}$$

in which next period's state y need not be the optimal choice of next period's state (in the deterministic problem) but all subsequent choices are optimal given y. If y is the optimal choice of next period's state, $A^{o}(x)$, then the alternative definitions (20) and (21) of W_g and W_n are equal to the earlier definitions in Eqs. (18) and (19):

$$\hat{W}_g[x,A^{o}(x)] = W_g(x)$$

$$\hat{W}_n[x,A^{o}(x)] = W_n(x)$$

By computing the appropriate derivatives of the first-order conditions (4) for optimality, the following decision rule expansion can be obtained:

$$i^{\epsilon}(x) = i^{o}(x) - \epsilon b(x)^{-1} \frac{\partial A(x,i)'}{\partial i}\bigg|_{i=i^{o}(x)} \left[\frac{\partial \hat{W}_g(x,y)}{\partial y} \;\; \frac{\partial \hat{W}_n(x,y)}{\partial y} \right] \begin{bmatrix} \sigma \\ 1 \end{bmatrix}\Bigg|_{y=A[x,i^{o}(x)]} + o(\epsilon) \tag{22}$$

where

$$b(x) = \left\{ \frac{\partial^2 u(x,i)}{\partial i \partial i'} + \beta \frac{\partial}{\partial i'} \left[\frac{\partial A(x,i)'}{\partial i} \frac{\partial W^o(y)}{\partial y} \bigg|_{y=A(x,i)} \right] \right\} \bigg|_{i=i^o(x)}$$

Recall that $i^o(x)$ is the optimal control law for the deterministic control problem.

In particular examples, alternative ways of deriving and expressing the decision rule expansions are often available. For example, in the stochastic growth model, an equivalent decision rule expansion can be obtained from the envelope condition

$$\frac{\partial W^\epsilon(x)}{\partial k} = M(x)[C^\epsilon(x)]^{-\gamma}$$

where the derivative is with respect to the first element of x, the log of capital, $k = \Gamma_k x$, and

$$C^\epsilon(x) = (1-\delta)\exp(\Gamma_k x) + \exp(P\Gamma_a x + \alpha \Gamma_k x) - \exp[i^\epsilon(x)]$$

$$M(x) = [(1-\delta)\exp(\Gamma_k x) + \alpha \exp(P\Gamma_a x + \alpha \Gamma_k x)]$$

The first-order expansion for the decision rule can be written

$$i^\epsilon(x) = i^o(x) + \epsilon \left[\frac{1}{\gamma M(x)(C^o[x])^{-\gamma-1} \exp(i^o[x])} \right] \left[\sigma \frac{\partial W_g(x)}{\partial k} + \frac{\partial W_n(x)}{\partial k} \right] + o(\epsilon)$$

This formula provides a simple way to characterize the amount of precautionary saving. Notice that in general the decision rule expansion involves first, second, and third derivatives of the deterministic value function whereas, when $\sigma = 0$, it would only involve third derivatives of the deterministic value function.

4. Computing the expansions

Section 4.1 describes an algorithm for computing the first-order expansion. Section 4.2 sketches an algorithm for computing higher order expansions. Section 4.3 describes a method for approximating unconditional expectations.

4.1. Computing first-order expansions

The expansions are easiest to compute when analytical functional forms for the value function and optimal decision rule in the deterministic problem, in which ϵ is zero, are known. Section 6 computes small noise expansions in two examples when this is the case.[2] In this section, we describe a numerical algorithm that is applicable more generally and is capable of computing the expansion when, as is usually the case, analytical solutions for the corresponding deterministic problem are not available.

In general, to compute the expansion we need to perform two steps. First, we need to solve a nonlinear deterministic problem numerically and second we need to compute W_g and W_n. If, in addition, we want the decision rule expansion, then $\hat{W}_g$ and $\hat{W}_n$ must be differentiated with respect to the date one state vector.

4.1.1. Solving the deterministic problem

We need to solve the deterministic problem that is obtained by setting ϵ to zero in the general problem described in Section 1. In the deterministic problem, the decision maker wants to maximize

$$\sum_{t=0}^{\infty} \beta^t u(x_t, i_t) \tag{23}$$

by choosing control vectors for all dates, $\{i_t\}_{t=0}^{\infty}$, subject to the constraints

$$x_{t+1} = A(x_t, i_t) \tag{24}$$

at all dates, $t \geq 0$. The initial value of the state, x_0, is known.

To solve this problem we use a finite horizon approximation that we formulate by writing the agent's problem as

$$W^o(x_0) = \max_{\{i_t\}_{t=0}^{s-1}} \left(\left[\sum_{t=0}^{s-1} \beta^t u(x_t, i_t) \right] + \beta^s Q(x_s) \right) \tag{25}$$

where the maximization is subject to equations in (24) for $t = 0, 1, \ldots s-1$. Here s is a large number and Q is a one of a number of possible approximations to the value function. For example, if it is known that x_t converges to a steady state, x^*,

[2] If it also is the case that analytical functional forms for the value function and optimal decision rule in the risk-sensitive *stochastic* problem are known then the expansion can be computed by straightforward differentiation with respect to ϵ.

one possible choice for Q is lifetime utility

$$Q(x)=\frac{u[x,\zeta(x)]}{1-\beta} \tag{26}$$

where $\zeta(x)$ is the decision that keeps the state constant at x.[3] In this case one could also impose, as an additional constraint, that $x_s=x^*$ in the optimization problem. However, a better selection for Q, when there is a steady state, is to let it be a second-order steady state approximation for the deterministic value function.

Special care should be exercised in selecting the value of s. The value of s should be large enough so that increasing or decreasing it by a small amount has little effect on the approximation. As we will see in the next subsection, there are additional requirements on s that it be large enough that derivatives of the deterministic value function can be well-approximated.[4]

There exist several good algorithms for solving deterministic problems including Fair and Taylor (1983) and Fair (2003). Fair and Taylor (1983) solve a nonlinear deterministic problem by iterating on first-order conditions and Fair (2003) solves a nonlinear deterministic problem by maximizing lifetime utility by simultaneously choosing optimal decisions at all dates. We found that these algorithms work well, provided that a good initial guess of the solution is available. We also experimented with a time-varying linear–quadratic approximation to the deterministic problem and found that this algorithm tends to work well and does not require a good initial guess when the deterministic problem is sufficiently concave.

4.1.2. Computing W_g and W_n

To find the derivatives in Eqs. (9) and (22), we differentiate the Bellman equation and the first-order condition with respect to the state vector x two times. This yields equations that expresses the derivatives of the value function and decision rules at time t as a function of the derivatives at time $t+1$. By iterating backwards along the optimal path found in step (1), we can find the required derivatives at all dates. The iterations are started at a guess of the derivatives at a date far in the future. When there is a unique stable steady state, a good guess can be found from a steady state approximation for the derivatives.

To compute $W_g(x_0)$ we need to find the first derivative of the value function with respect to the state along the optimal deterministic path. Differentiating the Bellman equation with respect to the state at time t, x_t, yields

$$\frac{\partial W^o(x_t)}{\partial x_t}=\frac{\partial u(x_t,i_t^o)}{\partial x_t}+\beta\left[\frac{\partial A(x_t,i_t^o)}{\partial x_t}\right]'\frac{\partial W^o(x_{t+1})}{\partial x_{t+1}} \tag{27}$$

where in the derivatives of the utility function and the state transition function, A, the optimal deterministic decision is treated as fixed and is not differentiated with respect to x_t.[5] Given $\partial W^o(x_{t+1})/\partial x_{t+1}$ and i_t^o, this formula can be used to find $\partial W^o(x_t)/\partial x_t$. If there is a steady state, then we could solve

$$\left.\frac{\partial W^o(x)}{\partial x}\right|_{x=x^*}=\left.\frac{\partial u(x,i^*)}{\partial x}\right|_{x=x^*}+\beta\left[\left.\frac{\partial A(x,i^*)}{\partial x}\right|_{x=x^*}\right]'\left.\frac{\partial W^o(x)}{\partial x}\right|_{x=x^*} \tag{28}$$

for $\partial W^o(x)/\partial x|_{x=x^*}$, where i^* is the optimal decision at the steady state, and set

$$\frac{\partial W^o(x_s)}{\partial x_s}=\left.\frac{\partial W^o(x)}{\partial x}\right|_{x=x^*} \tag{29}$$

for some large s and use Eq. (27) to iteratively compute $\partial W^o(x_t)/\partial x_t$ for $t=s-1,s-2,s-3,\ldots 0$.[6] However, when there is a steady state, a better choice for $\partial W^o(x_s)/\partial x_s$ could be found from a second (or higher) order steady state approximation for $\partial W^o(x_s)/\partial x_s$. When there is not a steady state, it often works well to set

$$\frac{\partial W^o(x_s)}{\partial x_s}=0 \tag{30}$$

for a large s and iteratively compute the first derivatives for previous dates. The value of s has to correspond to the value used in Section 4.1.1 because we need the optimal deterministic path between dates zero and s.

[3] $\zeta(x)$ solves $x=A[x,\zeta(x)]$. It may be necessary to place additional inequality constraints on x_s if x_s is not required to be x^* because it may not be possible to find $\zeta(x)$ for all x and, in addition, the utility function may not be defined for some solutions.

[4] Algorithms for computing the deterministic value function will also work when $Q(x)=0$ for all x as long as s is large enough, if there is a unique fixed point for the Bellman equation in the deterministic problem. In this case, one needs to be careful that inexactness in the computed optimal path for the state variables when t is close to s does not contaminate the derivative calculations described in the next section.

[5] If the decision rule was treated as a function of x_t then the expression for the first derivative of the value function would be identical to Eq. (27) due to optimality.

[6] See Judd and Guu (1993, 1997) and Judd (1996, 1998) for more on how to compute derivatives at a steady state.

To compute $W_n(x_t)$, we need to find the second derivative of the value function with respect to the state along the optimal deterministic path. We twice differentiate the Bellman equation with respect to x_t:

$$\frac{\partial^2 W^o(x_t)}{\partial x_t\,\partial x_t'} = \frac{\partial}{\partial x_t'}\left(\frac{\partial u(x_t,i_t^o[x])}{\partial x_t} + \beta\left[\frac{\partial A(x_t,i_t^o[x_t])}{\partial x_t}\right]'\frac{\partial W^o(x_{t+1})}{\partial x_{t+1}}\right) \tag{31}$$

to express the time t second derivative as a function of the time $t+1$ second derivative, where here the decision rule is differentiated with respect to x_t. In addition to the second derivative of the value function, the right hand side in general involves the first derivative of the time $t+1$ value function with respect to x_{t+1} as well as the first derivative of the time t decision rule with respect to x_t.[7]

The first derivative of the value function is known from the previous paragraph and the first derivative of the decision rule can be found by differentiating the first-order condition, as we now describe.

To find the first derivative of the decision rule, we write the first-order condition for the deterministic problem at time t as

$$F_t(x,i) = 0 \tag{32}$$

where

$$F_t(x,i) = \frac{\partial u(x,i)}{\partial i} + \beta\frac{\partial A(x,i)'}{\partial i}\frac{\partial W^o(y)}{\partial y}\bigg|_{y=A(x,i)} \tag{33}$$

Differentiating Eq. (32) with respect to $x = x_t$, viewing the optimal decision rule as a function of x_t, yields

$$\frac{\partial F_t(x_t,i^o[x_t])}{\partial x_t} = 0 \tag{34}$$

$$\left(\frac{\partial F_t(x,i)}{\partial x} + \frac{\partial F_t(x,i)}{\partial i}\frac{\partial i^o(x)}{\partial x}\right)\bigg|_{x=x_t,i=i^o(x_t)} = 0 \tag{35}$$

This is a linear equation that can be solved for $\partial i^o(x)/\partial x|_{x=x_t}$ provided that $\partial W^o(x_{t+1})/\partial x_{t+1}$ and $\partial^2 W^o(x_{t+1})/\partial x_{t+1}\,\partial x_{t+1}'$ are known.[8] Now we can iteratively compute $\partial^2 W^o(x_t)/\partial x_t\,\partial x_t'$ starting from an initial guess of $\partial^2 W^o(x_s)/\partial x_s\,\partial x_s'$ for some large s. The initial guess could be the steady state second derivative, a steady state second (or higher) order approximation for $\partial^2 W^o(x_s)/\partial x_s \partial x_s'$, or simply zero.

If we want the decision rule expansion in ϵ, then we differentiate $\hat{W}_g$ and $\hat{W}_n$ with respect to y. The derivative of $\hat{W}_g$ will involve first and second derivatives of the deterministic value function as well as the first derivative of the decision rule with respect to x along the optimal path. The derivative of $\hat{W}_n$ will involve the second and third derivatives of the deterministic value function as well as the first derivative of the decision rule with respect to x along the optimal path. The required derivatives of the value function were computed above except for the third derivative. To find the third derivative of the value function we can differentiate the Bellman equation three times with respect to x_t, which yields a recursive equation that can be solved for the third derivative at all dates, which allows third derivatives to be computed like the first and second derivatives of the value function were above. The third derivative depends on the second derivative of the deterministic decision rule, which also needs to be computed.

4.1.3. Summary of the algorithm

In summary the following algorithm can be used to compute the first order expansion in ϵ:

1. Pick some large s.
2. Choose the function Q and solve the deterministic problem for all dates between 0 and s to obtain the optimal deterministic path for the state variables and the value of the agent's problem at time zero. This gives us x_t for $t=0,1,\ldots s$ as well as $W^o(x_0)$.
3. Compute derivatives of the deterministic value function:
   (a) Guess the values of

   $$\frac{\partial W^o(x_s)}{\partial x_s} \quad \text{and} \quad \frac{\partial^2 W^o(x_s)}{\partial x_s\,\partial x_s'} \tag{36}$$

   for the s chosen in step 1. If the decision rule expansion is desired then a guess of the third derivative of the value function with respect to x_s is also needed.

   (b) Set $t=s-1$.

[7] In computing the derivatives on the right hand side of Eq. (31) one can exploit optimality to eliminate some, but not all, occurrences of the derivatives of the decision rules.

[8] The matrix $\partial F_t(x,i)/\partial i$ will be nonsingular if the deterministic objective is a strictly concave function of i.

(c) Compute

$$\frac{\partial W^o(x_t)}{\partial x_t} \quad \frac{\partial i^o(x_t)}{\partial x_t} \quad \text{and} \quad \frac{\partial^2 W^o(x_t)}{\partial x_t\,\partial x_t{}'} \tag{37}$$

using Eqs. (27), (35) and (31). If the decision rule expansion is desired the third derivative of the value function and the second derivative of the decision rule with respect to x_t also need to be computed.

(d) If $t=0$ then done else set $t \leftarrow (t-1)$ and go to step 3c.

4. Approximate the infinite sums in W_g and W_n with the finite sums:

$$W_g(x_0) = -\frac{\beta}{2}\sum_{t=1}^{s}\left[\beta^t \frac{\partial W^o(y)'}{\partial y} \Lambda[A^{o,t-1}(x_0)]\Lambda[A^{o,t-1}(x_0)]' \frac{\partial W^o(y)}{\partial y}\bigg|_{y=A^{0,t}(x_0)}\right] \tag{38}$$

$$W_n(x_0) = \frac{1}{2}\sum_{t=1}^{s}\left[\beta^t \mathrm{tr}\left(\Lambda[A^{o,t-1}(x_0)]\Lambda[A^{o,t-1}(x_0)]' \frac{\partial^2 W^o(y)}{\partial y \partial y'}\bigg|_{y=A^{o,t}(x_0)}\right]\right) \tag{39}$$

Placing these expressions into Eq. (9), along with the deterministic value function, obtained in step 2 yields the first-order expansion for the value function.

5. If we want the decision rule expansion, then let

$$\hat{W}_g(x,y) = -\frac{\beta^2}{2}\left[\frac{\partial W^o(y)'}{\partial y}\Lambda(x)\Lambda(x)'\frac{\partial W^o(y)}{\partial y}\right] + \beta W_g(y) \tag{40}$$

$$\hat{W}_n(x,y) = \frac{\beta}{2}\left[\mathrm{tr}\left(\Lambda(x)\Lambda(x)'\frac{\partial^2 W^o(y)}{\partial y\,\partial y'}\right]\right) + \beta W_n(y) \tag{41}$$

using the approximations for W_g and W_n. Derivatives of $\hat{W}_g(x,y)$ and $\hat{W}_n(x,y)$ with respect to y can then be computed and placed in Eq. (22) to obtain the decision rule expansion.

It is important to verify that the solution does not depend on the selection of s. If it does, then s was not chosen large enough. In practice, if good choices of Q (in step 2) and the derivatives of the value function at time s (in step 3a) are made, then s can usually be chosen to be much smaller than if the initial guesses for Q and the derivatives are zero.

4.2. Computing higher order expansions

Higher order expansions can be computed in similar ways to the first-order expansion presented in Sections 2 and 3 except that we begin with an nth-order expansion in the state variables for the deterministic value function (rather than a second-order as in Eq. (11)). To compute the higher order expansions described in this section, we presume that A, W and i are differentiable as many times as we like. The decision rule expansion and derivatives of $A(x,i)$ do affect higher order expansions. To illustrate the procedure we consider an example that computes an nth order expansion in ϵ for the value function.

In general, since the decision rule depends on ϵ, next period's state will depend on ϵ. We assume that we can write

$$y = z + \sum_{k=1}^{n} \epsilon^k h_{zk} + \sqrt{\epsilon} H w + o(\epsilon^n) \tag{42}$$

where z is next period's state in the deterministic problem and where h_{zk} for all k are correction terms involving derivatives of the state transition matrix, A, with respect to the decision rule and the decision rule with respect to ϵ. To compute h_{z1} one needs the first-order decision rule expansion described in Section 3, but not the second-order decision rule expansion. More generally, h_{zk} depends on the decision rule expansion of orders k and less. To keep the notation simple, we view the current period state x as fixed.[9] However, its important to remember that the correction terms, h_{zk}, do depend on x.[10]

Using an argument similar to our justification of Eq. (11), we can write

$$W^\epsilon(y) - \Omega^\epsilon(z) = \sum_{k=1/2,1,3/2,\ldots n} \epsilon^k \phi_k(z,w) + o(\epsilon^n) \tag{43}$$

where Ω^ϵ includes terms in W^ϵ that do not depend on the random variable w and the functions ϕ_k include terms that depend on both z and w. To keep the notation simple, we have suppressed the dependence of Ω^ϵ and ϕ_k on the current state x.

[9] The current period state is the value of the state the period before the state is y.

[10] The correction h_{zn} does not enter into the nth-order value function expansion and can be computed after the nth-order value function expansion. The terms h_{zk} for $k < n$ generally will enter the nth-order value function expansion and can be computed from value function expansions of orders less than n.

Exponentiating, it follows that

$$\exp[-\beta\sigma W^{\epsilon}(y)+\beta\sigma\Omega^{\epsilon}(z)] = 1+\sum_{k=1/2,1,3/2,\ldots n}\epsilon^{k}\tau_{k}(z,w)+o(\epsilon^{n}) \tag{44}$$

where the functions τ_k can be found by applying the power series for the exponential to the right hand side of Eq. (43):

$$\exp(\theta) = 1+\frac{\theta^2}{2}+\frac{\theta^3}{3!}+\cdots \tag{45}$$

where θ is the right hand side of Eq. (43) multiplied by $-\beta\sigma$.

The expected value of the exponential in Eq. (44) can be written as

$$\int \exp[-\beta\sigma W^{\epsilon}(y)+\beta\sigma\Omega^{\epsilon}(z)]f(w)\,dw = 1+\sum_{k=1}^{n}\epsilon^{k}\overline{\tau}_{k}(z)+o(\epsilon^{n}) \tag{46}$$

where the functions multiplying ϵ are the expected values of τ_k which we denote as $\overline{\tau}_k(z)$:

$$\overline{\tau}_k(z) = \int \tau_k(z,w)f(w)\,dw \tag{47}$$

The expected values of τ_k when k is not an integer are zero. Now taking a logarithm, we have

$$-\frac{1}{\sigma}\log\int \exp[-\beta\sigma W^{\epsilon}(y)]f(w)\,dw = \beta\Omega^{\epsilon}(z)+\sum_{k=1}^{n}\epsilon^{k}\varrho_{k}(z)+o(\epsilon^{n}) \tag{48}$$

for some functions ϱ_k. The functions ϱ_k can be found using a power series for the logarithm, centered at one. Combining Eq. (48) with an expansion for the current period utility function, using the first-order conditions to simplify, and solving forward (as we did in Section 2.3) yields an expansion for the value function of the form

$$W^{\epsilon}(x) = W^{o}(x)+\sum_{k=1}^{n}\epsilon^{k}V_{k}(x)+o(\epsilon^{n}) \tag{49}$$

for some functions V_k, where the dependence on x is explicit.

Like the first-order expansion, the higher order expansions can be interpreted as an expansion in the two parameters ϵ and σ. For example, a second-order value function expansion can be written as

$$W^{\epsilon}(x) = W^{o}(x)+\epsilon[W_g(x)\ \ W_n(x)]\begin{bmatrix}\sigma\\1\end{bmatrix}+\epsilon^2[W_{g2,2}(x)\ \ W_{g2,1}(x)\ \ W_{n2}(x)]\begin{bmatrix}\sigma^2\\ \sigma\\ 1\end{bmatrix}+o(\epsilon^2) \tag{50}$$

where W_g and W_n are defined in Eqs. (18) and (19); and $W_{g2,2}$, $W_{g2,1}$, and W_{n2} depend on the first through fourth derivatives of the deterministic value function and the first-order (in ϵ) decision rule expansion. Notice that the second-order value function correction involves terms multiplying σ raised to the first and second powers. We will see in later sections that the second-order expansions for many other variables also involve terms multiplying both σ and σ^2.

Once the nth-order value function expansion is known, the nth-order decision rule expansion can be found by substituting the nth-order value function expansion into the first-order conditions and differentiating, similar to the first-order analysis in Section 3.

To compute the expansion in (49) we need higher order derivatives of the value function with respect to the state along the optimal deterministic path. These can be computed by differentiating Eq. (27) the required number of times, yielding equations that express the derivatives of the value function and decision rules at time t as a function of the derivatives at time $t+1$. By working backwards along the optimal path found in step 1 (of the algorithm described in the previous section), we can find the required derivatives at all dates. The iterations are started at a guess of the derivatives at a date far into the future. Possible guesses include the derivatives at the steady state, a high order approximation for the derivatives, or zero.

The higher order expansions in ϵ become messy very quickly and are difficult to interpret, even in one-state-variable problems. A tractable approach for computing higher order expansions is to have a computer that automatically generates the expansions with respect to ϵ using the approach described in this section.

4.3. Computing unconditional expectations

Suppose that we want to compute the unconditional expectation of a variable m_t^{ϵ} that can be written as a function of the current state, x_t, so that

$$m_t^{\epsilon} = M^{\epsilon}(x_t)$$

for some function M^{ϵ}. We presume M^{ϵ} can be written as

$$M^{\epsilon}(x) = M^{o}(x)+\epsilon h_m(x)+o(\epsilon) \tag{51}$$

where h_m is continuous. Let

$$Q_{s,t}(x) = M^o[A^{o,t-s}(x)]$$

be the value of m_t^o in the deterministic problem when the state at time $s \le t$ is x. A first-order approximation to the expected value of m_t^ϵ conditioned on information available at time $s \le t$ is

$$Q_{s,t}(x) + \frac{\epsilon}{2}\sum_{j=s+1}^{t} \text{tr}\left[\Lambda[A^{o,j-s-1}(x)]\Lambda[A^{o,j-s-1}(x)]' \frac{\partial^2 Q_{j,t}(z)}{\partial z\,\partial z'}\bigg|_{z=A^{o,j-s}(x)} \right] + \epsilon h_m[A^{o,t-s}(x)] + o(\epsilon)$$

where the state at time s is x.[11] If the limit of this expression exists as t goes to ∞, then it is a candidate for the first-order approximation to the long run mean of m_t^ϵ conditioned on the value of x at time s. If the state transition probabilities are asymptotically stationary with a unique invariant distribution, then this mean will not depend on x and we call it the unconditional expectation of m_t^ϵ.[12] To simplify computations in this case, one can choose x to be a deterministic steady state, regardless of the actual initial value of x.[13]

Computing expansions for higher order moments is straightforward. For example, to compute an unconditional variance one can subtract the square of the expansion for the unconditional expectation of $M^\epsilon(x_t)$ from the expansion for the unconditional expectation of $[M^\epsilon(x_t)]^2$.

5. Asset prices

In this section, we review asset pricing formulas in the presence of risk sensitivity and heuristically derive first-order expansions for asset prices and expected returns. Chabi-Yo et al. (2006a,b) also discuss small noise expansions for asset prices in the absence of risk-sensitivity.

Consider an asset that pays $\exp[b_i^\epsilon(y)]$ as a function of next period's state y. Since the value of next period's state is a random variable from the point of view of today, the payoff is also a random variable whose value depends upon the realization of y. We assume that the payoff also depends on the amount of noise in the economy, ϵ, and that when $\epsilon = 0$ the asset's payoffs are known with complete certainty this period. To make the analysis simple we focus on an asset whose payoff is positive with probability one and assume that the log of payoffs $b_i^\epsilon(y)$ is twice differentiable with respect to y.[14]

Using arguments in Hansen et al. (1999) it can be shown that today's price of the asset is

$$P_i^\epsilon(x) = \beta \frac{\int \exp[-\sigma\beta W^\epsilon(y) + q^\epsilon(y) + b_i^\epsilon(y)] f(w)\,dw}{\int \exp[-\sigma\beta W^\epsilon(y) + q^\epsilon(x)] f(w)\,dw} \tag{54}$$

where

$$y = A[x, i^\epsilon(x)] + \sqrt{\epsilon}\Lambda(x) w \tag{55}$$

Today's price depends on today's state, x, as well as the amount of noise in the economy, ϵ. The functional form for q^ϵ depends upon the specification of the economy. In the stochastic growth model presented in Section 1.2, $q^\epsilon(x)$ is the log of the marginal utility of consumption:

$$q^\epsilon(x) = -\gamma \log[(1-\delta)\exp(\Gamma_k x) + \exp(P\Gamma_a x + \alpha\Gamma_k x) - \exp(i^\epsilon(x))]$$

where Γ_k and Γ_a are the selector vector and matrix defined in Section 1.2.

We assume we can write

$$W^\epsilon(z) = W^o(z) + \epsilon h(z) + o(\epsilon)$$

$$q^\epsilon(z) = q^o(z) + \epsilon h_q(z) + o(\epsilon)$$

$$b_i^\epsilon(z) = b_i^o(z) + \epsilon h_{bi}(z) + o(\epsilon)$$

where the h functions are small noise corrections presumed to be continuous.

In the appendix, we heuristically show that a first-order expansion for the log price is

$$\log P_i^\epsilon(x) = \log P_i^o(x) + \epsilon[\varphi_q(x) + \varphi_{bi}(x) + \varphi_{ei}(x)] + \epsilon\sigma[\varphi_w(x) + \varphi_{ci}(x)] + o(\epsilon) \tag{56}$$

---

[11] To verify this expansion use the result in Eq. (17) with $\sigma = 0$ to first derive an expansion for $E_{t-1}m_t^\epsilon$ and then to derive expansions for

$$E_{t-2}m_t^\epsilon = E_{t-2}[E_{t-1}m_t^\epsilon] \tag{52}$$

for

$$E_{t-3}m_t^\epsilon = E_{t-3}[E_{t-2}m_t^\epsilon] \tag{53}$$

and so on.

[12] See Ljungqvist and Sargent (2004) for discussions of asymptotic stationarity and invariant distributions.

[13] Although in this case the true expectation will not depend on x, we have not ruled out that the approximation depends on x.

[14] The analysis in this section can be adapted to assets which have both positive and negative payoffs. For such an asset the price will be sum of the price of the positive payoffs plus minus one times the price of the absolute value of the negative payoffs.

where

$$\varphi_q(x) = \frac{1}{2}\left[\frac{\partial q^o(z)}{\partial z}\right]' \Lambda(x)\Lambda(x)'\left[\frac{\partial q^o(z)}{\partial z}\right] + \frac{1}{2}\operatorname{tr}\left[\Lambda(x)\Lambda(x)'\frac{\partial^2 q^o(z)}{\partial z\,\partial z'}\right] + h_q(z) - h_q(x) \tag{57a}$$

$$\varphi_{bi}(x) = \frac{1}{2}\operatorname{tr}\left[\Lambda(x)\Lambda(x)'\frac{\partial^2 b_i^o(z)}{\partial z\,\partial z'}\right] + h_{bi}(z) \tag{57b}$$

$$\varphi_{ei}(x) = \frac{1}{2}\left[\frac{\partial b_i^o(z)}{\partial z}\right]' \Lambda(x)\Lambda(x)'\left[2\frac{\partial q^o(z)}{\partial z} + \frac{\partial b_i^o(z)}{\partial z}\right] \tag{57c}$$

$$\varphi_w(x) = -\beta\left[\frac{\partial W^o(z)}{\partial z}\right]' \Lambda(x)\Lambda(x)'\frac{\partial q^o(z)}{\partial z} \tag{57d}$$

$$\varphi_{ci}(x) = -\beta\left[\frac{\partial W^o(z)}{\partial z}\right]' \Lambda(x)\Lambda(x)'\frac{\partial b_i^o(z)}{\partial z} \tag{57e}$$

with $z = A^o[x, i^o(x)]$ and with P_i^o being the price of the asset when ϵ is zero:

$$\log P_i^o(x) = \log\beta + q^o(z) - q^o(x) + b_i^o(z)$$

The terms φ_q and φ_w are constant across assets, whereas the terms $\varphi_{bi}, \varphi_{ei}$, and φ_{ci} depend upon the payoff of the asset being priced. In the absence of risk sensitivity, only the terms φ_q, φ_{bi} and φ_{ei} are in play. The correction term h to the value function W^ϵ does not affect the first-order expansion for asset prices.

The term φ_q is a first-order correction to the log of the expected value of marginal utility growth; φ_{bi} is a first-order correction for the expected value of the *logarithm* of the payoff; φ_{ei} is a correction term that in conjunction with φ_{bi} gives the log of the expected value of the payoff itself (rather than the expected value of the logarithm of the payoff) and includes an additional correction term that takes into account the multiplication of payoffs by future marginal utility; φ_w and φ_{ci} include cross-product terms of the first derivative of the value function with q and b, respectively, and embody the additional contribution of risk sensitivity to asset prices.

We are also interested in the expected log return:

$$\overline{r}_i^\epsilon(x) = \int r_i^\epsilon(x,w) f(w)\,dw \tag{58}$$

where

$$r_i^\epsilon(x,w) = b_i^\epsilon(z + \sqrt{\epsilon}\Lambda(x)w) - \log P_i^\epsilon(x)$$

is the log return. In the appendix, we show that a first-order expansion for the expected log return is

$$\overline{r}_i^\epsilon(x) = \overline{r}_i^o(x) - \epsilon[\varphi_q(x) + \varphi_{ei}(x)] - \epsilon\sigma\left[\varphi_w(x) + \varphi_{ci}(x)\right] + o(\epsilon) \tag{59}$$

The first-order correction term is almost identical to the negative of the first-order term for the log price expansion. The only difference is that φ_{bi} does not enter the log return expansion because it cancels with the expansion for the expected value of log payoffs. In Eq. (59)

$$\overline{r}_i^o(x) = q^o(x) - q^o[A^o(x)] - \log\beta \tag{60}$$

is the log return in the deterministic economy. Since the right hand side of Eq. (60) does not depend on i, it must be the case that all assets have the same return in the deterministic problem.[15] At a deterministic steady state x^*, this formula simplifies to $\overline{r}_i^o(x^*) = -\log(\beta)$, since at a steady state $x^* = A^o(x^*)$.

To evaluate the impact of risk sensitivity on prices, we compute the expected market return and the risk-free rate. The market return is the return on an asset that pays consumption as dividends every period in the future. This can also be viewed as the return on asset that pays

$$C^\epsilon(y) + P_m^\epsilon(y)$$

next period as a function of next period's state y and nothing in later periods. Here $P_m^\epsilon(y)$ is the price of the market next period and

$$b_m^\epsilon(z) = b_m^o(z) + \epsilon h_{bm}(z) + o(\epsilon)$$

where

$$b_m^o(z) = \log(\exp[c^o(z)] + P_m^o(z)) \tag{61}$$

[15] If assets did have different returns in the deterministic economy there would be arbitrage opportunities.

and where c^o is log consumption in the deterministic problem. The value of h_{bm} will not enter the first-order expansion for the expected log return. Since the price of the market in the deterministic problem is

$$P_m^o(x) = \sum_{j=1}^{\infty} \beta^j \exp(q^o[A^{o,j}(x)] - q^o(x) + c^o[A^{o,j}(x)])$$

we can write

$$b_m^o(z) = \log\left(\exp[c^o(z)] + \sum_{j=1}^{\infty} \beta^j \exp(q^o[A^{o,j}(z)] - q^o(z) + c^o[A^{o,j}(z)])\right) = \log\left(\sum_{j=0}^{\infty} \beta^j \exp(q^o[A^{o,j}(z)] - q^o(z) + c^o[A^{o,j}(z)])\right)$$

Formulas (57a)–(57e) can be used in conjunction with this formula. The derivatives of b_m with respect to z can be messy, but at a steady state they are much simpler since if z is a steady state, call it x^*,

$$b_m^o(x^*) = c^o(x^*) - \log(1-\beta)$$

The risk-free rate is the return on asset that pays one unit of consumption next period in every possible state of the world. For this asset

$$b_f^\epsilon(x) = 0$$

so that $\exp[b_f^\epsilon(x)] = 1$. Because b_f^ϵ is constant, all of the asset specific terms in the first-order expansion equation (56) are zero. In particular,

$$\varphi_{bf}(x) = \varphi_{ef}(x) = \varphi_{cf}(x) = 0$$

so that the first-order expansion for the log price of a unit payoff is

$$\log P_f^\epsilon(x) = \log P_f^o(x) + \epsilon\varphi_q(x) + \epsilon\sigma\varphi_w(x) + o(\epsilon)$$

where

$$\log P_f^o(x) = \log\beta + q^o[A^o(x)] - q^o(x)$$

is the log price in the deterministic economy. A first-order expansion for the log risk-free rate is thus

$$\overline{r}_f^\epsilon(x) = \overline{r}_f^o(x) - \epsilon\varphi_q(x) - \epsilon\sigma\varphi_w(x) + o(\epsilon)$$

where $\overline{r}_f^o(x) = \overline{r}_i^o(x)$ is the log return in the deterministic economy (Eq. (60)). The reduction in the risk-free rate due to risk sensitivity is entirely captured by φ_w.

Higher order expansions can be found using a similar procedure starting with higher order expansions for the payoffs, value function, and decision rule rather than the first-order approximations that led to Eqs. (56) and (59). The higher order expansions become unwieldy very quickly, so in later sections we compute them numerically without providing analytical expressions.

6. Simple illustrations of the noise expansion

The main purpose of this section is to illustrate how the expansion works and to provide some preliminary analysis of its accuracy. We provide two simple examples of small noise expansions for which other, more accurate, solutions are available.

Both examples are special cases of a simplified version of the stochastic growth model described in Section 1.2 in which capital fully depreciates ($\delta = 1$) and the exogenous variables a are i.i.d. ($\Omega_a = 0$). The simplifications allow us to write output at time $t+1$ in terms of output and consumption at time t:

$$Y_{t+1} = \exp(a_{t+1})K_t^\alpha = \exp(\Omega_0 + \sqrt{\epsilon}w_{t+1})(Y_t - C_t)^\alpha$$

where we have set $P=1$ and $\Omega_v = 1$ and used Eqs. (6) and (7). To map this economy into our general framework, it is convenient to let the log of output be the single state variable and the log of output minus consumption be the single control variable:

$$x_t = \log Y_t, \quad i_t = \log(Y_t - C_t), \quad A(x,i) = \Omega_0 + \alpha i, \quad \Lambda(x) = 1$$

and

$$u(x,i) = \frac{[\exp(x) - \exp(i)]^{1-\gamma}}{1-\gamma}$$

Notice that this way of mapping the stochastic growth model into the general framework uses only one state variable whereas the mapping described in Section 1.2 would have used two state variables.

The rest of this section considers two special cases. In Section 6.1, households have logarithmic current period utility. In Section 6.2, we assume power utility and set $\alpha = 1$ in the production function.

6.1. Logarithmic preferences

In this subsection, following Tallarini (2000), we let the intertemporal substitution parameter $\gamma = 1$ and replace the current period utility function with the logarithmic function. We also assume $0 < \alpha\beta < 1$. With this current period utility function, analytical solutions for the value function and optimal decision rule are available. We shall compute the first-order noise expansion and show that it is identical to the analytical solution.

The analytical formulas for the value function and consumption decision rule in the risk-sensitive problem are

$$W^{\epsilon}(x) = Dx + \sigma\epsilon F + G \tag{62}$$

$$i^{\epsilon}(x) = x + \log \alpha\beta \tag{63}$$

where D,F and G are constants:

$$D = \frac{1}{1-\alpha\beta}$$

$$F = -\frac{1}{2}\left[\frac{\beta^2}{(1-\beta)(1-\alpha\beta)^2}\right]$$

$$G = \frac{\log(1-\alpha\beta)}{1-\beta} + \frac{\beta\Omega_0 + \alpha\beta \log(\alpha\beta)}{(1-\beta)(1-\alpha\beta)}$$

To compute the small noise expansion, we use analytical formulas for the value function and decision rule in the corresponding deterministic problem. These are special cases of Eqs. (62) and (63)

$$W^o(x) = Dx + G$$

$$i^o(x) = x + \log \alpha\beta$$

and are obtained by setting ϵ at zero. As is well-known, in the *absence* of risk sensitivity (when $\sigma = 0$), noise affects neither the decision rule nor the value function. The presence of both risk sensitivity and noise alters the value function because F is not zero but does not affect the decision rule.

We now differentiate the analytical functional forms for the deterministic problem to compute the small noise expansion. Since

$$\frac{\partial W^o(x)}{\partial x} = D, \quad \frac{\partial^2 W^o(x)}{\partial x^2} = 0$$

the first-order contribution of noise, in the absence of risk sensitivity, to the value function is

$$W_n(x_0) = \frac{1}{2}\sum_{t=1}^{\infty} \beta^t \frac{\partial^2 W^o(x_t)}{\partial x_t^2} = 0$$

and the additional contribution of risk sensitivity is

$$W_g(x_0) = -\frac{\beta}{2}\sum_{t=1}^{\infty} \beta^t \left[\frac{\partial W^o(x_t)}{\partial x_t}\right]^2 = -\frac{\beta}{2}\sum_{t=1}^{\infty} \beta^t D^2 = -\frac{\beta}{2}\left[\frac{\beta D^2}{1-\beta}\right] = -\frac{1}{2}\left[\frac{\beta^2}{(1-\beta)(1-\alpha\beta)^2}\right]$$

The first-order approximation to the risk-sensitive value function is thus

$$W^o(x) + \sigma\epsilon W_g(x) + \epsilon W_n(x) = W^o(x) - \frac{1}{2}\left[\frac{\beta^2}{(1-\beta)(1-\alpha\beta)^2}\right]\sigma\epsilon$$

which equals the exact value function $W^{\epsilon}(x)$ given in Eq. (62). This happens because the true value function is linear in ϵ. The decision rule approximation is also exact since the derivatives of W_g and W_n with respect to x are zero (which implies that the first order approximation equals the deterministic decision rule).

In this example, it is also easy to compute first-order expansions for the market return and risk-free rate. The log consumption decision rule is

$$c^{\epsilon}(x) = x + \log(1-\alpha\beta)$$

which implies that next period's state is

$$\Omega_0 + \alpha x + \alpha \log \alpha\beta + \sqrt{\epsilon} w$$

where x is this period's state and w is a standard normal random variable. Using the notation of Section 5, it follows that

$$q^o(x) = -x - \log(1-\alpha\beta) \quad (64)$$

$$\log P^o_m(x) = x + \log(1-\alpha\beta) + \log\beta - \log(1-\beta) \quad (65)$$

$$b^o_m(x) = x + \log(1-\alpha\beta) - \log(1-\beta) \quad (66)$$

$$\log P^o_f(x) = \log\beta + (1-\alpha)x - \Omega_0 - \alpha\log\alpha\beta \quad (67)$$

Thus,

$$\varphi_q(x) = \tfrac{1}{2}, \quad \varphi_w(x) = \beta D$$

and for the market

$$\varphi_{bm}(x) = 0, \quad \varphi_{em}(x) = -\tfrac{1}{2}, \quad \varphi_{cm}(x) = -\beta D$$

The formulas are simple in this case because the second derivative of $q^o(z)$ with respect to z is zero. For the market, since $\varphi_q + \varphi_{bm} + \varphi_{em} = 0$ and $\varphi_w + \varphi_{cm} = 0$, the first-order correction term is zero.[16] A first-order approximation to the price of the market is just the price of the market in the deterministic economy and a first-order approximation to the expected log market return is just the log market return in the deterministic economy:

$$\Omega_0 - (1-\alpha)x + \alpha\log\alpha\beta - \log\beta$$

It is straightforward to verify that these approximations are exact. The actual log price and the actual expected log market return do not depend upon the amount of noise in the economy or the amount of risk sensitivity.

A first-order approximation for the log risk-free rate is

$$r^\epsilon_f(x) = -\log P^o_f(x) - \epsilon(\tfrac{1}{2} + \sigma\beta D)$$

where $P^o_f(x)$ is the price of a unit payoff (Eq. (67)) in the deterministic economy. When ϵ and σ increase, the risk-free rate becomes smaller. It is straightforward to show that this approximation is exact since the actual log risk-free rate is linear in ϵ.

We now move on to compute approximations of unconditional expectations. In computing these, we have assumed that the initial value of x is finite because there is a steady state in which $\exp(x) = 0$. In this simple example, the unconditional expected value of x is the value of x at a deterministic steady state

$$\text{mean}(x) = \frac{\Omega_0 + \alpha\log\alpha\beta}{1-\alpha}$$

and the unconditional variance of x is

$$\text{var}(x) = \frac{\epsilon}{1-\alpha^2}$$

It follows that the unconditional mean and variance of the risk-free rate are

$$\text{mean}(r_f) = -\log\beta - \epsilon(\tfrac{1}{2} + \sigma\beta D)$$

$$\text{var}(r_f) = \left[\frac{(1-\alpha)^2}{1-\alpha^2}\right]\epsilon$$

The unconditional mean of the risk-free rate depends upon the amount of risk sensitivity but the unconditional variance does not. In this example the approximations for the unconditional means and variances are exact.

In this example, we see risk sensitivity can potentially have a large effect on the risk-free rate but that it never has an effect on the expected log market return. In later subsections, we shall show that when preferences are not logarithmic, risk sensitivity often has a much larger effect on the risk-free rate than it does on the market return.

In this example, the first-order approximations to the value function, decision rules, expected log market return, and the log risk-free rate are all exact. Because the actual functional forms are known, for any ϵ and σ, we could simply have differentiated these functions to obtain the first-order expansion. To illustrate how the small noise approximation works when the actual stochastic functions may not be available, we chose instead to follow the procedure described in earlier sections to compute the expansions without using the stochastic value functions and decision rules. The next section considers another example in which the small-noise expansion can be computed analytically.

[16] In this example $h_q(x) = 0$ and $h_{bm}(x) = 0$.

6.2. Consumption-savings

In this section, we assume that $0<\gamma<1$ and we set $\alpha=1$ so that the marginal product of capital is constant. We make the following assumption so that the infinite horizon deterministic problem is well-posed:

Assumption 1. $\beta\exp[(1-\gamma)\Omega_0]<1$.

There exists an analytical solution for the deterministic problem as well as for the stochastic problem in the absence of risk sensitivity. We will analytically compute the first-order small noise expansion and compare it to the true value function in the absence of risk sensitivity. We also display the expansion in the presence of risk sensitivity. We find that in the absence of risk sensitivity the decision rule expansion is exact and the value function expansion is accurate if $\epsilon\leq 0.001$. The effect of risk sensitivity on decisions is very sensitive to the values of γ,β, and Ω_0.

The analytical expressions for the value function and consumption decision in the deterministic problem are

$$W^o(x)=\frac{1}{1-\gamma}D\exp[(1-\gamma)x],\quad i^o(x)=x+\theta \tag{68}$$

where

$$\theta=\frac{1}{\gamma}\log\beta+\frac{1-\gamma}{\gamma}\Omega_0,\quad D=\left[\frac{1}{1-\exp(\theta)}\right]^\gamma$$

For the deterministic problem, assumption 1 is equivalent to the condition that $\theta<0$. From Eq. (68), given x_0, we know that the optimal value for capital at time t in the deterministic problem is

$$x_t=x_0+t(\Omega_0+\theta)=x_0+\frac{t}{\gamma}(\Omega_0+\log\beta) \tag{69}$$

which implies that

$$(1-\gamma)x_t=(1-\gamma)x_0+t(\theta-\log\beta) \tag{70}$$

The first two derivatives of the value function with respect to x_t evaluated along the optimal trajectory of the deterministic problem are

$$\frac{\partial W^o(x_t)}{\partial x_t}=D\exp[(1-\gamma)x_t]=D\exp[(1-\gamma)x_0+t\theta-t\log\beta],$$

$$\frac{\partial^2 W^o(x_t)}{\partial x_t^2}=(1-\gamma)D\exp[(1-\gamma)x_t]=(1-\gamma)D\exp[(1-\gamma)x_0+t\theta-t\log\beta]$$

where we have used Eq. (70) to replace x_t with x_0.

The first-order contribution of noise to the value function in the absence of risk sensitivity is

$$W_n(x_0)=\frac{1}{2}\sum_{t=1}^{\infty}\beta^t\frac{\partial^2 W^o(x_t)}{\partial x_t^2}=\frac{1}{2}(1-\gamma)D\exp[(1-\gamma)x_0]\sum_{t=1}^{\infty}\exp(t\theta)=N\exp[(1-\gamma)x_0]$$

where

$$N\equiv\frac{1}{2}\left[\frac{(1\text{-}\gamma)D\exp(\theta)}{1\text{-}\exp(\theta)}\right]$$

If $\exp(2\theta)<\beta$, we can write the incremental first-order contribution of risk sensitivity as

$$W_g(x_0)=-\frac{\beta}{2}\sum_{t=1}^{\infty}\beta^t\left[\frac{\partial^2 W^o(x_t)}{\partial x_t^2}\right]^2 \tag{71}$$

$$=-\frac{\beta D^2\exp[2(1-\gamma)x_0]}{2}\sum_{t=1}^{\infty}\exp[t(2\theta-\log\beta]) \tag{72}$$

$$=-\beta G\exp\,[2(1-\gamma)x_0] \tag{73}$$

where

$$G\equiv\frac{1}{2}\left[\frac{D^2\exp(2\theta)}{\beta\text{-}\exp(2\theta)}\right]$$

Clearly, when $\exp(2\theta)\geq\beta$, this contribution will be infinite and the last formula in Eq. (73) is not valid. We leave it to future work to discover whether $\exp(2\theta)<\beta$ is a necessary condition for the existence of a solution to the risk-sensitive problem and whether this is a condition that must be satisfied for the expansion to be valid. The small noise

Table 1
The constant terms in the small noise expansions for different parameter values.

β	$\exp(\Omega_0)$	γ	$\frac{D}{1-\gamma}$	N	G
0.95	1/0.95	0.5	8.9443	21.243	190.00
0.95	1/0.95	0.7	27.139	23.204	629.75
0.95	1/0.95	0.9	148.23	14.082	2087.3
0.95	1.02	0.5	11.826	10.277	51.985
0.95	1.02	0.7	23.154	15.566	296.25
0.95	1.02	0.9	46.633	12.42	1630.3

Table 2
Accuracy of the noise expansion, in the absence of risk sensitivity, for several values of ϵ for the parameter values $\beta=0.95$, $\exp(\Omega_0)=1/0.95$, and $\gamma=3$.

ϵ	$\frac{D^{\epsilon}_{\sigma=0}}{1-\gamma}$	$\frac{D}{1-\gamma}+\epsilon N$
0.00000	-4.00000×10^3	-4.00000×10^3
0.00001	-4.00152×10^3	-4.00152×10^3
0.00005	-4.00761×10^3	-4.00760×10^3
0.00010	-4.01524×10^3	-4.01520×10^3
0.00050	-4.07699×10^3	-4.07600×10^3
0.00100	-4.15599×10^3	-4.15200×10^3
0.00500	-4.86913×10^3	-4.76000×10^3
0.01000	-6.01383×10^3	-5.52000×10^3
0.05000	-8.86617×10^4	-1.16000×10^4

approximation to the risk-sensitive value function is

$$W^o(x)+W_g(x)\sigma\epsilon+W_n(x)\epsilon=\left(\frac{D}{1-\gamma}+\epsilon N\right)\exp[(1-\gamma)x]-\beta\sigma\epsilon G\exp[2(1-\gamma)x]$$

See Table 1 for the values of the constants D, N, and G in several examples.

In the presence of noise but without risk sensitivity, an analytical formula for the value function is[17]

$$W^{\epsilon}_{\sigma=0}(x)=\frac{D^{\epsilon}_{\sigma=0}}{1-\gamma}\exp[(1-\gamma)x]$$

where

$$D^{\epsilon}_{\sigma=0}=\left[\frac{1}{1-\exp(\theta^{\epsilon}_{\sigma=0})}\right]^{\gamma}$$

$$\theta^{\epsilon}_{\sigma=0}=\frac{1}{\gamma}\log\beta+\frac{1-\gamma}{\gamma}\Omega_0+\frac{\epsilon(1-\gamma)^2}{2\gamma}$$

The optimal decision rule is

$$i^{\epsilon}_{\sigma=0}(x)=x+\theta^{\epsilon}_{\sigma=0}$$

In the absence of risk sensitivity, the value function will be finite when $\theta^{\epsilon}_{\sigma=0}<0$ and x is finite. When $\theta^{\epsilon}_{\sigma=0}<0$, it must be the case that $\theta<0$, so that the corresponding deterministic problem discussed earlier in this section is finite whenever the stochastic value function is finite.

Table 2 displays the values of $D^{\epsilon}_{\sigma=0}/(1-\gamma)$ and the coefficient in the noise expansion on $\exp[(1-\gamma)x]$ for several values of ϵ. We see that the expansion does very well when ϵ is 0.001 and smaller, but poorly when ϵ is 0.05.

It can be shown that the decision rule expansion is

$$i^{\epsilon}(x)=x+\theta+\epsilon\left[\frac{(1-\gamma)^2}{2\gamma}\right]-\epsilon\beta\sigma\left(\frac{[1-\gamma]\exp[\theta]}{\gamma[(1-\exp(\theta)]^{\gamma-1}[\beta-\exp(2\theta)]}\right)\exp[(1-\gamma)x]+o(\epsilon)$$

Notice that the decision rule expansion is exact in the absence of risk sensitivity, since in this case the optimal decision rule is linear in ϵ. When we introduce risk sensitivity, this expansion is an approximation unless $\gamma=1$. The impact of risk sensitivity on precautionary savings depends crucially on the magnitude of γ and how close $\exp(2\theta)$ is to β.

[17] The subscript $\sigma=0$ indicates that a formula is only valid when $\sigma=0$.

If $\gamma > 1$ and $\sigma > 0$, then the risk-sensitive value function for this problem appears to be globally minus infinity. One could modify the problem by letting agents have the option of investing in a risk-free bond or by requiring that the distribution of the shock (w_{t+1}) be bounded from below. Either assumption would lead to a well formulated problem for many values of β and Ω_0 when $\gamma > 1$ and $\sigma > 0$. Another modification that would also lead to a well-formulated problem would scale the risk-sensitivity parameter by the value function as in Maenhout (2004).

Since the post World War II standard deviation of the return on the US market has been about 16%, a reasonable approximation for ϵ would be about 0.02 or 0.03. We see that for reasonable values of ϵ, the first-order value function expansion in the absence of risk sensitivity starts to break down. Including higher order terms would fix this problem.

7. Steady state expansions in the state variables and the shock standard deviation

Expansions that are typically used in economics are computed about a steady state and in terms of a parameter that scales the standard deviation of the shocks. The small noise expansions described above are around the optimal deterministic paths of the state variables and in terms of a parameter that scales the variance of the shocks. In this section, we discuss the relationship between these expansions.

In the economics literature on perturbations, it is more common to write the state evolution equation as

$$x_{t+1} = A(x_t, i_t) + \tau \Lambda(x_t) w_{t+1} \tag{74}$$

rather than as $x_{t+1} = A(x_t, i_t) + \sqrt{\epsilon}\Lambda(x_t) w_{t+1}$ and then to seek an expansion in $\tau = \sqrt{\epsilon}$ rather than in ϵ.[18] In this section, we discuss the equivalence between an expansion in τ and an expansion in ϵ. We write the value function, now denoted as Q^τ, as depending on τ:

$$Q^\tau(x) = \max_i \left(u(x,i) - \frac{1}{\sigma} \log E[\exp(\beta\sigma Q^\tau[A(x,i) + \tau\Lambda(x)w])] \right)$$

and show the equivalence between a first-order expansion of W^ϵ in ϵ and a second-order expansion of Q^τ in $\tau = \sqrt{\epsilon}$.

Since by construction $W^\epsilon(x) = Q^\tau(x)$, the chain rule implies

$$\frac{\partial Q^\tau(x)}{\partial \tau} = \frac{\partial W^\epsilon(x)}{\partial \epsilon}\frac{\partial \epsilon}{\partial \tau} = 2\tau \left[\frac{\partial W^\epsilon(x)}{\partial \epsilon}\right] \tag{75}$$

and

$$\frac{\partial^2 Q^\tau(x)}{\partial \tau^2} = 2\left[\frac{\partial W^\epsilon(x)}{\partial \epsilon}\right] + 2\tau \frac{\partial^2 W^\epsilon(x)}{\partial \epsilon^2}\frac{\partial \epsilon}{\partial \tau} = 2\left[\frac{\partial W^\epsilon(x)}{\partial \epsilon}\right] + 4\tau^2 \left[\frac{\partial^2 W^\epsilon(x)}{\partial \epsilon^2}\right] \tag{76}$$

Eq. (75) shows that the first derivative of Q^τ with respect to τ is zero when $\tau = 0$ and implies the well-known result that the first-order expansion in terms of τ is zero about the point $\tau = 0$. Since the second derivative of Q^τ with respect to τ is $2[\partial W^\epsilon(x)/\partial\epsilon]$, when $\tau = 0$ it must be the case that our first-order expansion in ϵ

$$W^\epsilon(x) = W^o(x) + \epsilon \left.\frac{\partial W^\epsilon(x)}{\partial \epsilon}\right|_{\epsilon=0} + o(\epsilon) \tag{77}$$

is equivalent to a second-order expansion in τ:

$$Q^\tau(x) = Q^o(x) + \tau^2 \left(\frac{1}{2}\right) \left.\frac{\partial^2 Q^\tau(x)}{\partial \tau^2}\right|_{\tau=0} + o(\tau^2)$$

Its possible to show that higher order expansions in τ are related to higher order expansions in ϵ in similar ways.[19]

In subsequent sections, we will sometimes refer to the expansion in Eq. (77) as a second-order path expansion in $\sqrt{\epsilon} = \tau$ because it facilitates comparisons with the rest of the economics literature and, in addition, it is more straightforward to implement expansions in $\sqrt{\epsilon} = \tau$ in computer programs.

In the economics literature (following Judd and Guu, 1993, 1997; Judd, 1996, 1998) it is common to seek expansions in $\sqrt{\epsilon}$ and the state variables around $\sqrt{\epsilon} = 0$ and the deterministic steady state. In subsequent sections, we will call this a steady-state expansion. A second-order steady state expansion is

$$W^\epsilon(x) \approx W^o(x^*) + (x - x^*)' \left[\frac{\partial W^o(y)}{\partial y}\right] + \frac{1}{2}(x - x^*)' \left[\frac{\partial^2 W^o(y)}{\partial y^2}\right](x - x^*) + \epsilon[\beta W_g(x^*) \;\; W_n(x^*)] \begin{bmatrix}\sigma \\ 1\end{bmatrix} \tag{78}$$

where $y = x^*$. The next section compares the accuracy of steady state expansions and path expansions.

[18] In the continuous time economics literature it is more common to seek expansions in parameters that scale the variance rather than the standard deviation.

[19] We interpret the derivatives of $W^\epsilon(x)$ with respect to ϵ as right-hand derivatives, since the value function is not defined when $\epsilon < 0$ due to the $\sqrt{\epsilon}$ term that appears in the law of motion for the state vector.

8. Numerical comparisons

In this section, we analyze the accuracy of solutions to the stochastic growth model described in Section 1.2 produced by the path perturbation method (described in Sections 2–4) and the steady state perturbation algorithm (described near the end of Section 7). In the stochastic growth model, we let the exogenous state a be a scalar. We set

$$\alpha=0.3, \quad \delta=0.09, \quad \epsilon=1, \quad P=1, \quad \Omega_a=0.50, \quad \Omega_v=0.02$$

and consider different values of Ω_0,β,γ and σ. To measure accuracy, we compare the solutions produced by the path and steady state perturbation methods to the solutions produced by a grid algorithm. Section 8.1 discusses the grid algorithm and measures of accuracy. Section 8.2 analyzes the accuracy of the path and steady state methods in the stochastic growth model.

8.1. Grid algorithm: a useful benchmark

To measure accuracy, we compare our expansions to a grid-based algorithm. To implement the grid, we use $(1+10\times 2^{15})=327,681$ points for the capital stock, k, and $(1+2^8)=257$ points for the exogenous state, a. The grid points for k are equally spaced between $[-4\log(10)]$ and $4\log(10)$. The grid points for a are determined by Gaussian quadrature using the approach of Tauchen and Hussey (1991).

The grid-based algorithm is known to be one of the more accurate approximations available for low dimensional problems. The error of the decision rule produced by the grid algorithm can be at least as large as

$$\frac{4\log(10)-[-4\log(10)]}{2g}=5.62\times 10^{-5}$$

where $g=327{,}681$ is the number of grid points for capital used.[20] Near the boundaries of the state space we would expect the grid algorithm to be less accurate than this because the decision maker might prefer to choose points outside of the bounds but we prevent him from doing so. Despite the inaccuracies in the grid algorithm, in the following analysis we treat the solution computed by the grid algorithm as the target solution for the path and steady state approximation methods.

To analyze accuracy, we look at computed solutions at a wide range of different values for the current state. Some of the values are perhaps unreasonably small or large for US data. We choose values for the state in order to display properties of algorithms rather than values that have been seen in recent US (detrended) data. It is important that algorithms work even for extreme values of the state in order to provide an analysis of hypothetical events. We choose values that are far from the edge of our grid because we suspect our grid algorithm is less accurate near the edges.

To provide precise measures of accuracy for the value function, we display the relative difference of the solution of perturbation algorithms from the solution produced by the grid algorithm:

$$\left|\frac{W^{\epsilon}(x)-\hat{W}^{\epsilon}(x)}{\hat{W}^{\epsilon}(x)}\right|$$

where W^{ϵ} is the approximate value function for the stochastic growth model computed by either the path or steady state perturbation algorithms and $\hat{W}^{\epsilon}$ is the approximate value function computed by the grid algorithm. We refer to this as the relative error of the value function.

For the log capital decision rule, we display the absolute difference of the solution of perturbation algorithms from the solution produced by the grid algorithm:

$$|i^{\epsilon}(x)-\hat{i}^{\epsilon}(x)|$$

where i^{ϵ} is the approximate decision rule computed by either the path or steady state perturbation algorithms and $\hat{i}^{\epsilon}$ is the approximate decision rule computed by the grid algorithm. We refer to this as the absolute error of the decision rule.[21] In our parameterization, investment, i, is equal to the choice of next period's log capital.

8.2. Accuracy of expansions for the stochastic growth model

In this section, we refer to a nth-order expansion in ϵ as a $2n$th-order expansion in $\sqrt{\epsilon}$ following the standard language used in economics (see Section 7 for further discussion). We consider three examples.

8.2.1. A non-robust example

We begin with a challenging example without robustness in which the coefficient of relative risk-aversion is $\gamma=20$ and the subjective discount factor is $\beta=0.90$.[22] In the top 3×3 plots in Fig. 1, we graph the accuracy of the value

[20] The actual error can be larger or smaller than this. This is only a rough approximation to what the error could be.

[21] It would also be interesting to consider the relative error in the decision rule. We choose not to do this because optimal decisions are close to zero near the deterministic steady state.

[22] We set $\Omega_0=-0.19996246$ so that the steady state value of log-capital $k^*=0$ and the steady state value of log productivity is $a^*=-0.39992493$.

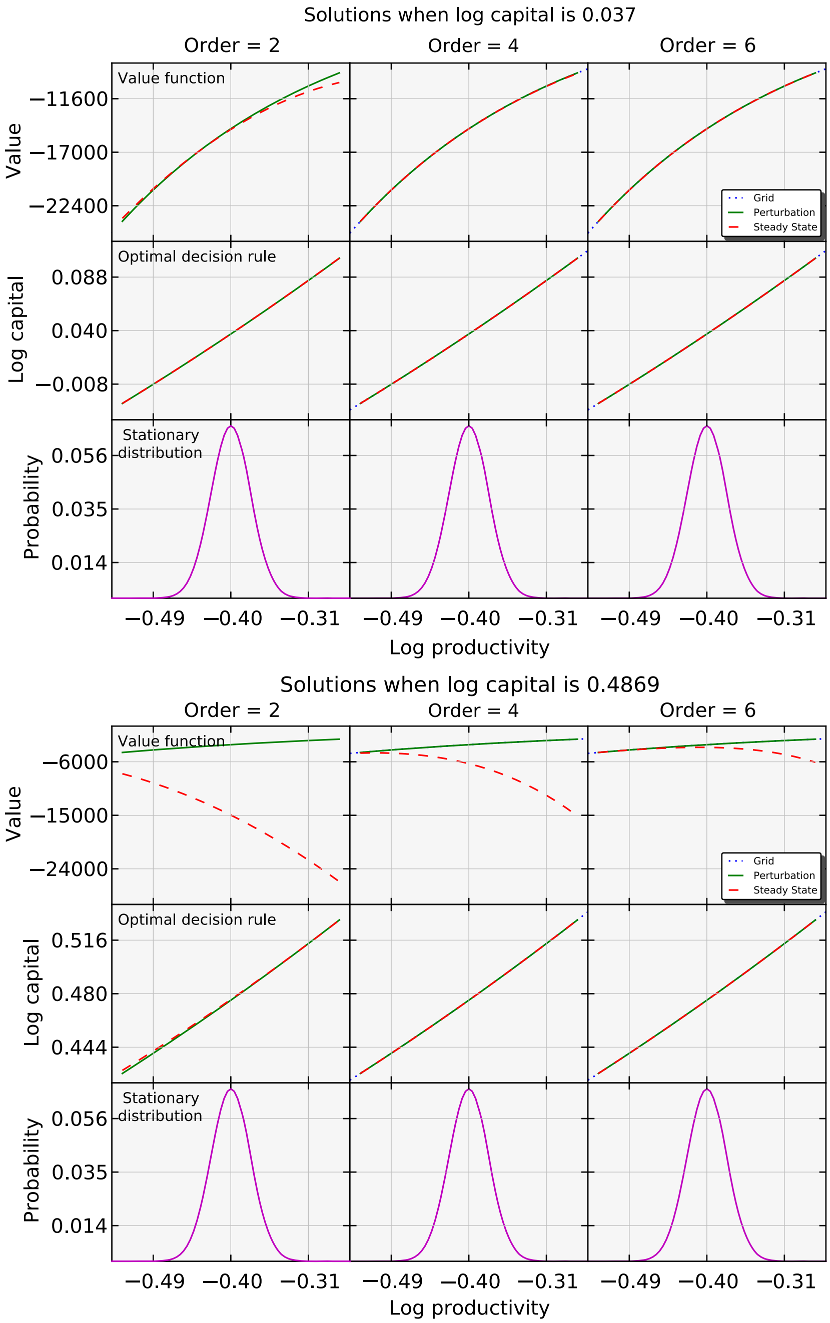

Fig. 1. *Non-robust example.* Decision rules, value functions, and the stationary distribution for the parameters given in Section 8.2.1. In this figure we plot the value function and optimal decision rule, as a function of log productivity holding capital fixed. We also plot the unconditional stationary distribution for log productivity. (The stationary distribution is not conditioned on the fixed value of log capital.) In the upper set of 3×3 plots, log capital is fixed at the mean of its stationary distribution, 0.037. In the bottom plot, log capital is fixed at an extreme value. In both panels, in each square of the first row, three solutions for the value function are displayed. In each square of the second row, three solutions for the optimal decision rule are plotted. In each square of the third row, the unconditional stationary distribution is displayed. In the first column, the second order (in $\sqrt{\epsilon}$) path and steady-state perturbation algorithms are displayed. In the second column, the fourth order (in $\sqrt{\epsilon}$) path and steady-state perturbation algorithms are displayed. In the third column the sixth order (in $\sqrt{\epsilon}$) path and perturbation algorithms are displayed. (The grid solution and stationary distribution are the same in each column.) The grid and path algorithms are visually indistinguishable for most values plotted.

function and decision rule as a function of log productivity when this period's log-capital is fixed at its long run mean, 0.037.[23] In the bottom 3×3 plots in Fig. 1, we examine the value functions and decision rules when current log capital is extremely large (0.4869), and far from its long-run mean.

In Table 3, we see that the second-order path decision rule is an order of magnitude more accurate than the second-order steady state method for most listed values of the current state. The eighth and tenth order path and steady state decision rule expansions are identical, for both methods. The fourth order path expansion for the decision rule is equal to the 10th order expansion for a majority of the values of the state displayed. The fourth order steady state expansion is slightly worse than the second order path expansion.

In Table 3, we also see that the second-order steady state value function expansion is widely inaccurate whereas the second order path value function expansion does fairly well. The fourth order path expansion is several orders of magnitude more accurate than the fourth order steady state expansion. The sixth order steady state expansion is very accurate near the mean of the stationary distribution but is still much worse than the second order path expansion far away from the steady state.

We see that both the steady state and path perturbation algorithms do an excellent job of approximating the decision rule. The path perturbation algorithm also does a good job of approximating the value function, even at low orders. The steady state perturbation algorithm does a poor job of approximating the value function for extreme values of log productivity and log capital at low orders but produces a very accurate solution when the current state is close to the mean of the stationary distribution.

We see that for the non-robust stochastic growth model, even when the utility function is highly non-linear, both the steady state and path methods produce solutions for the decision rule that are good enough for most applications. Part of the reason for the good performance of the steady-state method is that the long-run mean of the stationary distribution is very close to the deterministic steady state.

8.2.2. Homothetic robust example

In this example we let the utility function be logarithmic, the risk-sensitivity parameter be $\sigma = 400$, and the subjective discount factor be $\beta = 0.95$.[24] For these parameter values the long run mean of log-capital is 0.4914 which is far away from the deterministic steady state.

The top half of Fig. 2 graphs value functions and decision rules when current log capital is at the mean of its stationary distribution. The bottom half of Fig. 2 graphs value functions and decision rules for a large value of log capital. We see that both the steady state and path expansions perform poorly when the order is two, though the path approximation method is significantly better than the steady state method.

When the order of the expansion is increased to four, the path approximation does a very good job of approximating the decision rule whereas there still is a noticeable error with the steady state method. When the order is set to six both methods do a very good job of approximating the decision rule. For the value function, the path method does significantly better when the order is two and slightly better when the order is four. When the order is six, the steady state path and perturbation expansions are visually indistinguishable. When we fix capital at a value greater than its long run mean, the relative performance of the steady state method deteriorates.

Table 4 provides precise measures of accuracy. We see that the second order path expansion for the value function and decision rule is about twice as accurate (i.e., its differences with the grid algorithm are approximately half as large) as the steady state method.

In this example the path approximation method does significantly better than the steady state method at low orders. Part of the reason for the dramatic failure of the steady-state method is that the long-run mean of the stationary distribution is very far from the deterministic steady state.

8.2.3. Non-homothetic robust example

In this example we let $\gamma = 0.9$, let $\sigma = 400$, and let the rest of the parameters be identical to the non-robust case discussed in Section 8.2.1.

The top half of Fig. 3 displays value functions and decision rules when current log capital is at the mean of its stationary distribution. The bottom half of Fig. 3 graphs value functions and decision rules for a large value of log capital. The results are similar to the previous example. Table 5 confirms this.

8.2.4. Discussion

Generally, we find that in the absence of risk sensitivity the path method and steady state method both produce accurate approximations to decision rules. The second-order decision rule expansion produced by the path method is more accurate away from the steady state but both approximations are typically excellent near the deterministic steady state. In the presence of risk sensitivity, the second-order decision rule produced by the path method can be significantly more

[23] The long-run mean is approximated with the grid algorithm. In the presence of noise, the long-run mean is not identical to the deterministic steady state.

[24] We set $\Omega_0 = -0.37175877$ so that the steady state value of log-capital is $k^* = 0$ and the steady state value of log productivity is $a^* = -0.74351754$.

Table 3

Accuracy in the stochastic growth model for the non-robust example described in Section 8.2.1. The path method expands in $\sqrt{\epsilon}$ about the point $\sqrt{\epsilon}=0$. The steady state method expands in $\sqrt{\epsilon}$ about the point $\sqrt{\epsilon}=0$ and expands in the state variables about the deterministic steady state $k^{*}=0$, $a^{*}=-0.39992493$.

State		Path method					Steady state method				
		Order					Order				
k	a	2	4	6	8	10	2	4	6	8	10
Panel A: *Relative accuracy of the value function*											
−0.16	−0.53	5.52×10^{-4}	6.58×10^{-6}	1.71×10^{-7}	1.18×10^{-7}	1.17×10^{-7}	1.60×10^{-1}	1.90×10^{-2}	1.46×10^{-3}	8.13×10^{-5}	3.59×10^{-6}
0.037	−0.53	4.91×10^{-4}	5.65×10^{-6}	1.87×10^{-7}	1.43×10^{-7}	1.43×10^{-7}	1.30×10^{-2}	2.53×10^{-4}	3.17×10^{-6}	1.65×10^{-7}	1.43×10^{-7}
0.24	−0.53	4.31×10^{-4}	4.80×10^{-6}	2.29×10^{-7}	1.94×10^{-7}	1.94×10^{-7}	3.24×10^{-3}	1.32×10^{-5}	8.00×10^{-7}	2.24×10^{-7}	1.93×10^{-7}
−0.16	−0.4	5.20×10^{-4}	6.11×10^{-6}	1.84×10^{-7}	1.35×10^{-7}	1.35×10^{-7}	4.10×10^{-2}	1.82×10^{-3}	5.29×10^{-5}	1.24×10^{-6}	1.53×10^{-7}
0.037	−0.4	4.61×10^{-4}	5.23×10^{-6}	1.96×10^{-7}	1.56×10^{-7}	1.56×10^{-7}	4.55×10^{-3}	9.21×10^{-5}	1.10×10^{-6}	1.43×10^{-7}	1.56×10^{-7}
0.24	−0.4	4.05×10^{-4}	4.50×10^{-6}	2.84×10^{-7}	2.52×10^{-7}	2.52×10^{-7}	2.65×10^{-1}	2.11×10^{-2}	9.91×10^{-4}	3.17×10^{-5}	5.06×10^{-7}
−0.16	−0.27	4.84×10^{-4}	5.60×10^{-6}	1.87×10^{-7}	1.43×10^{-7}	1.43×10^{-7}	6.55×10^{-4}	1.22×10^{-5}	5.35×10^{-10}	1.44×10^{-7}	1.43×10^{-7}
0.037	−0.27	4.29×10^{-4}	4.81×10^{-6}	2.17×10^{-7}	1.81×10^{-7}	1.81×10^{-7}	1.08×10^{-1}	6.04×10^{-3}	2.06×10^{-4}	4.75×10^{-6}	9.30×10^{-8}
0.24	−0.27	3.78×10^{-4}	4.11×10^{-6}	2.63×10^{-7}	2.34×10^{-7}	2.34×10^{-7}	1.43×10^{-6}	3.03×10^{-1}	3.76×10^{-2}	3.22×10^{-3}	2.08×10^{-4}
Panel B: *Accuracy of the log capital decision rule*											
−0.16	−0.53	1.36×10^{-5}	1.42×10^{-5}	1.42×10^{-5}	1.42×10^{-5}	1.42×10^{-5}	3.30×10^{-4}	1.31×10^{-5}	1.42×10^{-5}	1.42×10^{-5}	1.42×10^{-5}
0.037	−0.53	1.02×10^{-5}	1.04×10^{-5}	1.04×10^{-5}	1.04×10^{-5}	1.04×10^{-5}	1.60×10^{-5}	1.05×10^{-5}	1.04×10^{-5}	1.04×10^{-5}	1.04×10^{-5}
0.24	−0.53	1.65×10^{-5}	1.65×10^{-5}	1.65×10^{-5}	1.65×10^{-5}	1.65×10^{-5}	6.65×10^{-4}	2.14×10^{-5}	1.65×10^{-5}	1.65×10^{-5}	1.65×10^{-5}
−0.16	−0.4	1.25×10^{-5}	1.22×10^{-5}	1.21×10^{-5}	1.21×10^{-5}	1.21×10^{-5}	1.61×10^{-4}	1.22×10^{-5}	1.21×10^{-5}	1.21×10^{-5}	1.21×10^{-5}
0.037	−0.4	7.44×10^{-7}	7.93×10^{-7}	8.03×10^{-7}	8.03×10^{-7}	8.03×10^{-7}	3.47×10^{-5}	8.37×10^{-7}	8.03×10^{-7}	8.03×10^{-7}	8.03×10^{-7}
0.24	−0.4	2.61×10^{-5}	2.60×10^{-5}	2.60×10^{-5}	2.60×10^{-5}	2.60×10^{-5}	1.95×10^{-4}	2.54×10^{-5}	2.60×10^{-5}	2.60×10^{-5}	2.60×10^{-5}
−0.16	−0.27	6.32×10^{-7}	4.95×10^{-7}	4.84×10^{-7}	4.84×10^{-7}	4.84×10^{-7}	3.59×10^{-4}	2.95×10^{-6}	4.70×10^{-7}	4.84×10^{-7}	4.84×10^{-7}
0.037	−0.27	1.83×10^{-6}	1.92×10^{-6}	1.91×10^{-6}	1.91×10^{-6}	1.91×10^{-6}	1.36×10^{-4}	1.25×10^{-6}	1.91×10^{-6}	1.91×10^{-6}	1.91×10^{-6}
0.24	−0.27	3.29×10^{-6}	3.09×10^{-6}	3.09×10^{-6}	3.09×10^{-6}	3.09×10^{-6}	2.58×10^{-4}	1.55×10^{-6}	3.22×10^{-6}	3.09×10^{-6}	3.09×10^{-6}

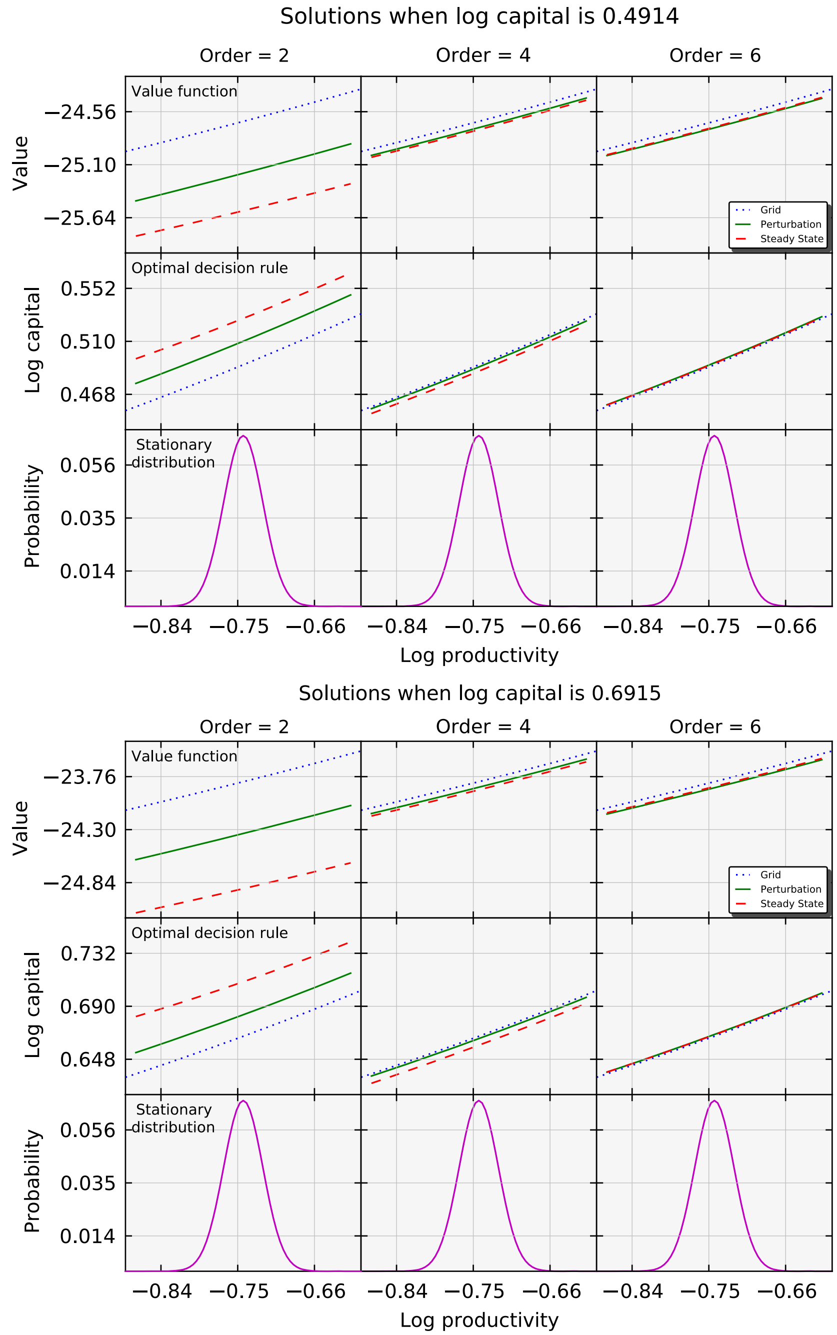

Fig. 2. *Homothetic robust example.* Decision rules, value functions, and the stationary distribution for the parameters given in Section 8.2.2. In this example, 0.4914 is the mean of the stationary distribution for log capital. The organization of the plots follows the structure in Fig. 1.

accurate than the second order expansion produced by the steady state method. Usually, when higher orders are included, both methods perform well.

In the presence of risk sensitivity, expansions for the value function are particularly interesting because the value function determines the magnitude of distortions. We see that the path method generally performs significantly better than the steady-state method at approximating risk-sensitive value functions.[25]

[25] The steady state value function approximations could perhaps be improved by expanding in a variable other than the state vector. See Fernandez-Villaverde and Rubio-Ramirez (2006) for an example of how nonlinear transformations can improve the accuracy of steady state approximations. The optimal transformation is problem specific and possibly parameter specific. One advantage of the path approximations described in this paper is that they do not depend on knowledge of the optimal transformation.

Table 4
Accuracy in the stochastic growth model for the homothetic robust example described in Section 8.2.2. The path method expands in $\sqrt{\epsilon}$ about the point $\sqrt{\epsilon}=0$. The steady state method expands in $\sqrt{\epsilon}$ about the point $\sqrt{\epsilon}=0$ and expands in the state variables about the deterministic steady state $k^*=0$, $a^*=-0.74351754$.

State		**Path method**					**Steady state method**				
		Order					Order				
k	a	2	4	6	8	10	2	4	6	8	10
Panel A: Relative accuracy of the value function											
0.29	−0.87	2.06×10^{-2}	2.88×10^{-3}	2.81×10^{-3}	2.99×10^{-3}	2.96×10^{-3}	2.80×10^{-2}	3.32×10^{-3}	2.64×10^{-3}	3.00×10^{-3}	2.97×10^{-3}
0.49	−0.87	2.12×10^{-2}	2.64×10^{-3}	2.71×10^{-3}	2.88×10^{-3}	2.85×10^{-3}	3.56×10^{-2}	3.41×10^{-3}	2.38×10^{-3}	2.91×10^{-3}	2.85×10^{-3}
0.69	−0.87	2.19×10^{-2}	2.41×10^{-3}	2.63×10^{-3}	2.78×10^{-3}	2.75×10^{-3}	4.44×10^{-2}	3.46×10^{-3}	2.10×10^{-3}	2.85×10^{-3}	2.75×10^{-3}
0.29	−0.74	2.08×10^{-2}	2.81×10^{-3}	2.77×10^{-3}	2.95×10^{-3}	2.92×10^{-3}	2.93×10^{-2}	3.34×10^{-3}	2.58×10^{-3}	2.97×10^{-3}	2.92×10^{-3}
0.49	−0.74	2.14×10^{-2}	2.58×10^{-3}	2.68×10^{-3}	2.84×10^{-3}	2.82×10^{-3}	3.69×10^{-2}	3.43×10^{-3}	2.32×10^{-3}	2.89×10^{-3}	2.82×10^{-3}
0.69	−0.74	2.22×10^{-2}	2.36×10^{-3}	2.61×10^{-3}	2.76×10^{-3}	2.73×10^{-3}	4.58×10^{-2}	3.48×10^{-3}	2.06×10^{-3}	2.84×10^{-3}	2.73×10^{-3}
0.29	−0.62	2.10×10^{-2}	2.74×10^{-3}	2.73×10^{-3}	2.91×10^{-3}	2.88×10^{-3}	3.08×10^{-2}	3.36×10^{-3}	2.51×10^{-3}	2.93×10^{-3}	2.88×10^{-3}
0.49	−0.62	2.17×10^{-2}	2.53×10^{-3}	2.65×10^{-3}	2.81×10^{-3}	2.79×10^{-3}	3.84×10^{-2}	3.45×10^{-3}	2.27×10^{-3}	2.86×10^{-3}	2.79×10^{-3}
0.69	−0.62	2.24×10^{-2}	2.32×10^{-3}	2.59×10^{-3}	2.74×10^{-3}	2.71×10^{-3}	4.72×10^{-2}	3.49×10^{-3}	2.02×10^{-3}	2.82×10^{-3}	2.70×10^{-3}
Panel B: Accuracy of the log capital decision rule											
0.29	−0.87	2.02×10^{-2}	1.10×10^{-3}	1.85×10^{-3}	2.10×10^{-3}	1.63×10^{-3}	3.16×10^{-2}	3.38×10^{-3}	1.73×10^{-3}	2.36×10^{-3}	2.41×10^{-3}
0.49	−0.87	1.85×10^{-2}	1.52×10^{-3}	1.54×10^{-3}	1.61×10^{-3}	1.66×10^{-3}	3.80×10^{-2}	5.38×10^{-3}	1.39×10^{-3}	2.05×10^{-3}	1.91×10^{-3}
0.69	−0.87	1.68×10^{-2}	1.72×10^{-3}	1.38×10^{-3}	1.30×10^{-3}	1.28×10^{-3}	4.54×10^{-2}	7.50×10^{-3}	1.41×10^{-3}	1.83×10^{-3}	1.57×10^{-3}
0.29	−0.74	2.04×10^{-2}	1.40×10^{-3}	1.75×10^{-3}	1.95×10^{-3}	1.17×10^{-3}	3.09×10^{-2}	3.21×10^{-3}	1.42×10^{-3}	2.27×10^{-3}	2.26×10^{-3}
0.49	−0.74	1.87×10^{-2}	1.71×10^{-3}	1.54×10^{-3}	1.55×10^{-3}	1.23×10^{-3}	3.68×10^{-2}	5.07×10^{-3}	1.20×10^{-3}	2.04×10^{-3}	1.85×10^{-3}
0.69	−0.74	1.70×10^{-2}	1.86×10^{-3}	1.40×10^{-3}	1.26×10^{-3}	9.05×10^{-4}	4.32×10^{-2}	7.07×10^{-3}	1.24×10^{-3}	1.85×10^{-3}	1.52×10^{-3}
0.29	−0.62	2.05×10^{-2}	1.68×10^{-3}	1.68×10^{-3}	1.81×10^{-3}	2.10×10^{-3}	3.07×10^{-2}	3.11×10^{-3}	1.15×10^{-3}	2.20×10^{-3}	2.12×10^{-3}
0.49	−0.62	1.88×10^{-2}	1.93×10^{-3}	1.50×10^{-3}	1.45×10^{-3}	1.57×10^{-3}	3.62×10^{-2}	4.92×10^{-3}	9.77×10^{-4}	2.01×10^{-3}	1.74×10^{-3}
0.69	−0.62	1.72×10^{-2}	2.06×10^{-3}	1.37×10^{-3}	1.18×10^{-3}	1.06×10^{-3}	4.19×10^{-2}	6.85×10^{-3}	1.06×10^{-3}	1.83×10^{-3}	1.43×10^{-3}

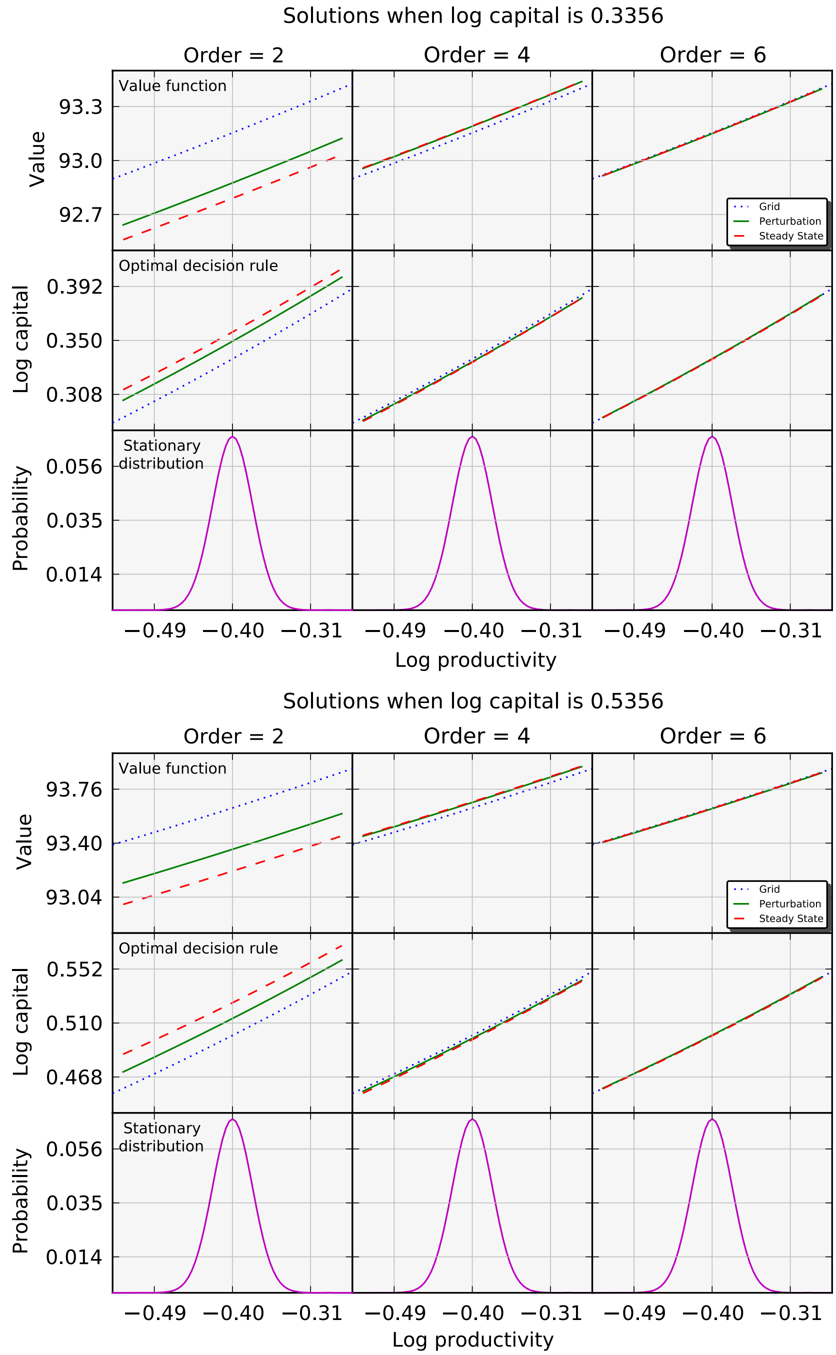

Fig. 3. *Non-homothetic robust example.* Decision rules, value functions, and the stationary distribution for the parameters given in Section 8.2.3. In this example, 0.3356 is the mean of the stationary distribution for log capital. The organization of the plots follows the structure in Fig. 1.

In summary, we find that the path method is generally reliable across a wide range of examples at low-orders. In most examples, practitioners can be confident that a fourth-order expansion (in $\sqrt{\epsilon}$) will be sufficiently accurate for all values of the state. The perturbation method sometimes, without warning, fails at lower orders. It often takes an expansion of order greater than four for the steady state method to do as well as the second order path method. A researcher who employs the steady state method needs to be careful in selecting the order because there is no guarantee that a second or fourth order expansion will be accurate. A researcher using the path method can, in most cases, use a fourth-order expansion. This is an

Table 5
Accuracy in the stochastic growth model for the non-homothetic robust example described in Section 8.2.3. The path method expands in $\sqrt{\epsilon}$ about the point $\sqrt{\epsilon}=0$. The steady state method expands in $\sqrt{\epsilon}$ about the point $\sqrt{\epsilon}=0$ and expands in the state variables about the deterministic steady state $k^*=0$, $a^*=-0.39992493$.

State		**Path method**					**Steady state method**				
		Order					Order				
k	a	2	4	6	8	10	2	4	6	8	10
Panel A: Relative accuracy of the value function											
0.14	−0.53	3.03×10^{-3}	3.83×10^{-4}	4.70×10^{-5}	5.57×10^{-6}	6.72×10^{-7}	3.34×10^{-3}	4.01×10^{-4}	4.55×10^{-5}	4.94×10^{-6}	5.84×10^{-7}
0.34	−0.53	2.99×10^{-3}	3.87×10^{-4}	4.86×10^{-5}	5.80×10^{-6}	6.61×10^{-7}	3.88×10^{-3}	4.20×10^{-4}	3.92×10^{-5}	2.94×10^{-6}	3.05×10^{-7}
0.54	−0.53	2.95×10^{-3}	3.89×10^{-4}	5.02×10^{-5}	6.10×10^{-6}	6.70×10^{-7}	4.48×10^{-3}	4.44×10^{-4}	3.22×10^{-5}	4.22×10^{-7}	1.20×10^{-7}
0.14	−0.4	3.03×10^{-3}	3.85×10^{-4}	4.75×10^{-5}	5.63×10^{-6}	6.62×10^{-7}	3.38×10^{-3}	3.93×10^{-4}	4.24×10^{-5}	4.30×10^{-6}	5.10×10^{-7}
0.34	−0.4	2.99×10^{-3}	3.88×10^{-4}	4.91×10^{-5}	5.88×10^{-6}	6.57×10^{-7}	3.91×10^{-3}	4.12×10^{-4}	3.64×10^{-5}	2.27×10^{-6}	2.03×10^{-7}
0.54	−0.4	2.95×10^{-3}	3.91×10^{-4}	5.07×10^{-5}	6.19×10^{-6}	6.69×10^{-7}	4.50×10^{-3}	4.37×10^{-4}	2.97×10^{-5}	2.82×10^{-7}	2.60×10^{-7}
0.14	−0.27	3.03×10^{-3}	3.87×10^{-4}	4.80×10^{-5}	5.69×10^{-6}	6.65×10^{-7}	3.44×10^{-3}	3.86×10^{-4}	3.92×10^{-5}	3.60×10^{-6}	4.22×10^{-7}
0.34	−0.27	2.99×10^{-3}	3.90×10^{-4}	4.96×10^{-5}	5.96×10^{-6}	6.55×10^{-7}	3.95×10^{-3}	4.06×10^{-4}	3.34×10^{-5}	1.52×10^{-6}	7.95×10^{-8}
0.54	−0.27	2.94×10^{-3}	3.93×10^{-4}	5.13×10^{-5}	6.30×10^{-6}	6.78×10^{-7}	4.53×10^{-3}	4.33×10^{-4}	2.71×10^{-5}	1.11×10^{-6}	4.31×10^{-7}
Panel B: Accuracy of the log capital decision rule											
0.14	−0.53	1.38×10^{-2}	1.99×10^{-3}	4.63×10^{-6}	6.67×10^{-5}	1.66×10^{-4}	1.73×10^{-2}	2.52×10^{-3}	3.59×10^{-5}	1.20×10^{-4}	1.85×10^{-5}
0.34	−0.53	1.34×10^{-2}	2.07×10^{-3}	1.29×10^{-4}	9.28×10^{-5}	1.16×10^{-4}	2.17×10^{-2}	2.95×10^{-3}	1.13×10^{-4}	2.50×10^{-4}	4.08×10^{-5}
0.54	−0.53	1.29×10^{-2}	2.18×10^{-3}	1.74×10^{-4}	4.56×10^{-5}	3.08×10^{-5}	2.67×10^{-2}	3.69×10^{-3}	1.90×10^{-4}	3.06×10^{-4}	1.70×10^{-5}
0.14	−0.4	1.41×10^{-2}	2.07×10^{-3}	5.24×10^{-5}	9.42×10^{-5}	1.49×10^{-5}	1.69×10^{-2}	2.18×10^{-3}	1.18×10^{-4}	1.67×10^{-4}	4.37×10^{-5}
0.34	−0.4	1.37×10^{-2}	2.22×10^{-3}	1.14×10^{-4}	5.37×10^{-5}	1.17×10^{-4}	2.09×10^{-2}	2.67×10^{-3}	2.58×10^{-4}	2.34×10^{-4}	1.24×10^{-6}
0.54	−0.4	1.32×10^{-2}	2.28×10^{-3}	2.00×10^{-4}	4.44×10^{-5}	1.11×10^{-5}	2.54×10^{-2}	3.35×10^{-3}	2.95×10^{-4}	3.33×10^{-4}	2.20×10^{-5}
0.14	−0.27	1.44×10^{-2}	2.17×10^{-3}	9.36×10^{-5}	1.12×10^{-4}	3.24×10^{-4}	1.67×10^{-2}	1.88×10^{-3}	2.15×10^{-4}	2.10×10^{-4}	5.77×10^{-5}
0.34	−0.27	1.40×10^{-2}	2.30×10^{-3}	1.61×10^{-4}	7.31×10^{-5}	3.85×10^{-5}	2.06×10^{-2}	2.38×10^{-3}	3.47×10^{-4}	2.84×10^{-4}	1.43×10^{-5}
0.54	−0.27	1.34×10^{-2}	2.40×10^{-3}	2.15×10^{-4}	2.85×10^{-5}	5.90×10^{-5}	2.47×10^{-2}	3.09×10^{-3}	4.15×10^{-4}	3.54×10^{-4}	4.35×10^{-5}

especially important feature for models with many state variables. When there are many state variables, the higher order path and perturbation methods can become intractable but the lower order methods still are feasible.[26]

9. Robustness

In this section we describe the link between risk-sensitivity and robustness, discuss a robust small noise expansion, and describe a simple method for assessing reasonable levels of robustness.

9.1. A stochastic dynamic game

In this section we describe the link between risk-sensitivity and a stochastic dynamic game. The equivalence is familiar from the literature on risk-sensitive and robust control (e.g. see Glover and Doyle, 1988; Campi and James, 1996). Also see Hansen et al. (1999) and Hansen and Sargent (2007).

Instead of introducing risk-sensitivity into the preference ordering of the decision-maker, we now allow the shock distribution to be altered from a standard normal distribution. Moreover, this distribution can depend on past history. The decision maker explores a family of such distributions subject to a penalty based on the *relative entropy* of the alternative distributions. In the recursive formulation we are led to study a two-player zero sum Markov game:

$$W^\epsilon(x) = \max_i \min_\varrho [u(x,i) + \beta E[\varrho(w) W^\epsilon(y)] + \theta E[\varrho(w) \log \varrho(w)] \tag{79}$$

where

$$y = A(x,i) + \sqrt{\epsilon} \Lambda(x) w \tag{80}$$

and where E integrates over w using the standard normal density and ϱ is an alternative shock density relative to the standard normal. We may think of ϱ as an "unmodeled" misspecification in the shock distribution. The aim of the robust control is to deduce a decision rule that is robust to this class of model perturbations. The parameter θ penalizes deviations of the shock distribution from the standard normal. By making θ arbitrarily large, we can approximate a standard control problem without a concern for robustness.

It is straightforward to show that

$$\varrho(w) = \frac{\exp\left[-\frac{\beta}{\theta} W^\epsilon(y)\right]}{\int \exp\left[-\frac{\beta}{\theta} W^\epsilon(y)\right] f(w)\, dw} \tag{81}$$

and

$$\min_\varrho \beta E[\varrho(w) W^\epsilon(y)] + \theta E[\varrho(w) \log \varrho(w)] = \mathcal{T}^\epsilon(\beta W^\epsilon)(x,i) \tag{82}$$

where $\theta = 1/\sigma$. This gives the simple link between robustness and risk sensitivity and allows our expansions to be applied directly to this class of robust control problems.

9.2. A limiting deterministic dynamic game

We now discuss a different small noise expansion for a risk-sensitive control problem. In this section, the limiting problem is a two-person deterministic dynamic game in which one player maximizes discounted utility by choice of a contingency rule for investment and the other minimizes this same objective by choice of a contingency rule for distortions to the state vector, g. The common lifetime utility for each agent can be written as

$$\sum_{t=0}^{\infty} \beta^t \left[u(x_t, i_t) + \frac{1}{2\nu} (g_t \cdot g_t) \right] \tag{83}$$

The inverse of ν is interpretable as a Lagrange multiplier on the specification-error constraint:

$$\sum_{t=0}^{\infty} \beta^t (g_t \cdot g_t) \le \lambda \tag{84}$$

where λ dictates the magnitude of the admissible specification error. λ can be chosen to match a pre-specified value of ν or vice-versa. We may think of g_t as "unmodeled" shifts in the mean of the shock process. The aim of the robust controller is to deduce a decision rule that is not sensitive to this class of model perturbations.

[26] When computing higher order terms we face a curse of dimensionality because a large number of cross-derivatives are needed.

Consider the following Bellman equation for a risk-sensitive problem:

$$V^{\epsilon}(x) = \max_i [u(x,i) + \mathcal{Q}^{\epsilon}(\beta V^{\epsilon})(x,i)]$$

where

$$\mathcal{Q}^{\epsilon}(\beta V)(x,i) \equiv -\frac{\epsilon}{v} \log E\left(\exp\left[-\frac{\beta v}{\epsilon} V(y) \right] \middle| x,i \right)$$

In this problem, the risk-sensitivity parameter is written as v/ϵ. We will drive ϵ to zero, holding v fixed. Effectively the risk-sensitive parameter $\sigma = v/\epsilon$ increases as ϵ tends toward zero. Campi and James (1996) show, under strong assumptions, that

$$V^{\epsilon}(x) = V^0(x) + \epsilon V_n(x) + o(\epsilon)$$

where $V^0(x)$ solves the following deterministic dynamic game:

$$V^0(x) = \max_i \min_g \left[u(x,i) + \beta V^0(y) + \frac{1}{2v} g \cdot g \right]$$

with

$$y = A(x,i) + \Lambda(x) g$$

and where V_n is analogous to W_n:

$$V_n(x) = \frac{1}{2} \sum_{j=1}^{\infty} \left[\beta^j \mathrm{tr}\left(\Lambda[A^{o,j-1}(x)] \Lambda[A^{o,j-1}(x)]' \frac{\partial^2 V^o(y)}{\partial y \partial y'} \bigg|_{y = A^{o,j}(x)} \right] \right)$$

This dynamic game is the Markov version of the game described in the previous paragraph. The noise correction is analogous to our earlier noise correction but uses the value function V^0 rather than W^o. The analog of the risk-sensitivity correction term W_g, described in earlier sections, does not appear.

9.3. Reasonable levels of robustness

In this section we use the solution to the stochastic dynamic game to calibrate reasonable levels of robustness. We take logs of the minimizing distortion given in Eq. (81) and explicitly indicate its dependence on ϵ via a superscript:

$$\log \varrho^{\epsilon}(w) = -\beta\sigma W^{\epsilon}(y) + \sigma \mathcal{T}^{\epsilon}(\beta W^{\epsilon})(x,i) \tag{85}$$

where

$$y = A(x,i) + \sqrt{\epsilon} \Lambda(x) w \tag{86}$$

The distribution of w in the perceived worst case that agents focus on is

$$\varrho^{\epsilon}(w) f(w) \tag{87}$$

We provide a simple method of calibrating the reasonable level of robustness that looks at the mean of $\sqrt{\epsilon} w$ under the worst case density:

$$\int \sqrt{\epsilon} w \varrho^{\epsilon}(w) f(w)\, dw \tag{88}$$

We heuristically derive a first-order expansion for this expected value. First note that

$$\log \varrho^{\epsilon}(w) = -\beta\sigma \left[W^o(z) + \sqrt{\epsilon} \frac{\partial W^o(z)'}{\partial z} \Lambda(x) w + \frac{\epsilon}{2} w' \Lambda(x)' \frac{\partial^2 W^o(z)}{\partial z\, \partial z'} \Lambda(x) w + \epsilon\sigma W_g(z) + \epsilon W_n(z) \right]$$
$$+ \sigma\beta W^o(z) + \epsilon\sigma^2 W_g(x) + \epsilon\sigma W_n(x) + o(\epsilon)$$

which implies

$$\log \varrho^{\epsilon}(w) = -\beta\sigma \left[\sqrt{\epsilon} \frac{\partial W^o(z)'}{\partial z} \Lambda(x) w + \frac{\epsilon}{2} w' \Lambda(x)' \frac{\partial^2 W^o(z)}{\partial z\, \partial z'} \Lambda(x) w \right]$$
$$- \frac{\epsilon\sigma^2\beta^2}{2} \frac{\partial W^o(z)'}{\partial z} \Lambda(x) \Lambda(x)' \frac{\partial W^o(z)}{\partial z} + \frac{\epsilon\sigma\beta}{2} \mathrm{tr}\left(\Lambda(x) \Lambda(x)' \frac{\partial^2 W^o(z)}{\partial z\, \partial z'} \right) + o(\epsilon)$$

where

$$z = A(x,i) \tag{89}$$

Since p^ϵ and $\log p^\epsilon$ have the same first order term in an ϵ expansion[27] it follows that

$$\sqrt{\epsilon} w \varrho^\epsilon(w) = \sqrt{\epsilon} w - \epsilon w \left[\beta\sigma \frac{\partial W^o(z)'}{\partial z} \Lambda(x) w \right] + o(\epsilon) \tag{90}$$

$$\epsilon = \sqrt{\epsilon} w - \epsilon\beta\sigma w w' \Lambda(x)' \frac{\partial W^o(z)}{\partial z} + o(\epsilon) \tag{91}$$

and hence

$$\int \sqrt{\epsilon} w \varrho^\epsilon(w) f(w)\, dw = -\epsilon\beta\sigma\Lambda(x)' \frac{\partial W^o(z)}{\partial z} + o(\epsilon) \tag{92}$$

In particular examples it is convenient to focus on transformations of this expected value. For example, in the stochastic growth model, the mean distortion to log productivity, namely,

$$P\Gamma_a \Lambda(x) \int \sqrt{\epsilon} w\ \varrho^\epsilon(w) f(w)\, dw = -\epsilon\beta\sigma P\Gamma_a \Lambda(x)\Lambda(x)' \frac{\partial W^o(z)}{\partial z} + o(\epsilon) \tag{93}$$

is a convenient scalar measure of the amount of distortions. To determine if the average distortion is reasonable one could compare it to beliefs about the accuracy of estimates of average productivity. Such beliefs could come from intuition, from future forecasts, or from standard errors of estimates of log productivity from historical data.[28] The end of the next section provides an example of average distortions in the stochastic growth model.

10. Optimal decisions and asset returns

In Section 8, we inspected numerical accuracy. In this section, we give an example of the effect of robustness on optimal decisions and asset prices in the stochastic growth model. In Table 6, we present second-order expansions for the unconditional means and variances for several financial and macroeconomic variables. In Panel A of Table 6, we set

$$\alpha = 0.3, \quad \delta = 0.09, \quad P = 1, \quad \Omega_a = 0.50, \quad \Omega_v = 0.02, \quad \beta = 0.90, \quad \gamma = 2.0$$

and leave ϵ and σ as free parameters. We also set $\Omega_0 = -0.19996246$ so that the steady state value of log-capital $k^* = 0$ and the steady state value of log productivity is $a^* = -0.39992493$.[29] In Panel B of Table 6 we revisit the non-homothetic example discussed in Section 8.

In Table 6, we present expansions for the unconditional means and unconditional variances of several key variables. In the absence of uncertainty, the unconditional mean of log capital (denoted mean(k)) is zero because steady state log capital is zero.[30] The unconditional mean of the log market return (mean(r_m)) and the log risk-free rate (mean(r_f)) are identical because the long run one-period log return on all assets is $-\log(\beta) \approx 0.105$ when there is no uncertainty. In the absence of uncertainty, the variances of all variables are by construction zero.

Noise in the exogenous variables puts noise into all reported variables and leads to a precautionary savings motive that increases the mean of log capital and reduces the log risk-free rate. Depending upon other parameters, risk-sensitivity may increase or decrease the unconditional expected log market return. For example, we see in Panel A of Table 6 that a second-order in ϵ (and fourth-order in $\sqrt{\epsilon}$) expansion for the unconditional expected log market return is

$$0.105 - 0.00000882\epsilon - 0.000274\epsilon\sigma + 0.00000000172\epsilon^2 + 0.0000000802\epsilon^2\sigma + 0.000000177\epsilon^2\sigma^2 \tag{94}$$

The introduction of a small amount of risk sensitivity affects the first order (in ϵ) expansion for the means of all other reported variables, but has no first-order effect on variances. Risk-sensitivity does have a small second order effect on variances.

Interestingly when $\gamma = 2$, the ratio of the decrease in the market return to the decrease in the risk-free rate is much larger for risk sensitivity than it is for noise. In Panel A of Table 6 we see that when $\sigma\epsilon$ increases by one, holding ϵ fixed, the mean log risk-free rate decreases by about 0.000928, and the mean log expected market return decreases by about 0.000274 (a ratio of approximately 3/10), whereas when ϵ increases, holding $\sigma\epsilon$ fixed, the mean log risk-free rate decreases by about 0.000101 and the expected log market return decreases by about 0.00000882 (a ratio of approximately 9/100). When $\gamma = 0.9$, risk-sensitivity increases the mean log market return (to the first order) whereas noise continues to decrease it. Both noise and risk-sensitivity decrease the log risk-free rate.

[27] They have the same first order term since $p^\epsilon(w) = 1 + \log p^\epsilon(w) + o(\epsilon)$.

[28] See Anderson et al. (2003) and Hansen and Sargent (2007) for a more formal approach (which uses detection probabilities) of determining reasonable levels of robustness.

[29] When computing the unconditional expectations in this section, we have again assumed that the initial values of k and a are finite because there is a steady state in which $\exp(k) = 0$.

[30] This is the unconditional mean of log capital as long as initial capital is guaranteed to be not zero (or equivalently as long as the initial log capital is not negative infinity).

Table 6
Second-order expansions for the unconditional expectation and variance of several variables in the stochastic growth model for two different settings of the parameters. Here k is log capital, c is log consumption, y is log output, r_m is the log market return, r_f is the log risk-free rate, and d is distortions introduced to log productivity by concerns for robustness. In both panels we set $\alpha=0.3$, $\delta=0.09$, $\epsilon=1$, $P=1$, $\Omega_a=0.50$, $\Omega_v=0.02$, $\beta=0.90$, and $\Omega_0=-0.1999624$. In Panel A, $\gamma=2.0$ and in Panel B, $\gamma=0.9$. A dash "–" indicates that the value is zero: 0.00×10^0. Recall that, by construction, the mean of log capital is zero when $\epsilon=0$.

Variable	Coefficient					
	0th order	**1st order**		**2nd order**		
	1	ϵ	$\sigma\epsilon$	ϵ^2	$\sigma\epsilon^2$	$\sigma^2\epsilon^2$
Panel A: Non-homothetic example for the parameters listed at the beginning of Section 10						
mean(k)	–	9.17×10^{-4}	5.46×10^{-3}	-2.74×10^{-7}	-3.06×10^{-6}	-1.28×10^{-6}
mean(c)	-5.44×10^{-1}	3.59×10^{-4}	1.05×10^{-3}	-4.83×10^{-8}	-6.68×10^{-7}	-1.55×10^{-6}
mean(y)	-4.00×10^{-1}	2.75×10^{-4}	1.64×10^{-3}	-8.22×10^{-8}	-9.18×10^{-7}	-3.85×10^{-7}
mean(r_m)	1.05×10^{-1}	-8.82×10^{-6}	-2.74×10^{-4}	1.72×10^{-9}	8.02×10^{-8}	1.77×10^{-7}
mean(r_f)	1.05×10^{-1}	-1.01×10^{-4}	-9.28×10^{-4}	5.14×10^{-8}	7.82×10^{-7}	2.26×10^{-6}
mean(d)	–	–	-1.30×10^{-3}	–	3.72×10^{-7}	8.18×10^{-7}
var(k)	–	1.18×10^{-3}	–	-7.00×10^{-7}	-6.66×10^{-6}	–
var(c)	–	3.39×10^{-4}	–	-1.07×10^{-7}	-1.09×10^{-6}	–
var(y)	–	7.71×10^{-4}	–	-1.15×10^{-7}	-9.73×10^{-7}	–
var(r_m)	–	1.14×10^{-4}	–	-5.64×10^{-8}	-5.54×10^{-7}	–
var(r_f)	–	1.33×10^{-5}	–	-2.08×10^{-8}	-2.15×10^{-7}	–
Panel B: Non-homothetic example for the parameters in Section 8.2.3						
mean(k)	–	3.89×10^{-4}	1.01×10^{-3}	-2.93×10^{-8}	-1.54×10^{-7}	-4.47×10^{-7}
mean(c)	-5.44×10^{-1}	2.52×10^{-4}	1.94×10^{-4}	-3.57×10^{-9}	3.11×10^{-9}	-1.30×10^{-7}
mean(y)	-4.00×10^{-1}	1.17×10^{-4}	3.03×10^{-4}	-8.78×10^{-9}	-4.63×10^{-8}	-1.34×10^{-7}
mean(r_m)	1.05×10^{-1}	-8.82×10^{-8}	1.50×10^{-5}	3.37×10^{-12}	-9.13×10^{-11}	-7.40×10^{-9}
mean(r_f)	1.05×10^{-1}	-2.59×10^{-5}	-2.58×10^{-4}	5.13×10^{-9}	6.42×10^{-8}	2.60×10^{-7}
mean(d)	–	–	-7.16×10^{-4}	–	9.53×10^{-9}	2.92×10^{-7}
var(k)	–	6.45×10^{-4}	–	-8.20×10^{-8}	-2.93×10^{-7}	–
var(c)	–	3.61×10^{-4}	–	-1.92×10^{-8}	-1.60×10^{-7}	–
var(y)	–	7.08×10^{-4}	–	-2.35×10^{-8}	-6.85×10^{-8}	–
var(r_m)	–	6.39×10^{-5}	–	-8.42×10^{-9}	-5.16×10^{-8}	–
var(r_f)	–	5.82×10^{-6}	–	-2.72×10^{-9}	-1.70×10^{-8}	–

What level of risk sensitivity is plausible? The mean distortion to log productivity provides one rough guide. We define d to be mean distortion to log productivity

$$d = P\Gamma_a\Lambda(x)\int\sqrt{\epsilon}w\varrho^{\epsilon}(w)f(w)\,dw = -\epsilon\beta\sigma P\Gamma_a\Lambda(x)\Lambda(x)'\frac{\partial W^o(z)}{\partial z}+o(\epsilon)$$

which was discussed in Section 9. Panel A of Table 6 shows that if $\epsilon=1$ and $\sigma=10$ then the unconditional mean distortion to log productivity is approximately -0.0130, which entails that agents are worried that output may be about 1.3% less than their best model based forecast. If we increase σ to 100 then the unconditional mean distortion is approximately -0.130 which entails that agents are worried that output may be about 13% less than their best forecast.[31] As discussed at the end of Section 9, there are several approaches available to help guide researchers choosing a reasonable value of σ. We believe most researchers would view a σ that leads to a mean distortion of 1.3% as small whereas a σ which leads to a mean distortion of 13% may be large or small depending upon one's faith in their model.

Panel B tells a similar story. Since the mean distortion increases at a slower rate when σ increases, slightly higher levels of σ are plausible when $\gamma=0.9$.

11. Kreps–Porteus CES specification

The expansions presented in this paper can be extended to other forms of recursive utility, including the Kreps–Porteus CES specification. Kreps and Porteus suggest a relaxation of the reduction of compound intertemporal lotteries. A commonly used CES version of their specification updates utility using:

$$U_t^* = [(1-\beta)[u(x_t,i_t)]^{1-\rho}+\beta[\mathcal{R}_t^*(U_{t+1}^*)]^{1-\rho}]^{1/(1-\rho)}$$

[31] The numbers reported in this paragraph are for a first-order expansion. We see from Table 6 that if a second order expansion is used the unconditional mean distortion when $\sigma=10$ decreases very slightly to -0.0129 and the unconditional mean distortion when $\sigma=100$ decreases slightly to -0.122.

where U_t^* is the continuation value of a plan and $\rho > 0$. The operator $\mathcal{R}_t^*$ is given by

$$\mathcal{R}_t^*(V_{t+1}) = (E[(V_{t+1})^{1-\gamma}|\mathcal{F}_t])^{1/(1-\gamma)}$$

Taking logarithms of the continuation value ($U_t = \log U_t^*$) gives

$$U_t = \frac{1}{1-\rho}\log((1-\beta)[u(x_t,i_t)]^{1-\rho} + \beta\exp[(1-\rho)\mathcal{R}_t(U_{t+1})])$$

where $\mathcal{R}_t$ is the risk-sensitivity operator with $\sigma = \gamma - 1$.

This leads us to compute the derivative of the function

$$U = \frac{1}{1-\rho}\log((1-\beta)[u^{1-\rho} + \beta\exp[(1-\rho)R])$$

with respect to R:

$$\frac{dU}{dR} = \beta\frac{\exp[(1-\rho)R]}{\exp[(1-\rho)U]}$$

This ratio is then used recursively in the derivative calculations and combined with our current expansion of the risk sensitivity operator. Along a deterministic path R is replaced by the next period continuation value.

12. Conclusion

This paper derives expansions for the decision rules, value functions, long-run means, and asset returns in economies with risk-sensitive preferences. Although steady-state approximations are very accurate near the steady state, they can provide inaccurate approximations for value functions away from a steady state. We show that path approximation methods are generally more accurate than steady state approximations and usually provide good approximations to both decision rules and value functions globally. For most applications, the second-order in ϵ path approximations are accurate enough and it is not necessary to compute higher order expansions. We provide examples which illustrate the effect of risk sensitivity on capital accumulation and asset prices.

Although Campi and James (1996) have provided formal proofs for risk-sensitive small noise expansions in finite horizon economies, we do not provide formal proofs for infinite horizon economies. As far as we are aware, there have not been any papers formally justifying the expansion in infinite horizon economies and there are many subtle issues that have to be addressed in order to provide a general proof. In the absence of risk sensitivity, Williams (2004) has provided a formal proof in the context of a stochastic growth model. In the presence of risk sensitivity, the limit results of Albertini et al. (2001) may be useful in formulating formal proofs.

Acknowledgments

We thank Ken Judd, Jinill Kim, Chao Wei, and an anonymous referee for helpful comments. We also thank seminar participants at The University of North Carolina, North Carolina State, and The University of Virginia as well as participants at the Society of Computational Economics Conference on Computing in Economics and Finance and 2007 Winter Meetings of the Econometric Society for helpful comments.

Appendix A. Expansions for asset prices and returns

In this appendix, we heuristically derive the expansions for log prices and returns given in Section 5. Taking logs of the pricing formula given in Eq. (54) yields

$$\log P_i^\epsilon(x) = \log\beta + \log\int\exp[-\sigma\beta W^\epsilon(y) + q^\epsilon(y) + b_i^\epsilon(y)]f(w)\,dw - \log\int\exp[-\sigma\beta W^\epsilon(y) + q^\epsilon(x)]f(w)\,dw \tag{A.1}$$

where $y = A[x, i^\epsilon(x)] + \sqrt{\epsilon}\Lambda(x)w$.

Using a result similar to Eq. (17), a first-order correction for the first logarithmic term in Eq. (A.1) is ϵ times:

$$\frac{1}{2}\left[-\beta\sigma\frac{\partial W^o(z)}{\partial z} + \frac{\partial q^o(z)}{\partial z} + \frac{\partial b_i^o(z)}{\partial z}\right]'\Lambda(x)\Lambda(x)'\left[-\beta\sigma\frac{\partial W^o(z)}{\partial z} + \frac{\partial q^o(z)}{\partial z} + \frac{\partial b_i^o(z)}{\partial z}\right]$$
$$+\frac{1}{2}\mathrm{tr}\left[\Lambda(x)\Lambda(x)'\left(-\beta\sigma\frac{\partial^2 W^o(z)}{\partial z\,\partial z'} + \frac{\partial^2 q^o(z)}{\partial z\,\partial z'} + \frac{\partial^2 b_i^o(z)}{\partial z\,\partial z'}\right)\right] - \beta\sigma h(z) + h_q(z) + h_{bi}(z) \tag{A.2}$$

and a first-order correction for the second logarithmic term is $(-\epsilon)$ times:

$$\frac{\beta^2\sigma^2}{2}\left[\frac{\partial W^o(z)}{\partial z}\right]'\Lambda(x)\Lambda(x)'\left[\frac{\partial W^o(z)}{\partial z}\right] - \frac{\beta\sigma}{2}\mathrm{tr}\left[\Lambda(x)\Lambda(x)'\frac{\partial^2 W^o(z)}{\partial z\,\partial z'}\right] - \beta\sigma h(z) + h_q(x) \tag{A.3}$$

where $z = A[x, i^{\epsilon}(x)]$. Subtracting the expression in (A.3) from the expression in (A.2) yields the coefficient on ϵ in the log price expansion presented in Eq. (56). Notice that the term $\beta\sigma h(z)$ cancels so that the first-order correction to the value function does not affect the first order expansion for asset prices.

Using a result similar to Eq. (17), a first-order correction for the expected log payoff is ϵ times

$$\frac{1}{2}\mathrm{tr}\left[\Lambda(x)\Lambda(x)'\left(\frac{\partial^2 b_i^o(z)}{\partial z\,\partial z'}\right)\right] + h_{bi}(z) \tag{A.4}$$

Subtracting the coefficient on ϵ in the log price expansion from the expression in (A.4) yields the coefficient on ϵ in the expected log return expansion presented in Eq. (59).

References

Albertini, F., Pra, P.D., Prior, C., 2001. Small parameter limit for ergodic, discrete-time, partially observed, risk-sensitive control problems. Mathematics of Control, Signals, and Systems 14, 1–28.
Anderson, E.W., Hansen, L.P., Sargent, T.J., 2003. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. Journal of the European Economic Association 1, 68–123.
Campi, M.C., James, M.R., 1996. Nonlinear discrete-time risk-sensitive optimal control. International Journal of Robust and Nonlinear Control 6, 1–19.
Chabi-Yo, F., Ghysels, E., Renault, E., 2006a. Disentangling the Effects of Heterogeneous Beliefs and Preferences on Asset Prices. Working Paper.
Chabi-Yo, F., Leisen, D.P., Renault, E., 2006b. Implications of Asymmetry Risk for Portfolio Analysis and Asset Pricing. Working Paper.
Chen, B., Zadrozny, P.A., 2003. Higher-moments in perturbation solution of the linear–quadratic exponential Gaussian optimal control problem. Computational Economics 21, 45–64.
Epstein, L.G., Zin, S.E., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica 57, 937–969.
Fair, R.C., 2003. Optimal control and stochastic simulation of large nonlinear models with rational expectations. Computational Economics 21, 245–256.
Fair, R.C., Taylor, J.B., 1983. Solution and maximum likelihood estimation of dynamic nonlinear rational expectation models. Econometrica 51, 1169–1185.
Fernandez-Villaverde, J., Rubio-Ramirez, J.F., 2006. Solving DSGE models with perturbation methods and a change of variables. Journal of Economic Dynamics and Control 30, 2509–2531.
Fleming, W., Souganidis, P., 1986. Asymptotic series and the method of vanishing viscosity. Indiana University Mathematics Journal 35, 425–447.
Fleming, W., Yang, J., 1994. Numerical methods for infinite horizon risk sensitive stochastic control. In: Proceedings of the 33rd Conference on Decision and Control, pp. 1945–1949.
Glover, K., Doyle, J., 1988. State-space formulae for all stabilizing controllers that satisfy an H^{∞}-norm bound and relations to risk sensitivity. Systems Control Letters 11, 17–172.
Hansen, L.P., Heaton, J., Lee, J., Roussanov, N., 2007. Intertemporal substitution and risk aversion. Part 1 of Handbook of Econometrics, vol. 6. Elsevier, pp. 3967–4056.
Hansen, L.P., Heaton, J.C., Li, N., 2008. Consumption strikes back? Measuring long-run risk. Journal of Political Economy 116, 260–302.
Hansen, L.P., Sargent, T.J., 1995. Discounted linear exponential quadratic Gaussian control. IEEE Transactions on Automatic Control 40, 968–971.
Hansen, L.P., Sargent, T.J., 2007. Robustness. Princeton University Press, Princeton, NJ.
Hansen, L.P., Sargent, T.J., Tallarini Jr., T.D., 1999. Robust permanent income and pricing. Review of Economic Studies 66, 873–907.
Jacobson, D., 1973. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. IEEE Transactions on Automatic Control AC-18, 124–131.
James, M.R., 1992. Asymptotic analysis of nonlinear stochastic risk-sensitive control and differential games. Mathematics of Control, Signals, and Systems 5, 401–417.
James, M.R., Baras, J.S., Elliott, R.J., 1994. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. IEEE Transactions on Automatic Control 39, 780–792.
Judd, K.L., 1996. Approximation, perturbation, and projection methods in economic analysis. In: Amman, H.M., Kendrick, D., Rust, J. (Eds.), Handbook of Computational Economics, Elsevier Science.
Judd, K.L., 1998. Numerical Methods in Economics. The MIT press.
Judd, K.L., Guu, S.M., 1993. Perturbation solution methods for economic growth models. In: Varian, H.R. (Ed.), Economic and Financial Modeling with Mathematica, TELOS.
Judd, K.L., Guu, S.M., 1997. Asymptotic methods for aggregate growth models. Journal of Economic Dynamics and Control 21, 1025–1042.
Kogan, L., Uppal, R., 2003. Risk Aversion and Dynamic Optimal Portfolio Policies. Working Paper.
Ljungqvist, L., Sargent, T.J., 2004. Recursive Macroeconomic Theory, second ed. The MIT press.
Maenhout, P.J., 2004. Robust portfolio rules and asset pricing. Review of Financial Studies 17, 951–983.
Pratt, J.W., 1964. Risk aversion in the small and in the large. Econometrica 32, 122–136.
Runolfsson, T., 1994. Optimal control of a stochastic system with an exponential-of-integral performance criterion. Systems and Control Letters 22, 451–456.
Santos, M., 1994. Smooth dynamics and computation in models of economic growth. Journal of Economic Dynamics and Control 18, 879–895.
Tallarini Jr., T.D., 2000. Risk-sensitive real business cycles. Journal of Monetary Economics 45, 507–532.
Tauchen, G., Hussey, R., 1991. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. Econometrica 59, 371–396.
Trojani, F., Vanini, P., 2002. Perturbative Solutions of Hamilton Jacobi Bellman Equations in Robust Decision Making. Working Paper.
Whittle, P., 1990. Risk-Sensitive Optimal Control. Wiley, New York.
Williams, N., 2004. Small noise asymptotics for a stochastic growth model. Journal of Economic Theory 119, 271–298.