Lecture Notes On Solution Methods for Microeconomic Dynamic Stochastic Optimization Problems

Christopher D. Carroll

February 16, 2011

Abstract

These notes sketch some techniques for solving microeconomic dynamic stochastic optimization problems. I make no attempt at a systematic overview of the many possible technical choices; instead, I present a specific set of methods that I have found useful in my own work. Paried with these notes is *Mathematica* and Matlab software that solves the problems described in the text. These notes were originally written for my Advanced Topics in Macroeconomic Theory class at Johns Hopkins University; instructors elsewhere are welcome to use them for teaching purposes.

PDF: `http://econ.jhu.edu/people/ccarroll/SolvingMicroDSOPs.pdf`
Web: `http://econ.jhu.edu/people/ccarroll/SolvingMicroDSOPs/`
Archive: `http://econ.jhu.edu/people/ccarroll/SolvingMicroDSOPs.zip`
(Contains LaTeX code for this document and software producing figures and results)

This draft improves on earlier versions in several respects, especially by revising the notation and software to be consistent with those in "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems" published in *Economics Letters*, available at `http://econ.jhu.edu/people/ccarroll/EndogenousArchive.zip`, and by including sample code for a method of simulated moments estimation of the life cycle model *a la* Gourinchas and Parker (2002) and Cagetti (2003). Background derivations, notation, and related subjects are treated in my class notes for first year macro, available at `http://econ.jhu.edu/people/ccarroll/public/lecturenotes/consumption`.

I am grateful to several generations of graduate students in helping me to refine these notes, to Marc Chan for help in updating the text and software to be consistent with Carroll (2006), to Kiichi Tokuoka for drafting the section on structural estimation, and to Damiano Sandri for exceptionally insightful help in revising and updating the method of simulated moments estimation section. All errors are my own.

Department of Economics, Johns Hopkins University, Baltimore, MD 21218-2685, 410-516-7602 (office), ccarroll@jhu.edu.

1 Introduction

These notes describe the solution of several dynamic stochastic optimization problems for a consumer facing uninsurable labor income risk, with detailed intuitive discussion of various tricks and techniques useful in making the problem computationally tractable. The problem is solved with and without liquidity constraints, and the infinite horizon solution can be obtained as the limit of the finite horizon solution. After the basic consumption/saving problem with a constant interest rate is described and solved, an extension with portfolio choice between a riskless and a risky asset is also solved. Finally, a simple example of how to use these methods to estimate parameters like the coefficient of relative risk aversion using life cycle data is presented (following the seminal papers by Gourinchas and Parker (2002) and Cagetti (2003)). The tricks and techniques used in solving these problems have broad applicability to many dynamic stochastic optimization problems.

2 The Problem

Consider a consumer whose goal in period t is to maximize discounted utility from consumption over the remainder of a lifetime that ends in period T:

$$\max\ \mathbb{E}_t \left[\sum_{n=0}^{T-t} \beth^n \mathrm{u}(C_{t+n}) \right], \tag{1}$$

and whose circumstances evolve according to the transition equations[1]

$$\begin{aligned} A_t &= M_t - C_t & (2)\\ Y_{t+1} &= \mathbf{P}_{t+1}\Theta_{t+1} & (3)\\ M_{t+1} &= \mathsf{R}_{t+1}A_t + Y_{t+1} & (4) \end{aligned}$$

where

$\beth$ – time discount factor
M_t – resources available for consumption ('cash-on-hand')
A_t – assets after all actions have been taken in period t
C_t – consumption in period t
R_t – interest factor $(1+\mathsf{r})$ from period $t-1$ to t
$\mathrm{u}(C)$ – utility derived from consumption
Y_t – noncapital income in period t
$\mathbf{P}_t$ – 'permanent labor income' in period t.

[1]These transition equations are often combined into a single expression, but it is useful to disarticulate them so that it is clear that several distinct processes (stochastic shocks, intertemporal returns, income growth) are involved.

The exogenous variables evolve as follows:

$$\mathsf{R}_t = \mathsf{R} \quad \text{- Constant interest factor} = 1 + \mathsf{r} \tag{5}$$

$$\mathbf{P}_{t+1} = \Gamma_{t+1}\mathbf{P}_t \quad \text{- Permanent labor income dynamics} \tag{6}$$

$$\log\ \Theta \sim \mathcal{N}(-\sigma_\theta^2/2, \sigma_\theta^2) \quad \text{- Lognormally distributed transitory shocks} \tag{7}$$

Using the fact that if $\log \Phi \sim \mathcal{N}(\phi, \sigma_\phi^2)$ then $\log \mathbb{E}[\Phi] = \phi + \sigma_\phi^2/2$, the assumption about the distribution of Θ guarantees that $\log \mathbb{E}[\Theta] = 0$ which means that $\mathbb{E}[\Theta] = 1$.[2]

Equation (6) indicates that we are assuming that the average profile of income growth over the lifetime Γ_t is nonstochastic (to allow, for example, for typical career wage paths).[3]

Finally, we will assume that the utility function is of the CRRA form, $\mathrm{u}(C) = C^{1-\rho}/(1-\rho)$.

As is well known, this problem can be rewritten in the recursive (Bellman equation) form

$$\mathrm{V}_t(M_t, \mathbf{P}_t) = \max_{C_t}\ \mathrm{u}(C_t) + \beth\,\mathbb{E}_t[\mathrm{V}_{t+1}(M_{t+1}, \mathbf{P}_{t+1})] \tag{8}$$

subject to the Dynamic Budget Constraint (DBC) (2)-(4) as above, where V_t measures total expected discounted utility from behaving optimally now and henceforth.

3 Renormalization

Probably the single most important method for speeding the solution of such models is to see if you can redefine the problem in a way that reduces the number of state variables. In the consumption problem at hand the obvious idea is to see whether the problem can be rewritten in terms of the ratio of various variables to permanent labor income $\mathbf{P}_t$.

Consider the problem in the last period of life. Since there is no future, $\mathrm{V}_{T+1} = 0$, and the optimal plan in the last period of life is to consume everything, so that

$$\mathrm{V}_T(M_T, \mathbf{P}_T) = \frac{M_T^{1-\rho}}{1-\rho}. \tag{9}$$

Now define lower-case variables as the upper-case variable divided by the level of permanent income in the same period, so that, for example, $m_T = M_T/\mathbf{P}_T$, and define

[2] This fact is referred to as `[ELogNorm]` in the handout `MathFacts`, in the references as Carroll (Current); further citation to facts in that handout will be referenced simply by the name used in the handout for the fact in question, e.g. `[LogELogNorm]` is the name of the fact that implies that $\log \mathbb{E}[\Theta] = 0$.

[3] This assumes that there are no shocks to permanent income. A large literature finds that, in reality, permanent shocks to income exist and are quite large; such shocks therefore need to be incorporated into any 'serious' model (that is, one that hopes to match and explain empirical data), but the treatment of such shocks clutters the exposition without adding much to the intuition, so they have been omitted (until the last section of the notes, which shows how to match the model with empirical micro data). For the full treatment of the theory including permanent shocks, see Carroll (2011).

lower-case $\mathrm{v}_T(m_T) = \mathrm{u}(m_T)$. Using the fact that $\mathrm{u}(xy) = x^{1-\rho}\mathrm{u}(y)$, equation (9) can be rewritten as

$$\mathrm{V}_T(M_T, \mathbf{P}_T) \quad = \quad \mathbf{P}_T^{1-\rho}\frac{m_T^{1-\rho}}{1-\rho} = \mathbf{P}_{T-1}^{1-\rho}\Gamma^{1-\rho}m_T^{1-\rho} = \mathbf{P}_{T-1}^{1-\rho}\Gamma^{1-\rho}\mathrm{v}_T(m_T). \tag{10}$$

Now suppose we define a new optimization problem:

$$\begin{aligned} \mathrm{v}_t(m_t) &= \max_{\mathrm{c}_t} \ \mathrm{u}(c_t) + \underbrace{\Gamma_{t+1}^{1-\rho}\beth}_{\equiv\beta_{t+1}}\mathbb{E}_t[\mathrm{v}_{t+1}(m_{t+1})], && (11)\\ &\text{s.t.} \\ a_t &= m_t - c_t \\ m_{t+1} &= a_t \underbrace{(\mathsf{R}/\Gamma_{t+1})}_{\equiv\mathcal{R}_{t+1}} + \Theta_{t+1} && (12) \end{aligned}$$

where the accumulation equation can be seen to be the normalized version of the transition equation for M_{t+1}[4]. Then it is easy to show that for $t = T-1$, this expression for $\mathrm{v}_{T-1}(m_{T-1})$ implies

$$\mathrm{V}_{T-1}(M_{T-1}, \mathbf{P}_{T-1}) \quad = \quad \mathbf{P}_{T-1}^{1-\rho}\mathrm{v}_{T-1}(m_{T-1}) \tag{16}$$

and so on back to all earlier periods. Hence, if we solve the problem (11) which has only a single state variable, we can obtain the levels of the value function, consumption, and all other variables of interest simply by multiplying the results from this optimization problem by the appropriate function of $\mathbf{P}_t$, e.g. $\mathrm{C}_t(M_t, \mathbf{P}_t) = \mathbf{P}_t\mathrm{c}_t(M_t/\mathbf{P}_t)$ or $\mathrm{V}_t(M_t, \mathbf{P}_t) = \mathbf{P}_t^{1-\rho}\mathrm{v}(m_t)$. Hence we have reduced a problem with two state variables to a single-state-variable problem.

For some problems it will not be as obvious as for this one that there is an appropriate 'normalizing' variable, but many problems can be normalized if sufficient thought is given. For example, Valencia (2006) shows how a bank's optimization problem can be normalized by the level of the bank's productivity.

Henceforth we will take the single-state-variable problem defined in (11) as the problem under consideration, and to simplify matters further we will assume that permanent income remains constant $\Gamma_t = 1 \; \forall \; t$, yielding a constant value for β so that we will drop the time subscript from now on.

[4] Derivation:

$$\begin{aligned} M_{t+1}/\mathbf{P}_{t+1} &= (M_t - C_t)\mathsf{R}/\mathbf{P}_{t+1} + Y_{t+1}/\mathbf{P}_{t+1} && (13)\\ m_{t+1} &= \left(\frac{M_t}{\mathbf{P}_t} - \frac{C_t}{\mathbf{P}_t}\right)\mathsf{R}\frac{\mathbf{P}_t}{\mathbf{P}_{t+1}} + \frac{Y_{t+1}}{\mathbf{P}_{t+1}} && (14)\\ &= \underbrace{(m_t - c_t)}_{a_t}(\mathsf{R}/\Gamma_{t+1}) + \Theta_{t+1}. && (15) \end{aligned}$$

4 The Usual Theory, and Some Notation

The usual theoretical analysis of this problem proceeds as follows.

The first order condition for (11) with respect to c_t is

$$\mathrm{u}'(c_t) = \beta_{t+1}\,\mathbb{E}_t[\mathcal{R}\mathrm{v}'_{t+1}(m_{t+1})] \tag{17}$$

and because the `Envelope` theorem tells us that

$$\mathrm{v}'_t(m_t) = \beta_{t+1}\,\mathbb{E}_t[\mathcal{R}\mathrm{v}'_{t+1}(m_{t+1})] \tag{18}$$

we can substitute the LHS of (18) for the RHS of (17) to get

$$\mathrm{u}'(c_t) = \mathrm{v}'_t(m_t) \tag{19}$$

and rolling this equation forward one period yields

$$\mathrm{u}'(c_{t+1}) = \mathrm{v}'_{t+1}(a_t\mathcal{R} + y_{t+1}) \tag{20}$$

and substituting in equation (17) gives us the Euler equation for consumption

$$\mathrm{u}'(c_t) = \beta_{t+1}\,\mathbb{E}_t[\mathcal{R}\mathrm{u}'(c_{t+1})]. \tag{21}$$

Now note that in equation (20) neither m_t nor c_t has any *direct* effect on $\mathbb{E}_t[\mathrm{v}_{t+1}]$ - it is only the difference between them (i.e. unconsumed market resources or 'savings' a_t) that matters. It is therefore possible (and will turn out to be convenient) to define a function

$$\hat{\mathfrak{v}}_t(a_t) = \beta_{t+1}\,\mathbb{E}_t[\mathrm{v}_{t+1}(\mathcal{R}a_t + \Theta_{t+1})] \tag{22}$$

which returns the expected $t+1$ value associated with ending period t with any given amount of assets. Note also that this definition implies that

$$\hat{\mathfrak{v}}'_t(a_t) = \beta_{t+1}\,\mathbb{E}_t[\mathcal{R}\mathrm{v}'_{t+1}(\mathcal{R}a_t + \Theta_{t+1})] \tag{23}$$

or, substituting from equation (20),

$$\hat{\mathfrak{v}}'_t(a_t) = \beta_{t+1}\,\mathbb{E}_t\left[\mathcal{R}\mathrm{u}'\left(\mathrm{c}_{t+1}(\mathcal{R}a_t + \Theta_{t+1})\right)\right]. \tag{24}$$

Finally, note that the first order condition (17) can be rewritten as

$$\mathrm{u}'(c_t) = \hat{\mathfrak{v}}'_t(m_t - c_t). \tag{25}$$

5 Solving the Next-To-Last Period

The problem in the second-to-last period of life is:

$$\mathrm{v}_{T-1} = \max_{\mathrm{c}_{T-1}} \mathrm{u}(c_{T-1}) + \beta\,\mathbb{E}_{T-1}\left[\mathrm{v}_T((m_{T-1} - c_{T-1})\mathcal{R} + \Theta_T)\right],$$

and substituting from the budget constraint, the definition of $\mathrm{u}(c)$, and the definition of the expectations operator, this becomes

$$\mathrm{v}_{T-1}(m_{T-1}) = \max_{\mathrm{c}_{T-1}} \frac{c_{T-1}^{1-\rho}}{1-\rho} + \beta \int_0^\infty \frac{((m_{T-1} - c_{T-1})\mathcal{R} + \Theta)^{1-\rho}}{1-\rho} d\mathcal{F}(\Theta).$$

where $\mathcal{F}$ is the cumulative distribution function for Θ.

In principle, this implicitly defines a function $\mathrm{c}_{T-1}(m_{T-1})$ that yields the optimal value of consumption in period $T-1$ for any given level of resources m_{T-1}. Unfortunately, however, there is no analytical solution to this maximization problem, and so for any given value of m_{T-1} we must use numerical computational tools to find the c_{T-1} that maximizes the expression. But this is excruciatingly slow because for every potential c_{T-1} to be considered, the integral must be calculated numerically, and numerical integration is *very* slow.

5.1 Discretizing the Distribution

The first time-saving step is therefore to construct a discrete approximation to the lognormal distribution that can be used in place of numerical integration of the true lognormal. An n-point approximation is calculated as follows.

Define a set of points on the $[0, 1]$ interval as $\sharp = \{0, 1/(n-1), 2/(n-1), \ldots, 1\}$. Call the inverse of the lognormal distribution $\mathcal{F}^{-1}_{\sharp}$, and define the points $\sharp^{-1}_i = \mathcal{F}^{-1}(\sharp_i)$. Then define

$$\Theta_i \quad = \quad \int_{\sharp^{-1}_{i-1}}^{\sharp^{-1}_i} \Theta \; d\mathcal{F}(\Theta). \tag{26}$$

The Θ_i represent the mean values of Θ in each of the regions bounded by the $\sharp^{-1}_i$ endpoints. The method is illustrated in figure 1. The curve represents the CDF of $\mathcal{F}_\Theta(\Theta)$, a lognormal distribution such that $\mathbb{E}[\Theta] = 1$ and $\sigma_\theta = 0.1$, and the dots represent the n equiprobable values of Θ_i which are used to approximate this distribution.[5]

The maximization problem can now be rewritten

$$\mathrm{v}_{T-1}(m_{T-1}) \quad = \quad \max_{c_{T-1}} \left\{ \frac{c_{T-1}^{1-\rho}}{1-\rho} + \beta \left(\frac{1}{n}\right) \sum_{i=1}^{n} \frac{((m_{T-1} - c_{T-1})\mathcal{R} + \Theta_i)^{1-\rho}}{1-\rho} \right\} \tag{27}$$

Analogously, recalling our definition of $\hat{\mathfrak{v}}_t(a_t)$, we now define

$$\mathfrak{v}_{T-1}(a_{T-1}) \quad = \quad \beta \left(\frac{1}{n}\right) \sum_{i=1}^{n} \frac{(\mathcal{R} a_{T-1} + \Theta_i)^{1-\rho}}{1-\rho} \tag{28}$$

so we can rewrite (27) in approximate form as

$$\mathrm{v}_{T-1}(m_{T-1}) \quad = \quad \max_{c_{T-1}} \left\{ \frac{c_{T-1}^{1-\rho}}{1-\rho} + \mathfrak{v}_{T-1}(m_{T-1} - c_{T-1}) \right\}. \tag{29}$$

[5] There are more sophisticated methods available (e.g. Gauss-Hermite quadrature), but the method described in the text is easy to understand, quick to calculate, and performs about as well as the sophisticated methods.

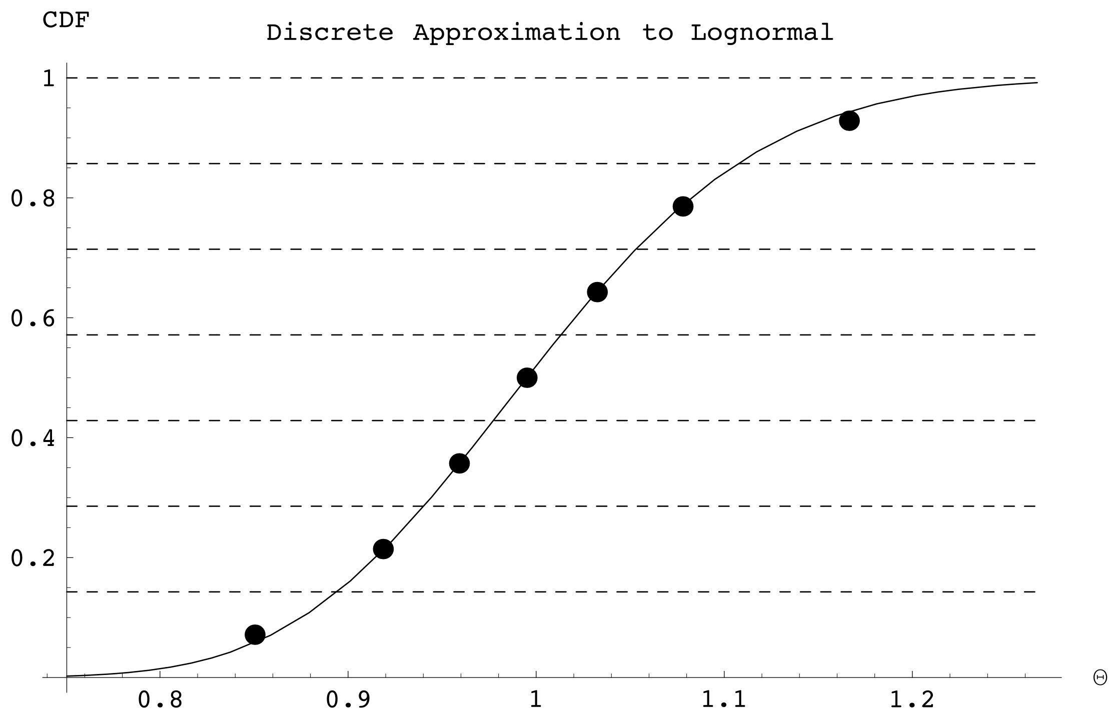

Figure 1 Discrete Approximation to Lognormal Distribution

5.2 The Approximate Consumption Function and the Approximate Value Function

Given a particular value of m_{T-1}, a numerical maximization routine can now be expected to find the c_{T-1} that maximizes (29) in a reasonable amount of time. A *Mathematica* program that solves exactly this problem called `2period.nb`. (The archive also contains parallel Matlab programs, but these notes will dwell on the specifics of the *Mathematica* implementation, which is shorter and simpler.)

The structure of the program is as follows. `setup_params.nb` reads in values for parameters like the coefficient of relative risk aversion and the time preference rate. `setup_shocks.nb` calculates the values for the Θ_i defined above (and puts those values, and the probability associated with each of them, in the vector variables `ΘVec` and `ΘVecProb`.) Finally, `setup_grids.nb` constructs a list of potential values of cash-on-hand and puts them in the vector variable μ`Vec` $= \{0, 1, 2, 3, 4, 5\}$.

Next, the program constructs behavior for the last period of life. Some commands deserve detailed discussion. `PDVMinY = {Min[ΘVec]}` is a list with the single element `Min[ΘVec]`, the minimum possible discounted human wealth at the beginning of period T. This is useful in determining the search range for the optimal level of consumption in the maximization problem. Due to the precautionary saving motive, the agent will never consume more than the minimum possible realization of his discounted total wealth.

The perfect foresight consumption function is also constructed. In *Mathematica*, functions can be saved as objects using the commands `#` and `&`. The former denotes the argument of the function, while the latter, placed at the end of the function, tells *Mathematica* that the function should be saved as an object. In the program, the last period perfect foresight consumption function is saved as an element in the list `cPF = {(# - 1 + PDVExpY[[1]]) MPCPF[[1]] &}`, where `PDVExpY[[1]] = 1` gives the expected discounted human wealth at the beginning of period T, and `MPCPF[[1]] = 1` gives the perfect foresight marginal propensity to consume in period T. Since `#` represents m in the model, the discounted total wealth is decomposed into discounted non-human wealth `# - 1` and discounted human wealth `PDVExpY`. The resulting formula is then $\bar{\mathrm{c}}_T = m_T$, indicating that the agent will consume all cash in the last period of life.

The infinite-horizon marginal propensity to save parameter

$$\lambda = (1/\mathcal{R})(\mathcal{R}\beta)^{1/\rho} \tag{30}$$

is also defined. This parameter is useful in generating perfect foresight MPCs in earlier periods of life. Detailed discussion can be found in Carroll (2011).

The program then constructs behavior for one iteration back from the last period of life. Using the *Mathematica* command `AppendTo`, the lists are redefined to include an additional element representing the relevant formulas in the second to last period of life. For example, `MPCPF` now has two elements. The second element, given by `1/(1 + λ/Last[MPCPF])`, is the perfect foresight marginal propensity to consume in $t = T - 1$. A proof given in Carroll (2011) shows that this is also a recurring formula that extends inductively to earlier periods.

Next, the program defines a function `𝔳[at_]` which is the exact implementation of (22): It returns the expectation of the value of behaving optimally in period T given any specific amount of assets at the end of $T-1$, a_{T-1}.

The heart of the program is the next expression. This expression loops over the values of the variable `μVec`, solving the maximization problem[6]

$$\max_{\texttt{c}} \quad \texttt{u[c] + }\mathfrak{v}\texttt{[}\mu\texttt{Vec[[i]]-c]} \tag{31}$$

for each of the i values of `μVec` (henceforth let's call these points $\mu_{i,T-1}$). The maximization routine returns two values: the maximized value, and the value of c which yields that maximized value. When the loop (the `Table` command) is finished, the variable `vAndcList` contains two lists, one with the values $v_{i,T-1}$ and the other with the consumption levels $\chi_{i,T-1}$ associated with the $\mu_{i,T-1}$.

5.3 Interpolating the Consumption Function

Now we use the first of the really convenient built-in features of *Mathematica*. Given data of the form $\{\{m_1, y_1\}, \{m_2, y_2\} \ldots, \{m_3, y_3\} \ldots\}$ *Mathematica* can create an object called an `InterpolatingFunction` which, when applied to an input m will yield the value of y which corresponds to a linear interpolation of the value of y from the points in the `InterpolatingFunction` object. We can therefore define a function $\hat{\mathrm{c}}_{T-1}(m_{T-1})$

[6]Actually, *Mathematica* implements minimization rather than maximization, so we find the minimum of the negative of this expression.

which, when called with an m_{T-1} that is equal to one of the points in $\mu\texttt{Vec}_i$ returns the associated value of $\chi_{i,T-1}$, and when called with a value of m_{T-1} that is not exactly equal to one of the $\mu\texttt{Vec}_i$, returns the value of c that represents a linear interpolation between the $\chi_{i,T-1}$ associated with the two $\mu\texttt{Vec}_i$ points nearest to m_{T-1}. Thus if the function is called with $m_{T-1} = 1.75$ and the nearest gridpoints are $\mu_{j,T-1} = 1$ and $\mu_{k,T-1} = 2$ then the value of c_{T-1} to be returned by the function would be $(0.25\chi_{j,T-1} + 0.75\chi_{k,T-1})$. We can define a numerical approximation to the value function $\hat{\mathrm{v}}_{T-1}(m_{T-1})$ in an exactly analogous way.

When asked to evaluate an `InterpolatingFunction` object at a gridpoint outside the set of gridpoints that define it, *Mathematica* constructs an extrapolation that simply continues the function at the same slope it had in the interval between the last two gridpoints. However, we can do better than this: Theory tells us as the level of m_{T-1} approaches infinity, the level of c_{T-1} approaches arbitrarily close to the level of consumption that would obtain if there were no labor income risk. Thus, designating the perfect foresight solution as $\bar{c}_{T-1}$ we know that

$$\lim_{m_{T-1}\to\infty} \mathrm{c}_{T-1}(m_{T-1}) \;=\; \bar{\mathrm{c}}_{T-1}(m_{T-1}) = (m_{T-1} + 1/\mathcal{R})\,\kappa_{T-1}, \tag{32}$$

where κ_{T-1} is the marginal propensity to consume in the perfect foresight problem in the second-to-last period of life:

$$\kappa_{T-1} \;=\; \left(\frac{1}{1+\mathsf{R}^{-1}(\mathsf{R}\beta)^{1/\rho}}\right) = \frac{1}{1+\lambda/\kappa_T} \tag{33}$$

where $\kappa_T = 1$ and λ was defined in equation (30).

This means we can define some very large value in the variable $\mu\texttt{Huge}$, augment $\mu\texttt{Vec}$ with $\mu\texttt{Huge}$ and augment the set of $\chi_{i,T-1}$ values with $\bar{\mathrm{c}}_{T-1}(\mu\texttt{Huge})$. Similarly, it can be shown that the value function of the perfect foresight problem is

$$\bar{\mathrm{v}}_{T-1}(m_{T-1}) \;=\; \left(\frac{\bar{\mathrm{c}}_{T-1}^{1-\rho}}{(1-\rho)\kappa_{T-1}}\right), \tag{34}$$

so we can similarly augment the set of $v_{i,T-1}$ points to include v_{T-1} at $\mu_{i,T-1} = \mu\texttt{Huge}$, and *Mathematica* will not need to extrapolate if we graph the function over a reasonable interval.

Figures 2 and 3 show plots of the `cTm1` and `VTm1` `InterpolatingFunctions` that are generated by the program. While the `cTm1` function looks very smooth, the fact that the `VTm1` function is a set of line segments is very evident.

5.4 Interpolating Expectations

The program `2period.nb` works fine in the sense that it generates a good approximation to the true optimal consumption function. However, there is a clear inefficiency in the program: Because it uses equation (27) to solve, for every value of m_{T-1}, in order to find the optimal c_{T-1} the program must calculate the utility consequences of various possible choices of c_{T-1}. But for any given value of a_{T-1}, there is a good chance that the program may end up calculating the corresponding value many times while maximizing utility

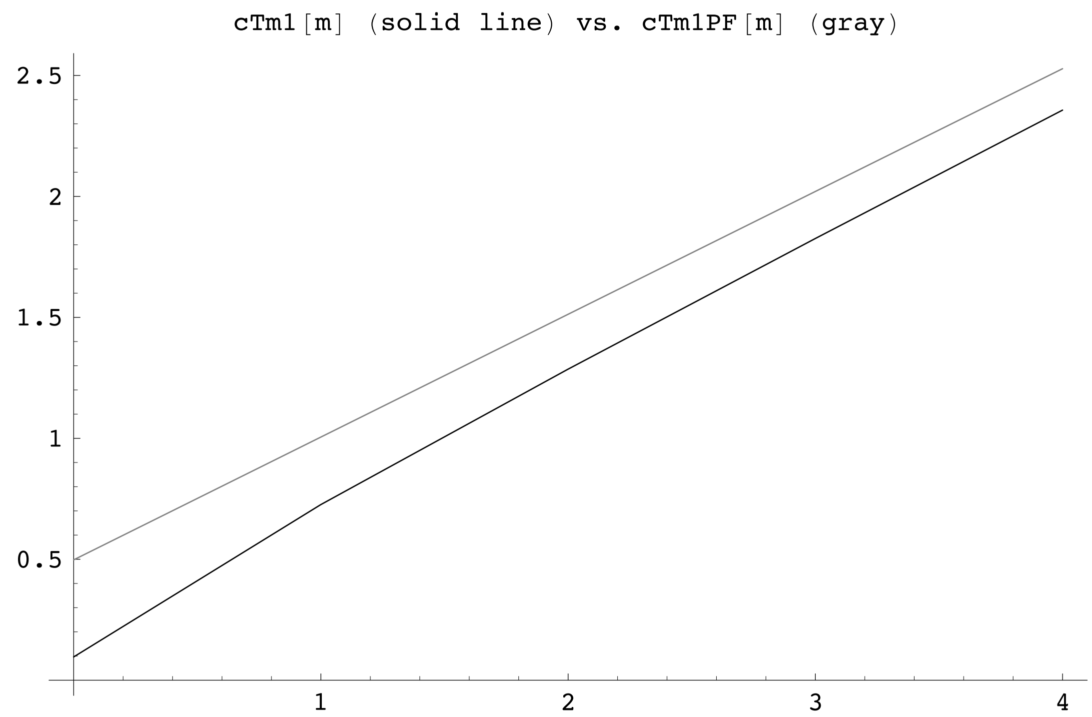

Figure 2 $\hat{\mathrm{c}}_{T-1}(m_{T-1})$

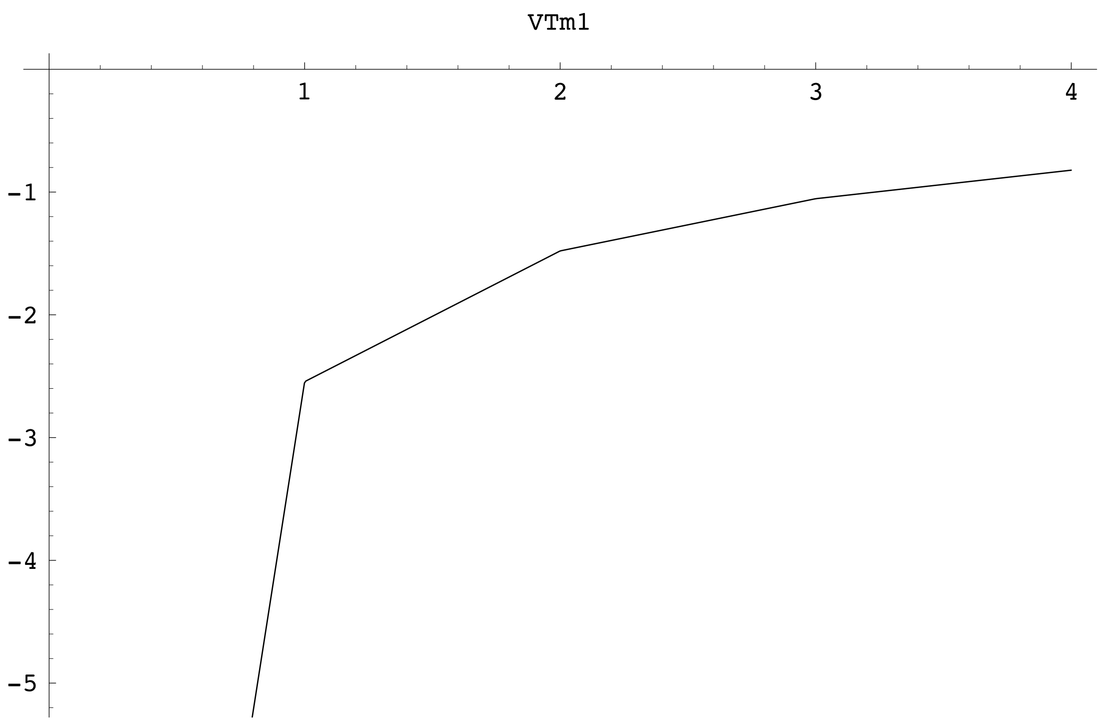

Figure 3 $\hat{\mathrm{v}}_{T-1}(m_{T-1})$

from different m_{T-1}'s. For example, it is quite likely that the program will calculate the value of saving exactly zero dozens of times. It would be much more efficient if the program could make that calculation once and then merely recall the value when it is needed again.

This can be achieved using the same interpolation technique used above to construct a numerical value function: construct a grid of possible values for saving at time $T-1$, α`Vec`, designating the specific points $\alpha_{i,T-1}$; and for each of these values of $\alpha_{i,T-1}$, calculate $\mathfrak{v}_{i,T-1} = \mathfrak{v}_{T-1}(\alpha_{i,T-1})$ using equation (22); and construct an `InterpolatingFunction` object $\hat{\mathfrak{v}}_{T-1}(a_{T-1})$ from the list of values $\{\{\alpha_{1,T-1}, \mathfrak{v}_{1,T-1}\}, \{\alpha_{2,T-1}, \mathfrak{v}_{2,T-1}\} \ldots\}$. Finally, for extrapolation beyond the predefined α`Vec`, defining a point α`Huge` that is much larger than the largest gridpoint we know that

$$\lim_{a_{T-1}\rightarrow\infty} \mathfrak{v}_{T-1}(a_t) = \beta \mathrm{v}_T(\mathbb{E}_{T-1}[m_T]) \tag{35}$$

$$= \beta \left(\frac{(\mathcal{R}a_{T-1}+1)^{1-\rho}}{1-\rho} \right) \tag{36}$$

so we augment the $\mathfrak{v}_{i,T-1}$ points with the value of this function at $\alpha_{i,T-1} =\alpha$`Huge`.

Thus, we are now interpolating for the expected value of saving; the program `2periodIntexp.nb` solves this problem. Figure 4 compares the true value function to the `InterpolatingFunction` approximation; the two are of course identical at the gridpoints chosen for a_{T-1} and they appear reasonably close except in the region below $m_{T-1} = 1$.[7]

Nevertheless, the resulting consumption rule obtained when $\hat{\mathfrak{v}}(a_{T-1})$ is used instead of $\mathfrak{v}_{T-1}(a_{T-1})$ are surprisingly bad, as shown in figure 5. For example, when m_{T-1} goes from 2 to 3, $\hat{c}_{T-1}$ goes from about 1 to about 2, yet when m_{T-1} goes from 3 to 4, $\hat{c}_{T-1}$ goes from about 2 to about 2.05. The function fails even to be strictly concave, which is problematic because Carroll and Kimball (1996) prove that the correct consumption function is strictly concave in problems like this one.

5.5 Value Function Versus First Order Condition

Loosely speaking, the problem reflects the fact that *behavior* is determined by the *marginal* value function, not by the level of the value function. To see this, recall that a quadratic utility function exhibits risk aversion because

$$\mathbb{E}[-(c-\cancel{c})^2] < -(\mathbb{E}[c]-\cancel{c})^2. \tag{37}$$

for c less than the 'bliss point' of $\cancel{c}$ even though the consumption/saving *behavior* of consumers with quadratic utility is unaffected by risk. The reason behavior is unaffected by risk is that behavior is determined by the first order condition, which depends on

[7] This is one implementation of the basic idea which has been called 'the method of parameterized expectations.' Here we are using a nonparametric method, but the basic idea is the same. The technique would be better labelled the method of approximated expectations.

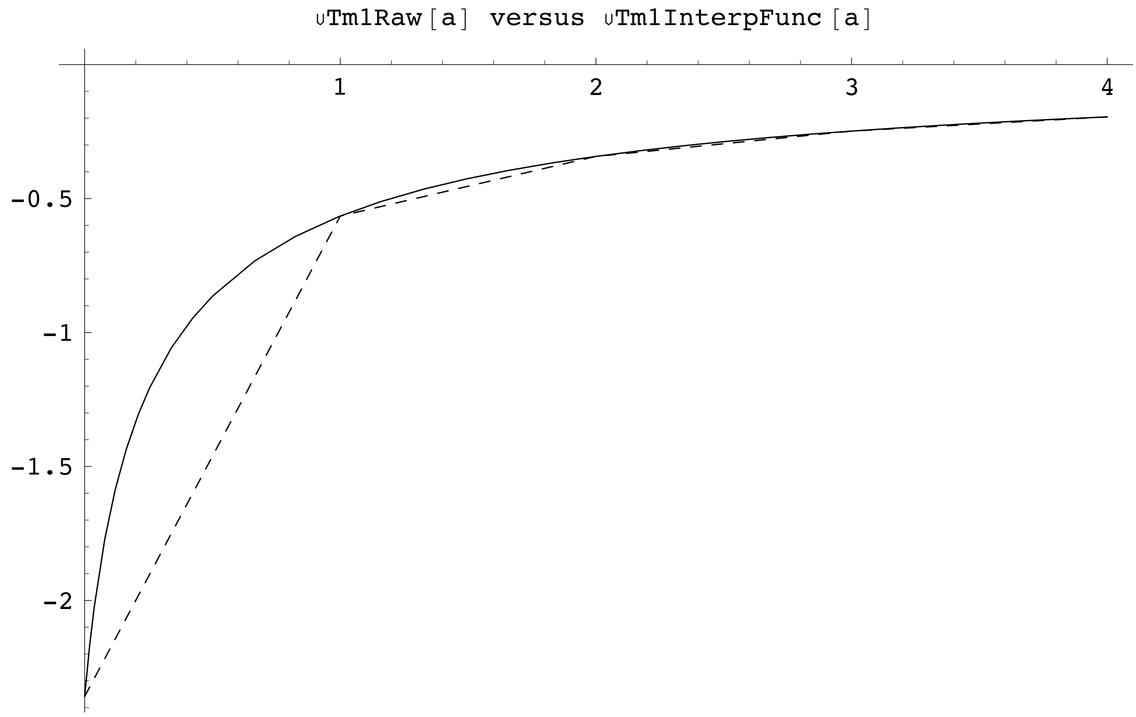

Figure 4 $\mathfrak{v}_{T-1}(a_{T-1})$ (solid) versus $\hat{\mathfrak{v}}_{T-1}(a_{T-1})$ (dashing)

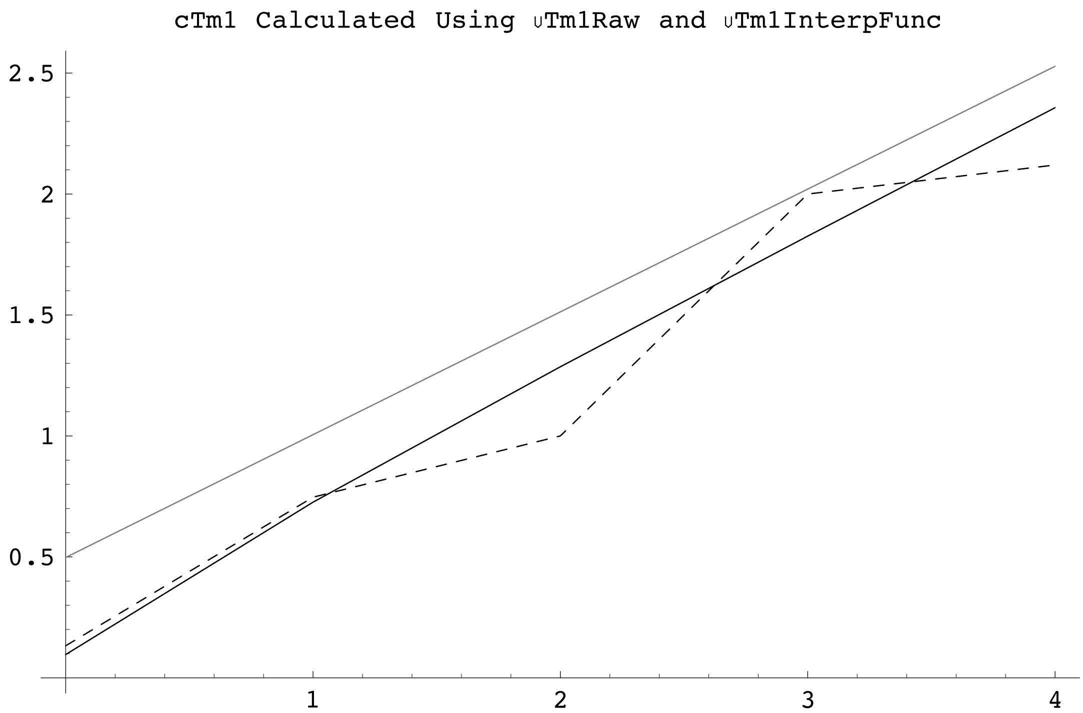

Figure 5 $\mathrm{c}_{T-1}(m_{T-1})$ (solid) versus $\hat{\mathrm{c}}_{T-1}(m_{T-1})$ (dashing)

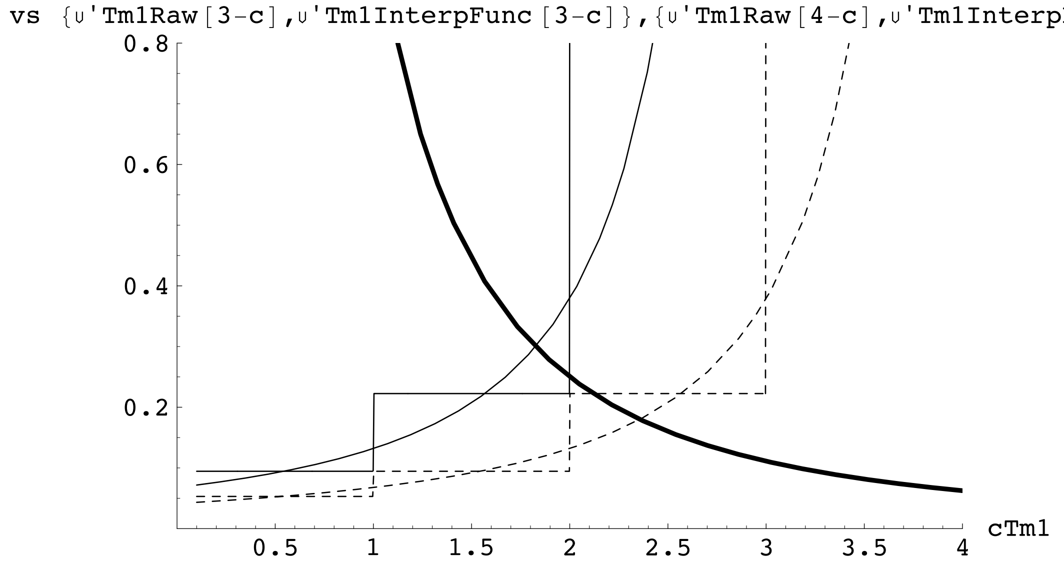

Figure 6 $\mathrm{u}'(c)$ versus $\mathfrak{v}'_{T-1}(3-c), \mathfrak{v}'_{T-1}(4-c), \hat{\mathfrak{v}}'_{T-1}(3-c), \hat{\mathfrak{v}}'_{T-1}(4-c)$

marginal utility, and marginal utility is unaffected by risk:

$$\mathbb{E}[-2(c-\cancel{c})] = -2(\mathbb{E}[c]-\cancel{c}). \tag{38}$$

Intuitively speaking, if one's goal is to accurately capture behavior that is governed by marginal utility or the marginal value function, numerical techniques that approximate the *marginal* value function are likely to lead to a more accurate approximation to optimal behavior than techniques that operate on the level of the value function.

The first order condition of the maximization problem in period $T-1$ is that:

$$\mathrm{u}'(c_{T-1}) = \beta\,\mathbb{E}_{T-1}[\mathcal{R}\mathrm{u}'(c_T)] \tag{39}$$

$$c_{T-1}^{-\rho} = \beta\mathcal{R}\left(\frac{1}{n}\right)\sum_{i=1}^{n}\left[\mathcal{R}(m_{T-1}-c_{T-1})+\Theta_i\right]^{-\rho}. \tag{40}$$

The downward-sloping curve in figure 6 shows the value of $c_{T-1}^{-\rho}$ for our baseline parameter values for $0 \le c_{T-1} \le 4$ (the horizontal axis). The solid upward-sloping curve shows the value of the RHS of (40) as a function of c_{T-1} under the assumption that $m_{T-1} = 3$. Constructing this figure is rather time-consuming, because for every value of c_{T-1} plotted we must calculate the RHS of (40). The value of c_{T-1} for which the RHS and LHS of (40) are equal is the optimal level of consumption given that $m_{T-1} = 3$, so the intersection of the downward- sloping and the upward-sloping curves gives the optimal value of c_{T-1}. As we can see, the two curves intersect just below $c_{T-1} = 2$. Similarly, the upward-sloping dashing curve shows the expected value of the

RHS of (40) under the assumption that $m_{T-1} = 4$, and the intersection of this curve with $\mathrm{u}'(c_{T-1})$ yields the optimal level of consumption if $m_{T-1} = 4$. These two curves intersect slightly below $c_{T-1} = 2.5$. Thus, increasing m_{T-1} from 3 to 4 increases optimal consumption by about 0.5.

Now consider the derivative of our function $\hat{\mathfrak{v}}_{T-1}(a_{T-1})$. Because we have constructed $\hat{\mathfrak{v}}_{T-1}$ as a linear interpolation, the slope of $\hat{\mathfrak{v}}_{T-1}(a_{T-1})$ between any two adjacent points $\{\alpha_{i,T-1}, \alpha_{i+1,T-1}\}$ is constant. The level of the slope immediately below any particular gridpoint is different, of course, from the slope above that gridpoint, a fact which implies that the derivative of $\hat{\mathfrak{v}}_{T-1}(a_{T-1})$ follows a step function.

The solid-line step function in figure 6 depicts the actual value of $\hat{\mathfrak{v}}'_{T-1}(3-c_{T-1})$. When we attempt to find optimal values of c_{T-1} given m_{T-1} using $\hat{\mathfrak{v}}_{T-1}(a_{T-1})$, the numerical optimization routine will return the c_{T-1} for which $\mathrm{u}'(c_{T-1}) = \hat{\mathfrak{v}}'_{T-1}(m_{T-1}-c_{T-1})$. Thus, for $m_{T-1} = 3$ the program will return the value of c_{T-1} for which the downward-sloping $\mathrm{u}'(c_{T-1})$ curve intersects with the $\hat{\mathfrak{v}}'_{T-1}(3 - c_{T-1})$; as the diagram shows, this value is exactly equal to 2. Similarly, if we ask the routine to find the optimal c_{T-1} for $m_{T-1} = 4$, it finds the point of intersection of $\mathrm{u}'(c_{T-1})$ with $\hat{\mathfrak{v}}'_{T-1}(4 - c_{T-1})$; and as the diagram shows, this intersection is only slightly above 2. Hence, this figure illustrates why the numerical consumption function plotted earlier returned values very close to $c_{T-1} = 2$ for both $m_{T-1} = 3$ and $m_{T-1} = 4$.

We would obtain much better estimates of the point of intersection between $\mathrm{u}'(c_{T-1})$ and $\mathfrak{v}'_{T-1}(m_{T-1} - c_{T-1})$ if our estimate of $\hat{\mathfrak{v}}'_{T-1}$ were not a step function. In particular, we already know how to construct linear interpolations to functions, so the next idea we pursue is to construct a linear interpolating approximation to the *expected marginal value of saving* function $\mathfrak{v}'$. That is, we calculate the value of

$$\mathfrak{v}'_{T-1}(a_{T-1}) \quad = \quad \beta\mathcal{R}\left(\frac{1}{n}\right)\sum_{i=1}^{n}\left[\mathcal{R}a_{T-1} + \Theta_i\right]^{-\rho} \tag{41}$$

at the points in α`Vec` yielding $\{\{\alpha_{1,T-1}, \mathfrak{v}'_{1,T-1}\}, \{\alpha_{2,T-1}, \mathfrak{v}'_{2,T-1}\}\ldots\}$ and construct $\hat{\mathfrak{v}}'_{T-1}(a_{T-1})$ as the linear interpolating function that fits this set of points.

The program file `2periodIntExpFOC.nb` therefore defines a function `𝔳P[at_]` as the embodiment of equation (41), and constructs the `InterpolatingFunction` as described above. The results are shown in figure 7. The linear interpolating approximation looks roughly as good (or bad) for the marginal value function as it was for the level of the value function. However, figure 8 shows that the new consumption function (long dashes) is a considerably better approximation of the true consumption function (solid) than was the consumption function obtained by approximating the level of the value function (short dashes).

5.6 Transformation

However, even the new function diverges noticeably from the optimum. That is because the linear interpolation does an increasingly poor job of capturing the nonlinearity of the $\mathfrak{v}'_{T-1}(a_{T-1})$ function at lower and lower levels of a.

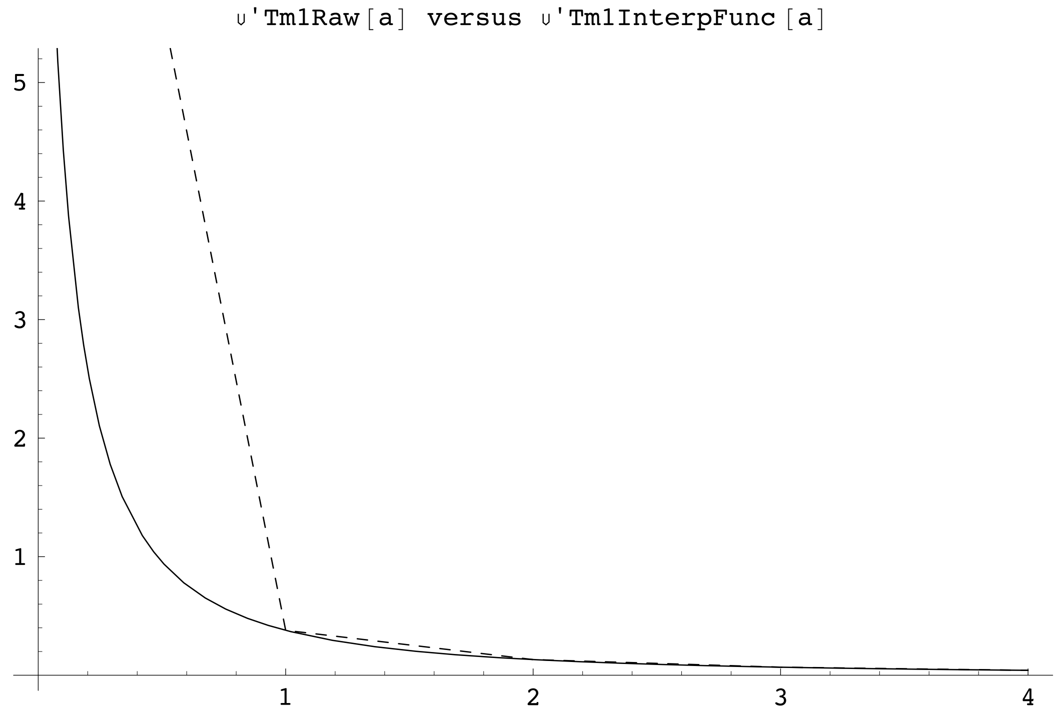

Figure 7 $\mathfrak{v}'_{T-1}(a_{T-1})$ versus $\hat{\mathfrak{v}}'_{T-1}(a_{T-1})$

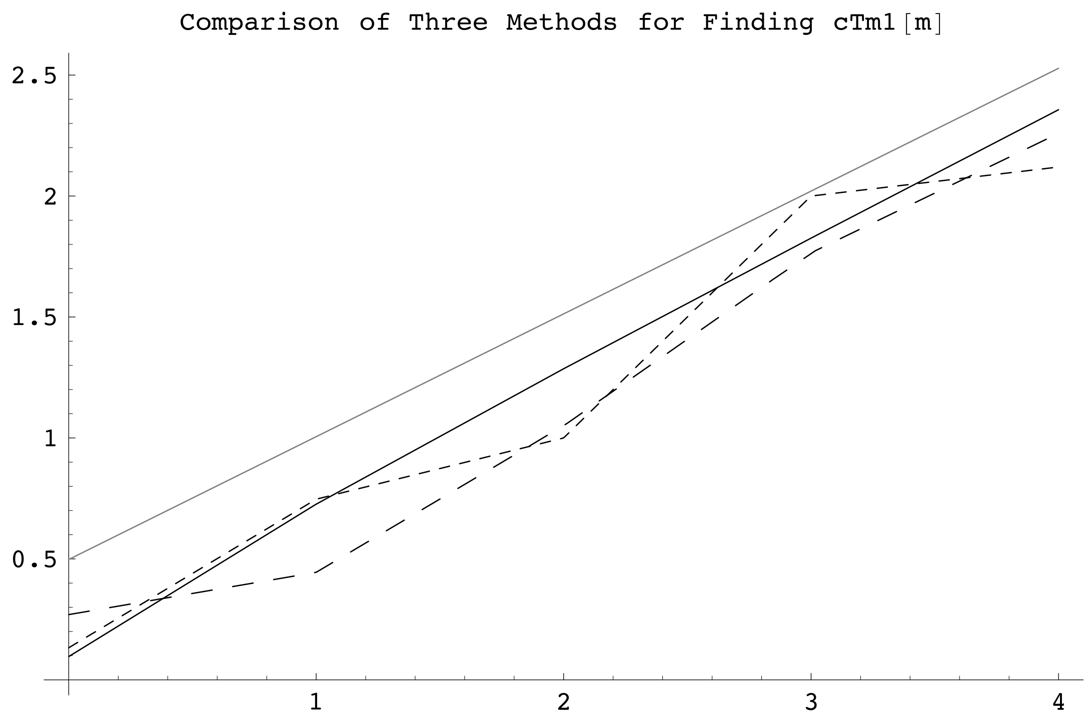

Figure 8 $\mathrm{c}_{T-1}(m_{T-1})$ (solid) Versus Two Methods for Constructing $\hat{\mathrm{c}}_{T-1}(m_{T-1})$

This is where we bring our next trick into play.

Consider the case where $\mathcal{R} = \beta = 1$ and there is no uncertainty (that is, we know for sure that income next period will be $\mathbb{E}[\Theta] = 1$). The Euler equation linking the marginal utility of consumption in period $T-1$ to the marginal utility in period T is:

$$c_{T-1}^{-\rho} \quad = \quad c_T^{-\rho} \tag{42}$$

Now recall that in the case of the problem with no uncertainty and with $\beta = \mathcal{R} = 1$, the optimal solution is to spend half of total lifetime resources in period $T-1$ and the remainder in period T. Since total resources are known with certainty to be $m_{T-1} + 1$, and since we know that $\mathrm{v}'_{T-1}(m_{T-1}) = \mathrm{u}'(c_{T-1})$ this implies that

$$\mathrm{v}'_{T-1}(m_{T-1}) = \left(\frac{m_{T-1} + 1}{2}\right)^{-\rho}. \tag{43}$$

Note that for ρ much above 1 this becomes a highly nonlinear function. However, if we raise both sides of the equation to the power $(-1/\rho)$ it becomes a linear function:

$$[\mathrm{v}'_{T-1}(m_{T-1})]^{-1/\rho} \quad = \quad \frac{m_{T-1} + 1}{2}. \tag{44}$$

This is a specific example of a general phenomenon: Theoretical results cited in Carroll and Kimball (1996) establish that under perfect certainty, if the period-by-period marginal utility function is of the form $c_t^{-\rho}$, the marginal value function will be of the form $(\alpha m_t + \zeta)^{-\rho}$. This means that if we were solving the perfect certainty problem numerically, we could always calculate a numerically exact interpolation. To put this in intuitive terms, the problem we are facing is that the marginal utility function is highly nonlinear. But we have a great solution to that problem, because we know that much of that nonlinearity springs from the fact that we are raising something to the power $-\rho$. In effect, we can 'unwind' all of the nonlinearity owing to that operation and the remaining nonlinearity will not be nearly so great (in the perfect foresight case, no nonlinearity at all will remain). Specifically, we can define

$$\mathfrak{c}_t(a_t) \quad \equiv \quad [\mathfrak{v}'_t(a_t)]^{-1/\rho}. \tag{45}$$

which will be linear.

Thus, our procedure is to calculate the values of $\mathfrak{c}_{i,T-1}$ at each of the a_{T-1} gridpoints and to construct $\hat{\mathfrak{c}}_{T-1}(a_{T-1})$ as the interpolating function connecting these points.

Indeed, this transformation solves another problem that I have glossed over to this point. To see the nature of the problem, designate $\underline{\Theta}$ as the lowest possible realization of income in the last period of life. Now consider what happens to $\mathfrak{v}_{T-1}(a_{T-1})$ as a_{T-1} approaches the value $\underline{a} = -\underline{\Theta}/\mathcal{R}$.

From (41) we have

$$\lim_{a_{T-1} \downarrow -\underline{a}} \mathfrak{v}'_{T-1}(a_{T-1}) \quad = \quad \lim_{a_{T-1} \downarrow -\underline{a}} \beta\mathcal{R}\left(\frac{1}{n}\right)\sum_{i=1}^{n}\left(\mathcal{R}a_{T-1} + \Theta_i\right)^{-\rho}. \tag{46}$$

But exactly at $a_{T-1} = \underline{a}$ the first term in the summation would be $1/0^{\rho}$ which is infinity. The reason for this is simple: $-\underline{a}$ is the PDV of the minimum possible realization of income in period T. Thus, if the consumer borrows an amount greater than or equal to $\underline{a}$ and then draws the worst possible income shock in period T, he will have to consume zero (or a negative amount), which yields $-\infty$ utility and ∞ marginal utility.

This feature of the problem means that the consumer will face what can be thought of as a 'self-imposed' liquidity constraint that results from the precautionary saving motive. The consumer will never borrow an amount greater than or equal to $\underline{a}$. This is a 'self-imposed' constraint in the sense that if the utility function were different (say, Constant Absolute Risk Aversion), the consumer would be willing to borrow more than $\underline{a}$ because having zero or negative consumption does not yield negative infinite utility.

The self-imposed constraint, however, cannot be captured well when the $\mathfrak{v}'_{T-1}$ function is approximated by a piecewise linear function like $\hat{\mathfrak{v}}'_{T-1}$, because a linear approximation can never reach the correct gridpoint for $\mathfrak{v}'_{T-1}(-\underline{a}) = \infty$. To see what will happen instead, note that if the smallest value in α`Vec` is greater than $-\underline{a}$, then when the function is evaluated at some value less than $-\underline{a}$ the function will return a positive finite number; in particular, even evaluating at $\hat{\mathfrak{v}}'(-\underline{a})$ or below will yield a finite number. This means that the precautionary saving motive is understated by an arbitrarily large amount as the level of assets approaches its true theoretical minimum possible amount.

But if we define

$$\hat{\hat{\mathfrak{v}}}'_{T-1}(a_{T-1}) = [\hat{\mathfrak{c}}_{T-1}(a_{T-1})]^{-\rho}, \tag{47}$$

then our approximation to $\mathfrak{v}'_{T-1}$ has the property that if evaluated at any value of a_{T-1} greater than $-\underline{a}$ it yields a finite number, but that number approaches arbitrarily close to infinity as a approaches $-\underline{a}$.

These observations motivate a further change in the program; previously there was no compelling way to chose the grid values of a. Now, however, there is a plausible candidate at least for the lowest value of a in the grid: `Min[`α`Vec] =` $-\underline{a}$.

Figure 10 shows that when we calculate $\hat{\hat{\mathfrak{v}}}'_{T-1}(a_{T-1})$ as $[\hat{\mathfrak{c}}_{T-1}(a_{T-1})]^{-\rho}$ (dashing line) we obtain a *much* closer approximation to the true function $\mathfrak{v}'_{T-1}(a_{T-1})$ (solid line) than we did in the previous program which did not do the transformation (figure 7).

The obvious next step would be to use our constructed $\hat{\mathfrak{v}}'_{T-1}(a_{T-1})$ function to find the value of the consumption rule at our previously-chosen set of gridpoints μ`Vec` via the first order condition

$$\mathrm{u}'(c_{T-1}) \quad = \quad \mathfrak{v}'_{T-1}(m_{T-1} - c_{T-1}). \tag{48}$$

This can certainly be done. However, finding the value of c_{T-1} that solves this equation for arbitrary values of m_{T-1} is computationally expensive, because it involves finding the root of a numerically-defined function; this is a slow, slow operation (computationally expensive).

Conveniently, there is an alternative that lets us skip this burdensome step. The alternative can be understood by noting that any arbitrary value of α_i (greater than its minimum feasible value $-\underline{a}$) will be associated with some marginal valuation as of the end of period $T-1$, and the further observation that it is trivial to find the value of c

that yields the same marginal valuation from the first order condition,

$$\begin{aligned}
\mathrm{u}'(\chi_{i,T-1}) &= \mathfrak{v}'_{T-1}(\alpha_{i,T-1}) && (49)\\
\chi_{i,T-1} &= \mathrm{u}'^{-1}(\mathfrak{v}'_{T-1}(\alpha_{i,T-1})) && (50)\\
&= (\mathfrak{v}'_{T-1}(\alpha_{i,T-1}))^{-1/\rho} && (51)\\
&= \mathfrak{c}_{T-1}(\alpha_{i,T-1}) && (52)\\
&= \mathfrak{c}_{i,T-1} && (53)\\
\mu_{i,T-1} &= \chi_{i,T-1} + \alpha_{i,T-1}. && (54)
\end{aligned}$$

Thus, we can generate a set of $\mu_{i,T-1}$ and $\chi_{i,T-1}$ pairs that can be interpolated between in order to yield $\hat{c}(m_{T-1})$ at virtually zero computational cost once we have the $\hat{\mathfrak{c}}_{i,T-1}$ values in hand![8] One might worry about whether the $\{m, c\}$ points arrived at this way will provide a good representation of the consumption function as a whole, but in practice they work very well.

As illustration, look at figure 9 generated by the program which solves the optimization problem using these transformations, `2periodIntexpFOCInv.nb.` (We use the text `\[GothicC]` to stand for $\mathfrak{c}$ in the programs). The solid line calculates the exact numerical value of $\mathfrak{c}_{T-1}(a_{T-1})$ while the dashing line is the linear interpolating approximation $\hat{\mathfrak{c}}_{T-1}(a_{T-1})$. This figure illustrates the value of the transformation: the true figure is very close to linear, and so the linear approximation is almost indistinguishable from the true value except at the very lowest values of a_{T-1}.

Figure 11 shows that the consumption function that emerges from this program is now very close to the 'true' consumption function all the way down to $m_{T-1} = 0$. Although you can still see small deviations, the function is now being constructed using literally thousands of times fewer computations than would have been required without all these tricks. We could increase the density of the gridpoints for a_{T-1} by a factor of ten and still be able to solve the problem vastly more quickly than it can be solved without these techniques.

A further trick is motivated by the observation that the $\hat{\mathfrak{c}}_{T-1}$ and consumption functions are farthest off from the true values at low values of a and m respectively. This suggests that it would have been more efficient to use an α`Vec` that had points that were not equally spaced but rather concentrated on the low end of possible values. This can be accomplished by modifying the `setup_grids.nb` program. One approach that seems to work well in practice is to specify a minimum and a maximum value of a and then to construct the individual points using a triple-exponential growth rate between these two values. Results from a version of the programs that does this are presented in the next figures.

Note that the appropriate transformation for v_{T-1} is different from that for v'_{T-1}. The idea is to transform the object in such a way that, in a perfect-certainty world, the resulting function would be linear. It turns out that in the perfect certainty world, the value function corresponding to a maximization problem with CRRA period utility functions is itself a CRRA function with the same parameter; that is, $\mathrm{v}_t(m_t) = \frac{(\alpha m_t+\zeta)^{1-\rho}}{1-\rho}$

[8] This is the essential point of Carroll (2006).

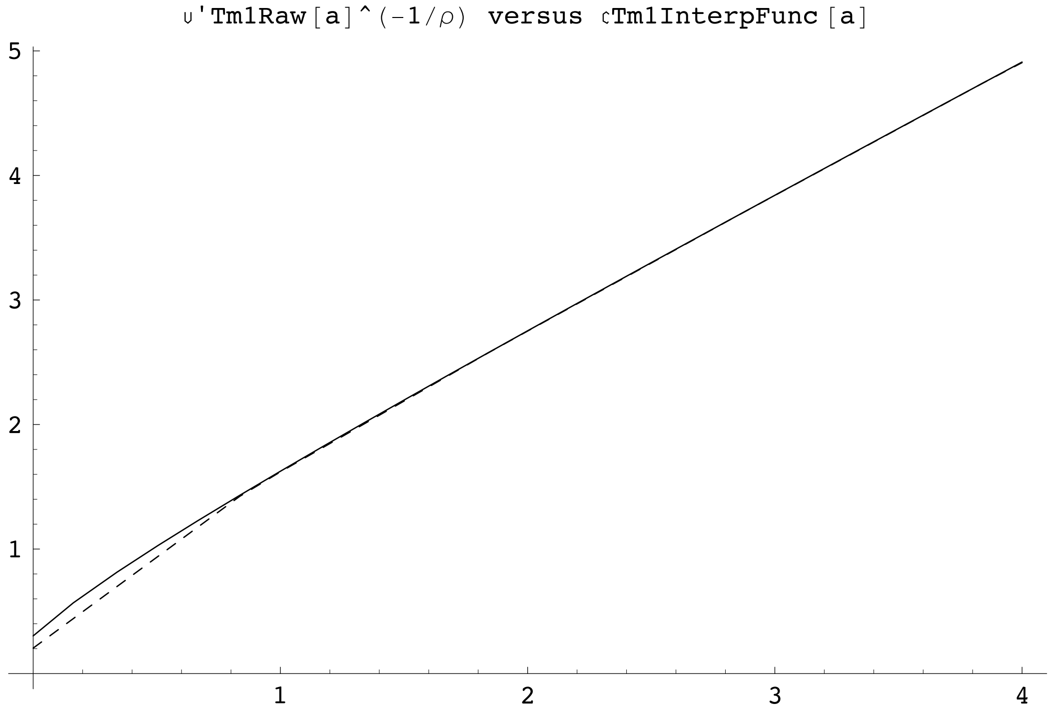

Figure 9 $\mathfrak{c}_{T-1}(a_{T-1})$ versus $\hat{\mathfrak{c}}_{T-1}(a_{T-1})$

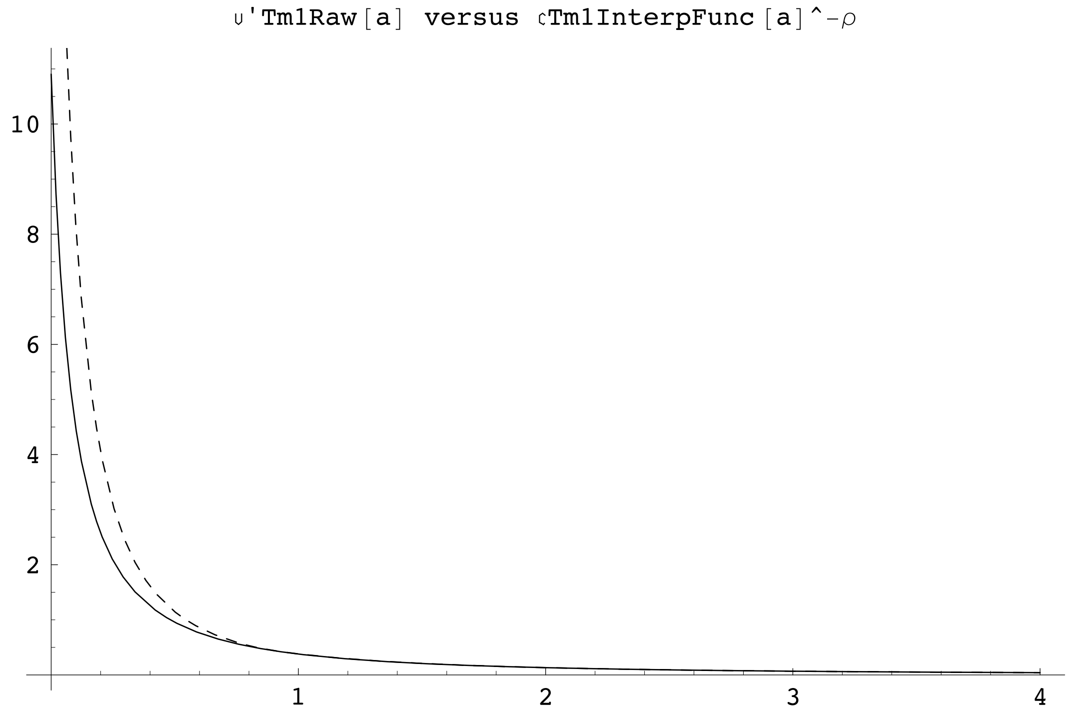

Figure 10 $\mathfrak{v}'_{T-1}(a_{T-1})$ vs. $\hat{\hat{\mathfrak{v}}}'_{T-1}(a_{T-1})$ Constructed Using $\hat{\mathfrak{c}}_{T-1}(a_{T-1})$

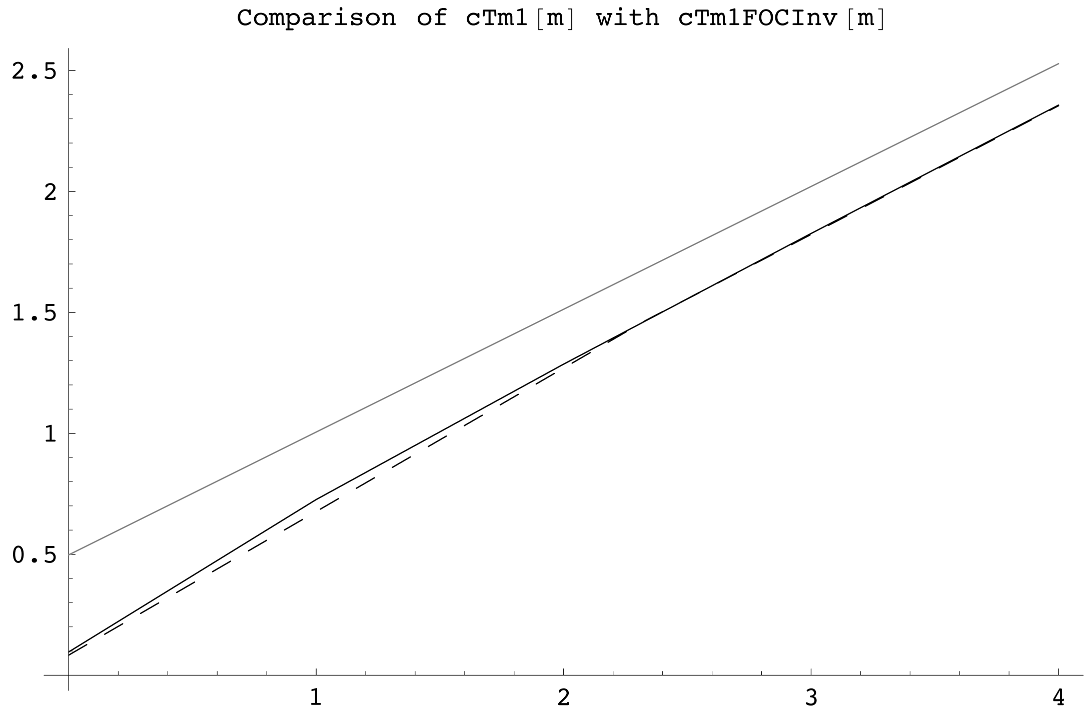

Figure 11 $\mathrm{c}(m_{T-1})$ vs. $\hat{\mathrm{c}}(m_{T-1})$ Constructed Using $\hat{\mathfrak{c}}_{T-1}(a_{T-1})$

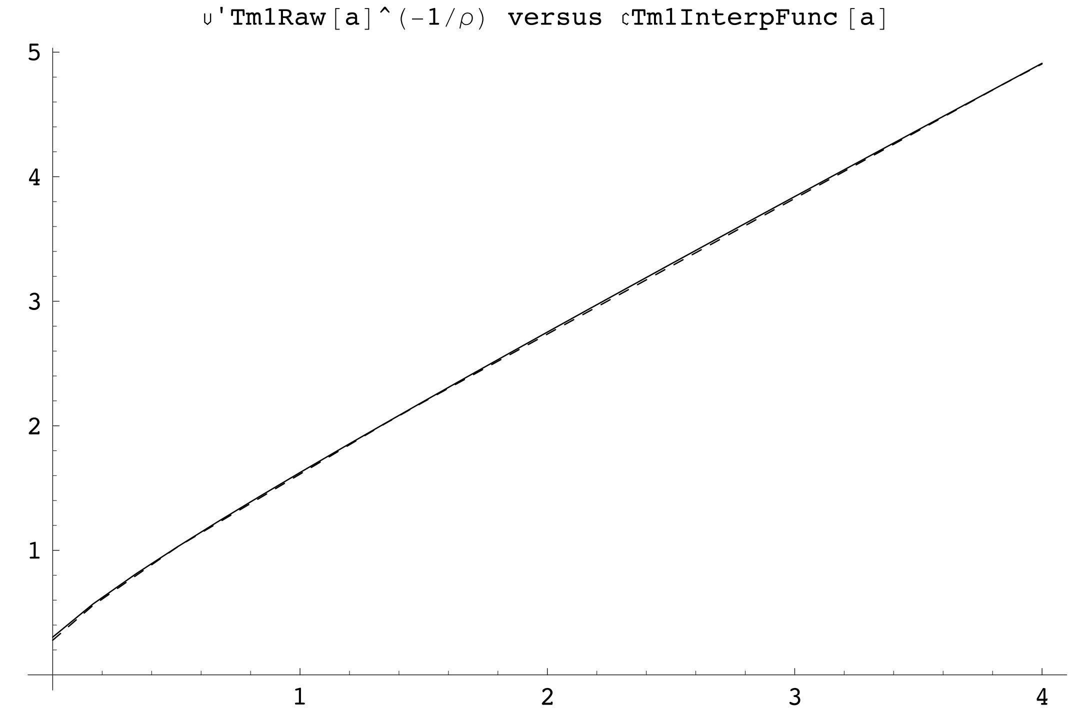

Figure 12 $\mathfrak{c}_{T-1}(a_{T-1})$ versus $\hat{\mathfrak{c}}_{T-1}(a_{T-1})$, Triple Exponential αVec

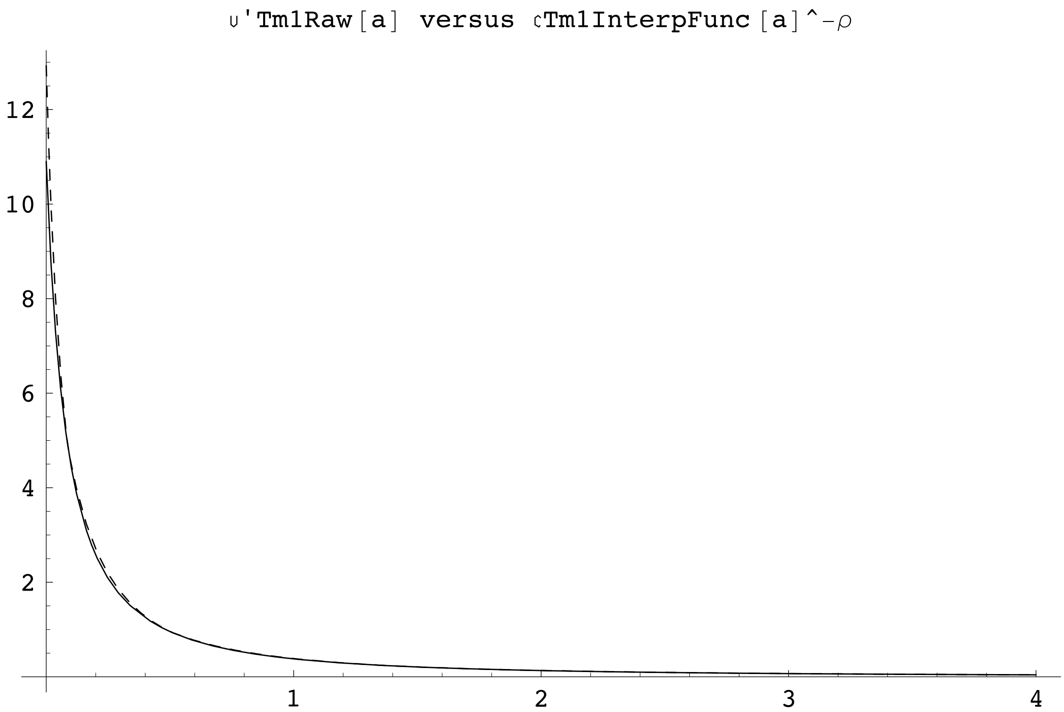

Figure 13 $\mathfrak{v}'_{T-1}(a_{T-1})$ vs. $\hat{\hat{\mathfrak{v}}}'_{T-1}(a_{T-1})$ Constructed Using $\hat{\mathfrak{c}}_{T-1}(a_{T-1})$, Triple Exponential αVec

for some α. The appropriate transformation is therefore $\not{\mathrm{v}}_t(m_t) = ((1-\rho)\mathrm{v}_t(m_t))^{1/(1-\rho)}$ and similarly $\not{\mathfrak{v}}_t(m_t) = ((1-\rho)\mathfrak{v}_t(m_t))^{1/(1-\rho)}$.

5.7 Constraints

Problems of this type often come with additional constraints that must be satisfied. The most common type of constraint is a liquidity constraint that prevents the consumer's net worth from falling below some value, often zero.

With such an additional constraint, the problem can be rewritten

$$\begin{aligned} \mathrm{v}_{T-1}(m_{T-1}) &= \max_{c_{T-1}} \ \mathrm{u}(c_{T-1}) + \beta\,\mathbb{E}_{T-1}\,\mathrm{v}_T(m_T) && (55) \\ &\text{s.t.} \\ a_{T-1} &= m_{T-1} - c_{T-1} && (56) \\ m_T &= \mathcal{R}a_{T-1} + Y_T && (57) \\ a_{T-1} &\geq 0. && (58) \end{aligned}$$

By definition, the constraint will bind in those circumstances where the unconstrained consumer would behave in such a way as to violate the constraint. In this case, that means that the constraint binds if the level of consumption that satisfies the unconstrained FOC

$$c_{T-1}^{-\rho} \;=\; \mathfrak{v}'_{T-1}(m_{T-1} - c_{T-1}) \tag{59}$$

is greater than m_{T-1}. Call the function returning the level of c_{T-1} that satisfies this equation $\check{\mathrm{c}}_{T-1}$. Then the constrained optimal level of consumption will be

$$\mathrm{c}_{T-1}(m_{T-1}) \;=\; \min[m_{T-1}, \check{\mathrm{c}}_{T-1}(m_{T-1})]. \tag{60}$$

The introduction of the constraint also introduces a sharp nonlinearity in all of the functions at the point where the constraint begins to bind. As a result, to get solutions that are anywhere close to numerically accurate it is useful to augment the grid of values of the state variable to include the exact value at which the constraint becomes binding. Fortunately, the value of this point is easy to calculate. We know that when the constraint is binding the consumer is saving nothing. We know the marginal value of saving at the point of zero saving is given by $\mathfrak{v}'_{T-1}(0)$. Finally, we know that when the constraint is binding, $c_{T-1} = m_{T-1}$. Thus, the largest value of consumption for which the constraint is binding will be the point for which the marginal utility of consumption is exactly equal to the (expected, discounted) marginal value of saving 0. We know this because the marginal utility of consumption is a downward-sloping function and so if the consumer were to consume ϵ more, the marginal utility of that extra consumption would be *below* the (discounted, expected) marginal utility of saving, and thus the consumer would engage in positive saving and the constraint would no longer be binding. Thus the level of m_{T-1} at which the constraint begins to bind is:

$$\begin{aligned} \mathrm{u}'(m_{T-1}) &= \mathfrak{v}'_{T-1}(0) \\ m_{T-1} &= (\mathfrak{v}'_{T-1}(0))^{(-1/\rho)}. \\ m_{T-1} &= \mathfrak{c}_{T-1}(0). \end{aligned} \tag{61}$$

Note that once we have constructed the interpolating function $\hat{\mathfrak{c}}_{T-1}$ this expression is fast and easy to calculate. Once the value that solves this equation is calculated, we simply add that value of m to the vector μ`Vec` and treat the new point just like any other point in μ`Vec`.

The constrained problem is solved by `2periodIntexpFOCInvEEECon.nb`; the resulting consumption rule is shown in figure 14. For comparison purposes, the approximate consumption rule from figure 11 is reproduced here as the solid line. As expected, the liquidity constraint only causes a divergence between the two functions at the point where the optimal unconstrained consumption rule runs into the 45 degree line.

6 Recursion

6.1 Theory

Before we begin solving the problem for periods earlier than $T-1$, we assume a liquidity constraint of the kind just discussed that prevents the agent from consuming more than his/her net worth. This simplifies matters a bit because we can now only consider an

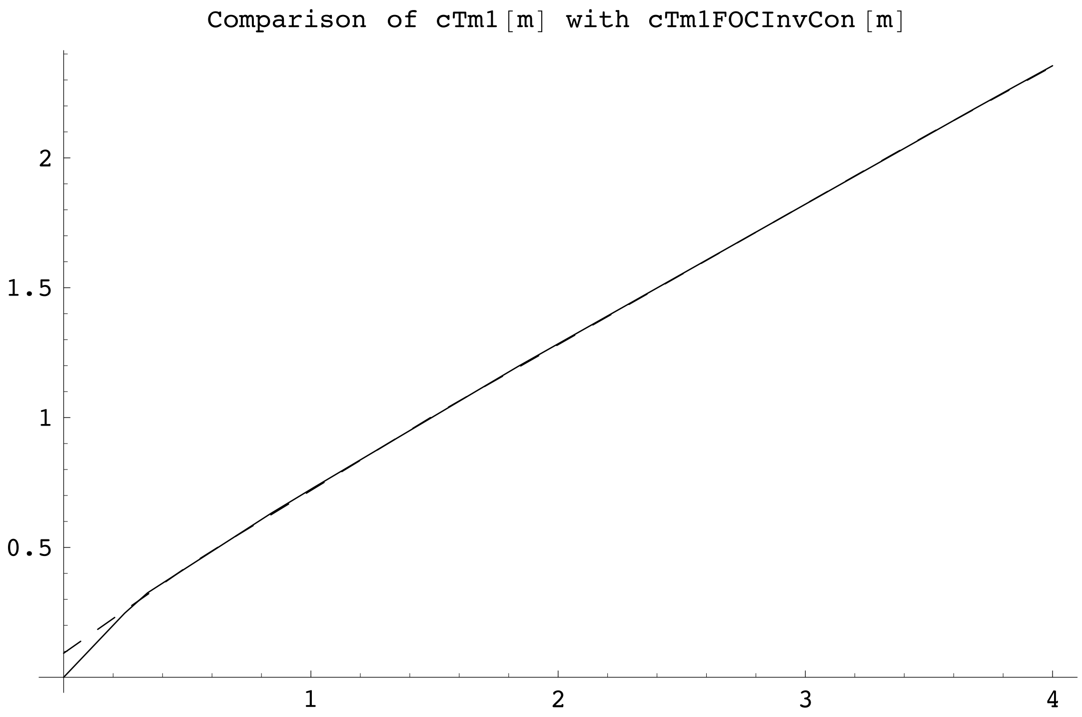

Figure 14 Constrained (solid) and Unconstrained (dashing) Consumption

α`Vec` that starts with zero as the smallest element.[9] We have now learned how to construct an approximation to the value function $\mathrm{v}_{T-1}(m_{T-1})$ and optimal consumption function $\mathrm{c}_{T-1}(m_{T-1})$. How do we proceed back to earlier periods of life?

Recall equations (24) and (25):

$$\begin{aligned} \mathfrak{v}'_t(a_t) &= \beta\,\mathbb{E}_t[\mathsf{R}\mathrm{u}'(\mathrm{c}_{t+1}(\mathcal{R}a_t + \Theta_{t+1}))] && (62) \\ \mathrm{u}'(c_t) &= \mathfrak{v}'_t(m_t - c_t). && (63) \end{aligned}$$

Assuming the problem has been solved up to period $t+1$ (and thus assuming that we have a numerical function $\hat{\mathrm{c}}_{t+1}(m_{t+1})$), the first of these tells us how to calculate $\mathfrak{v}_t(a_t)$, and the second tells us how to calculate $\hat{\mathrm{c}}_t(m_t)$ given $\mathfrak{v}_t(a_t)$. Our solution method essentially involves using these two equations in succession to work back progressively from period $T-1$ to the beginning of life. Stated generally, the method is as follows.

1. For the grid of values $\alpha_{i,t} = \alpha_i$ in α`Vec`, numerically calculate the value of $\mathfrak{c}_t(\alpha_{i,t})$,

$$\begin{aligned} \mathfrak{c}_t(\alpha_{i,t}) &= \left(\mathfrak{v}'_t(\alpha_{i,t})\right)^{-1/\rho}, && (64) \\ &= \left(\beta\,\mathbb{E}_t\left[\mathcal{R}(\hat{\mathrm{c}}_{t+1}(\mathcal{R}\alpha_{i,t} + \Theta_{t+1}))^{-\rho}\right]\right)^{-1/\rho}, && (65) \end{aligned}$$

   generating a list of values $\mathfrak{c}_{i,t}$. (The last point in α`Vec` should be α`Huge`; when this value is evaluated, return the solution for the perfect foresight finite horizon consumption problem).

[9] If there are no liquidity constraints, then α`Vec` needs to be augmented by $\underline{a}$, the minimum possible PDV of future labor income. Further discussions can be found in Carroll (2011).

2. Construct an interpolating function $\hat{\mathfrak{c}}_t(a_t)$ that 'connects the dots' of $\{\{\alpha_{1,t}, \mathfrak{c}_{1,t}\}, \{\alpha_{2,t}, \mathfrak{c}_{2,t}\}, \ldots\}$.
3. Construct a corresponding list of values of $\chi_{i,t}$ and $\mu_{i,t}$ from $\chi_i = \mathfrak{c}_i$ and $\mu_i = \chi_i + \alpha_i$.
4. Augment the resulting μ`Vec` with the point μ`Huge`, and the consumption vector with the perfect foresight solution for the point μ`Huge`.
5. If there is a liquidity constraint, augment χ_i and μ_i with the point $\{0, 0\}$.
6. Interpolate between the $\{\mu_i, \chi_i\}$ to obtain $\hat{\mathrm{c}}_t(m_t)$.

Once we have $\hat{\mathrm{c}}_t(m_t)$ we can continue the backwards recursion to period $t-1$ and so on back to the beginning of life.

Note that this loop does not contain steps for constructing $\hat{\mathfrak{v}}'_t(a_t)$ or $\hat{\mathrm{v}}'_t(m_t)$. This is because with $\hat{\mathfrak{c}}_t(a_t)$ and $\hat{\mathrm{c}}_t(m_t)$ in hand, we simply *define* $\hat{\mathfrak{v}}'_t(a_t) = [\hat{\mathfrak{c}}_t(a_t)]^{-\rho}$ and $\hat{\mathrm{v}}'_t(m_t) = \mathrm{u}'(\hat{\mathrm{c}}_t(m_t))$ so there is no need to construct interpolating functions for these functions - they arise 'free' (or nearly so) from our constructed $\hat{\mathfrak{c}}_t(a_t)$ and $\hat{\mathrm{c}}_t(m_t)$.

The program `multiperiod.nb` presents a fairly general and flexible approach to solving problems of this kind. The essential structure of the program is a loop which simply works its way back from an assumed last period of life, using the command `AppendTo` to record the interpolated consumption functions in the earlier time periods back from the end. For a realistic life cycle problem, it would also be necessary at a minimum to allow for a nonconstant path of income growth over the lifetime, which can be done easily by adding a variable `GPath` which contains age-specific expected income growth rates.

6.2 Program Structure

After the usual initializations, the heart of the program works like this.

6.2.1 Iteration

After setting up a variable `PeriodsToSolve` which defines the total number of periods that the program will solve, the program sets up a "`Do[SolveAnotherPeriod,{PeriodsToSolve}]`" loop that runs the function `SolveAnotherPeriod` the number of times corresponding to `PeriodsToSolve`. Every time `SolveAnotherPeriod` is run, the interpolated consumption function for one period of life earlier is calculated. The structure of the `SolveAnotherPeriod` function is as follows:

1. The `Table` command sets up a loop that computes and records the elements necessary for the construction of the `InterpolatingFunction`. For each α in α`Vec`, construct χ from `c(`α`)`, which is defined as

$$\mathfrak{c}_t(a_t) \;=\; \left(\beta\,\mathbb{E}_t\left[\mathcal{R}(\hat{\mathrm{c}}_{t+1}(\mathcal{R}a_t + \Theta_{t+1}))^{-\rho}\right]\right)^{-1/\rho} \tag{66}$$

$$= \left(\beta\frac{1}{n}\sum_{i=1}^{n} R\left[(\hat{c}_{t+1}(\mathcal{R}a_t+\Theta_i))^{-\rho}\right]\right)^{-1/\rho}. \quad (67)$$

Obviously, $\mathfrak{c}$(α) depends on the interpolated consumption function from one period later of life. It is captured by `Last[cInterpFunc]`, the updated `InterpolatingFunction` obtained in the last "`Do`" iteration. Finally, the program sets $\mu = \chi+\alpha$.

2. In order to implement the liquidity constraint, add $\{0., 0.\}$ to the beginning of the list of $\{\mu,\chi\}$ generated by the loop described above. By construction, the $\{\mu_0, \chi_0\}$ point that corresponds to `Min[`α`Vec]=0` lies on the 45 degree line in the consumption function diagram. It is also the leftmost of all $\{\mu, \chi\}$ points calculated from corresponding αVec's. While for μ larger than μ_0 the liquidity constraint does not bind, for any μ smaller than μ_0 the constraint *does* bind, i.e. $\chi = \mu$ for all $\mu \leq \mu_0$. Linear interpolation between $\{0., 0.\}$ and $\{\mu_0, \chi_0 = \mu_0\}$ conveniently captures this binding portion of the consumption function.

3. Construct a piecewise linear interpolation among the points from the list above.

4. Add the new `InterpolatingFunction` to the end of the existing list of interpolation functions `cInterpFunc`.

6.3 Results

As written, the program creates `cInterpFunc` for $\hat{c}_t(m_t)$. All other functions that define the solution can be defined using this function or $\hat{\mathfrak{c}}_t(a_t)$. These functions can be evaluated in any period for any value of m. For values of m outside of the grid encompassed in μ`Vec` and α`Vec`, the program extrapolates from the relationship between the nearest two points in the grids.

As an illustration, figure 15 shows $\hat{c}_t(m_t)$ for $t = \{20, 15, 10, 5, 1\}$. At least one feature of this figure is encouraging: the consumption functions converge as time recedes from the end of life, something that Carroll (2011) shows must be true under certain parametric conditions that are satisfied by the baseline parameter values being used here.

7 Multiple Control Variables

We now consider how to solve problems with multiple control variables.

7.1 Theory

The new control variable that we will assume the consumer can choose is the portion of the portfolio to invest in risky versus safe assets. Designating the gross return on the risky asset between period t and $t + 1$ as $\mathfrak{R}_{t+1}$, and using ς_t to represent the proportion of the portfolio invested in equities between t and $t + 1$, and continuing to use R for

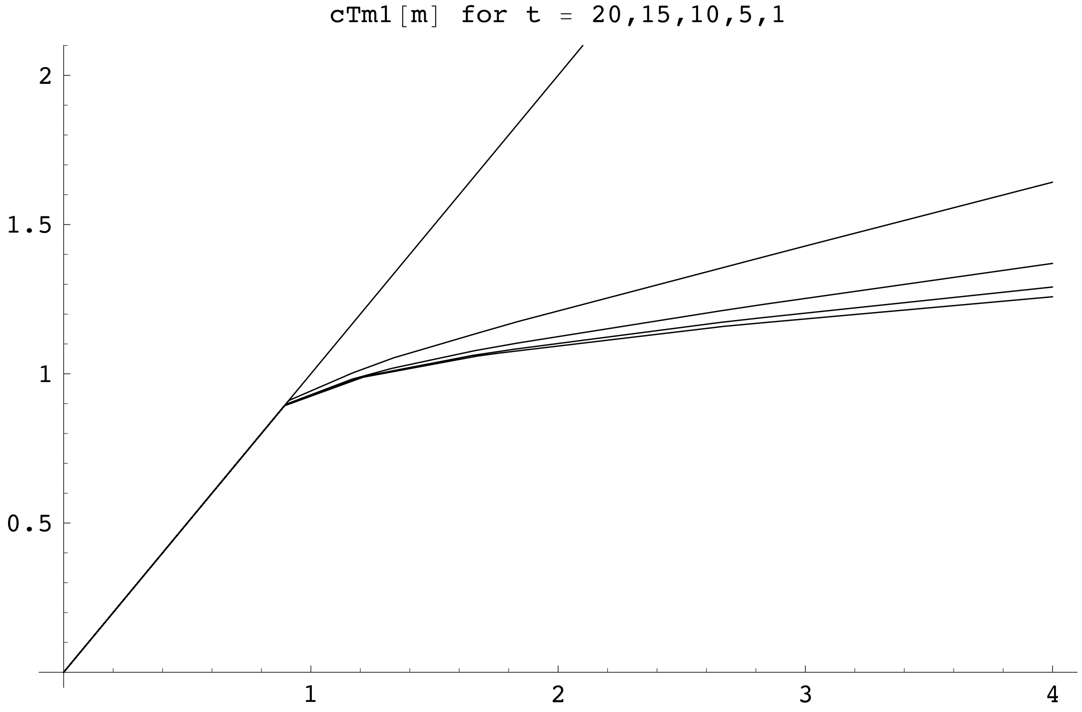

Figure 15 Converging $c_{T-n}(m_{T-n})$ Functions as n Increases

the rate of return on the riskless asset, the overall return on the consumer's portfolio between t and $t+1$ will be (in an almost unforgivable abuse of notation, designed to capture the fact that multiple rates of return are combined into a single object):

$$\begin{aligned} \mathbb{R}_{t+1} &= \mathcal{R}(1-\varsigma_t) + \mathfrak{R}_{t+1}\varsigma_t && (68) \\ &= \mathcal{R} + (\mathfrak{R}_{t+1} - \mathcal{R})\varsigma_t && (69) \end{aligned}$$

and the maximization problem is

$$\begin{aligned} \mathrm{v}_t(m_t) &= \max_{\{c_t,\varsigma_t\}} \{\mathrm{u}(c_t) + \beta\,\mathbb{E}_t[\mathrm{v}_{t+1}(m_{t+1})]\} && (70) \\ &\text{s.t.} \\ \mathbb{R}_{t+1} &= \mathcal{R} + (\mathfrak{R}_{t+1} - \mathcal{R})\varsigma_t && (71) \\ m_{t+1} &= (m_t - c_t)\mathbb{R}_{t+1} + \Theta_{t+1} && (72) \\ 0 \leq \quad \varsigma_t &\leq 1, && (73) \end{aligned}$$

or

$$\begin{aligned} \mathrm{v}_t(m_t) &= \max_{\{c_t,\varsigma_t\}} \; \mathrm{u}(c_t) + \beta\,\mathbb{E}_t[\mathrm{v}_{t+1}(\mathbb{R}_{t+1}(m_t - c_t) + \Theta_{t+1})] \\ &\text{s.t.} \\ 0 \leq \quad \varsigma_t &\leq 1. \end{aligned}$$

The first order condition with respect to c_t is almost identical to that in the single-control problem, equation (17), with the only difference being that the nonstochastic interest rate $\mathcal{R}$ is now replaced by $\mathbb{R}_{t+1}$,

$$\mathrm{u}'(c_t) \quad = \quad \beta\,\mathbb{E}_t[\mathbb{R}_{t+1}\mathrm{v}'_{t+1}(m_{t+1})], \tag{74}$$

and the Envelope theorem derivation remains the same so that we still have

$$\mathrm{u}'(c_t) \quad = \quad \mathrm{v}'_t(m_t) \tag{75}$$

implying the Euler equation for consumption

$$\mathrm{u}'(c_t) \quad = \quad \beta\,\mathbb{E}_t[\mathbb{R}_{t+1}\mathrm{u}'(c_{t+1})]. \tag{76}$$

The first order condition with respect to the risky portfolio share is

$$0 \quad = \quad \mathbb{E}_t[\mathrm{v}'_{t+1}(m_{t+1})(\mathfrak{R}_{t+1}-\mathcal{R})a_t] \tag{77}$$

$$\quad = \quad a_t\,\mathbb{E}_t\left[\mathrm{u}'\left(\mathrm{c}_{t+1}(m_{t+1})\right)\left(\mathfrak{R}_{t+1}-\mathcal{R}\right)\right]. \tag{78}$$

As before, it will be useful to define $\mathfrak{v}_t$ as a function that yields the expected value as of $t+1$ of ending period t in a given state. However, now that there are two control variables, the expectation must be defined as a function of the choices of both of those variables, because the expectation as of time t of value as of time $t+1$ will depend not just on how much the agent saves, but also on how those assets are allocated between the risky and riskless assets. Thus we define

$$\mathfrak{v}_t(a_t,\varsigma_t) \quad = \quad \beta\,\mathbb{E}_t[\mathrm{v}_{t+1}(m_{t+1})]$$

which has derivatives

$$\mathfrak{v}^a_t \quad = \quad \beta\,\mathbb{E}_t[\mathbb{R}_{t+1}\mathrm{v}^m_{t+1}(m_{t+1})]$$

$$\mathfrak{v}^{\varsigma}_t \quad = \quad \beta\,\mathbb{E}_t[(\mathfrak{R}_{t+1}-\mathcal{R})\mathrm{v}^m_{t+1}(m_{t+1})]a_t$$

implying that the first order conditions (76) and (78) and can be rewritten

$$\mathrm{u}'(c_t) \quad = \quad \mathfrak{v}^a_t(m_t-c_t,\varsigma_t), \tag{79}$$

$$0 \quad = \quad \mathfrak{v}^{\varsigma}_t(a_t,\varsigma_t). \tag{80}$$

7.2 Application

Our first step is to specify the stochastic process for $\mathfrak{R}_{t+1}$. We follow the common practice of assuming that returns are lognormally distributed, $\log\mathfrak{R} \sim \mathcal{N}(\phi+\mathsf{r},\sigma^2_e)$ where ϕ is the equity premium over the returns r available on the riskless asset.

As with labor income uncertainty, it is necessary to discretize the rate-of-return risk in order to have a problem that is soluble in a reasonable amount of time. We follow the same procedure as for labor income uncertainty, generating a set of m equiprobable values of $\mathfrak{R}$ which we will index by j, $\mathfrak{R}_{j,t+1}$; the portfolio-weighted return (contingent on the portfolio share in equity) will be designated $\mathbb{R}_{j,t+1}$.

Now let's rewrite the expressions for the derivatives of $\mathfrak{v}_t$ explicitly:

$$\mathfrak{v}^a_t(a_t,\varsigma_t) \quad = \quad \beta\left(\frac{1}{mn}\right)\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\mathbb{R}_{j,t+1}\left(\mathrm{c}_{t+1}(\mathbb{R}_{j,t+1}a_t+\Theta_i)\right)^{-\rho}\right] \tag{81}$$

$$\mathfrak{v}_t^{\varsigma}(a_t, \varsigma_t) \;=\; \beta \left(\frac{1}{mn}\right) \sum_{i=1}^{n} \sum_{j=1}^{m} \left[(\Re_{j,t+1} - \mathcal{R}) \left(\mathrm{c}_{t+1}(\mathbb{R}_{j,t+1} a_t + \Theta_i)\right)^{-\rho} \right]. \quad (82)$$

Writing these equations out explicitly makes a problem very apparent: For every different combination of $\{a_t, \varsigma_t\}$ that the routine wishes to consider, it must perform two double-summations of $m \times n$ terms. Once again, there is an inefficiency if it must perform these same calculations many times for the same or nearby values of $\{a_t, \varsigma_t\}$, and again the solution is to construct an approximation to the derivatives of the $\mathfrak{v}$ function.

Details of the construction of the interpolating approximation are given below; assume for the moment that we have the approximations $\hat{\mathfrak{v}}_t^a$ and $\hat{\mathfrak{v}}_t^{\varsigma}$ in hand and we want to proceed. As noted above, nonlinear equation solvers (including those built into *Mathematica*) can find the solution to a set of simultaneous equations. Thus we could ask *Mathematica* to solve

$$c_t^{-\rho} \;=\; \hat{\mathfrak{v}}_t^a(m_t - c_t, \varsigma_t) \quad (83)$$

$$0 \;=\; \hat{\mathfrak{v}}_t^{\varsigma}(m_t - c_t, \varsigma_t) \quad (84)$$

simultaneously for the set of potential m_t values defined in μ`Vec`. However, multidimensional constrained maximization problems are difficult and sometimes quite slow to solve. There is a better way. Define the problem

$$\bar{\mathfrak{v}}_t(a_t) \;=\; \max_{\{\varsigma_t\}} \; \mathfrak{v}_t(a_t, \varsigma_t) \quad (85)$$

$$\text{s.t.}$$

$$0 \leq \; \varsigma_t \; \leq 1 \quad (86)$$

where the bar accent on $\mathfrak{v}$ indicates that this is the $\mathfrak{v}$ that has been optimized with respect to all of the arguments other than the one still present (a_t). We solve this problem for the set of gridpoints in α`Vec` and use the results to construct the interpolating function $\hat{\bar{\mathfrak{v}}}_t^a(a_t)$. With this function in hand, we can use the first order condition from the single-control problem

$$c_t^{-\rho} \;=\; \hat{\bar{\mathfrak{v}}}_t^a(m_t - c_t)$$

to solve for the optimal level of consumption as a function of m_t. Thus we have transformed the multidimensional optimization problem into a sequence of two simple optimization problems for which solutions are much easier and more reliable.

Note the parallel between this trick and the fundamental insight of dynamic programming: dynamic programming techniques transform a multi-period (or infinite-period) optimization problem into a sequence of two-period optimization problems which are individually much easier to solve; we have done the same thing here, but with multiple dimensions of controls rather than multiple periods.

7.3 Implementation

The program which solves the problem with multiple control variables is `multicontrol.nb`.

Some of the functions defined in `multicontrol.nb` correspond to the derivatives of $\mathfrak{v}_t(a_t, \varsigma_t)$.

The first function definition that does not resemble anything in `multiperiod.nb` is `wRaw[at_]`. This function, for its input value of a_t, calculates the value of the portfolio share ς_t which satisfies the first order condition (84), tests whether the optimal portfolio share would violate the constraints, and if so resets the portfolio share to the constrained optimum. The function returns the optimal value of the portfolio share itself, ς_t^*, from which the functions $\bar{\mathfrak{v}}_t^a(a_t)$ and $\hat{\varsigma}_t(a_t)$ will be constructed.

As $\hat{\varsigma}_t(a_t)$ can be constructed by `wRaw[at_]`, $\bar{\mathfrak{v}}_t^a(a_t)$ can be constructed by a new definition in the file `𝔳POpt[at_]`, where the naming convention is obviously that 'Opt' stands for 'Optimized.' With $\bar{\mathfrak{v}}_t^a(a_t)$ in hand the analysis is essentially identical to that for the standard multiperiod problem with a single control.

The structure of the program in detail is as follows. First, perform the usual initializations. Then initialize `wVec` and the other variables specific to the multiple control problem.[10] In particular, there are now three kinds of functions: those with both a_t and ς_t as arguments, those with just a_t, and those with m_t.

Once the setup is complete, the heart of the program is the following.

1. Construct $\mathfrak{v}_t^{\varsigma}(a_t, \varsigma_t)$ using the usual calculation over the tensor defined by the combinations of the elements of α`Vec` and `wVec`.

2. For any level of saving `at`, the function `wRaw[at_]` performs a rootfinding operation[11]

$$\begin{aligned} 0 &= \mathfrak{v}_t^{\varsigma}(a_t, \varsigma_t) && (87) \\ &\text{s.t.} \\ 0 \leq \ &\varsigma_t \leq 1 && (88) \end{aligned}$$

   and generates the corresponding optimal portfolio share $\bar{\varsigma}_t$.

3. Construct the function `𝔳POpt[at_]`

$$\bar{\mathfrak{v}}_t^a(a_t) \equiv \mathfrak{v}_t^a(a_t, \hat{\varsigma}_t(a_t)) \tag{89}$$

   where $\hat{\varsigma}_t(a_t)$ is computed by `wRaw[at_]`.

4. Using $\bar{\mathfrak{v}}_t^a(a_t) \equiv$ `𝔳POpt[at_]` (in place of $\mathfrak{v}_t^a(a_t) \equiv$ `𝔳P[at_]` in `multiperiod.nb`), follow the same procedures as in `multiperiod.nb` to generate $\hat{c}_t(m_t)$.

[10] Note the choice of a coefficient of relative risk aversion of 6, in contrast with the choice of 2 made for the previous problems. This choice reflects the well-known 'stockholding puzzle,' which is the microeconomic equivalent of the equity premium puzzle: for plausible descriptions of income uncertainty, rate of return risk, and the equity premium, the typical consumer should hold all or nearly all of their portfolio in equities. Thus we choose a high value for the coefficient of relative risk aversion in order to generate portfolio structure behavior more interesting than a choice of 100 percent equities in every period for every level of wealth.

[11] Alternatively, the rootfinding operation would be $0 = \hat{\mathfrak{v}}_t^{\varsigma}(a_t, \varsigma_t)$, where the interpolation function of $\mathfrak{v}_t^{\varsigma}(a_t, \varsigma_t)$ is used instead. However, the results obtained (especially $\hat{\varsigma}_t(a_t)$) are much less satisfactory.

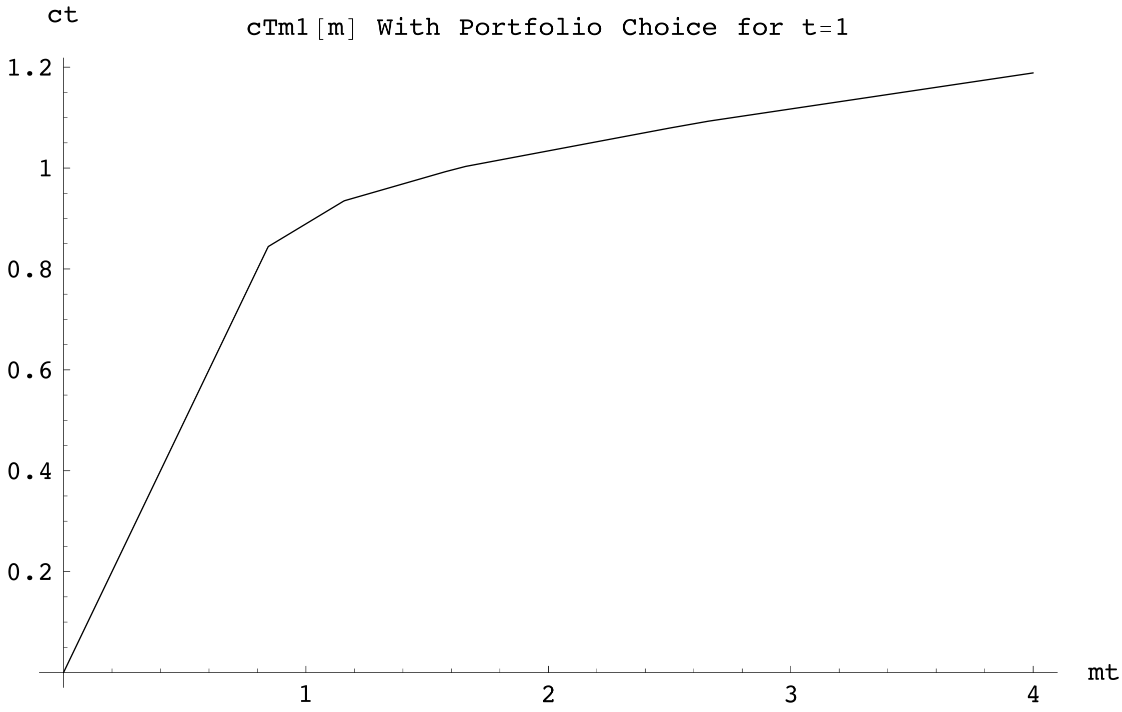

Figure 16 $\mathrm{c}_1(m_1)$ With Portfolio Choice

7.4 Results

Figure 16 plots the first-period consumption function generated by the program; qualitatively it does not look much different from the consumption functions generated by the program without portfolio choice. Figure 17 plots the optimal portfolio share as a function of the level of assets. This figure exhibits several interesting features. First, even with a coefficient of relative risk aversion of 6, an equity premium of only 4 percent, and an annual standard deviation in equity returns of 15 percent, the average level of the portfolio share kept in stocks is 100 percent at values of a_t less than about 2. Second, the proportion of the portfolio kept in stocks is *declining* in the level of wealth - i.e., the poor should hold all of their meager resources in stocks, while the rich should be more cautious. This bizarre prediction is a consequence of the assumption about labor income risk. Those consumers who are poor in measured financial wealth are likely to derive a high proportion of future consumption from their labor income. Since by assumption labor income risk is uncorrelated with rate-of-return risk, the covariance between their future consumption and future stock returns is relatively low. By contrast, persons with large amounts of current physical wealth will be financing a large proportion of future consumption out of that wealth, and hence their consumption will have a high covariance with stock returns if they invest too much in stocks. Consequently, they reduce that correlation by holding some of their wealth in the riskless form.

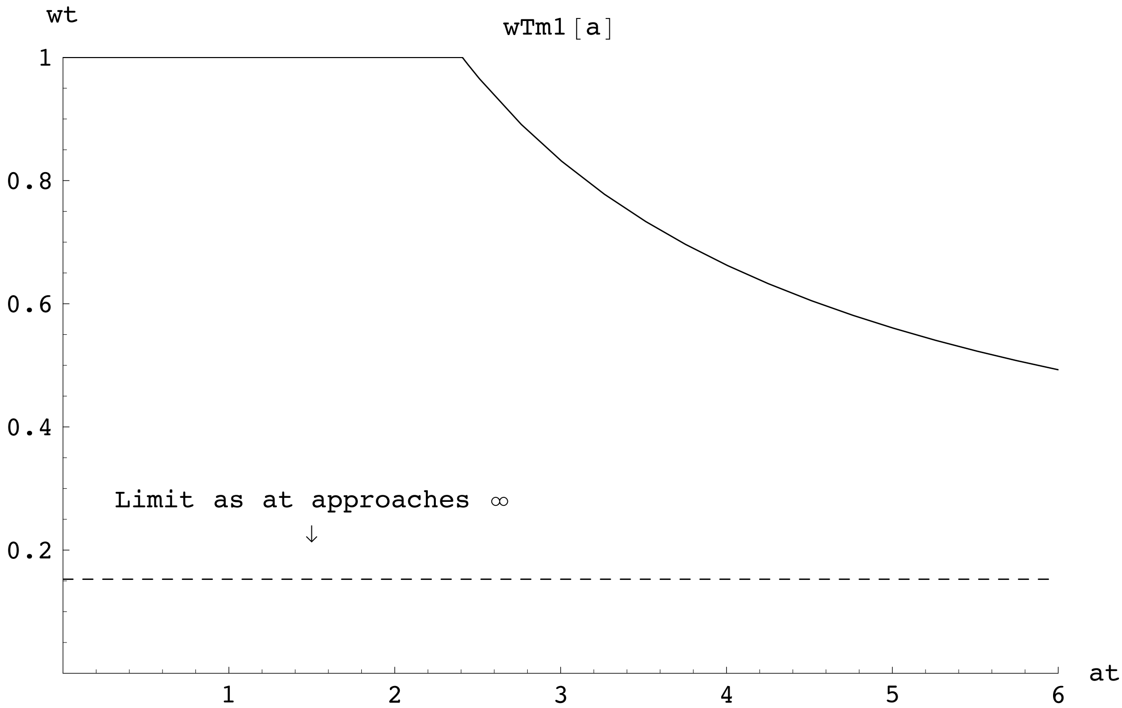

Figure 17 Portfolio Share in Risky Assets in First Period $\varsigma_1(a_1)$

8 The Infinite Horizon

All of the solution methods presented so far have involved period-by-period iteration from an assumed last period of life, as is appropriate for life cycle problems. However, if the parameter values for the problem satisfy certain conditions (detailed in Carroll (2011)), the consumption rules (and the rest of the problem) will converge to a fixed rule as the horizon (remaining lifetime) gets large, as illustrated in figure 15. Furthermore, Deaton (1991), Carroll (1992; 1997) and others have argued that the 'buffer-stock' saving behavior that emerges under some further restrictions of parameter values is a good approximation of the behavior of typical consumers over much of the lifetime. Methods for finding the converged functions are therefore of interest, and are dealt with in this section.

8.1 Convergence

In solving an infinite-horizon problem, it is necessary to have some metric for when to stop. Generally speaking, this is a matter of judgment, but it is usually wise to define two criteria both of which must be met before a solution is declared to have converged: One criterion referring to the average deviation between successive consumption rules, and one referring to the maximum deviation. For the simple one-dimensional consumption problem the criteria used here will be

$$\left(\frac{1}{n}\right)\sum_{i=1}^{n}|(\chi_{i,t+1}-\chi_{i,t})| \;<\; 0.001 \tag{90}$$

and

$$\max_i |(\chi_{i,t+1}-\chi_{i,t})| \;\leq\; 0.01. \tag{91}$$

Similar criteria can obviously be specified for other problems. However, it is always wise to plot successive function differences and to experiment a bit with convergence criteria to verify that the function has converged for all practical purposes.

8.2 The Last Period

For the last period of a finite-horizon lifetime, in the absence of a bequest motive it is obvious that the optimal policy is to spend everything. However, in an infinite-horizon problem there is no last period, and the policy of spending everything is obviously very far from optimal. Generally speaking, it is much better to start off with a 'last-period' consumption rule and value function equal to those corresponding to the infinite-horizon solution to the perfect foresight problem (assuming such a solution is known).

For the perfect foresight infinite horizon consumption problem, the solution is

$$\bar{\mathrm{c}}(m_t) \;=\; \overbrace{(1-\mathcal{R}^{-1}(\mathcal{R}\beta)^{1/\rho})}^{\equiv\underline{\kappa}}\left[m_t - 1 + \left(\frac{1}{1-1/\mathcal{R}}\right)\right] \tag{92}$$

where $\underline{\kappa}$ is the MPC in the infinite-horizon perfect foresight problem. In our baseline problem, we set $G = p_t = 1$. It is straightforward to show that the infinite-horizon perfect-foresight value function and marginal value function are given by

$$\bar{\mathrm{v}}(m_t) \;=\; \left(\frac{\bar{\mathrm{c}}(m_t)^{1-\rho}}{(1-\rho)\underline{\kappa}}\right) \tag{93}$$

$$\bar{\mathrm{v}}'(m_t) \;=\; (\bar{\mathrm{c}}(m_t))^{-\rho} \tag{94}$$

$$\bar{\mathfrak{v}}'(a_t) \;=\; \beta\mathcal{R}\bar{\mathrm{v}}'(\mathcal{R}a_t+1). \tag{95}$$

8.3 Coarse Then Fine αVec

The speed of each iteration is directly proportional to the number of gridpoints at which the problem must be solved. Therefore reducing the number of points in α`Vec` can increase the speed of solution greatly. Of course, this also decreases the accuracy of the solution. However, once the converged solution is obtained for a coarse α`Vec`, the density of the grid can be increased and iteration can continue until a convered solution is found for the finer α`Vec`.

8.4 Coarse then Fine ΘVec

The speed of solution is also roughly proportionate to the number of points used in approximating the distribution of shocks. At least 3 gridpoints should probably be used as an initial minimum, and my experience is that increasing the number of gridpoints beyond 7 generally yields only very small changes in the resulting function. The program `multiperiod_infhor.nb` begins with three gridpoints, solves for successively finer αVecs, then densifies the ΘVec vector as the last step.

9 Structural Estimation

This section describes how to use the methods developed above to structurally estimate a life-cycle consumption model, following closely the work of Cagetti (2003).[12] The key idea of structural estimation is to look for the parameter values (for the time preference rate, relative risk aversion, or other parameters) which lead to the best possible match between simulated and empirical moments. (The code for the structural estimation is in the self-contained subfolder `StructuralEstimation` in the Matlab and Mathematica directories; as of this writing (June 2008) the Matlab code is a bit out of date, and refers to an older version of this document).

9.1 Life Cycle Model

The decision problem for the household at age t is:

$$\max \left\{ \mathrm{u}(C_t) + \mathbb{E}_t \left[\sum_{s=t+1}^{T} \beth^{s-t} \left(\Pi_{i=t+1}^{s} \hat{\beta}_i \not{D}_i \right) \mathrm{u}(C_s) \right] \right\} \tag{96}$$

subject to the constraints

$$\begin{aligned} A_s &= M_s - C_s \\ M_{s+1} &= \mathsf{R} A_s + Y_{s+1} \\ Y_{s+1} &= \mathbf{P}_{s+1} \Theta_{s+1} \\ \mathbf{P}_{s+1} &= \Gamma_{s+1} \mathbf{P}_s \Psi_{s+1} \end{aligned}$$

where

$\not{D}_s$: probability of being alive (not dead) until age s given being alive at age $s-1$

$\hat{\beta}_s$: time-varying discount factor between age s and $s-1$

Ψ_s : mean-one shock to permanent income.

and all the other variables are defined as in section 2.

Households start life at age $s = 25$ and live with probability 1 until retirement ($s = 65$). Thereafter the survival probability shrinks every year and agents are dead by $s = 91$ as

[12] Similar structural estimation exercises have been also performed by Palumbo (1999) and Gourinchas and Parker (2002).

assumed by Cagetti. Note that in addition to a typical time-invariant discount factor $\beth$, there is a time-varying discount factor $\hat{\beta}_s$ in (96) which captures the effect of timevarying demographic variables (e.g. changes in family size).

Transitory and permanent shocks are distributed as follows:

$$\Xi_s = \begin{cases} 0 & \text{with probability } \wp > 0 \\ \Theta/\wp & \text{with probability } (1-\wp), \text{ where } \log\Theta \sim \mathcal{N}(-\sigma_\Theta^2/2, \sigma_\Theta^2) \end{cases} \tag{97}$$

$$\log\Psi_s \sim \mathcal{N}(-\sigma_\Psi^2/2, \sigma_\Psi^2) \tag{98}$$

where $\wp$ is the probability of unemployment (and unemployment shocks are turned off after retirement).

The parameter values for the shocks are taken from Carroll (1992), $\wp = 0.5/100$, $\sigma_\Theta = 0.1$, and $\sigma_\Psi = 0.1$.[13] The income growth profile Γ_s is from Carroll (1997) and the values of $\not{\mathcal{D}}_s$ and $\hat{\beta}_s$ are obtained from Cagetti (2003) (figure 18).[14] The interest rate is assumed to equal 1.03. The model parameters are included in table 1.

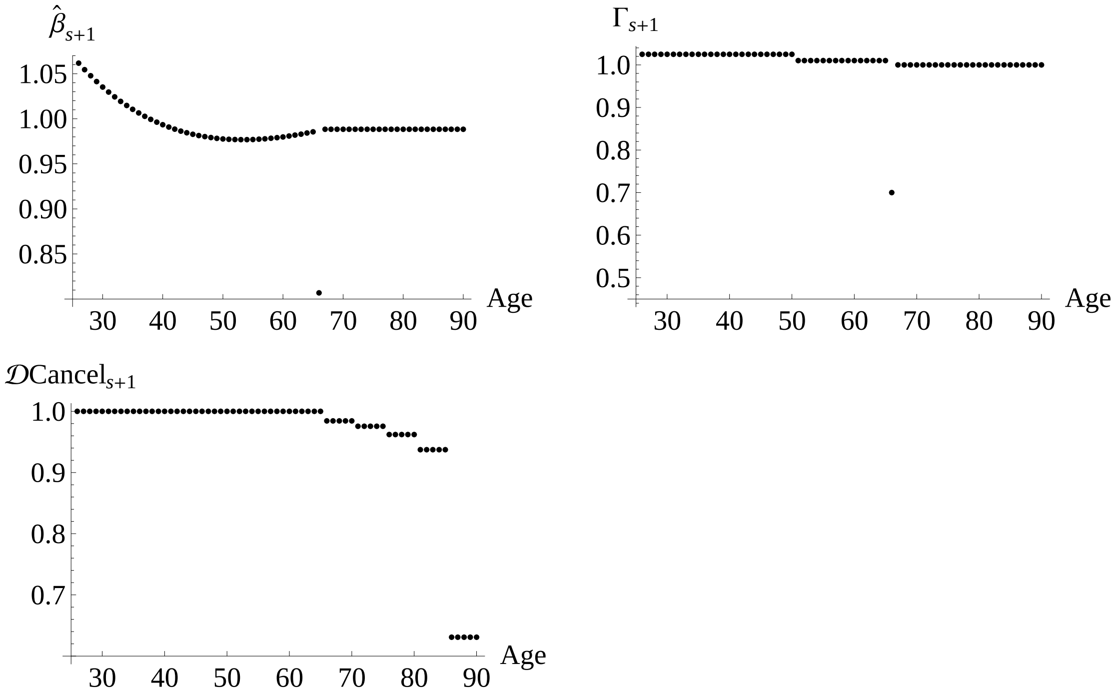

Figure 18 Time varying parameters

[13]Note that $\sigma_\Theta = 0.164$ is smaller than the estimate for college graduates estimated in Carroll and Samwick (1997) ($= 0.197 = \sqrt{0.039}$) which is used by Cagetti (2003). The reason for this choice is that Carroll and Samwick (1997) themselves argue that their estimate of σ_Θ is almost certainly increased by measurement error.

[14]The income growth profile is the one used by Caroll for operatives. Cagetti computes the time-varying discount factor by educational groups using the methodology proposed by Attanasio et al. (1999) and the survival probabilities from the 1995 Life Tables (National Center for Health Statistics 1998).

Table 1 Parameter Values

σ_Θ	0.1	Carroll (1992)
σ_Ψ	0.1	Carroll (1992)
$\wp$	0.005	Carroll (1992)
Γ_s	figure 18	Carroll (1997)
$\hat{\beta}_s, \not{D}_s$	figure 18	Cagetti (2003)
R	1.03	Cagetti (2003)

The parameters $\beth$ and ρ are structurally estimated following the procedure described below.

9.2 Estimation

When economists say that they are performing "structural estimation" of a model like this, what they mean is that they have devised a formal procedure for searching for values for the parameters $\beth$ and ρ at which some measure of the model's outcome (like "median wealth by age") is as close as possible to an empirical measure of the same thing. In particular, we want to match the median of the wealth to permanent income ratio across 7 age groups, from age $26-30$ up to $56-60$.[15] The choice of matching the medians rather the means is motivated by the fact that the wealth distribution is much more concentrated than the model is capable of explaining using a single set of parameter values. This means that in practice one must pick some portion of the population who one wants to match well; since the model has little hope of capturing the behavior of Bill Gates, but might conceivably match the behavior of Homer Simpson, we chose to match medians rather than means.

As explained in section 3, it is convenient to work with the normalized version the model which can be written as:

$$\begin{aligned}
\bar{\mathrm{v}}_t(m_t) &= \max_{c_t} \left\{ \mathrm{u}(c_t) + \beth \not{D}_{t+1} \hat{\beta}_{t+1}\, \mathbb{E}_t[(\Psi_{t+1}\Gamma_{t+1})^{1-\rho} \bar{\mathrm{v}}_{t+1}(m_{t+1})] \right\} \\
& \text{s.t.} \\
a_t &= m_t - c_t \\
m_{t+1} &= a_t \left(\frac{\mathsf{R}}{\Psi_{t+1}\Gamma_{t+1}} \right) + \Theta_{t+1}
\end{aligned}$$

with the first order condition:

$$\mathrm{u}'(c_t) = \beth \not{D}_{t+1} \hat{\beta}_{t+1} \mathsf{R}\, \mathbb{E}_t \left[\mathrm{u}' \left(\Psi_{t+1}\Gamma_{t+1} \mathrm{c}_{t+1} \left(a_t \left(\frac{\mathsf{R}}{\Psi_{t+1}\Gamma_{t+1}} \right) + \Theta_{t+1} \right) \right) \right] \quad (99)$$

[15] Cagetti (2003) matches wealth levels rather than wealth to income ratios. We believe it is more appropriate to match ratios both because the ratios are the state variable in the theory and because empirical moments for ratios of wealth to income are not influenced by the method used to remove the effects of inflation and productivity growth.

The first step is to solve for the consumption functions at each age using the routines included in the `setup_ConsFn.nb` file. We need to discretize the shock distribution (`DiscreteMeanOneLogNormal`) and solve for the policy functions by backward induction using equation (99) following the procedure in sections 5 and 6 (`ConstructcInterpFunc`). The latter routine is slightly complicated by the fact that we are considering a life-cycle model and therefore the growth rate of permanent income, the probability of death, the time-varying discount factor and the distribution of shocks will be different across the years. We thus have to make sure that at each backward iteration the right parameters are considered.

Once we have the age varying consumption functions, we can proceed to generate the simulated data and compute the simulated medians using the routines defined in the `setup_Sim.nb` file. We first have to draw the shocks for each agent and period. This involves discretizing the shock distribution in as many points as the agents we want to simulate (`ConstructShockDistribution`). We then randomly permute this shock vector as many times as we need to simulate the model for, thus obtaining a time varying shock for each agent (`ConstructSimShocks`). Note that this is much more time efficient than drawing at each time from the shock distribution a shock for each agent and also ensures a stable distribution of shocks across the simulation periods even for a small number of agents. We then initialize the wealth to income ratio of agents at age 25 as in Cagetti (2003) by randomly assigning the equal probability values 0.17, 0.50 and 0.83 and run the simulation (`Simulate`). In particular we consider a population of agents at age 25 and follow their consumption and wealth accumulation dynamics as they reach the age of 60. It is again crucial to use the correct age specific consumption functions and the age-varying parameters. The simulated medians are obtained by taking the medians of the wealth to income ratio of the 7 age groups.

Given these simulated medians, we could estimate the model by calculating empirical medians and then measuring the model's success by calculating the difference between the empirical median and the actual median. Specifically, defining ξ as the set of parameters to be estimated (in the current case $\xi = [\rho, \beth]$), we could search for the parameter values which solve

$$\min_{\xi} \sum_{\tau=1}^{7} |\varsigma^{\tau} - \mathbf{s}^{\tau}(\xi)|. \tag{100}$$

where ς^{τ} and $\mathbf{s}^{\tau}$ are respectively the empirical and simulated medians of the wealth to permanent income ratio for age group τ.

A drawback of proceeding in this way is that it treats the empirically estimated medians essentially as absolute perfectly measured truth. Imagine, however, that one of the age groups had four times as many data observations as another age group; then we would expect the median to be more precisely estimated for the age group with more observations; yet (100) assigns equal importance to a deviation between the model and the data for all age groups, whereas it should actually be less upset by a deviation from the poorly measured median than from the well measured median.

We can get around this problem (and a variety of others) by instead minimizing the

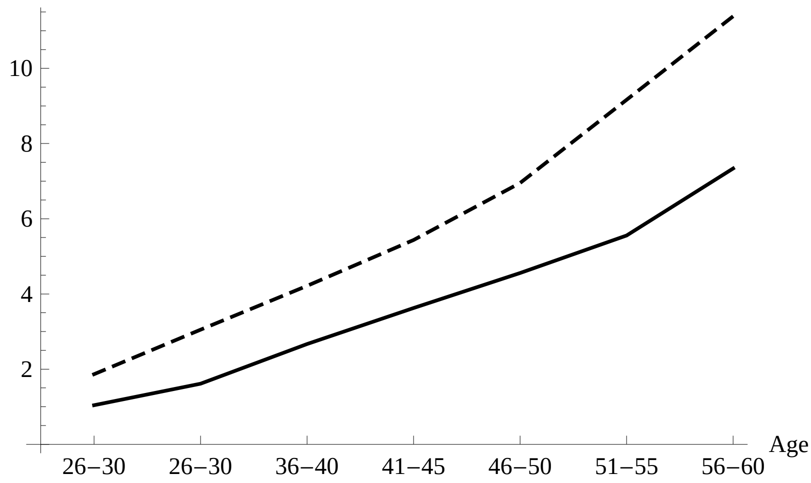

Figure 19 Wealth to Permanent Income Ratios from SCF (means (dashed) and medians (solid))

slightly more complex object

$$\min_{\xi} \sum_{i}^{N} \omega_i \left| \varsigma_i^{\tau} - \mathbf{s}^{\tau}(\xi) \right| \tag{101}$$

where ω_i is the weight of household i in the entire population,[16] and ς_i^{τ} is the empirical wealth to permanent income ratio of household i whose head belongs to age group τ. ω_i is needed because unequal weight is assigned to each observation in the Survey of Consumer Finances (SCF). The absolute value is used since the formula is based on the fact that the median is the value which minimizes the sum of the absolute deviations from itself.

The actual data are taken from several waves of the SCF and the medians and means for each age category are plotted in figure 19. More details on the SCF data are included in appendix (A).

The key function to perform structural estimation is defined in the `setup_Estimation.nb` file as follows:

$$\texttt{GapEmpiricalSimulatedMedians}[\rho, \beth]\texttt{:=} \quad \left[\begin{array}{l} \texttt{ConstructcInterpFunc}\{\rho, \beth\}; \\ \texttt{Simulate}; \\ \sum_{i}^{N} \omega_i \left| \varsigma_i^{\tau} - \mathbf{s}^{\tau}(\xi) \right| \end{array} \right.$$
$$];$$

[16]The Survey of Consumer Finances includes many more high-wealth households than exist in the population as a whole; therefore if one wants to produce population-representative statistics, one must be careful to weight each observation by the factor that reflects its "true" weight in the population.

For a given pair of the parameters to be estimated, the `GapEmpiricalSimulatedMedians` routine therefore:

1. solves for the consumption functions by calling `ConstructcInterpFunc`
2. simulates the data and computes the simulated medians by calling `Simulate`
3. returns the value of equation (101)

We can delegate the task of finding the coefficients which minimize the `GapEmpiricalSimulatedMedians` function to the numerical minimizers built into Mathematica, such as `FindMinimum` and `NMinimize` respectively searching for local and global minima. This task can be quite time demanding and rather problematic if the `GapEmpiricalSimulatedMedians` function has very flat regions or sharp features. It is thus wise to verify the accuracy of the solution for example using different starting values.

Finally the standard errors are computed by bootstrap using the routines in the `setup_Bootstrap.nb` file.[17] This involves:

1. drawing new shocks for the simulation
2. drawing a random sample (with replacement) of actual data from the SCF
3. obtaining new estimates for ρ and $\beth$

We repeat the above procedure several times (`Bootstrap`) and take the standard deviation for each of the estimated parameters across the various bootstrap iterations.

The file `StructEstimation.nb` produces the ρ and $\beth$ estimates with standard errors using 10,000 simulated agents. Results are reported in table 2.[18] Figure 20 shows the contour plot of the `GapEmpiricalSimulatedMedians` function and the parameter estimates. The contour plot shows equally spaced isoquants of the `GapEmpiricalSimulatedMedians` function, i.e. the pairs of ρ and $\beth$ which lead to the same deviations between simulated and empirical medians. We can thus interestingly see that there is a large rather flat region, or more formally speaking there exists a broad set of parameter pairs which leads to similar simulated wealth to income ratios. Intuitively, the flatter and larger is this region, the harder it is for the structural estimation procedure to precisely identify the parameters.

10 Conclusion

There are many alternative choices that can be made for solving microeconomic dynamic stochastic optimization problems. The set of techniques, and associated programs,

[17] For a treatment of the advantages of the bootstrap see Horowitz (2001)

[18] Differently from Cagetti (2003) who estimates a different set of parameters for college graduates, high school graduates and high school dropouts graduates, we perform the structural estimation on the full population.

Table 2 Estimation results

ρ	$\beth$
4.68	1.00
(0.13)	(0.00)

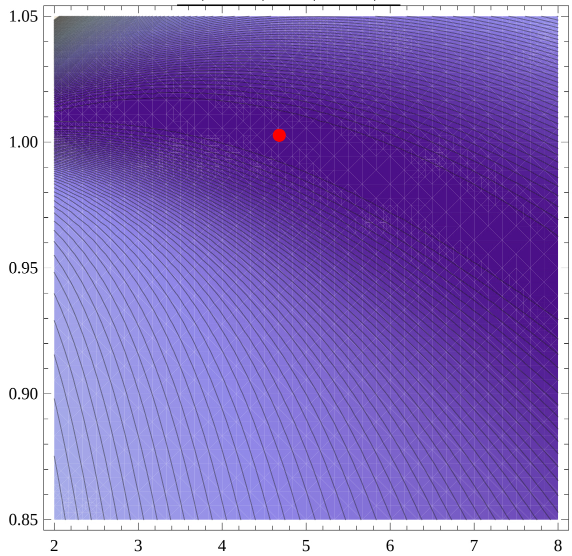

Figure 20 Contour plot (larger values are shown lighter) with $\{\rho, \beth\}$ estimates (red dot).

described in these notes represents an approach that I have found to be powerful, flexible, and efficient, but other problems may require other techniques. For a much broader treatment of many of the issues considered here, see Judd (1998).

Appendices

A Further Details on SCF data

Data used in the estimation is constructed using the SCF 1992, 1995, 1998, 2001 and 2004 waves. The definition of wealth is net worth including housing wealth, but excluding pensions and social securities. The data set contains only households whose heads are aged 26-60 and excludes singles, following Cagetti (2003).[19] Furthermore, the data set contains only households whose heads are college graduates. The total sample size is 4,774.

In the waves between 1995 and 2004 of the SCF, levels of *normal* income are reported. The question in the questionnaire is "About what would your income have been if it had been a normal year?" We consider the level of normal income as corresponding to the model's theoretical object P, permanent noncapital income. Levels of normal income are not reported in the 1992 wave. Instead, in this wave there is a variable which reports whether the level of income is normal or not. Regarding the 1992 wave, only observations which report that the level of income is normal are used, and the levels of income of remaining observations in the 1992 wave are interpreted as the levels of permanent income.

Normal income levels in the SCF are before-tax figures. These before-tax permanent income figures must be rescaled so that the median of the rescaled permanent income of each age group matches the median of each age group's income which is assumed in the simulation. This rescaled permanent income is interpreted as after-tax permanent income. This rescaling is crucial since in the estimation empirical profiles are matched with simulated ones which are generated using after-tax permanent income (remember the income process assumed in the main text). Wealth / permanent income ratio is computed by dividing the level of wealth by the level of (after-tax) permanent income, and this ratio is used for the estimation.[20]

[19] Cagetti (2003) argues that younger households should be dropped since educational choice is not modeled. Also, he drops singles, since they include a large number of single mothers whose saving behavior is influenced by welfare.

[20] Please refer to the archive code for details of how these after-tax measures of P are constructed.

References

Attanasio, O.P., J. Banks, C. Meghir, and G. Weber (1999): "Humps and Bumps in Lifetime Consumption," *Journal of Business & Economic Statistics*, 17(1), 22–35.

Cagetti, Marco (2003): "Wealth Accumulation Over the Life Cycle and Precautionary Savings," *Journal of Business and Economic Statistics*, 21(3), 339–353.

Carroll, Christopher D. (1992): "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 1992(2), 61–156, Available at http://econ.jhu.edu/people/ccarroll/BufferStockBPEA.pdf.

——— (1997): "Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, CXII(1), 1–56, Available at http://econ.jhu.edu/people/ccarroll/BSLCPIH.zip.

——— (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, pp. 312–320, Paper and software available at
http://econ.jhu.edu/people/ccarroll/EndogenousArchive.zip or
http://dx.doi.org/10.1016/j.econlet.2005.09.013.

——— (2011): "Theoretical Foundations of Buffer Stock Saving," *Revise and Resubmit, Review of Economic Studies*, Latest version available at http://econ.jhu.edu/people/ccarroll/BufferStockTheory.pdf.

——— (Current): "Math Facts Useful for Graduate Macroeconomics," Online Lecture Notes, Available at http://econ.jhu.edu/people/ccarroll/public/lecturenotes/mathfacts/.

Carroll, Christopher D., and Miles S. Kimball (1996): "On the Concavity of the Consumption Function," *Econometrica*, 64(4), 981–992, Available at http://econ.jhu.edu/people/ccarroll/concavity.pdf.

Carroll, Christopher D., and Andrew A. Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40(1), 41–71, Available at http://econ.jhu.edu/people/ccarroll/nature.pdf.

Deaton, Angus S. (1991): "Saving and Liquidity Constraints," *Econometrica*, 59, 1221–1248, Available at http://ideas.repec.org/a/ecm/emetrp/v59y1991i5p1221-48.html.

Gourinchas, Pierre-Olivier, and Jonathan Parker (2002): "Consumption Over the Life Cycle," *Econometrica*, 70(1), 47–89.

Horowitz, Joel L. (2001): "The Bootstrap," *Handbook of Econometrics*.

JUDD, KENNETH L. (1998): *Numerical Methods in Economics*. The MIT Press, Cambridge, Massachusetts.

PALUMBO, MICHAEL G (1999): "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle," *Review of Economic Studies*, 66(2), 395–421, Available at http://ideas.repec.org/a/bla/restud/v66y1999i2p395-421.html.

VALENCIA, FABIAN (2006): "Banks' Financial Structure and Business Cycles," Ph.D. thesis, Johns Hopkins University.