Existence, Uniqueness, and Computational Theory for Time Consistent Equilibria: A Hyperbolic Discounting Example

Kenneth L. Judd
Hoover Institution
Stanford, CA 94305
judd@hoover.stanford.edu
National Bureau of Economic Research

April, 2003
This revision: November, 2004*

Abstract

We present an asymptotically valid analysis of a simple optimal growth model with hyperbolic discounting. We use the implicit function theorem for Banach spaces to show that for small deviations from exponential discounting there is a unique solution near the exponential discounting solution in the Banach space of consumption functions with bounded derivatives. The proof is constructive and produces both an infinite series characterization and a perturbation method for solving these problems. The solution uses only the contraction properties of the exponential discounting case, suggesting that the techniques can be used for a wide variety of time consistency problems. We also compare the computational procedure implied by our asymptotic analysis to previous methods. Finally, we present a simple tax policy example that illustrates how to apply the method more generally.

1 Introduction

Many dynamic decision problems lead to problems of time inconsistency. These include government policy problems as well as the sale of durable goods by a monopolist and consumption decisions under hyperbolic discounting. In general, such problems can be treated as dynamic games. This

*I would like to thank Mordecai Kurz, conference participants at the 2003 meeting of the Society for Computational Economics, and seminar participants at the University of Wisconsin, Duke University, Stanford University, and the Board of Governors of the Federal Reserve for their comments, and Paul Klein, Per Krusell, and Tony Smith for useful discussions on the KKS procedure.

paper uses the hyperbolic discounting model in Krusell, Kuruscu, and Smith (2002) (KKS) to address issues of existence and uniqueness of time consistent equilibria in general. The analysis is constructive and leads directly to a perturbation method of solution. While we analyze only a hyperbolic discounting example, the analysis uses no special properties of the example and uses an abstract, dynamic game theoretic formulation of the problem. This indicates that the solution technique is applicable in a variety of dynamic strategic contexts.

Multiplicity of equilibria is a common problem in dynamic games. One strategy has been to focus on equilibria with continuous strategies. This has proven particularly powerful in at least one problem of time consistency. Stokey (1981) showed that there exists a continuum of time consistent solutions to the problem of the durable good monopolist, but showed that the Coase solution is the unique solution with continuous expectations. More generally, many have explicitly use continuity as a selection criterion. Furthermore, many others have implicitly made continuity restrictions. In particular, numerical solutions to time consistency problems and feedback Nash equilibria of dynamic games typically examine only continuous strategies; see, for example, Wright and Williams (1982), Kotlikoff et al. (1988), Rui and Miranda (1996), Ha and Sibert (1997), Krusell et al. (1997), Vedenov and Miranda (2001), Klein et al. (2002), and Doraszelski (2003). We will use the hyperbolic discounting model as a laboratory for examining theoretical and computational properties of this selection criterion.

The multiplicity problem arises often in dynamic games. For example, Fudenberg and Tirole (1983) showed that there is a continuum of feedback equilibria in a simple duopoly game of investment. A key feature of many of the equilibria in Fudenberg-Tirole (1983) is that each firm's strategy is discontinuous with respect to the state, with each firm discontinuously increasing investment if its opponent violates a tacitly agreed limit on its capital stock. Krusell-Smith (2003) use similar arguments to show that the hyperbolic discounting model of growth often has a continuum of solutions with discontinuous consumption functions. In both problems, the steady state stock of capital is often indeterminate. The focus on equilibria with continuous consumption functions will rule out discontinuous solutions but there is no reason to believe that there are continuous equilibria nor that there is exactly one. This paper shows that continuity will select a unique differentiable equilibrium for small deviations from exponential discounting, and that this unique equilibrium is as differentiable as the underlying tastes and technology. Our analysis has two implications for computational approaches to dynamic equilibria. First, our results justify the typical numerical focus on continuous (and differentiable) solutions for at least an open set of problems. Second, the constructive nature of our analysis itself suggests three computational approaches.

The basic approach of this paper is familiar. We begin with a particular case, exponential discounting, where we know there exists a unique solution. We then examine how the solution changes as we change a parameter representing the deviation from exponential discounting. This approach is the same as used in comparative statics, comparative dynamics, and determinacy theory for general equilibrium (see Debreu, 1976, and Shannon, 1999). It is also the approach to dynamic games taken in Judd (2003). Differentiability plays a key role in those analyses, and will be equally important here. However, we have an infinite-dimensional problem since we must compute savings functions. The key tools in this paper come from calculus in Banach spaces. The major mathematical challenges involve finding an appropriate topology for the analysis and then checking the conditions for the implicit function theorem. Along the way we must solve an unfamiliar functional equation of a differential composition character. This approach is similar to the analysis in Judd (2003) except here the choice of Banach space is less obvious since our operators are not as simple as the integral operators in Judd (2003). However, the tools are very general. The key fact is that the problem with exponential discounting reduces to analysis of a contraction map. We provide a condition which implies that a modified contraction property is inherited by problems with nearly exponential discounting. Since the key elements of the analysis are common features of dynamic economic problems, we suspect that the ideas are directly applicable to a wide range of time consistency problems.

Time consistency problems present special numerical challenges, particularly in the context of hyperbolic discounting. For example, many of the problems in Laibson and Harris (2002) appear to have only discontinuous solutions. We will examine the standard solution methods that have been used by agricultural economists and public finance economists to find time consistent equilibria of policy games, as well as a recent procedure proposed in KKS. We will show that these methods, some of which are essentially projection methods as defined in Judd (1992), have difficulties that point either to multiplicity of true solutions or the presence of extraneous solutions to the numerical approximations.

These numerical problems indicate that computational approaches to solving dynamic strategic problems need to be very careful. We use our asymptotic theory to present a perturbation method for solving the hyperbolic discounting problem that addresses both the existence and uniqueness issues. This procedure is limited in its applicability, but is promising since it is based on solid mathematical foundations. Furthermore, we show that it can be used to solve a wide range of hyperbolic discounting problems, and, presumably, many other dynamic strategic problems.

2 A Model of Growth with Hyperbolic Discounting

We will examine an optimal growth problem where the planner discounts future utility in a hyperbolic fashion[1]. Suppose that c_t is consumption in period t. The planner at time $t = 0$ values the

[1] See Krusell, Kuruscu, and Smith (2002) for a more complete description of this model, and Harris and Laibson (2003) for a more general discussion of hyperbolic discounting problems.

future stream of utility according to the infinite sum

$$U_0 = u(c_0) + \beta(\delta u(c_1) + \delta^2 u(c_2) + \delta^3 u(c_3) + \cdots)$$

whereas the agent at $t = 1$ values future utility according to the sum

$$U_1 = u(c_1) + \beta(\delta u(c_2) + \delta^2 u(c_3) + \cdots).$$

where $\delta < 1$ and β is usually taken to be less than one to represent a myopia on the part of the decisionmaker. In general, the planner at time t discounts utility between $t+1$ and $t+s+1$ at rate δ^s but discounts utility between time t and $t+s$ at rate $\beta\delta^s$. If $\beta = 1$ we have the standard discounted utility function.

2.1 Smooth Feedback Nash Equilibrium

We will examine only feedback Nash equilibria[2]; that is, we assume that the time t planner believes that future savings follow the process

$$k_{t+1} = h(k_t) \tag{1}$$

for some function h of the current capital stock[3]. We also define the consumption function, $C(k) \equiv f(k) - h(k)$. By the feedback Nash assumption, we need only consider the problem of the time $t = 0$ personality. At time $t = 0$, the time $t = 0$ self chooses current consumption to solve

$$h(k) \equiv \arg\max_x u(f(k) - x) + \beta\delta V(x) \tag{2}$$

where $V(k)$ is the value to the time $t = 0$ self of the utility flow of consumption from period $t = 1$ and after if the capital stock at time $t = 1$ is k. Under the assumption that future selves will follow (1), the value function $V(k)$ is the solution to the equation

$$V(k) = u(f(k) - h(k)) + \delta V(h(k)) \tag{3}$$

which, for any $h(k)$, has a unique solution since the right-hand side of (3) is a contraction operator on value functions V. Furthermore, the solution to (2) satisfies the first-order condition

$$u'(c) = \beta\delta V'(f(k) - c).$$

However, in a feedback Nash equilibrium, when capital is k gross savings must equal $h(k) = f(k) - c$. We use these equations to define our concept of equilibrium.

[2] The concept of feedback Nash equilibria from the dynamic games literature (see, for example, Basar and Olsder, 1982) is equivalent to the term "Markov perfect equilibrium" later favored by some economists.

[3] We also assume that the feedback rule is the same at all times t. We conjecture that this remains true even if we allow feedback rules of the form $h(k, t)$ since this is true of the $\beta = 1$ case. The proof of this would require us to formulate a similar function space for nonautonomous functions, and would take us away from the essential points of our analysis, so we leave it for future work.

Definition 1 *A continuously differentiable feedback Nash equilibrium will be a pair of functions* $V(k)$ *(*C^2*) and* $h(k)$ *(*C^1*) that satisfy the value function equation*

$$V(k) = u(f(k) - h(k)) + \delta V(h(k)), \tag{4}$$

the first-order condition

$$u'(f(k) - h(k)) = \beta\delta V'(h(k)), \tag{5}$$

and the global optimality condition

$$h(k) \equiv \arg\max_x u(f(k) - x) + \beta\delta V(x) \tag{6}$$

Existence and uniqueness problems arise in this model as they typically do in dynamic games, even when we restrict ourselves to feedback Nash equilibria. Krusell and Smith (2003) prove that there is a continuum of distinct solutions to the equilibrium pair (2, 4). Our definition of a C^1 feedback Nash equilibrium precisely formulates our equilibrium selection criterion by focussing on smooth value functions and savings functions. This is the assumption explicitly made and defended in Stokey (1981) and implicitly made (generally without discussion) by KKS and many other analyses of time consistent equilibria and dynamic games in general. Stokey (1981) proves that there is a unique continuous solution in its model, but KKS provides no proof of either existence of a continuous solution nor a uniqueness result.

Harris and Laibson (2001) examine a similar savings problem with hyperbolic discounting and prove existence of smooth solutions for small amounts of hyperbolic discounting. However, there are substantial differences between their analysis and the analysis presented below. First, their existence result assumes income uncertainty. This uncertainty is critical to smoothing out their problem and avoiding mathematical difficulties. Since deterministic problems are of substantial interest in general in time consistency problems, we will proceed with developing the tools necessary to analyze this deterministic problem. Also, they prove only that the set of solutions is a semicontinuous correspondence in hyperbolic discounting whereas we construct a smooth manifold of solutions, one for each value of hyperbolic discounting. The techniques used are also different with Harris and Laibson using techniques from the theory of functions of bounded variation whereas we use calculus methods in Banach spaces.

2.2 Operator Expression of Equilibrium

We will follow KKS and reduce the analysis to a single equation in $h(k)$. This will simplify the exposition but will not affect any substantive result since we could proceed in the same manner with the pair of equations (4,5). Differentiating (4) with respect to k implies

$$V'(k) = u'(f(k) - h(k))\left(f'(k) - h'(k)\right) + \delta V'(h(k))\, h'(k) \tag{7}$$

which also, by substituting $h(k)$ for k, implies

$$V'(h(k)) = u'(C(h(k)))\left(f'(h(k)) - h'(h(k))\right) + \delta V'(h(h(k)))\, h'(h(k)) \tag{8}$$

where $C(h(k)) = f(h(k)) - h(h(k))$. The first-order condition (5) when capital stock is $h(k)$ implies

$$u'(C(h(k))) = \beta\delta V'(h(h(k))). \tag{9}$$

Combining (7) and (8), using (9) to eliminate $V'(h(h(k)))$, implies the single equation[4]

$$u'(f(k) - h(k)) = \beta\delta u'(f(h(k)) - h(h(k)))\left(f'(h(k)) + \left(\frac{1}{\beta} - 1\right) h'(h(k))\right) \tag{10}$$

KKS call equation (10) the Generalized Euler Equation since it eliminates the value function[5]. Note that if $\beta = 1$, the case of exponential discounting, (10) does reduce to the usual Euler equation. I shall work with the GEE. It is a simplification of the equilibrium conditions for the dynamic game to a single equation in a single unknown function and helps keep our exposition simple. However, one could proceed with our analysis with the value function formulation; therefore, the methods below likely apply even when there is no GEE formulation.

We will rearrange the terms and express equilibrium in terms of a general function $G : \mathbb{R}^5 \rightarrow \mathbb{R}$

$$\begin{aligned} 0 &= u'(f(k) - h(k)) - \left(\frac{\delta}{1+\varepsilon}\right) u'(f(h(k)) - h(h(k))) \qquad (11) \\ &\quad \times \left(f'(h(k)) + \varepsilon h'(h(k))\right) \\ &\equiv G\left(k, h(k), h(h(k)), \varepsilon h'(h(k)), \varepsilon\right) \end{aligned}$$

where $\varepsilon = \beta^{-1} - 1$ represents the deviation from exponential discounting. When $\varepsilon = 0$ we have ordinary exponential discounting at rate δ, and the unique solution is the conventional optimal consumption function that solves

$$\begin{aligned} 0 &= u'(C(k)) - \beta\delta u'(C(h(k)))\, f'(h(k)) \\ C(k) &= f(k) - h(k) \end{aligned}$$

[4] A more complete derivation is

$$\begin{aligned} u'(f(k) - h(k)) &= \delta\beta V'(h(k)) \\ &= \delta\beta\left(u'(C(h(k)))\left(f'(h(k)) - h'(h(k))\right) + \delta V'(h(h(k)))\, h'(h(k))\right) \\ &= \delta\beta\left(u'(C(h(k)))\left(f'(h(k)) - h'(h(k))\right) + \frac{1}{\beta} u'(C(h(k)))\, h'(h(k))\right) \\ &= \beta\delta u'(C(h(k)))\left(f'(h(k)) + \left(\frac{1}{\beta} - 1\right) h'(h(k))\right) \end{aligned}$$

[5] This simplification is a special case of a more general fact. Rincon-Zapatero et al. (1998) have shown that a similar simplification is possible in a broad range of differential games. This indicates that the techniques in this paper are broadly applicable.

which has a unique bounded stable solution. Let $\bar{h}(k)$ denote this solution and let k_* denote the steady state of the $\varepsilon = 0$ solution; that is, $\bar{h}(k_*) = k_*$.

2.3 The Formal Singular Perturbation

The parameter ε appears in two distinct places and our form for G recognizes this. The parameter ε appears as part of the discounting term $\beta\delta$ and in the $\varepsilon h'(h(k))$ term. The appearance in the $\beta\delta$ term is unremarkable, but the other occurrence is quite important. Note that when $\varepsilon = 0$ the term of h of highest complexity is the $h(h(k))$ term but when $\varepsilon \neq 0$ the highest order term is $h'(h(k))$. The fact that a change in ε changes the fundamental nature of the operator equation implies that we have a (formal) singular perturbation as ε moves away from $\varepsilon = 0$. This should cause immediate alarm in any analysis of this problem. We shall see below that it demands careful attention and that ad hoc methods that ignore this singular perturbation can lead to incoherent "results."

3 Mathematical Preliminaries

We will need to use nonlinear functional analysis to analyze equilibrium in the hyperbolic discounting problem. This section will review the basic definitions and theorems we will use[6].

We will work with a Banach spaces of functions $h : I \to \mathbb{R}$ where $I = (a, b)$, $0 < a < k_* < b < \infty$, is an open interval. We need to specify a space of such functions and a norm appropriate for our purposes. We want to focus on continuous solutions for h but the presence of $h'(h(k))$ in (10) implies that we also require differentiability. This implies that conventional spaces and norms such as L_1, L_2, or L_∞ are not appropriate for this problem. One approach for dealing with the presence of h' in applied mathematics is to work in a Sobolev space where the notion of a generalized (or weak) derivative is used. We will not take that approach since we do not want to burden this paper with generalized derivatives, and we probably would not get our strong uniqueness results since the step function solutions found in Krusell-Smith (2003) lie in the standard Sobolev spaces.

We use the standard generalization of the supremum norm. Let $C^m(U, V)$ denote the space of C^m functions f with domain $U \subset \mathbb{R}$ and range in $V \subset \mathbb{R}$. On this space, the norm $\|\cdot\|_m$ is defined by

$$\|f\|_m = \max_{0 \leq i \leq m} \sup_{x \in U} \left\| D^i f(x) \right\|. \tag{12}$$

$C^m(U, V)$ is a Banach space with the norm $\|f\|_m$ but is not a Hilbert space. A Hilbert space approach would replace the supremum norm in (12) with an inner product in an L^p space, and

[6] We take many of the critical definitions and theorems from Abraham et al. (1983). See Abraham et al. (1983), Joshi and Bose (1985), or any one of many monographs on nonlinear functional analysis for a more thorough discussion of the relevant theorems from calculus on Banach spaces.

would lead to a Sobolev space. Since this Sobolev space is neither needed nor desired here, we use the Banach space defined by $\|f\|_m$.

We need to extend calculus to mappings of $C^m(U,V)$ into $C^{m-1}(U,\mathbb{R})$. The key fact is that all of the basic results from ordinary calculus generalizes to these mappings. The reader can skip the rest of this section if he is not worried about the details. The first key concept is tangency.

Definition 2 *Suppose f and g are functions*

$$f,g : U \to F$$

where U is an open subset of E, V is an open subset of F, where E and F are Banach spaces each with a norm $\|\cdot\|$. The functions f and g are tangent at $x_0 \in U$ if

$$\lim_{x\to x_0} \frac{\|f(x)-g(x)\|}{\|x-x_0\|} = 0.$$

This notion of tangency implies an important uniqueness property.

Definition 3 *Let $L(E,F)$ denote the space of linear maps from E to F with the norm topology. Also, the spaces of linear maps $L^m(E,F)$ are defined inductively by the identities $L^m(E,F) = L(E,L^{m-1}(E,F))$, $m = 2,3,....$*

The following fact allows us to define differentiation. See Abraham et al. for a proof.

Lemma 4 *For $f : U \subset E \to F$ and $x_0 \in U$ there is at most one linear map $L \in L(E,F)$ such that the map $g(x) = f(x_0) + L(x-x_0)$ is tangent to f at x_0.*

We now use tangents to define differentiation, first and higher orders.

Definition 5 *If there is an $L \in L(E,F)$ such that $f(x_0) + L(x-x_0)$ is tangent to f at x_0, then we say f is differentiable (a.k.a., Fréchet differentiable) at x_0, and define the derivative of f at x_0 to be $Df(x_0) = L$.*

Definition 6 *If f is differentiable at each $x_0 \in U$, then the derivative of f is a map from U to the space of linear maps*

$$\begin{aligned} Df \quad &: \quad U \to L(E,F) \\ x \quad &\longmapsto \quad Df(x) \end{aligned}$$

Definition 7 *If $Df : U \to L(E,F)$ is a continuous map then f is $C^1(U,F)$ (e.g., continuously differentiable). As long as the derivatives exist, we define higher derivatives by the inductive formula*

$$D^m f = D(D^{m-1}f) : U \subset E \to L^m(E,F)$$

If $D^m f$ exists and is norm continuous we say f is $C^m(U,F)$.

The directional derivative is a related concept.

Definition 8 *Let $f : U \subset E \to F$ where E and F are Banach spaces, and let $x \in U$. We say that f has a derivative in the direction $e \in E$ at x if*

$$\lim_{t\to 0} \frac{d}{dt} f(x+te)$$

exists, in which case it is called the directional derivative.

Sometimes a function may have a directional derivative for all directions, (that is, it is *Gâteaux differentiable)* but may not be differentiable. The key fact is that the directional derivative is the intuitive way to compute derivatives of differentiable functions.

Lemma 9 *If f is differentiable at x, then the directional derivatives of f exist at x and are given by*

$$\lim_{t\to 0} \frac{d}{dt} f\left(x+te\right) = Df\left(x\right) \cdot e.$$

In general, we will just use the Gâteaux approach to compute our derivatives but our theorems will guarantee that the operators are Frechet differentiable.

The GEE contains a $h'\left(h\left(k\right)\right)$ term. Its presence rules out the usefulness of an L^p space since L^p spaces contain nondiffentiable functions. The main advantage of $C^m\left(U,V\right)$ is that the derivative map is differentiable. We shall express the differentiability result for the case of $C^m(I,E)$, $I, E \subset \mathbb{R}$, but it holds for more general spaces of differentiable functions.

Lemma 10 ***(Differentiability of the derivative map)*** *The map*

$$\begin{array}{rcl} D\left(f\right) & : & C^m\left(U,V\right) \to C^{m-1}\left(U,\mathbb{R}\right) \\ f & \longmapsto & f' \end{array}$$

is C^{m-1}.

One novel feature of the operator we will encounter is the presence of the evaluation map. The evaluation map is the map

$$\text{ev}: C^m\left(U,V\right) \times U \to V$$

defined by

$$\text{ev}(f,t) = f(t).$$

Lemma 11 ***(Evaluation Map Lemma).*** *The evaluation map is C^m and the derivatives are defined by the chain rule and equal*

$$\begin{aligned} & D^k ev(f,t) \cdot ((g_1, s_1), ..., (g_k, s_k)) \\ = \; & D^k f(t) \cdot (s_1, ..., s_k) + \sum_{i=1}^{k} D^{k-1} g_i(t) \cdot (s_1, ..., s_{i-1}, s_{i+1}, ..., s_k) \end{aligned}$$

for

$$(g_i, s_i) \in C^m(I, \mathbb{R}) \times \mathbb{R}, \qquad i = 1, ..., k.$$

We will use the following lemma on compositions. It is proved by applying the converse to the Taylor theorem (see Abraham et al.).

Lemma 12 ***(Composition Map Lemma)*** *Suppose $g : W \to V$ and $f : U \to W$ are C^m maps. Then the map*

$$\begin{aligned} T\left(g,f\right) \quad & : \quad C^m\left(U,W\right) \times C^m\left(W,V\right) \to C^m\left(U,V\right) \\ \left(f,g\right) \quad & \longmapsto \quad f \circ g \end{aligned}$$

is C^m.

The chain rule will be important in our problem. It follows from the general result on composite maps.

Lemma 13 ***(C^m Composite Mapping Lemma).*** *Suppose $g : W \to V$ and $f : U \to W$ are C^m maps between Banach spaces. Then the composite $g \circ f : U \to V$ is also C^m and*

$$D(g \circ f)\left(x\right) \cdot e = Dg(f\left(x\right)) \cdot \left(Df\left(x\right)\right) \cdot e)$$

See Abraham et al. (Box 2.4.A) for the formula for $D^\ell(g \circ f)$ for $\ell > 1$.

The final tool we need is the implicit function theorem. This states that if the linearization of the equation $f\left(x\right) = y$ is uniquely invertible then locally so is f; i.e., we can uniquely solve $f\left(x\right) = y$ for x as a function of y. This just the generalization of the familiar implicit function theorem in $\mathbb{R}^n$.

Theorem 14 ***(Implicit Function Theorem)*** *Let $U \subset E$, and $V \subset F$ be open and $f : U \times V \to G$ be $C^m, r \geqslant 1$. For some $x_0 \in U, y_0 \in V$ assume $D_2 f(x_0, y_0) : F \to G$ is an isomorphism. Then there are neighborhoods U_0 of x_0 and a unique C^m map $g : U_0 \to V$ such that for all $x \in U_0$,*

$$f(x, g(x)) = 0.$$

Applying the chain rule to the relation $f(x, g(x)) = 0$, one can explicitly compute the derivatives of g :

$$D_1 g(x) = -\left[D_2 f(x, g(x))\right]^{-1} \circ D_1 f(x, g(x)) \tag{13}$$

These formulas look familiar from ordinary calculus. However, they may be quite different in practice. In particular, the derivatives in (13) are linear operators in a function space, not just Jacobian matrices, and the inversions involve solutions to linear functional equations, not just inversion of Jacobian matrices. However, the key fact is that we only need to solve linear functional equations instead of nonlinear equations. Therefore, the implicit function theorem in Banach spaces does have the same simplifying properties of the normal implicit function theorem. The exact details for our hyperbolic discounting problem will be presented below.

4 Local Analysis of the Hyperbolic Discounting Problem

We now establish the critical mathematical facts about the hyperbolic discounting problem. We saw that any equilibrium savings function h satisfies the functional equation

$$0 = G\left(k, h(k), h(h(k)), \varepsilon h'(h(k)), \varepsilon\right)$$

where $G : \mathbb{R}^5 \rightarrow \mathbb{R}$ was defined in (11). We restate the problem as a functional one. Let $I \subset \mathbb{R}$ be an open, convex set containing the steady state k_*. We assume that the deterministic equilibrium $\bar{h}(k)$ is locally asymptotically stable. Therefore, $\bar{h}(I) \subset I$.

Define the operator

$$\begin{aligned} \mathcal{N} &: X \times E \rightarrow C^{m-1}(I, \mathbb{R}) \\ \mathcal{N}(h, \varepsilon)(k) &= G\left(k, h(k), h(h(k)), \varepsilon h'(h(k)), \varepsilon\right) \end{aligned}$$

where $X \subset C^m(I, \mathbb{R})$ and $E = (-\varepsilon_0, \varepsilon_0)$ for some ε_0. $\mathcal{N}$ is the critical operator for us. We view $\mathcal{N}$ as a mapping taking a continuous function h of k and a scalar ε to another function of k. The operator $\mathcal{N}$ is not defined for all functions $h \in C^m(I, I)$. For example, if $h(k) > f(k)$ then the current period's consumption is negative, rendering the Euler equation undefined. However, if $h - \bar{h}$ is sufficiently small, $f(k) - h(k)$ will always be positive. More specifically, the subset $X \subset C^m(I, \mathbb{R})$ will be a ball of radius r for some $r > 0$:

$$X^r = \left\{h \middle| \left\|h - \bar{h}\right\|_m < r\right\}$$

Lemma 15 *Assume G is C^m and that $\bar{h}$ is C^m. If $r > 0$ is sufficiently small then $X^r \subset C^m(I, I)$ and $\mathcal{N} : X^r \times E \rightarrow C^{m-1}(I, \mathbb{R})$.*

Proof. Clearly, $G\left(k, \bar{h}(k), \bar{h}\left(\bar{h}(k)\right), \varepsilon\bar{h}\left(\bar{h}(k)\right)\right)$ exists since $\bar{h}(k) > 0$ for $k \in I$ and is C^{∞}. $G(k, h(k), h(h(k)), \varepsilon h'(h(k)))$ exists if $h(k)$ and $h(h(k))$ are positive for all $k \in I$. The order m derivatives of $G(k, h(k), h(h(k)), \varepsilon h'(h(k)))$ with respect to ε and k exist as long as G is C^m and h is C^m. Therefore, if $\left\|h - \bar{h}\right\|_m$ is sufficiently small then $G(k, h(k), h(h(k)), \varepsilon h'(h(k)))$ exists and is C^{m-1} in (k, ε). ■

When $\varepsilon = 0$ the problem in (10) is just the ordinary optimal growth problem with exponential discounting, and there is a locally unique $\bar{h}(k)$ such that $\mathcal{N}\left(\bar{h}, 0\right) = 0$. The task is to show that there is a unique map $\mathcal{Y} : (-\varepsilon_0, \varepsilon_0) \to C^{m-1}(I, \mathbb{R})$ such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, $\mathcal{N}(\mathcal{Y}(\varepsilon), \varepsilon) = 0$. We also want $\mathcal{Y}(\varepsilon)$ to be differentiable in ε thereby allowing us to compute $\mathcal{Y}(\varepsilon)$ via Taylor series expansions. To accomplish this we must apply the implicit function theorem for Banach spaces of functions to $\mathcal{N}$. We need to show that $\mathcal{N}$ satisfies the conditions for the IFT.

We next need to show that $\mathcal{N}(h, \varepsilon)$ is (Frechet) differentiable with respect to h at $h = \bar{h}$ and $\varepsilon = 0$. Rewrite $\mathcal{N}$ as

$$\begin{aligned}
\mathcal{N}(h, \varepsilon)(k) &= G\left(k, h(k), h(h(k)), \varepsilon h'(h(k)), \varepsilon\right) \\
&= G\left(k, \mathrm{ev}(h, k), \mathrm{ev}(h, h(k)), \varepsilon \ \mathrm{ev}(D(h), h(k)), \varepsilon\right)
\end{aligned}$$

The chain rule, composition theorem, the omega lemma, and the smoothness of differentiation in the $\|\cdot\|_m$ norm prove the following result.

Lemma 16 *$\mathcal{N}$ is C^m in the $\|\cdot\|_m$ norm.*

We now compute the derivative of $\mathcal{N}$ with respect to h. This is where the specific structure of our problem becomes important. The following result follows directly from computing the directional derivatives.

Lemma 17 *$\mathcal{N}_h\left(\bar{h}, 0\right)$ is the linear operator $\mathcal{N}_h\left(\bar{h}, 0\right) : C^m(I, I) \times \{0\} \to C^{m-1}(I, \mathbb{R})$ defined by*

$$\begin{aligned}
\left(\mathcal{N}_h\left(\bar{h}, 0\right) \cdot \psi\right)(k) &= A(k)\psi(k) + B(k)\psi\left(\bar{h}(k)\right) \qquad (14) \\
A(k) &\equiv G_2\left(k, \bar{h}(k), \bar{h}\left(\bar{h}(k)\right), 0, 0\right) + G_3\left(k, \bar{h}(k), \bar{h}\left(\bar{h}(k)\right), 0, 0\right)\bar{h}'\left(\bar{h}(k)\right) \\
B(k) &\equiv G_3\left(k, \bar{h}(k), \bar{h}\left(\bar{h}(k)\right), 0, 0\right)
\end{aligned}$$

and $\mathcal{N}_\varepsilon\left(\bar{h}, 0\right)$ is the linear operator $\mathcal{N}_\varepsilon\left(\bar{h}, 0\right) : \left\{\bar{h}\right\} \times E \to C^{m-1}(I, \mathbb{R})$ defined by

$$\begin{aligned}
\left(\mathcal{N}_\varepsilon\left(\bar{h}, 0\right) \cdot \varepsilon\right)(k) &= \varepsilon G_4\left(k, \bar{h}(k), \bar{h}\left(\bar{h}(k)\right), 0, 0\right)\bar{h}'\left(\bar{h}(k)\right) + \varepsilon G_5\left(k, \bar{h}(k), \bar{h}\left(\bar{h}(k)\right), 0, 0\right) \\
&\equiv \varepsilon C(k)
\end{aligned}$$

The last step is to show that the derivative of $\mathcal{N}(h, \varepsilon)$ with respect to h is invertible at neighborhood of $\left(\bar{h}, 0\right)$. That is, we want to solve the linear operator equation

$$0 = \mathcal{N}_h\left(\bar{h}, 0\right) \cdot h_\varepsilon + \mathcal{N}_\varepsilon\left(\bar{h}, 0\right)$$

for the unknown function h_ε. The formal expression for the solution is

$$h_\varepsilon = -\mathcal{N}_h\left(\bar{h}, 0\right)^{-1}\mathcal{N}_\varepsilon\left(\bar{h}, 0\right)$$

but we need to check that $\mathcal{N}_h\left(\bar{h}, 0\right)^{-1}$ exists and is unique. That is, we need to show that for every C^{m-1} function $C(k)$ there is a function $\psi(k)$ such that

$$0 = \mathcal{N}_h\left(\bar{h}, 0\right) \cdot \psi + C(k)$$

We will approach this in an intuitive manner. Our operator has the form

$$\mathcal{N}(h, \varepsilon) = G(k, h, h(h), \varepsilon h'(h), \varepsilon)$$

We want to find a path of solutions of the form

$$h(k, \varepsilon) = \bar{h}(k) + \varepsilon \frac{\partial h}{\partial \varepsilon}(k, 0) + \frac{\varepsilon^2}{2}\frac{\partial^2 h}{\partial \varepsilon^2}(k, 0) + \cdots \tag{15}$$

such that

$$\mathcal{N}(h(k, \varepsilon), \varepsilon) = 0, \ \forall \varepsilon \tag{16}$$

Note that we require the Taylor series to converge uniformly for all k in on the interval I. We want to compute the functions $h_\varepsilon(k, 0)$, $h_{\varepsilon\varepsilon}(k, 0)$, etc., and form the Taylor series in (15). To reduce notational complexity, define

$$\psi(k) = \frac{\partial h}{\partial \varepsilon}(k, 0)$$

We solve for h in (16) for small ε by implicit differentiation. When we substitute (15) into $G(k, h, h(h), \varepsilon h'(h), \varepsilon)$ and differentiate (16) with respect to ε at $\varepsilon = 0$, we find that $\psi(k)$ must satisfy, for all k, the functional equation

$$0 = A(k)\psi(k) + B(k)\psi\left(\bar{h}(k)\right) + C(k) \tag{17}$$

where $A(k)$, $B(k)$, and $C(k)$ are known functions defined above. Equation (17) looks unusual at first. However, it is really quite familiar. We first note that it is linear in the function ψ. To see this define the operator

$$S(\psi)(k) = A(k)\psi(k) + B(k)\psi\left(\bar{h}(k)\right)$$

and note that $S(\alpha_1\psi_1 + \alpha_2\psi_2) = \alpha_1 S(\psi_1) + \alpha_2 S(\psi_2)$ for arbitrary scalars α_1 and α_2. This is reassuring since linear problems are always easier than nonlinear problems.

Second, assume that $A(k)$ is nonzero for all k and define $J(k) = A(k)^{-1}B(k)$. Then (17) has the form[7]

$$\psi(k) = J(k)\psi\left(\bar{h}(k)\right) + C(k) \tag{18}$$

[7] Since $C(k)$ is an arbitrary function, we let $C(k)$ absorb the $A(k)^{-1}$ function.

This form reveals the iterative nature to the problem and suggests a natural infinite series solution. Consider the recursion

$$\begin{aligned}
\psi(k) &= J(k)\,\psi\left(\bar{h}(k)\right) + C(k) \\
\psi(k) &= J(k)\left[J\left(\bar{h}(k)\right)\psi\left(\bar{h}\left(\bar{h}(k)\right)\right) + C\left(\bar{h}(k)\right)\right] + C(k) \\
&\ldots \\
&= C(k) + \sum_{i=1}^{T}\left(\prod_{j=0}^{i-1} J\left(\bar{h}^{j}(k)\right)\right) C\left(\bar{h}^{i}(k)\right) \qquad (19) \\
&\quad + \left(\prod_{j=0}^{T} J\left(\bar{h}^{j}(k)\right)\right)\psi\left(\bar{h}^{T}(k)\right)
\end{aligned}$$

where $\bar{h}^{i}(k)$ is defined inductively by

$$\begin{aligned}
\bar{h}^{0}(k) &= k \\
\bar{h}^{1}(k) &= \bar{h}(k) \\
\bar{h}^{i+1}(k) &= \bar{h}\left(\bar{h}^{i}(k)\right)
\end{aligned}$$

This shows that our problem has a natural recursive structure and suggests an infinite series solution. The critical issue is whether J and $\bar{h}$ interact in a manner which produces a convergent series in (19). We now state the critical theorem.

Theorem 18 *Consider the functions A, B, and C in (14). If (i) $A(k)$ is positive for all k, and (ii) the magnitude of $A(k)^{-1} B(k)$ is uniformly less than $\theta < 1$ for all k, then $\mathcal{N}_h\left(\bar{h}, 0\right) : X \to C^{m-1}(I, \mathbb{R})$ in an invertible C^{m-1} operator.*

Proof. Transform the equation

$$A(k)\,\psi(k) + B(k)\,\psi\left(\bar{h}(k)\right) + C(k) = 0.$$

into the equivalent equation

$$\psi(k) = J(k)\,\psi\left(\bar{h}(k)\right) + C(k). \qquad (20)$$

where $J(k) = A(k)^{-1} B(k)$ and, without loss of generality, we have replaced $C(k)$ with $-A(k)^{-1} C(k)$.

We first show that there is a unique solution in $C^{0}(I, \mathbb{R})$. Define

$$(T\psi)(k) = J(k)\,\psi\left(\bar{h}(k)\right) + C(k)$$

By *(i)* and *(ii)* $A(k)^{-1} B(k)$ exists and has magnitude less than 1.Since $\bar{h}(I) \subset I$, we conclude that

$$\max_{k\in I}\left|\psi_1\left(\bar{h}(k)\right) - \psi_2\left(\bar{h}(k)\right)\right| \leqslant \max_{k\in I}\left|\psi_1(k) - \psi_2(k)\right|.$$

Furthermore, T is a contraction mapping because

$$\begin{aligned} |T\psi_1 - T\psi_2| &= \max_k \left| J(k)\,\psi_1\left(\bar{h}(k)\right) - J(k)\,\psi_2\left(\bar{h}(k)\right)\right| \\ &\leqslant \left(\max_{k\in I} J(k)\right) \max_k \left|\psi_1\left(\bar{h}(k)\right) - \psi_2\left(\bar{h}(k)\right)\right| \\ &\leqslant \theta \max_k |\psi_1(k) - \psi_2(k)| \end{aligned}$$

and the iterates $\psi^0 = 0$ and $\psi^{i+1} = T\psi^i$ converge uniformly to the solution ψ^∞.

If $\psi'(k)$ exists, (20) implies

$$\begin{aligned} \psi'(k) &= J'(k)\,\psi\left(\bar{h}(k)\right) + J(k)\,\psi'\left(\bar{h}(k)\right)\bar{h}'(k) + C'(k) \\ &= J(k)\,\bar{h}'(k)\,\psi'\left(\bar{h}(k)\right) + \left(J(k)\,\psi\left(\bar{h}(k)\right) + C'(k) + J'(k)\,\psi\left(\bar{h}(k)\right)\right) \\ &= J(k)\,\bar{h}'(k)\,\psi'\left(\bar{h}(k)\right) + \tilde{C}(k) \end{aligned}$$

where $\tilde{C}(k)$ is $C^0(I,\mathbb{R})$. We need to prove that $\psi'(k)$ exists. Define the operator on C^{m-1}

$$\left(T^1\phi\right)(k) = J(k)\,\bar{h}'(k)\,\phi\left(\bar{h}(k)\right) + \tilde{C}(k)\,. \tag{21}$$

Note that $\phi\left(\bar{h}(k)\right)$ has the coefficient $J(k)\,\bar{h}'(k)$ which has magnitude less than $\theta < 1$ since $|h'(k)| < 1$ and $|J(k)| < \theta$ for $k \in I$. Furthermore, $J(k)\,\bar{h}'(k)$ is $C^0(I,\mathbb{R})$ by assumption. Therefore, the sequence $\phi^0 = 0$, and $\phi^{i+1} = T^1\phi^i$ converges to the unique fixed point $\phi^\infty(k)$ of T^1. Furthermore, since $\phi^i = \frac{d}{dk}\psi^i$ and the convergence of the ϕ^i is uniform, we can conclude that $\phi^\infty = \frac{d}{dk}\psi^\infty = \frac{d}{dk}\psi(k)$, proving that $\psi'(k)$ exists and satisfies (21). This step can be repeated as long as $\bar{h}$, J, and C have the necessary derivatives. Therefore, the solution ψ^∞ is C^{m-1}. ■

The global contraction properties assumed in Theorem 18 are strong. We next prove a local version of the same result.

Corollary 19 *If (i) $A(k_*)$ is nonsingular, and (ii) the magnitude of $A(k_*)^{-1} B(k_*)$ is less than one, then $\mathcal{N}_h\left(\bar{h},0\right) : X \to C^{m-1}(I,\mathbb{R})$ is an invertible C^{m-1} operator for some neighborhood of $\bar{h}$ in $C^m(I,I)$.*

Proof. Since A, B, and J are m-times differentiable, there is a neighborhood of k_* such that the assumptions of Theorem 18 hold. The conclusions then follow from Theorem 18 ■

The last result is the infinite series representation of the asymptotic terms.

Corollary 20 *Under the assumptions of Theorem 18 or Corollary 19 , the solution to (18) has the infinite series representation*

$$\psi(k) == \sum_{i=1}^{\infty} \left(\prod_{j=0}^{i-1} J\left(\bar{h}^j(k)\right)\right) C\left(\bar{h}^i(k)\right) + C(k) \tag{22}$$

Proof. This follows directly from the contraction map arguments in the proof of Theorem 18. This series also holds for small neighborhoods around k_* under the assumptions of Corollary 19. The infinite series representation holds globally since $\bar{h}$ is a strictly increasing function with fixed point at k_* ■

We will also state the multidimensional version of the theorem since that will be important in future generalizations. The proof is the same as above.

Corollary 21 *Consider the equation*

$$A(k)\,\psi(k) + B(k)\,\psi\left(\bar{h}(k)\right) + C(k) = 0 \tag{23}$$

for C^m functions $A, B, C : I^n \to \mathbb{R}^n$. If (i) $k_* \in I^n$ and $\bar{h}(k_*) = k_*$, (ii) $\bar{h}(I^n) \subset I^n$, (iii) $A(k_*)$ is nonsingular, and (iv) the spectral radius of $A(k_*)^{-1} B(k_*)$ is less than one, then 23 has a unique C^m solution $\psi : I^n \to \mathbb{R}^n$.

4.1 Global Multiplicity and Selection

We have discussed how there may be multiplicity problems for these models. Theorem 18 presents a local uniqueness result. Since it is a local result, it does not say anything about uniqueness for any particular parameters. For example, consider the relation $xe^{-x^2} = 0$ that defines $x \in \mathbb{R}$ implicitly as a function of $\varepsilon \in \mathbb{R}$. At $\varepsilon = 0$, the unique solution is $x = 0$. However, for any other value of ε there are three solutions for x but only one branch contains the $(x, \varepsilon) = (0, 0)$ solution.

Figure 1 displays a possible multiplicity problem consistent with Theorem 18. The vertical axis represents possible values of the scalar $\varepsilon = 1/\beta - 1$, and the horizontal axis represents the infinite-dimensional space of permissible savings functions. We know that when $\varepsilon = 0$ the unique solution, h^0, is the solution to the optimal growth problem. Theorem 18 shows that for ε close to zero, there is a unique solution close to $\bar{h}(k)$. However, at some positive ε_1 there may appear multiple solutions. That multiplicity may continue to hold until $\varepsilon = \varepsilon_2$, at which point we have a catastrophe[8]. The perturbation method implicitly makes a selection for $\varepsilon \in (\varepsilon_1, \varepsilon_2)$, assuming that the Taylor series is convergent for such ε. The selection is the smooth manifold of solutions connecting h^0 to the leftmost solution at $\varepsilon = \varepsilon_2$. This selection rule is consistent with many common selection arguments in game theory.

[8] We could perhaps compute the manifold beyond the catastrophe by expressing both h and ε as functions of the arc length parameter s often used in homotopy methods. We leave this possibility for future investigations.

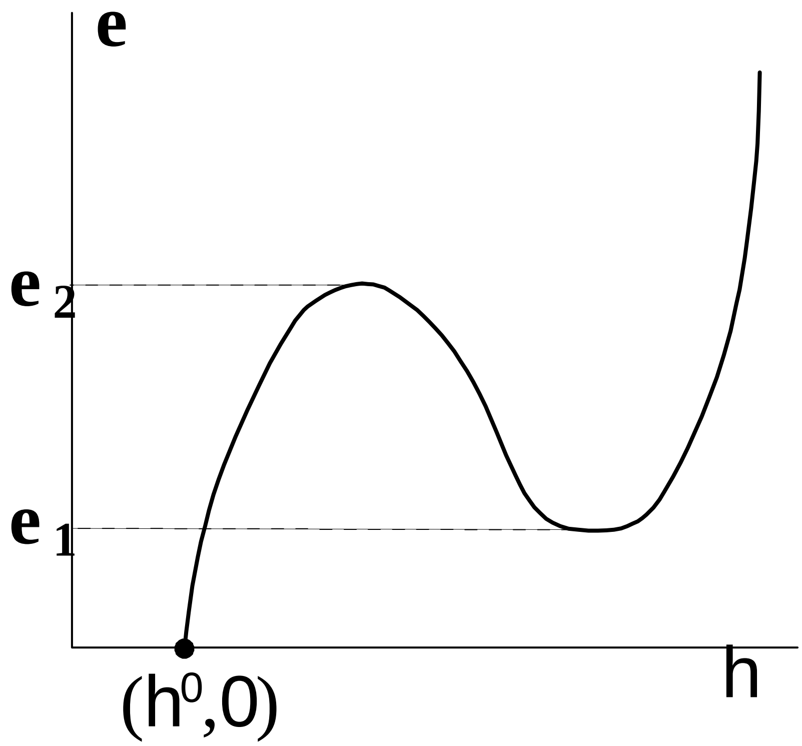

Figure 1: Possible equilibrium manifold

5 Existence and Value Function Iteration

These results may initially appear to be rather technical but they do reflect is simple intuition. Consider the hyperbolic discounting problem (10) with $\varepsilon = 0$. In that case there is a surely convergent algorithm to compute h implied by the contraction properties of standard dynamic programming. To see this, first note that when $\varepsilon = 0$ the policy function $h(k)$ is the optimal policy function to the dynamic programming problem

$$\begin{aligned} V(k) &\equiv \max_x u(f(k) - x) + \delta V(x) \\ h(k) &\equiv \arg\max_x u(f(k) - x) + \delta V(x) \end{aligned} \tag{24}$$

Value function iteration in (24) is equivalent to the time iteration scheme implied by the Euler equation. That is, h is the limit of

$$h^{i+1} = T\left(h^i\right) \tag{25}$$

where we define the time iteration T operator, $g = T(h)$, implicitly by

$$\begin{aligned} 0 &= u'(f(k) - g(k)) - \beta\delta u'(C(g(k)))\, f'(g(k)) \\ C(k) &= f(k) - h(k) \end{aligned}$$

Consider the directional derivative of T. Suppose that we change h in the direction δ and want to know how $T(h)$ changes as we change h. To do so, we parameterize this movement by looking at $h+\varepsilon\delta$ and compute the directional derivative

$$\frac{dg}{d\varepsilon}=\frac{d}{d\varepsilon}\left(T\left(h+\varepsilon\delta\right)\right)$$

by implicit differentiation

$$0=\left(G_2+G_3h'\left(g\right)\right)g_\varepsilon(k)+G_3\delta(g(k))$$

Given $g=Th$ we can directly compute

$$g_\varepsilon(k)=-(G_2+G_3h'\left(g\right))^{-1}G_3\delta(g(k))$$

For $\left\|h-\bar{h}\right\|<\varepsilon$, G_2+G_3h' is close to $G_2+G_3\bar{h}'$ at $k=k_*$. But, $G_2+G_3\bar{h}'$ is $N_h\left(\bar{h},0\right)$. This insight proves the following corollary

Corollary 22 *$\mathcal{N}_h\left(\bar{h},0\right):X\rightarrow C^{m-1}\left(I,\mathbb{R}\right)$ in an invertible C^{m-1} operator if and only if time iteration in (25) is locally convergent in the $C^{m-1}\left(I,\mathbb{R}\right)$ topology.*

This result establishes the intuition that if value function iteration in a dynamic programming problem converges to a unique solution, then there exists a unique Nash equilibrium for "nearby" games and time iteration is a convergent algorithm. However, we need to note that the topological approach we took to the problem was necessary. Value function iteration just mimics the convergence of a contraction mapping in the Banach space defined by the sup norm. Games do not generally have contraction mapping representations. Furthermore, we need to deal directly with strategy functions when we analyze games. The contraction mapping representation of dynamic programming is a fixed point problem in a Banach space, whereas our equilibrium formulation of a dynamic game is an example of the more general class of problems of finding a zero of a mapping from one Banach space to another Banach space.

6 Previous Computational Approaches

Computing equilibrium savings functions in the hyperbolic discounting model presents some difficulties. One can use value function iteration, but that will often take a long time to converge. Convergence of value function iteration is not assured since the problem does not have a contraction property. One would like to linearize around the steady state, a common approach in dynamic economics. Unfortunately, we do not know what the steady state is except for the special case of $\varepsilon=0$. In this section, we review some previous methods and their strengths and weaknesses.

For specificity, we will examine one particular case of the hyperbolic savings model. We will use the same example used in KKS. They assumed $u(c) = \log c$ and $\delta = .95$. They also assumed that the production function had a capital share of 0.36 and that the capital stock depreciated at a rate of 0.10 per unit of time; this implies that $f(k) = \frac{144}{342}k^{\alpha} + .9k$ where we have chosen units so that the steady state capital stock is $k_* = 1$ if $\beta = 1$, that is, exponential discounting. We will examine five cases of β: $\beta \in \{1, .95, .9, .85, .8\}$. We focus on changes in β since that is the hyperbolic discounting parameter. The results are similar for other choices of utility and production functions.

6.1 Polynomial Approximation Methods

There have been many papers in the public finance and resource economics literature which have solved for time consistent equilibria and for Nash equilibrium policy games. For example, Wright and Williams (1984) computed the impact of the strategic oil reserve when a government is known to impose price controls when oil prices get high. Kotlikoff et al. (1988) compute equilibrium bequest policies. Ha and Sibert (1997) compute Nash equilibrium tax policies between competing countries. Rui and Miranda (1996) compute Nash equilibrium commodity stockpiling policies. Judd (1998) examines a simple problem of time consistent tax policy. These papers all used flexible polynomial methods for computing equilibrium policies. Since they use polynomial approximations, they were searching only for continuous equilibria. Our approach shares that objective.

The problem with these methods is that they are subject to a curse of dimensionality. Our perturbation method does not suffer from as bad a curse of dimensionality. On the other hand, our approach will be local in contrast to the more global approach in many previous studies.

The polynomial approach can be easily and reliably applied to the hyperbolic savings problem[9]. More specifically, we first hypothesize that the solution is approximated by

$$\hat{h}(k) = \sum_{i=0}^{n} a_i \psi_i(k)$$

where $\psi_i(k)$ is a degree i polynomial (the ψ_i should be an orthogonal system, such as Chebyshev polynomials) and the a_i are unknown coefficients. We then fix the $n+1$ coefficients by solving the system of equations

$$\int G\left(k, \hat{h}(k), \hat{h}\left(\hat{h}(k)\right), \varepsilon \hat{h}'\left(\hat{h}(k)\right)\right) \phi_j(k), \quad j = 0, .., n \tag{26}$$

where the $\phi_j(k)$ are linearly independent functions. Essentially, we fix a by projecting the Generalized Euler Equation in $n+1$ directions, and the $\phi_j(k)$ represent those directions. Specifically,

[9] KKS assert that standard polynomial approximation methods will not solve their hyperbolic discounting problem, and offer their solution as an alternative. However, they offer no evidence to document their negative claims regarding the methods used since Wright and Williams (1982).

we let both the basis functions $\psi_i(k)$ and the test functions $\phi_j(k)$ be Chebyshev polynomials, and use Chebyshev quadrature for the integral in (26), producing a Chebyshev collocation method (see Judd, 1992, for details). For each problem, we easily found[10] a degree 31 polynomial for which the maximum Euler equation error was 10^{-13} for capital stocks between .25 and 1.75. For the case of exponential discounting, the steady state capital stock is $k_* = 1$. The deviations from $k_* = 1$ in the other cases give us some idea about the economic significance of the hyperbolic discounting. The steady state for each problem is listed in Table 1. We see that a value of $\beta = .8$ produces very different long-run dynamics.

Table 1: Steady State Capital Stock from Projection Method

β :	1.00	.950	.900	.850	.800
steady state k :	1.00	.904	.809	.716	.625

6.2 The KKS Procedure and the Projection Method

The steady state capital stock under hyperbolic discounting, defined by

$$h(k_*) = k_*$$

is easily computed if $\beta = 1$, in which case k_* is fixed by the equation $1 = \delta f'(k_*)$, but is not easily computed otherwise. In general, we cannot find the steady state without knowing the solution $h(k)$ more generally. KKS propose a procedure to find the unknown steady state, and then build a more global approximation to $h(k)$ around it. The KKS procedure begins with the Generalized Euler equation

$$0 = G\left(k, h(k), h(h(k)), \varepsilon h'(h(k)), \varepsilon\right) \tag{27}$$

They want to solve for the steady state k_*, which must solve

$$0 = G\left(k_*, k_*, k_*, \varepsilon h'(k_*), \varepsilon\right) \tag{28}$$

Unfortunately, (28) has two unknowns: k_* and $h'(k_*)$. They need another equation to pin down the unknowns. Let

$$\begin{aligned} & \mathfrak{G}^n\left(k_*, \varepsilon h'(k_*), h''(k_*), ..., h^{(n+1)}(k_*), \varepsilon\right) \\ \equiv\ & D_k^n G\left(k, h(k), h(h(k)), \varepsilon h'(h(k)), \varepsilon\right)|_{k=k_*=h(k_*)} \end{aligned}$$

be the n'th total derivative of the GEE. They differentiate (27) with respect to k and impose the steady state conditions to arrive at

$$0 = G^{(1)}\left(k_*, k_*, k_*, \varepsilon h'(k_*), \varepsilon h''(k_*), \varepsilon\right) \tag{29}$$

[10] A Mathematica program on a 1 GHz Pentium machine found a solution for each problem in less than five seconds.

The new equation (29) does add a condition but it also produces a new unknown, $h''(k_*)$. They continue this differentiation until they arrive at a list of $n+1$ equations

$$\begin{aligned} 0 &= \mathfrak{G}^0 = G\left(k_*, k_*, k_*, \varepsilon h'(k_*), \varepsilon\right) \\ 0 &= \mathfrak{G}^1\left(k_*, h'(k_*), h''(k_*), \varepsilon\right) \\ &\ldots \\ 0 &= \mathfrak{G}^n\left(k_*, h'(k_*), h''(k_*), ..., h^{(n+1)}(k_*), \varepsilon\right) \end{aligned} \tag{30}$$

with $n+2$ unknowns, whereupon they append the condition

$$0 = h^{(n+1)}(k_*) \tag{31}$$

This now produces a system of $n+2$ equations with $n+2$ unknowns. They are, however, nonlinear. To solve this system they form the least squares criterion

$$KKS = h^{(n+1)}(k_*)^2 + \sum_{i=0}^{n}\left(\mathfrak{G}^i\right)^2$$

and then choose k_* and the various derivatives of $h(k)$ at k_* to minimize KKS. Krusell et a. (2002) report "We have implemented this algorithm using polynomial decision rules up to order 3. We define $\hat{h}_\psi$ using ordinary polynomials $\hat{h}_\psi\left(\tilde{k}\right) = \sum_{i=0}^{n-1} a_i \tilde{k}^i$. We find that the numerical results change only to a small degree when increasing the order of the polynomial from 2 to 3." They also say "Moreover, our computational experiments that generalize the utility function and consider less than full depreciation show no indication of multiplicity."

Unfortunately, there are problems with the KKS procedure. Despite their claims, there are multiple solutions and the solution set changes significantly as we increase degree. Table 2 displays solutions for k_* for the example studied in KKS and various orders of approximation. For example, if the order is 1, then we set $h''(k_*) = 0$ to fix the steady state. The multiplicity of solutions in Table 2 is not due to numerical error. This is because we reduce (30,31) to one equation in the unknown k_*, which was then computed with 256 digits of decimal precision. Each of the results in Table 2 can be proven to lie within within 10^{-4} of a root by application of the intermediate value theorem. Since each solution in Table 1 is at least 10^{-4} away from the others, each reported solution in Table 2 represents a distinct solution to the KKS equations.

Table 2 displays numerous problems with the KKS procedure. First, there is an increasing number of solutions to the KKS procedure as we go to higher orders of approximation and that they are spread over a wide range of values. Second, many of them appear to persist. The one that KKS identifies does persist and is close to the solution found using Chebyshev collocation and displayed in Table 1. However, many other solutions appear also to persist. The solutions in Table 2 are displayed in a manner that highlights how some solutions appear, disappear, then reappear as we go to higher orders. There is no compelling reason to take the solution that KKS found.

Table 2: KKS Solutions

Approx. Order	Steady state k for stable solutions									
1	0.81									
2	0.81						0.87			
3	0.81				0.84					0.94
4	0.81			0.83				0.89		
5	0.81			0.83		0.86				0.93
6	0.81		0.822			0.85		0.89		
7	0.81		0.820		0.84		0.87		0.92	
8	0.81	0.818		0.83		0.86		0.89		
9	0.81	0.817		0.83		0.85	0.88			0.93
10	0.81	0.817		0.83	0.84	0.86		0.90		0.93
11	0.81	0.816	0.825		0.84	0.86	0.88		0.91	
12	0.81	0.816	0.824	0.83		0.85	0.87	0.89		0.93
13	0.81	0.816	0.823	0.83	0.84	0.86	0.88		0.91	0.94
14	0.81	0.816	0.823	0.83	0.84	0.85	0.87	0.89	0.92	
15					(10 solutions)					
16					(10 solutions)					
17					(11 solutions)					
18					(12 solutions)					
19					(12 solutions)					

7 Perturbation Methods for Problems with Small (and Large) Hyperbolic Deviations

We next use our asymptotic results to construct a perturbation method for solving the hyperbolic discounting problem. More precisely, we define the function $h(k,\varepsilon)$ to satisfy the Generalized Euler equation

$$u'(f(k)-h(k,\varepsilon)) = \beta\delta u'(f(h(k,\varepsilon)) - h(h(k,\varepsilon),\varepsilon))\left(f'(h(k,\varepsilon)) + \varepsilon h'(h(k,\varepsilon),\varepsilon)\right)$$

We first use standard perturbation methods to compute the Taylor series approximation of the exponential discounting problem

$$
\begin{aligned}
h(k,0) &\doteq h(k_*,0)+h_k(k_*,0)(k-k_*) \\
&+h_{kk}(k_*,0)(k-k_*)^2/2 \\
&+h_{kkk}(k_*,0)(k-k_*)^3/6 \\
&+...
\end{aligned}
$$

up to degree 12. This uses the procedure described in Judd and Guu (1992).

We next check the conditions of Theorem 18. For our model, the $J(k)$ term in Theorem 18 reduces to

$$
\frac{\delta}{1+\delta^2\frac{u'(c_*)}{u''(c_*)}f''(k_*)+\delta\left(1-\bar{h}'(k_*)\right)}<\delta
$$

which is less than one in magnitude for any concave f and concave u. Therefore, concavity in preferences and technology gives us the critical contraction condition and proves that there exists a unique smooth solution to hyperbolic discounting problem for small ε.

We now can solve nontrivial cases of the hyperbolic discounting terms. There are two ways we could proceed. The first approach would directly implement the power series solutions in (22) to compute , $h_\varepsilon(k,0)$, $h_{\varepsilon\varepsilon}(k,0)$, etc. over a significant range of k. Notice that the contraction operator associated with computing h_ε is more strongly contractive than the contraction factor δ for the original optimal growth problem. This fact would help computation of the perturbed terms. We do not do that here.

7.1 Local Taylor Series Approach

The second approach is to use the fact that Theorem 18 implies that $h(k,\varepsilon)$ is a smooth function of (k,ε) near $(k_*,0)$. We compute the Taylor series

$$
\begin{aligned}
h(k,\varepsilon) &\doteq h(k_*,0)+h_k(k_*,0)(k-k_*)+\frac{1}{2}h_{kk}(k_*,0)(k-k_*)^2+... \\
&+h_\varepsilon(k_*,0)\varepsilon+h_{k\varepsilon}(k_*,0)(k-k_*)\varepsilon+\frac{1}{2}h_{kk\varepsilon}(k_*,0)\varepsilon(k-k_*)^2 \\
&+\frac{1}{2}h_{\varepsilon\varepsilon}(k_*,0)\varepsilon^2+\frac{1}{2}h_{k\varepsilon\varepsilon}(k_*,0)(k-k_*)\varepsilon^2+\frac{1}{4}h_{kk\varepsilon\varepsilon}(k_*,0)\varepsilon^2(k-k_*)^2+..
\end{aligned}
$$

We accomplish this by differentiating the Generalized Euler equation with respect to ε to arrive at

$$
0=\frac{d}{d\varepsilon}\left(G\left(k,h(k),h(h(k)),\varepsilon h'(h(k))\right)\right)\Big|_{\varepsilon=0,k=k_*}
$$

which produces a linear equation for $h_\varepsilon(k_*, 0)$. We continue with differentiations of GEE with respect to k and ε until we have a Taylor series expansion for $h(k, \varepsilon)$ of degree 12 in (k, ε).

We applied this method to the four cases of hyperbolic discounting examined in Table 1. Figure 2 displays the net savings functions of the solutions. They are all stable.

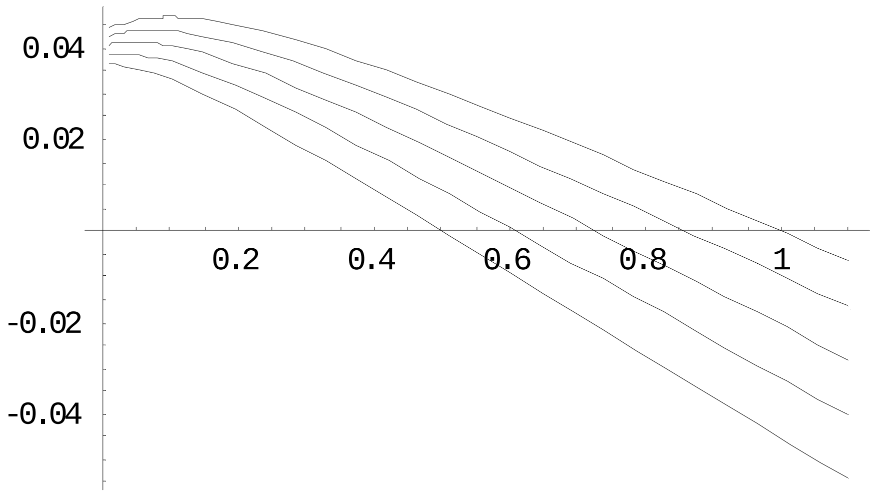

Figure 2: Net savings functions

Table 3 displays the steady states of the solutions from the perturbation method. Note that Tables 1 and 3 are identical. The Euler equation errors for each problem are displayed in Figure 3. Note that they are practically identical except for capital stocks near $k = 1$ where they are essentially zero for each problem. The Euler equation errors for the solutions in Table 1 were much smaller, but those solutions should be better since they used degree 31 polynomials. It is doubtful that perturbation methods could produce such high order solutions. However, they may produce good initial guesses for other methods, a fact which may be particularly important if we have multiple solutions to the true problem and/or the numerical procedure. Furthermore, even when $\beta = .8$ and the steady state is $k_* = .625$, the Euler equation errors near the steady state are no more than 10^{-8}. Therefore, all tests say that the perturbation solutions are excellent approximate solutions.

Table 3: Steady State Capital Stock from Perturbation Method

β :	1.00	.950	.900	.850	.800
steady state k :	1.00	.904	.809	.716	.625

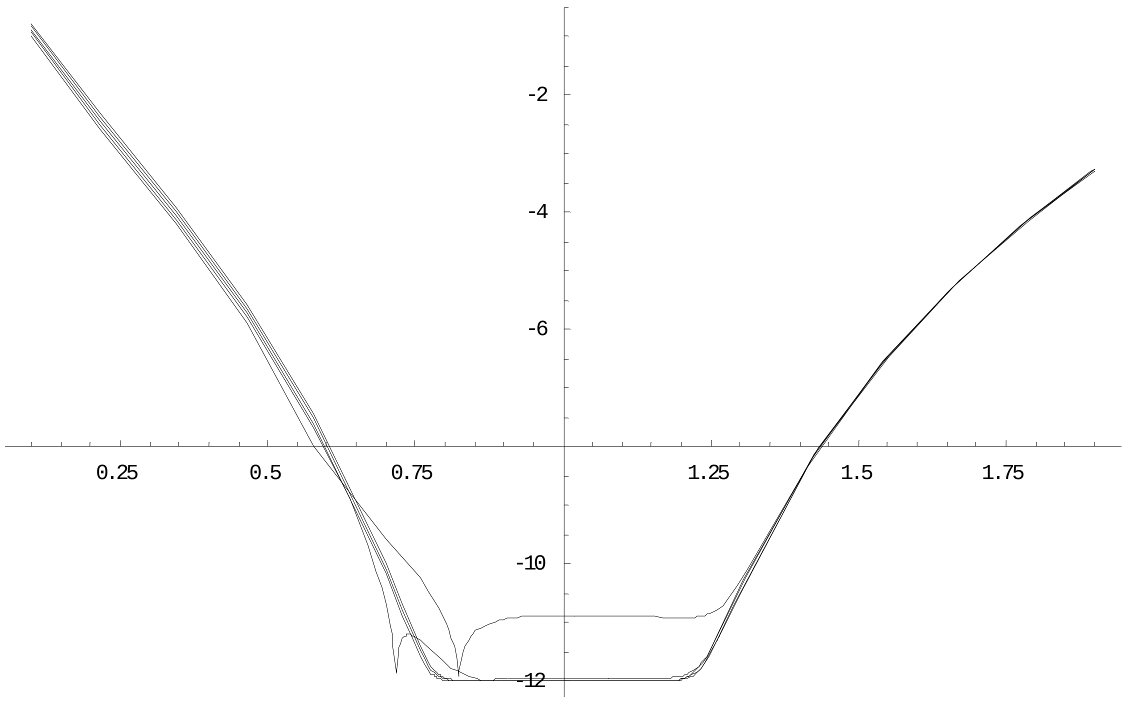

Figure 3: Log_{10} Euler equation errors

8 Time-Consistent Government Tax and Spending

We next briefly outline how these techniques can be used to analyze government tax and spending policies. Suppose $u(c, g)$ is the utility function of the representative individual over consumption c and government expenditure g. Suppose output is $F(k)$ where k is the capital stock, and the capital accumulation follows $k_{t+1} = F(k_t) - c_t - g_t$. We assume that there is a gross income tax, $\tau F(k)$, no bonds, and that spending equals revenue in each period, $g = \tau F(k)$. We assume that governments and consumers make simultaneous decisions at the beginning of each period, the consumers choosing c and the government choosing g and τ. This makes private investment, $F(k_t) - c_t - g_t$ the residual element that ensures material and financial balance.

The equilibrium analysis will follow recursive forms. The consumer follows the decision rule $c = C(k)$, and the government follows the tax rule $\tau = T(k)$. Define

$$\begin{aligned} G(k) &= T(k) F(k) \\ h(k) &= F(k) - C(k) - G(k). \end{aligned}$$

The consumer's Euler equation will be, as always,

$$u_c(c_t, g_t) = \beta u_c(c_{t+1}, g_{t+1}) F'(k_{t+1})(1 - T(k_{t+1}))$$

which implies the recursive functional equation

$$u_c(C(k), G(k)) = \beta u_c(C(h(k)), G(h(k))) F'(h(k))(1 - T(h(k))).$$

The government's problem is more complex. The value function for government is

$$V(k) = u(C(k), G(k)) + \beta V(h(k)).$$

The current government faces the choice problem

$$\tau \equiv \arg\max_{\tau} u(C(k), \tau F(k)) + \beta V(F(k) - C(k) - \tau F(k))$$

which implies the first-order condition

$$0 = u_g(C(k), G(k)) - \beta V'(F(k) - C(k) - \tau F(k)).$$

The combination of consumer and government problems implies that the equilibrium system for the functions $V(k)$, $C(k)$ and $T(k)$ are

$$\begin{aligned} u_c(C(k), G(k)) &= \beta u_c(C(k_+), G(k_+)) F'(k_+)(1 - T(k_+)) \\ V(k) &= u(C(k), G(k)) + \beta V(k_+) \\ 0 &= u_g(C(k), G(k)) - \beta V'(k_+) \end{aligned}$$

where

$$\begin{aligned} G(k) &\equiv T(k) F(k) \\ k_+ &\equiv F(k) - C(k) - G(k) \end{aligned}$$

We next need to find a special case where we know the answer. Economic intuition suggests that we begin with a case where the first-best optimal g is zero, and $T(k) = 0$. Under these conditions, $C(k)$ solves the simple Euler equation from optimal growth theory

$$u_c(C(k), 0) = \beta u_c(C(h(k)), 0) F'(h(k)).$$

We want to parameterize the problem so $g = 0$ is optimal when $\varepsilon = 0$. This is not trivial and requires some care. For example, a bad approach would be to assume $u(c, g) = u(c) + \varepsilon g$. In this case, the optimal g at $\varepsilon = 0$ is $-\infty$ since negative g would augment output. Similarly, $u(c, g) = u(c) + \varepsilon \ln g$ is not good since it is not locally analytic for (ε, g) near $(0, 0)$.

One good example would be

$$u(c, g) = u\left(c + (1 + \varepsilon) g - g^2\right)$$

In this case, $u_c = u_g$ at $\varepsilon = 0 = g$ and the optimal g is zero at $\varepsilon = 0$. Furthermore, $u(c, g)$ is locally analytic for (ε, g) near $(0, 0)$. Another choice would be

$$u(c, g) = u\left(c + (1 - \varepsilon)\left(g - g^2\right)\right) + \varepsilon\lambda \ln(1 - \varepsilon + g)$$

This choice is also analytic for (ε, g) near $(0, 0)$. Moreover, it is a homotopy construction in the sense that utility is a $u(c) + \lambda \ln(g)$ at $\varepsilon = 1$, which is a more conventional utility function. In this case, we would want the Taylor series to work at $\varepsilon = 1$; this is a reasonable conjecture if λ close to zero since optimal g is close to zero.

9 Conclusion

We have proved a local existence and uniqueness theorems for smooth solutions to a hyperbolic savings problems with small amounts of hyperbolic discounting. The analysis used general tools from nonlinear functional analysis and the dynamic stability of the exponential discounting case. This indicates that the techniques are applicable to a much wider range of dynamic strategic problems. Also, the proofs were constructive and lead directly to a stable and reliable numerical procedure for solving such problems.

References

[1] Abraham, Ralph, Jerrold E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications.* Addison-Wesley, Reading, 1983.

[2] Basar, T. and G. J. Olsder. *Dynamic Noncooperative Game Theory.* Academic Press, London/New York, 1982.

[3] Debreu, G. (1976), "Regular Differentiable Economies" *American Economic Review* 66: 280-287.

[4] Doraszelski, U. "An R&D Race with Knowledge Accumulation," *Rand Journal of Economics*, 34 (2003), pp. 19-41

[5] Friedman, J. and L. Samuelson, 1991, "Subgame perfect equilibrium with continuous reaction functions," *Games and Economic Behavior* **2** (1991), 304-324.

[6] Fudenberg, D., and J. Tirole. "Capital as a Commitment: Strategic Investment to Deter Mobility", *Journal of Economic Theory* 31, 2 (1983): 227-250

[7] Ha, Jiming, and Anne Sibert. "Strategic Capital Taxation in Large Open Economies with Mobile Capital", *International Tax and Public Finance*, **39** (July 1997), 243-262

[8] Harris, Christopher, and David Laibson. “Dynamic Choices of Hyperbolic Consumers”, *Econometrica*, **69**(2001), 935-957.

[9] Harris, Christopher, and David Laibson. “Hyperbolic Discounting and Consumption”. In Mathias Dewatripont, Lars Peter Hansen, and Stephen Turnovsky, eds., *Advances in Economics and Econometrics* (Cambridge University Press, 2003).

[10] Joshi, Mohan C., and Ramendra K. Bose. *Some topics in nonlinear functional analysis.* John Wiley & Sons, Inc., New York, 1985.

[11] Judd, Kenneth L. “Projection Methods for Solving Aggregate Growth Models,” *Journal of Economic Theory* 58 (December, 1992), 410-452.

[12] Judd, Kenneth L, and Sy-Ming Guu. “Perturbation Solution Methods for Economic Growth Models,” in *Economic and Financial Modelling with Mathematica*, edited by Hal Varian, pages 80-103. Springer-Verlag Publishers: New York. 1992.

[13] Judd, Kenneth L. “Numerical Methods in Economics,” MIT Press: Cambridge, MA, 1998.

[14] Judd, Kenneth L. "Closed-Loop Equilibrium in a Multi-Stage Innovation Race," *Economic Theory*, 21 (2003), 673-695.

[15] Klein, Paul, Per Krusell, Jose-Vıctor Rıos-Rull. “Time-Consistent Public Expenditures,” mimeo, March 15, 2002

[16] Kotlikoff, Laurence. J., John Shoven, and Avi Spivak. “The Effect of Annuity Insurance on Savings and Inequality,” *Journal of Labor Economics*, **4**(#3/pt.2, July, 1986), S183-S207.

[17] Krusell, Per, Vincenzo Quadrini and José-Víctor Ríos-Rull “Politico-Economic Equilibrium and Economic Growth”. *Journal of Economic Dynamics and Control* , **21** (1997), 243-272.

[18] Krusell, Per, Burhanettin Kuruscu, and Anthony J. Smith. “Equilibrium Welfare and Government Policy with Quasi-geometric Discounting, *Journal of Economic Theory* **105** (2002), 42-72.

[19] Krusell, Per, and Anthony Smith, Jr. “Consumption-Savings Decisions with Quasi-geometric Discounting,” *Econometrica* **71** (2003), 365–375.

[20] Rincon-Zapatero, J.P., Martinez, J., Martin-Herran, G. “New method to characterize subgame perfect Nash equilibria in differential games.” *Journal of Optimization Theory and Applications* 96 (1998), 377–395.

[21] Rogers, Carol Ann. "Expenditure taxes, income taxes, and time-inconsistency," *Journal of Public Economics*, **32** (1987), 215-230

[22] Rui, Xui W., and Mario J. Miranda. "Solving nonlinear dynamic games via orthogonal collocation: An application to international commodity markets," *Annals of Operations Research*, **68** (1996) 89-108.

[23] Shannon, C. "Determinacy of Competitive Equilibria in Economies with Many Commodities," *Economic Theory* **14** (1999), 29-87.

[24] Stanford, W. "Subgame perfect reaction function equilibria in discounted duopoly supergames are trivial," *Journal of Economic Theory* **39** (1986), 226-232.

[25] Stokey, Nancy L. "Rational Expectations and Durable Goods Pricing," *The Bell Journal of Economics* **12** (1981): 112-128.

[26] Vedenov, Dmitry V., and Mario J. Miranda. "Numerical solution of dynamic oligopoly games with capital investment," *Economic Theory*, **18** (2001) pp. 237-261.

[27] Williams, Jeffrey, and Brian Wright. The Roles of Public and Private Storage in Managing Oil Import Disruptions. *Bell Journal of Economics* **13** (1982), 341–353.