

Solution Methods for Microeconomic Dynamic Stochastic Optimization Problems

November 20, 2011

Christopher D. Carroll¹

Abstract

These notes describe some tools for solving microeconomic dynamic stochastic optimization problems, and show how to use those tools for estimating a standard life cycle consumption/saving model using microeconomic data. No attempt is made at a systematic overview of the many possible technical choices; instead, I present a specific set of methods that have proven useful in my own work (and explain why other popular methods, such as value function iteration, are a bad idea). Paired with these notes is *Mathematica* and Matlab software that solves the problems described in the text.

Keywords Dynamic Stochastic Optimization, Method of Simulated Moments, Structural Estimation

JEL codes E21, F41

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)

¹Carroll: Department of Economics, Johns Hopkins University, Baltimore, MD, <http://econ.jhu.edu/people/ccarroll/>, ccarroll@jhu.edu, Phone: (410) 516-7602

The 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. This draft incorporates several improvements related to new results in the paper “Theoretical Foundations of Buffer Stock Saving” (especially tools for approximating the consumption and value functions). Like the last major draft, it also builds on material 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, to Damiano Sandri for exceptionally insightful help in revising and updating the method of simulated moments estimation section, and to Weifeng Wu and Metin Uyanik for revising to be consistent with the ‘method of moderation’ and other improvements. All errors are my own.

Contents

1	Introduction	3
2	The Problem	4
3	Normalization	5
4	The Usual Theory and Some Notation	6
5	Solving the Next-to-Last Period	7
5.1	Discretizing the Distribution	7
5.2	The Approximate Consumption and Value Functions	8
5.3	An Interpolated Consumption Function	10
5.4	Interpolating Expectations	10
5.5	Value Function versus First Order Condition	13
5.6	Transformation	16
5.7	The Self-Imposed ‘Natural’ Borrowing Constraint and the a_{T-1} Lower Bound	17
5.8	The Method of Endogenous Gridpoints	20
5.9	Improving the a Grid	20
5.10	Approximating Precautionary Saving	21
5.11	The Method of Moderation	23
5.12	Approximating the Slope As Well As the Level	25
5.13	Imposing ‘Artificial’ Borrowing Constraints	27
5.14	Value	29
6	Recursion	31
6.1	Theory	31
6.2	<i>Mathematica</i> Background	32
6.3	Program Structure	33
6.3.1	Iteration	33
6.4	Results	33
7	Multiple Control Variables	34
7.1	Theory	34
7.2	Application	35
7.3	Implementation	37
7.4	Results	38
8	The Infinite Horizon	38
8.1	Convergence	40
8.2	Coarse then Fine θ Vec	40
9	Structural Estimation	40
9.1	Life Cycle Model	41

9.2 Estimation	42
10 Conclusion	46
A Further Details on SCF Data	48

1 Introduction

Economists have displayed impressive ingenuity over the years in finding ways to avoid having to solve plausibly-specified intertemporal choice problems under uncertainty (where by ‘plausible’ I mean problems with empirically realistic uncertainty and theoretically sensible utility). Budding graduate students are exposed (often with little motivation) to a host of tricks: Quadratic or Constant Absolute Risk Aversion utility, perfect markets, perfect insurance, perfect foresight, the “timeless” perspective, the restriction of uncertainty to very special kinds,¹ and more.

Explicit or not, the motivation is always to exchange an intractable general problem for a tractable specific alternative. But the burgeoning literature on numerical solutions has shown that the features that yield tractability also profoundly change the solution. A critic might say that the “tricks” are excuses to solve a problem that has been redefined to dodge the central difficulty: Understanding the role played by uncertainty in the solution to the ‘plausible’ model (as defined above).

Fortunately, the era when such tricks were necessary is drawing to a close, thanks to advances in mathematical analysis, increasing computing power, and the growing capabilities of numerical computation software. Together, these tools permit today’s desktop or even laptop computers to solve the kinds of plausible problems that required supercomputers a decade ago (and, before that, could not be solved at all).

These lecture notes provide a gentle introduction to a particular set of such tools and show how they can be used to solve some canonical problems in consumption choice and portfolio allocation. Specifically, the notes describe and solve optimization problems for a consumer facing uninsurable idiosyncratic risk to non-financial income (e.g. labor or transfer income),² with detailed intuitive discussion of the various mathematical and computational techniques that, together, speed the solution by many orders of magnitude compared to “brute force” methods. 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 deterministic 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 is presented of how to use these methods (via the statistical ‘method of simulated moments’ or MSM; sometimes called ‘simulated method of moments’ or SMM) to estimate structural parameters like the coefficient of relative risk aversion (*a la* Gourinchas and Parker (2002) and Cagetti (2003)).

¹E.g., lognormally distributed rate-of-return risk – but no labor income risk – under CRRA utility (the Merton (1969)-Samuelson (1969) model).

²Expenditure shocks (such as for medical needs, or to repair a broken automobile) are usually treated in a manner similar to labor income shocks. See Merton (1969) and Samuelson (1969) for a solution to the problem of a consumer whose only risk is rate-of-return risk on a financial asset; the combined case (both financial and nonfinancial risk) is solved below, and much more closely resembles the case with only nonfinancial risk than it does the case with only financial risk.

2 The Problem

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

$$\max \mathbb{E}_t \left[\sum_{n=0}^{T-t} \beta^n u(\mathbf{c}_{t+n}) \right], \quad (1)$$

and whose circumstances evolve according to the transition equations³

$$\mathbf{a}_t = \mathbf{m}_t - \mathbf{c}_t \quad (2)$$

$$\mathbf{b}_{t+1} = \mathbf{a}_t \mathbf{R}_{t+1} \quad (3)$$

$$\mathbf{y}_{t+1} = \mathbf{p}_{t+1} \theta_{t+1} \quad (4)$$

$$\mathbf{m}_{t+1} = \mathbf{b}_{t+1} + \mathbf{y}_{t+1} \quad (5)$$

where

β — pure time discount factor

\mathbf{a}_t — assets after all actions have been accomplished in period t

\mathbf{b}_{t+1} — ‘bank balances’ (nonhuman wealth) at the beginning of $t + 1$

\mathbf{c}_t — consumption in period t

\mathbf{m}_t — ‘market resources’ available for consumption (‘cash-on-hand’)

\mathbf{p}_{t+1} — ‘permanent labor income’ in period $t + 1$

\mathbf{R}_{t+1} — interest factor ($1 + \mathbf{r}_{t+1}$) from period t to $t + 1$

\mathbf{y}_{t+1} — noncapital income in period $t + 1$.

The exogenous variables evolve as follows:

$$\mathbf{R}_t = \mathbf{R} \quad - \text{constant interest factor} = 1 + \mathbf{r} \quad (6)$$

$$\mathbf{p}_{t+1} = \mathcal{G}_{t+1} \mathbf{p}_t \quad - \text{permanent labor income dynamics} \quad (7)$$

$$\log \theta \sim \mathcal{N}(-\sigma_\theta^2/2, \sigma_\theta^2) \quad - \text{lognormally distributed transitory shocks.} \quad (8)$$

Using the fact **ELogNorm**⁴ that if $\log \Phi \sim \mathcal{N}(\phi, \sigma_\phi^2)$ then $\log \mathbb{E}[\Phi] = \phi + \sigma_\phi^2/2$, assumption (8) guarantees that $\log \mathbb{E}[\theta] = 0$ which means that $\mathbb{E}[\theta] = 1$ (the mean value of the transitory shock is 1).

Equation (7) indicates that we are assuming that the average profile of income growth over the lifetime $\{\mathcal{G}\}_0^T$ is nonstochastic (allowing, for example, for typical career wage paths).⁵

³The usual analysis of dynamic programming problems combines these equations into a single expression; here, they are disarticulated to highlight the important point that several distinct processes (intertemporal choice, stochastic shocks, intertemporal returns, income growth) are involved in the transition from one period to the next.

⁴This fact is referred to as **ELogNorm** in the handout **MathFactsList**, 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 $\log \mathbb{E}[\theta] = 0$.

⁵This equation assumes that there are no shocks to permanent income. A large literature finds that, in reality, permanent shocks 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 permanent shocks clutters the exposition without adding much to the intuition, so permanent shocks have been omitted from the analysis (until the last section of the notes, which shows how to match the model with empirical micro data). For a full treatment of the theory including permanent shocks, see **Carroll (2011)**.

Finally, we assume that the utility function is of the Constant Relative Risk Aversion (CRRA), form, $u(\bullet) = \bullet^{1-\rho}/(1-\rho)$.

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

$$\mathbf{v}_t(\mathbf{m}_t, \mathbf{p}_t) = \max_{\mathbf{c}_t} u(\mathbf{c}_t) + \mathbb{E}_t[\beta \mathbf{v}_{t+1}(\mathbf{m}_{t+1}, \mathbf{p}_{t+1})] \quad (9)$$

subject to the Dynamic Budget Constraint (DBC) (2)-(5) given above, where \mathbf{v}_t measures total expected discounted utility from behaving optimally now and henceforth.

3 Normalization

The single most powerful method for speeding the solution of such models is to redefine the problem in a way that reduces the number of state variables (if possible). In the consumption problem the obvious idea is to see whether the problem can be rewritten in terms of the ratio of various variables to permanent noncapital ('labor') income \mathbf{p}_t .

In the last period of life, there is no future, $\mathbf{v}_{T+1} = 0$, so the optimal plan is to consume everything, implying that

$$\mathbf{v}_T(\mathbf{m}_T, \mathbf{p}_T) = \frac{\mathbf{m}_T^{1-\rho}}{1-\rho}. \quad (10)$$

Now define nonbold variables as the bold variable divided by the level of permanent income in the same period, so that, for example, $m_T = \mathbf{m}_T/\mathbf{p}_T$; and define $v_T(m_T) = u(m_T)$.⁶ Using the fact that $u(xy) = x^{1-\rho}u(y)$, equation (10) can be rewritten as

$$\mathbf{v}_T(\mathbf{m}_T, \mathbf{p}_T) = \mathbf{p}_T^{1-\rho} \frac{m_T^{1-\rho}}{1-\rho} = \mathbf{p}_{T-1}^{1-\rho} \mathcal{G}_T^{1-\rho} \frac{m_T^{1-\rho}}{1-\rho} = \mathbf{p}_{T-1}^{1-\rho} \mathcal{G}_T^{1-\rho} v_T(m_T). \quad (11)$$

Define a new optimization problem:

$$\begin{aligned} v_t(m_t) &= \max_{c_t} u(c_t) + \mathbb{E}_t[\beta_{t+1} \mathcal{G}_{t+1}^{1-\rho} v_{t+1}(m_{t+1})], \\ \text{s.t.} \\ a_t &= m_t - c_t \\ m_{t+1} &= \underbrace{(\mathcal{R}/\mathcal{G}_{t+1}) a_t}_{\equiv \mathcal{R}_{t+1}} + \theta_{t+1} \end{aligned} \quad (12)$$

where for generality we now allow for time variation in the discount factor β_{t+1} . The accumulation equation is the normalized version of the transition equation for

⁶Nonbold value is bold value divided by $\mathbf{p}^{1-\rho}$ rather than \mathbf{p} .

\mathbf{m}_{t+1} .⁷ Then it is easy to see that for $t = T - 1$,

$$\mathbf{v}_{T-1}(\mathbf{m}_{T-1}, \mathbf{p}_{T-1}) = \mathbf{p}_{T-1}^{1-\rho} \mathbf{v}_{T-1}(m_{T-1}) \quad (13)$$

and so on back to all earlier periods. Hence, if we solve the problem (12) which has only a single state variable (m_t), we can obtain the levels of the value function, consumption, and all other variables of interest simply by multiplying the results by the appropriate function of \mathbf{p}_t , e.g. $\mathbf{c}_t(\mathbf{m}_t, \mathbf{p}_t) = \mathbf{p}_t c_t(m_t/\mathbf{p}_t)$ or $\mathbf{v}_t(\mathbf{m}_t, \mathbf{p}_t) = \mathbf{p}_t^{1-\rho} v_t(m_t)$. We have thus reduced the problem from two continuous state variables to one (and thus enormously simplified our lives).

For some problems it will not be obvious 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.

4 The Usual Theory and Some Notation

Dropping time subscripts on β to reduce clutter, the first order condition for (12) with respect to c_t is

$$\begin{aligned} u'(c_t) &= \mathbb{E}_t[\beta \mathcal{R}_{t+1} \mathcal{G}_{t+1}^{1-\rho} v'_{t+1}(m_{t+1})] \\ &= \mathbb{E}_t[\beta \mathcal{R} \mathcal{G}_{t+1}^{-\rho} v'_{t+1}(m_{t+1})] \end{aligned} \quad (14)$$

and because the **Envelope** theorem tells us that

$$v'_t(m_t) = \mathbb{E}_t[\beta \mathcal{R} \mathcal{G}_{t+1}^{-\rho} v'_{t+1}(m_{t+1})] \quad (15)$$

we can substitute the LHS of (15) for the RHS of (14) to get

$$u'(c_t) = v'_t(m_t) \quad (16)$$

and rolling this equation forward one period yields

$$u'(c_{t+1}) = v'_{t+1}(a_t \mathcal{R}_{t+1} + \theta_{t+1}) \quad (17)$$

while substituting the LHS in equation (14) gives us the Euler equation for consumption

$$u'(c_t) = \mathbb{E}_t[\beta \mathcal{R} \mathcal{G}_{t+1}^{-\rho} u'(c_{t+1})]. \quad (18)$$

Now note that in equation (17) neither m_t nor c_t has any *direct* effect on v_{t+1} - only the difference between them (i.e. unconsumed market resources or ‘assets’ a_t) matters. It is therefore possible (and will turn out to be convenient) to define a

⁷Derivation:

$$\begin{aligned} \mathbf{m}_{t+1}/\mathbf{p}_{t+1} &= (\mathbf{m}_t - \mathbf{c}_t)\mathcal{R}/\mathbf{p}_{t+1} + \mathbf{y}_{t+1}/\mathbf{p}_{t+1} \\ m_{t+1} &= \left(\frac{\mathbf{m}_t}{\mathbf{p}_t} - \frac{\mathbf{c}_t}{\mathbf{p}_t} \right) \mathcal{R} \frac{\mathbf{p}_t}{\mathbf{p}_{t+1}} + \frac{\mathbf{y}_{t+1}}{\mathbf{p}_{t+1}} \\ &= \underbrace{(m_t - c_t)(\mathcal{R}/\mathcal{G}_{t+1})}_{a_t} + \theta_{t+1}. \end{aligned}$$

function⁸

$$\mathbf{v}_t(a_t) = \mathbb{E}_t[\beta \mathcal{G}_{t+1}^{1-\rho} \mathbf{v}_{t+1}(\mathcal{R}_{t+1}a_t + \theta_{t+1})] \quad (19)$$

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

$$\mathbf{v}'_t(a_t) = \mathbb{E}_t[\beta \mathcal{R} \mathcal{G}_{t+1}^{-\rho} \mathbf{v}'_{t+1}(\mathcal{R}_{t+1}a_t + \theta_{t+1})] \quad (20)$$

or, substituting from equation (17),

$$\mathbf{v}'_t(a_t) = \mathbb{E}_t[\beta \mathcal{R} \mathcal{G}_{t+1}^{-\rho} u'(c_{t+1}(\mathcal{R}_{t+1}a_t + \theta_{t+1}))]. \quad (21)$$

Finally, note for future use that the first order condition (14) can be rewritten as

$$u'(c_t) = \mathbf{v}'_t(m_t - c_t). \quad (22)$$

5 Solving the Next-to-Last Period

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

$$\mathbf{v}_{T-1}(m_{T-1}) = \max_{c_{T-1}} u(c_{T-1}) + \beta_T \mathbb{E}_{T-1} [\mathcal{G}_T^{1-\rho} \mathbf{v}_T((m_{T-1} - c_{T-1})\mathcal{R}_T + \theta_T)],$$

and using (1) the fact that $\mathbf{v}_T = u(c)$; (2) the definition of $u(c)$; (3) the definition of the expectations operator, and (4) the fact that \mathcal{G}_T is nonstochastic, this becomes

$$\mathbf{v}_{T-1}(m_{T-1}) = \max_{c_{T-1}} \frac{c_{T-1}^{1-\rho}}{1-\rho} + \beta_T \mathcal{G}_T^{1-\rho} \int_0^\infty \frac{((m_{T-1} - c_{T-1})\mathcal{R}_T + \theta)^{1-\rho}}{1-\rho} d\mathcal{F}(\theta)$$

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

In principle, the maximization implicitly defines a function $c_{T-1}(m_{T-1})$ that yields optimal 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 m_{T-1} we must use numerical computational tools to find the c_{T-1} that maximizes the expression. 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

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

Define a set of points from \sharp_0 to \sharp_n on the $[0, 1]$ interval as the elements of the set $\sharp = \{0, 1/n, 2/n, \dots, 1\}$.⁹ Call the inverse of the θ distribution \mathcal{F}^{-1} , and define the points $\sharp_i^{-1} = \mathcal{F}^{-1}(\sharp_i)$. Then the conditional mean of θ in each of the intervals numbered 1 to n is:

$$\theta_i \equiv \mathbb{E}[\theta | \sharp_{i-1}^{-1} \leq \theta < \sharp_i^{-1}] = \int_{\sharp_{i-1}^{-1}}^{\sharp_i^{-1}} \vartheta d\mathcal{F}(\vartheta). \quad (23)$$

⁸The peculiar letter designating our new function is pronounced ‘Gothic v’.

⁹These points define intervals that constitute a partition of the domain of \mathcal{F} .

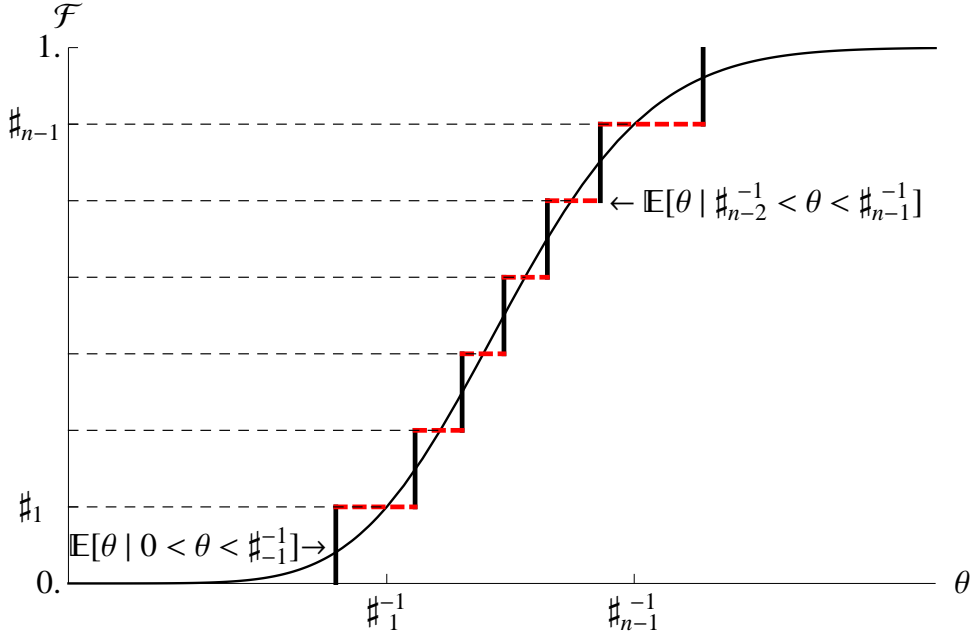


Figure 1 Discrete Approximation to Lognormal Distribution \mathcal{F}

The method is illustrated in Figure 1. The solid continuous curve represents the “true” CDF $\mathcal{F}(\theta)$ for a lognormal distribution such that $\mathbb{E}[\theta] = 1$, $\sigma_\theta = 0.1$. The short vertical line segments represent the n equiprobable values of θ_i which are used to approximate this distribution.¹⁰

Recalling our definition of $\mathbf{v}_t(a_t)$, for $t = T - 1$

$$\mathbf{v}_{T-1}(a_{T-1}) = \beta_T \mathcal{G}_T^{1-\rho} \left(\frac{1}{n} \right) \sum_{i=1}^n \frac{(\mathcal{R}_T a_{T-1} + \theta_i)^{1-\rho}}{1-\rho} \quad (24)$$

so we can rewrite the maximization problem as

$$\mathbf{v}_{T-1}(m_{T-1}) = \max_{c_{T-1}} \left\{ \frac{c_{T-1}^{1-\rho}}{1-\rho} + \mathbf{v}_{T-1}(m_{T-1} - c_{T-1}) \right\}. \quad (25)$$

5.2 The Approximate Consumption and Value Functions

Given a particular value of m_{T-1} , a numerical maximization routine can now find the c_{T-1} that maximizes (25) in a reasonable amount of time. The *Mathematica* program that solves exactly this problem called `2period.m`. (The archive also contains parallel Matlab programs, but these notes will dwell on the specifics of the *Mathematica* implementation, which is superior in many respects.)

The first thing `2period.m` does is to read in the file `functions.m` which contains definitions of the consumption and value functions; `functions.m` also defines the function `SolveAnotherPeriod` which, given the existence in memory of a solution for period $t + 1$, solves for period t .

¹⁰More sophisticated approximation methods exist (e.g. Gauss-Hermite quadrature; see Kopecky and Suen (2010) for a discussion of other alternatives), but the method described here is easy to understand, quick to calculate, and has additional advantages briefly described in the discussion of simulation below.

The next step is to run the programs `setup_params.m`, `setup_grids.m`, `setup_shocks.m`, respectively. `setup_params.m` sets values for the parameter values like the coefficient of relative risk aversion. `setup_shocks.m` calculates the values for the θ_i defined above (and puts those values, and the (identical) probability associated with each of them, in the vector variables \thetaVals and \thetaProb). Finally, `setup_grids.m` constructs a list of potential values of cash-on-hand and saving, then puts them in the vector variables $mVec = aVec = \{0, 1, 2, 3, 4\}$ respectively. Then `2period.m` runs the program `setup_lastperiod.m` which defines the elements necessary to determine behavior in the last period, in which $c_T(m) = m$ and $v_T(m) = u(m)$.

After all the setup, the only remaining step in `2period.m` is to invoke `SolveAnotherPeriod`, which constructs the solution for period $T - 1$ given the presence of the solution for period T (constructed by `setup_lastperiod.m`).

Because we will always be comparing our solution to the perfect foresight solution, we also construct the variables required to characterize the perfect foresight consumption function in periods prior to T . In particular, we construct the list `yExpPDV` (which contains the PDV of expected income – ‘expected human wealth’), and `yMinPDV` which contains the minimum possible discounted value of future income at the beginning of period $T - 1$ (‘minimum human wealth’).¹¹

The perfect foresight consumption function is also constructed (`setup_PerfectForesightSolut`). This program uses the fact that, in *Mathematica*, functions can be saved as objects using the commands `#` and `&`. The `#` denotes the argument of the function, while the `&`, 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 `cF = {(# - 1 + Last[yExpPDV]) Last[κMin] &}`, where `Last[yExpPDV]` gives the just-constructed PDV of human wealth at the beginning of T (equal to 1, since current income is included in h_T), and `Last[κMin]` gives the perfect foresight marginal propensity to consume (equal to 1, since it is optimal to spend all resources in the last period). Since `#` in the code stands in for what was called m in the model, the discounted total wealth is decomposed into discounted non-human wealth `# - 1` and discounted human wealth `Last[yExpPDV]`. The resulting formula then corresponds to $\bar{c}_T = (m_T - 1 + h_T)\underline{\kappa}_T$, which translates to $\bar{c}_T = m_T$ for $h_T = \underline{\kappa}_T = 1$.

The perfect foresight marginal propensity to save parameter

$$\lambda_t = (1/R)(R\beta)^{1/\rho} \quad (26)$$

is also defined (λ_t is useful in generating perfect foresight MPCs in earlier periods of life).¹²

The program then constructs behavior for one iteration back from the last period of life by using the function `AddNewPeriodToParamLifeDates`. Using the *Mathematica* command `AppendTo`, various existing lists (which characterized the solution for period T) are redefined to include an additional element representing the relevant formulas in the second to last period of life. For example, `κMin` now

¹¹This is useful in determining the search range for the optimal level of consumption in the maximization problem.

¹²Detailed discussion can be found in Carroll (2011).

has two elements. The second element, given by $1/(1 + \text{Last}[\lambda]/\text{Last}[\kappa\text{Min}])$, is the perfect foresight marginal propensity to consume in $t = T - 1$.¹³

Next, the program defines a function `v[at_]` (in `functions_stable.m`) which is the exact implementation of (19): 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 (in `functions.m`). This expression loops over the values of the variable `mVec`, solving the maximization problem (given in equation (25)):

$$\max_c \quad u[c] + v[\text{mVec}[[i]] - c] \quad (27)$$

for each of the i values of `mVec` (henceforth let's call these points $m_{T-1,i}$). 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 $v_{T-1,i}$ and the other with the consumption levels $c_{T-1,i}$ associated with the $m_{T-1,i}$.

5.3 An Interpolated Consumption Function

Now we use the first of the really convenient built-in features of *Mathematica*. Given data of the form $\{\{m_1, y_1\}, \{m_2, y_2\} \dots\}$ *Mathematica* can create an object called an `InterpolatingFunction` which when applied to an input m will yield the value of y that corresponds to a linear interpolation of the value of y from the points in the `InterpolatingFunction` object. We can therefore define a function $\hat{c}_{T-1}(m_{T-1})$ which, when called with an m_{T-1} that is equal to one of the points in `mVec[[i]]` returns the associated value of $c_{T-1,i}$, and when called with a value of m_{T-1} that is not exactly equal to one of the `mVec[[i]]`, returns the value of c that reflects a linear interpolation between the $c_{T-1,i}$ associated with the two `mVec[[i]]` points nearest to m_{T-1} . Thus if the function is called with $m_{T-1} = 1.75$ and the nearest gridpoints are $m_{j,T-1} = 1$ and $m_{k,T-1} = 2$ then the value of c_{T-1} returned by the function would be $(0.25c_{j,T-1} + 0.75c_{k,T-1})$. We can define a numerical approximation to the value function $\hat{v}_{T-1}(m_{T-1})$ in an exactly analogous way.

Figures 2 and 3 show plots of the \hat{c}_{T-1} and \hat{v}_{T-1} `InterpolatingFunctions` that are generated by the program `2PeriodInt.m`. While the \hat{c}_{T-1} function looks very smooth, the fact that the \hat{v}_{T-1} function is a set of line segments is very evident. This figure provides the beginning of the intuition for why trying to approximate the value function directly is a bad idea (in this context).¹⁴

5.4 Interpolating Expectations

`2period.m` works well in the sense that it generates a good approximation to the true optimal consumption function. However, there is a clear inefficiency in the program: Since it uses equation (25), for every value of m_{T-1} the program must calculate the

¹³A proof given in Carroll (2011) shows that this is also a recurring formula that extends inductively to earlier periods.

¹⁴For some problems, especially ones with discrete choices, value function approximation is unavoidable; nevertheless, even in such problems, the techniques sketched below can be very useful across much of the range over which the problem is defined.

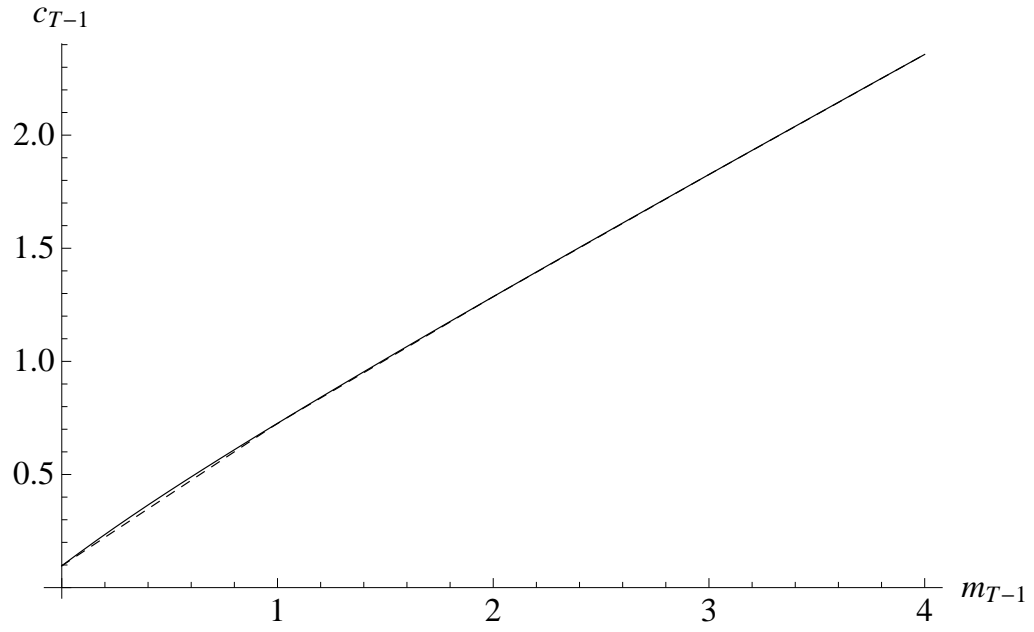


Figure 2 $c_{T-1}(m_{T-1})$ (solid) versus $\hat{c}_{T-1}(m_{T-1})$ (dashed)

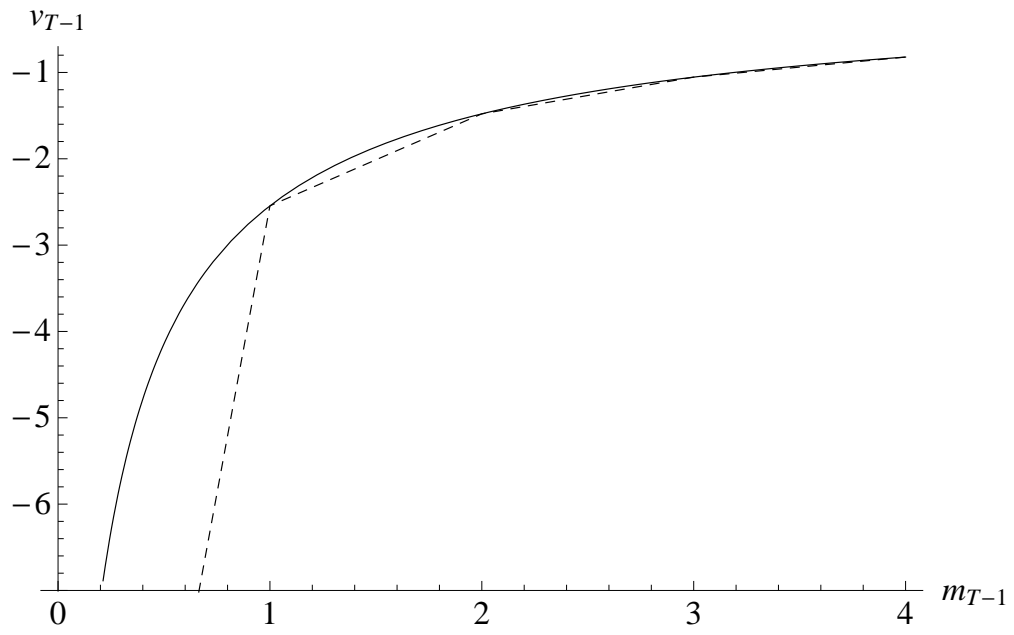


Figure 3 v_{T-1} (solid) versus $\hat{v}_{T-1}(m_{T-1})$ (dashed)

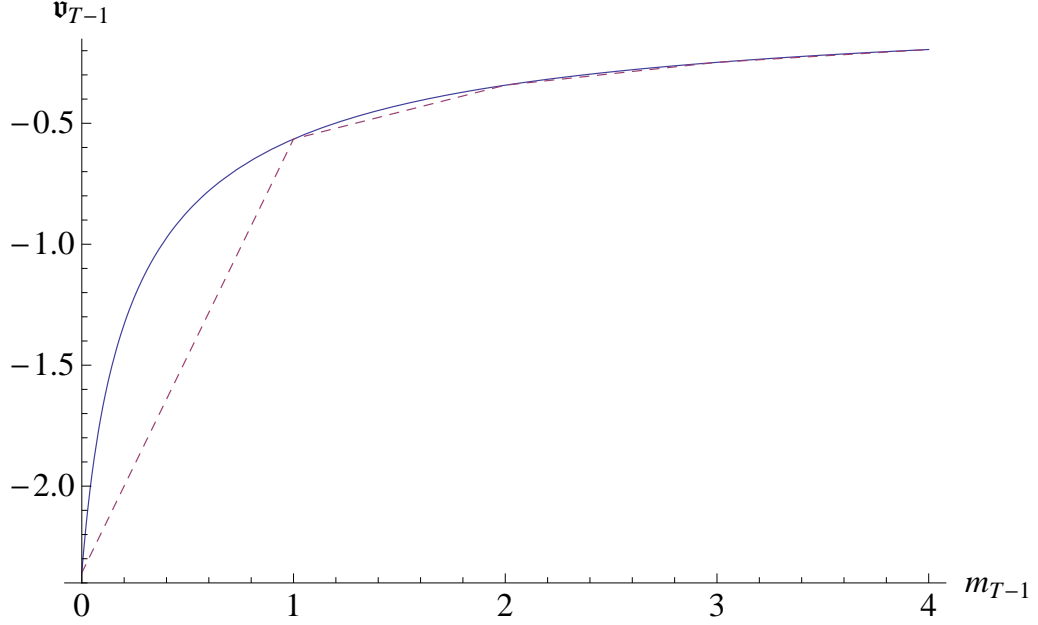


Figure 4 End-Of-Period Value $\mathbf{v}_{T-1}(a_{T-1})$ (solid) versus $\hat{\mathbf{v}}_{T-1}(a_{T-1})$ (dashed)

utility consequences of various possible choices of c_{T-1} as it searches for the best choice. But for any given value of a_{T-1} , there is a good chance that the program may end up calculating the corresponding \mathbf{v} many times while maximizing utility from different m_{T-1} 's. For example, it is possible that the program will calculate the value of ending the period with $a_{T-1} = 0$ 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 direct numerical approximation to the value function: Define a grid of possible values for saving at time $T-1$, \vec{a}_{T-1} (`aVec` in `setup_grids.m`), designating the specific points $a_{T-1,i}$; for each of these values of $a_{T-1,i}$, calculate the vector $\vec{\mathbf{v}}_{T-1}$ as the collection of points $\mathbf{v}_{T-1,i} = \mathbf{v}_{T-1}(a_{T-1,i})$ using equation (19); then construct an `InterpolatingFunction` object $\hat{\mathbf{v}}_{T-1}(a_{T-1})$ from the list of values $\{\vec{a}_{T-1}, \vec{\mathbf{v}}_{T-1}\}^\top = \{\{a_{T-1,1}, \mathbf{v}_{T-1,1}\}, \{a_{T-1,2}, \mathbf{v}_{T-1,2}\} \dots\}$.

Thus, we are now interpolating for the expected value of saving.¹⁵ The program `2periodIntExp.m` solves this problem. Figure 4 compares the true value function to the `InterpolatingFunction` approximation; the functions are of course identical at the gridpoints chosen for a_{T-1} and they appear reasonably close except in the region below $m_{T-1} = 1$.

Nevertheless, the resulting consumption rule obtained when $\hat{\mathbf{v}}_{T-1}(a_{T-1})$ is used instead of $\mathbf{v}_{T-1}(a_{T-1})$ is 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 distressing because Carroll and Kimball (1996) prove that

¹⁵What we are doing here is closely related to ‘the method of parameterized expectations’ of den Haan and Marcet (1990); the only difference is that our method is essentially a nonparametric version.

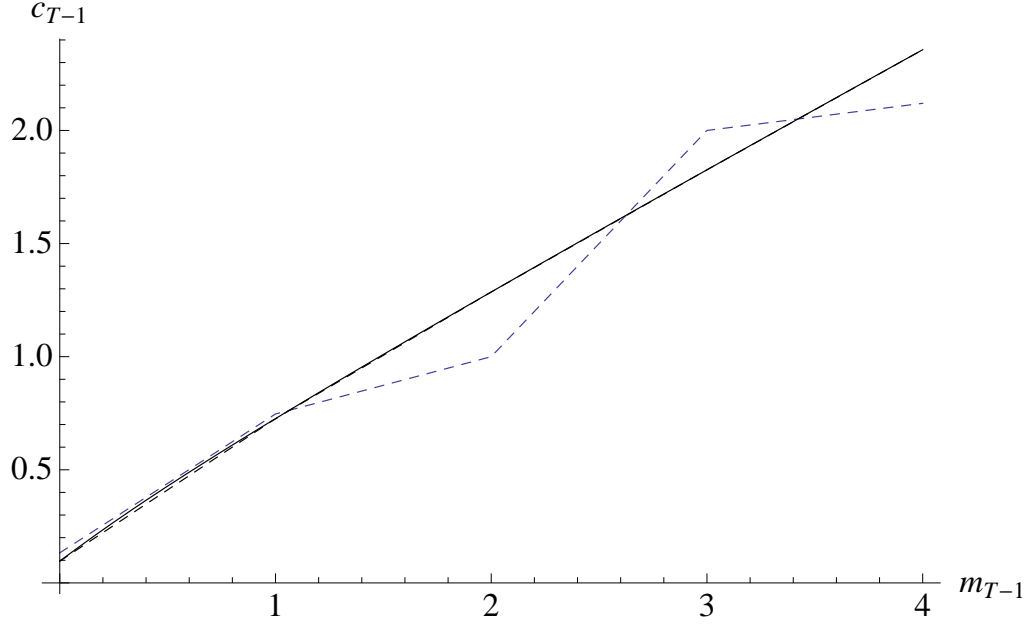


Figure 5 $c_{T-1}(m_{T-1})$ (solid) versus $\hat{c}_{T-1}(m_{T-1})$ (dashed)

the correct consumption function is strictly concave in a wide class of problems that includes this problem.

5.5 Value Function versus First Order Condition

Loosely speaking, our difficulty is caused by the fact that the consumption choice is governed by the *marginal* value function, not by the *level* of the value function (which is what we approximated). To see this, recall that a quadratic utility function exhibits risk aversion because with a stochastic c ,

$$\mathbb{E}[-(c - \ell)^2] < -(\mathbb{E}[c] - \ell)^2 \quad (28)$$

where ℓ is the ‘bliss point’. However, unlike the CRRA utility function, with quadratic utility the consumption/saving *behavior* of consumers is unaffected by risk since behavior is determined by the first order condition, which depends on *marginal* utility, and when utility is quadratic, marginal utility is unaffected by risk:

$$\mathbb{E}[-2(c - \ell)] = -2(\mathbb{E}[c] - \ell). \quad (29)$$

Intuitively, if one’s goal is to accurately capture choices that are governed by marginal value, numerical techniques that approximate the *marginal* value function will presumably lead to a more accurate approximation to dsoprtimal behavior than techniques that approximate the *level* of the value function.

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

$$u'(c_{T-1}) = \beta_T \mathbb{E}_{T-1}[\mathcal{G}_T^{-\rho} R u'(c_T)] \quad (30)$$

$$c_{T-1}^{-\rho} = \beta_T \left(\frac{1}{n} \right) \sum_{i=1}^n \mathcal{G}_T^{-\rho} (R(m_{T-1} - c_{T-1}) + \theta_i)^{-\rho}. \quad (31)$$

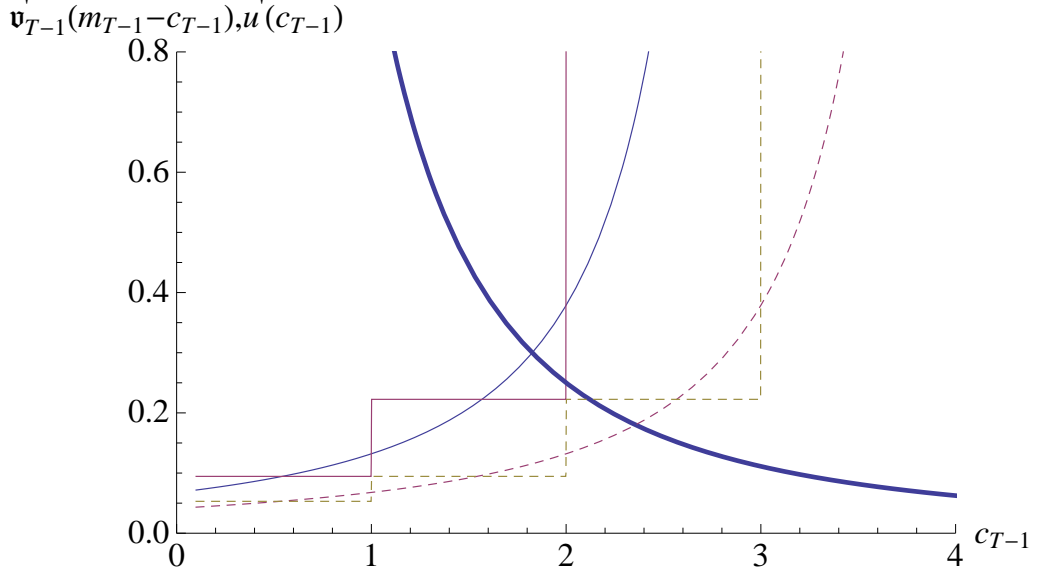


Figure 6 $u'(c)$ versus $\mathbf{v}'_{T-1}(3 - c)$, $\mathbf{v}'_{T-1}(4 - c)$, $\hat{\mathbf{v}}'_{T-1}(3 - c)$, $\hat{\mathbf{v}}'_{T-1}(4 - c)$

The downward-sloping curve in Figure 6 shows the value of $c_{T-1}^{-\rho}$ for our baseline parameter values for $0 \leq c_{T-1} \leq 4$ (the horizontal axis). The solid upward-sloping curve shows the value of the RHS of (31) 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 (31). The value of c_{T-1} for which the RHS and LHS of (31) 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 dashed curve shows the expected value of the RHS of (31) under the assumption that $m_{T-1} = 4$, and the intersection of this curve with $u'(c_{T-1})$ 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 $\hat{\mathbf{v}}_{T-1}(a_{T-1})$. Because we have constructed $\hat{\mathbf{v}}_{T-1}$ as a linear interpolation, the slope of $\hat{\mathbf{v}}_{T-1}(a_{T-1})$ between any two adjacent points $\{a_{T-1,i}, a_{i+1,T-1}\}$ 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 $\hat{\mathbf{v}}_{T-1}(a_{T-1})$ follows a step function.

The solid-line step function in Figure 6 depicts the actual value of $\hat{\mathbf{v}}'_{T-1}(3 - c_{T-1})$. When we attempt to find optimal values of c_{T-1} given m_{T-1} using $\hat{\mathbf{v}}_{T-1}(a_{T-1})$, the numerical optimization routine will return the c_{T-1} for which $u'(c_{T-1}) = \hat{\mathbf{v}}'_{T-1}(m_{T-1} - c_{T-1})$. Thus, for $m_{T-1} = 3$ the program will return the value of c_{T-1} for which the downward-sloping $u'(c_{T-1})$ curve intersects with the $\hat{\mathbf{v}}'_{T-1}(3 - c_{T-1})$; 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

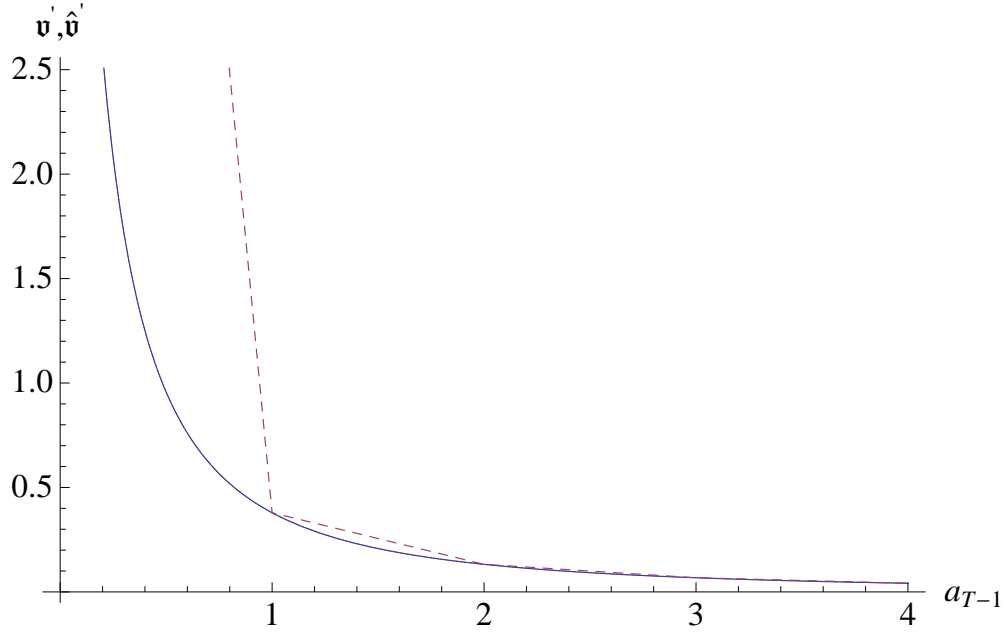


Figure 7 $v'_{T-1}(a_{T-1})$ versus $\hat{v}'_{T-1}(a_{T-1})$

of $u'(c_{T-1})$ with $\hat{v}'_{T-1}(4 - c_{T-1})$; 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 obviously obtain much better estimates of the point of intersection between $u'(c_{T-1})$ and $v'_{T-1}(m_{T-1} - c_{T-1})$ if our estimate of \hat{v}'_{T-1} were not a step function. In fact, we already know how to construct linear interpolations to functions, so the obvious next step is to construct a linear interpolating approximation to the *expected marginal value of end-of-period assets function* v' . That is, we calculate

$$v'_{T-1}(a_{T-1}) = \beta_T R \mathcal{G}_T^{-\rho} \left(\frac{1}{n} \right) \sum_{i=1}^n (\mathcal{R}_T a_{T-1} + \theta_i)^{-\rho} \quad (32)$$

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

The program file `functionsIntExpFOC.m` therefore uses the function `va[at_]` defined in `functions_stable.m` as the embodiment of equation (32), 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).

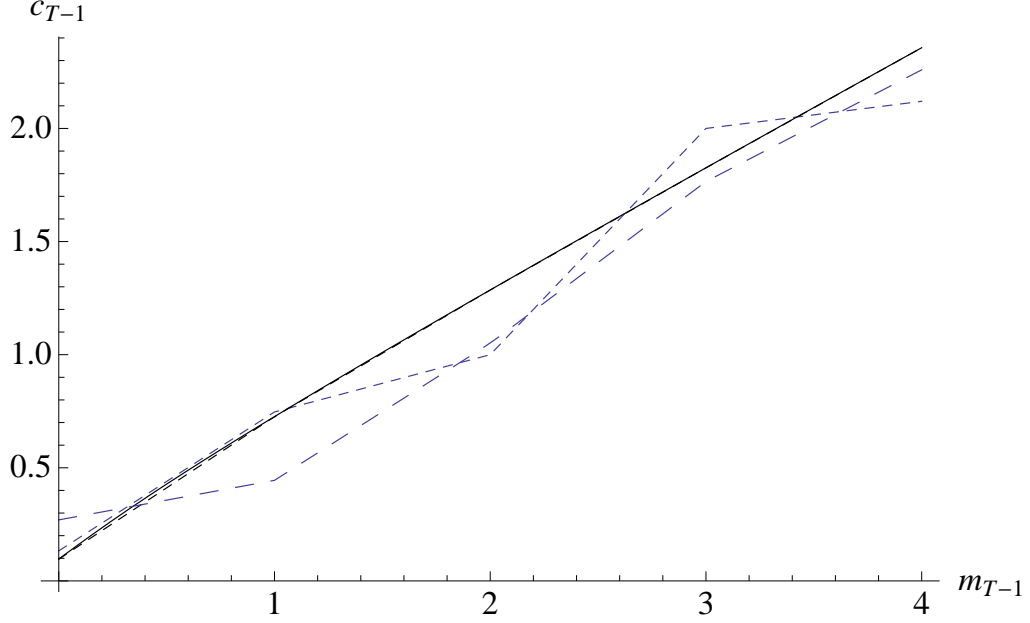


Figure 8 $c_{T-1}(m_{T-1})$ (solid) Versus Two Methods for Constructing $\hat{c}_{T-1}(m_{T-1})$

5.6 Transformation

However, even the new-and-improved consumption function diverges appallingly from the true solution, especially at lower values of m . That is because the linear interpolation does an increasingly poor job of capturing the nonlinearity of $\mathbf{v}'_{T-1}(a_{T-1})$ at lower and lower levels of a .

This is where we unveil our next trick. To understand the logic, start by considering the case where $\mathcal{R}_T = \beta_T = \mathcal{G}_T = 1$ and there is no uncertainty (that is, we know for sure that income next period will be $\theta_T = 1$). The final Euler equation is then:

$$c_{T-1}^{-\rho} = c_T^{-\rho}. \quad (33)$$

In the case we are now considering with no uncertainty and no liquidity constraints, the optimizing consumer does not care whether a unit of income is scheduled to be received in the future period T or the current period $T - 1$; there is perfect certainty that the income will be received, so the consumer treats it as equivalent to a unit of current wealth. Total resources therefore are comprised of two types: current market resources m_{T-1} and ‘human wealth’ (the PDV of future income) of $\mathfrak{h}_{T-1} = 1$ (where we use the Gothic font to signify that this is the expectation, as of the END of the period, of the income that will be received in future periods; it does not include current income, which has already been incorporated into m_{T-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 $m_{T-1} + \mathfrak{h}_{T-1} = m_{T-1} + 1$, and since $\mathbf{v}'_{T-1}(m_{T-1}) = u'(c_{T-1})$ this implies that

$$\mathbf{v}'_{T-1}(m_{T-1}) = \left(\frac{m_{T-1} + 1}{2} \right)^{-\rho}. \quad (34)$$

Of course, this is a highly nonlinear function. However, if we raise both sides of (34) to the power $(-1/\rho)$ the result is a linear function:

$$[v'_{T-1}(m_{T-1})]^{-1/\rho} = \frac{m_{T-1} + 1}{2}. \quad (35)$$

This is a specific example of a general phenomenon: A theoretical literature cited in Carroll and Kimball (1996) establishes 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 $(\gamma m_t + \zeta)^{-\rho}$ for some constants $\{\gamma, \zeta\}$. This means that if we were solving the perfect foresight problem numerically, we could always calculate a numerically exact (because linear) interpolation. To put this in intuitive terms, the problem we are facing is that the marginal value function is highly nonlinear. But we have a compelling solution to that problem, because the nonlinearity springs largely 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. Specifically, applying the foregoing insights to the end-of-period value function v_{T-1} , we can define

$$c_{T-1}(a_{T-1}) \equiv [v'_{T-1}(a_{T-1})]^{-1/\rho} \quad (36)$$

which would be linear in the perfect foresight case. Thus, our procedure is to calculate the values of $c_{T-1,i}$ at each of the $a_{T-1,i}$ gridpoints, with the idea that we will construct \hat{c}_{T-1} as the interpolating function connecting these points.

5.7 The Self-Imposed ‘Natural’ Borrowing Constraint and the a_{T-1} Lower Bound

This is the appropriate moment to ask an awkward question that we have neglected until now: How should a function like \hat{c}_{T-1} be evaluated outside the range of points spanned by $\{a_{T-1,1}, \dots, a_{T-1,n}\}$ for which we have calculated the corresponding $c_{T-1,i}$ gridpoints used to produce our linearly interpolating approximation \hat{c}_{T-1} (as described in section 5.3)?

The natural answer would seem to be linear extrapolation; for example, we could use

$$\hat{c}_{T-1}(a_{T-1}) = \hat{c}_{T-1}(a_{T-1,1}) + \hat{c}'_{T-1}(a_{T-1,1})(a_{T-1} - a_{T-1,1}) \quad (37)$$

for values of $a_{T-1} < a_{T-1,1}$. Unfortunately, this approach will lead us into difficulties. To see why, consider what happens to the true (not approximated) $v_{T-1}(a_{T-1})$ as a_{T-1} approaches the value $\underline{a}_{T-1} = -\underline{\theta}\mathcal{R}_T^{-1}$. From (32) we have

$$\lim_{a_{T-1} \downarrow \underline{a}_{T-1}} v'_{T-1}(a_{T-1}) = \lim_{a_{T-1} \downarrow \underline{a}_{T-1}} \beta_T \mathcal{R} \mathcal{G}_T^{-\rho} \left(\frac{1}{n} \right) \sum_{i=1}^n (a_{T-1} \mathcal{R}_T + \theta_i)^{-\rho}. \quad (38)$$

But since $\underline{\theta} = \theta_1$, exactly at $a_{T-1} = \underline{a}_{T-1}$ the first term in the summation would be $(-\underline{\theta} + \theta_1)^{-\rho} = 1/0^\rho$ which is infinity. The reason is simple: $-\underline{a}_{T-1}$ is the PDV, as of $T-1$, of the minimum possible realization of income in period T ($\mathcal{R}_T \underline{a}_{T-1} = -\theta_1$). Thus, if the consumer borrows an amount greater than or equal to $\underline{\theta}\mathcal{R}_T^{-1}$ (that is, if the consumer ends $T-1$ with $a_{T-1} \leq -\underline{\theta}\mathcal{R}_T^{-1}$) and then draws the worst possible income shock in period T , he will have to consume zero in period T (or a negative

amount), which yields $-\infty$ utility and ∞ marginal utility (or undefined utility and marginal utility).

These reflections lead us to the conclusion that the consumer faces a ‘self-imposed’ liquidity constraint (which results from the precautionary motive): He will never borrow an amount greater than or equal to $\underline{\theta}\mathcal{R}_T^{-1}$ (that is, assets will never reach the lower bound of \underline{a}_{T-1}).¹⁶ The constraint is ‘self-imposed’ in the sense that if the utility function were different (say, Constant Absolute Risk Aversion), the consumer would be willing to borrow more than $\underline{\theta}\mathcal{R}_T^{-1}$ because a choice of zero or negative consumption in period T would yield some finite amount of utility (though it is very unclear what a proper economic interpretation of negative consumption might be – this is an important reason why CARA utility, like quadratic utility, is increasingly not used for serious quantitative work, though it is still useful for illustrative and pedagogical purposes).

This self-imposed constraint cannot be captured well when the \mathbf{v}'_{T-1} function is approximated by a piecewise linear function like $\hat{\mathbf{v}}'_{T-1}$, because a linear approximation can never reach the correct gridpoint for $\mathbf{v}'_{T-1}(\underline{a}_{T-1}) = \infty$. To see what will happen instead, note first that if we are approximating \mathbf{v}'_{T-1} the smallest value in `aVec` must be greater than \underline{a}_{T-1} (because the expectation for any gridpoint $\leq \underline{a}_{T-1}$ is undefined). Then when the approximating \mathbf{v}'_{T-1} function is evaluated at some value less than the first element in `aVec[1]`, the approximating function will linearly extrapolate the slope that characterized the lowest segment of the piecewise linear approximation (between `aVec[1]` and `aVec[2]`), a procedure that will return a positive finite number, even if the requested a_{T-1} point is below \underline{a}_{T-1} . This means that the precautionary saving motive is understated, and by an arbitrarily large amount as the level of assets approaches its true theoretical minimum \underline{a}_{T-1} .

The foregoing logic demonstrates that the marginal value of saving approaches infinity as $a_{T-1} \downarrow \underline{a}_{T-1} = -\underline{\theta}\mathcal{R}_T^{-1}$. But this implies that $\lim_{a_{T-1} \downarrow \underline{a}_{T-1}} \mathbf{c}_{T-1}(a_{T-1}) = (\mathbf{v}'_{T-1}(a_{T-1}))^{-1/\rho} = 0$; that is, as a approaches its minimum possible value, the corresponding amount of c must approach *its* minimum possible value: zero.

The upshot of this discussion is a realization that all we need to do is to augment each of the \vec{a}_{T-1} and \vec{c}_{T-1} vectors with an extra point so that the first element in the list used to produce our `InterpolatingFunction` is $\{a_{T-1,0}, c_{T-1,0}\} = \{\underline{a}_{T-1}, 0\}$.

Figure 9 plots the results (generated by the program `2periodIntExpFOCInv.m`). The solid line calculates the exact numerical value of $\mathbf{c}_{T-1}(a_{T-1})$ while the dashed line is the linear interpolating approximation $\hat{\mathbf{c}}_{T-1}(a_{T-1})$. This figure well illustrates the value of the transformation: The true function is close to linear, and so the linear approximation is almost indistinguishable from the true function except at the very lowest values of a_{T-1} .

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

¹⁶Another term for a constraint of this kind is the ‘natural borrowing constraint.’

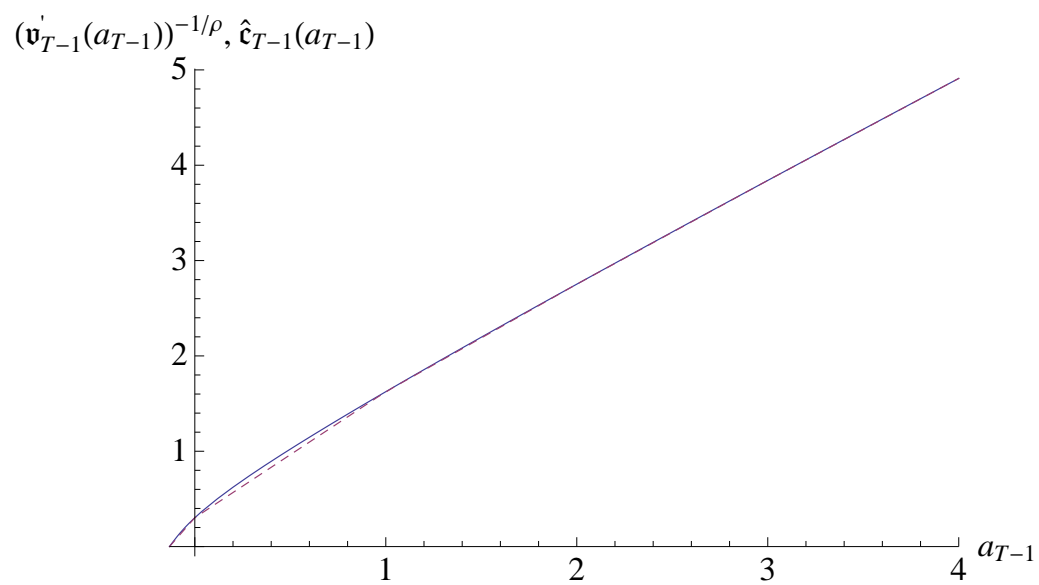


Figure 9 $\mathbf{c}_{T-1}(a_{T-1})$ versus $\hat{\mathbf{c}}_{T-1}(a_{T-1})$

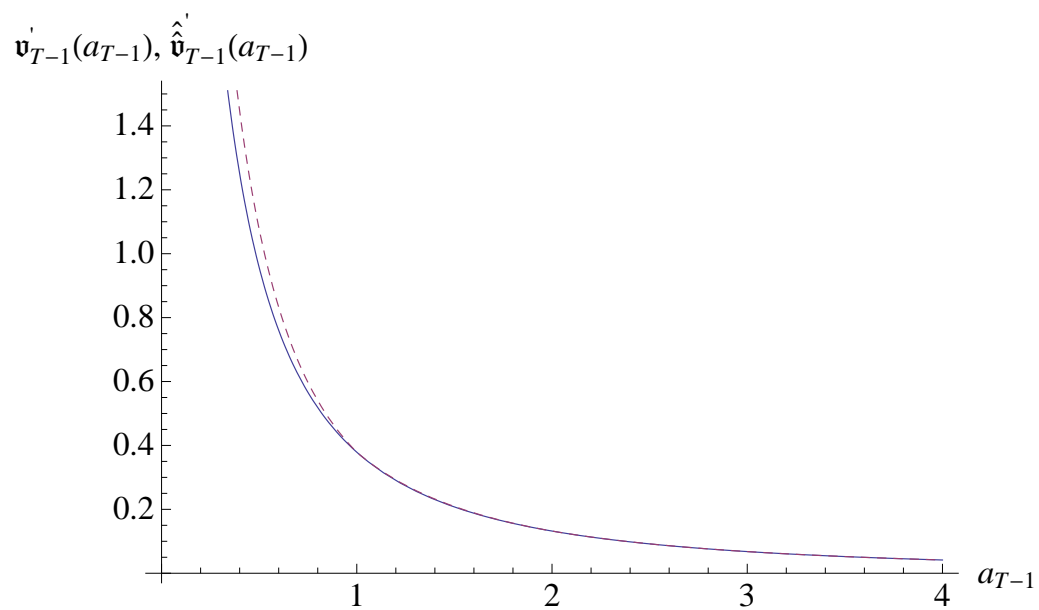


Figure 10 $\mathbf{v}'_{T-1}(a_{T-1})$ vs. $\hat{\mathbf{v}}'_{T-1}(a_{T-1})$ Constructed Using $\hat{\mathbf{c}}_{T-1}(a_{T-1})$

5.8 The Method of Endogenous Gridpoints

Our solution procedure for c_{T-1} still requires us, for each point in \vec{m}_{T-1} (`mVect` in the code), to use a numerical rootfinding algorithm to search for the value of c_{T-1} that solves $u'(c_{T-1}) = v'_{T-1}(m_{T-1} - c_{T-1})$. Unfortunately, rootfinding is a notoriously slow operation.

Fortunately, our next trick lets us completely skip this computationally burdensome step. The method can be understood by noting that any arbitrary value of $a_{T-1,i}$ (greater than its lower bound value \underline{a}_{T-1}) 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, using the first order condition,

$$u'(c_{T-1,i}) = v'_{T-1}(a_{T-1,i}) \quad (39)$$

$$c_{T-1,i} = u'^{-1}(v'_{T-1}(a_{T-1,i})) \quad (40)$$

$$= (v'_{T-1}(a_{T-1,i}))^{-1/\rho} \quad (41)$$

$$\equiv \mathbf{c}_{T-1}(a_{T-1,i}) \quad (42)$$

$$\equiv \mathbf{c}_{T-1,i}. \quad (43)$$

But with mutually consistent values of $c_{T-1,i}$ and $a_{T-1,i}$ (consistent, in the sense that they are the unique optimal values that correspond to the solution to the problem in a single state), we can obtain the $m_{T-1,i}$ that corresponds to both of them from

$$m_{T-1,i} = c_{T-1,i} + a_{T-1,i}. \quad (44)$$

These m_{T-1} gridpoints are “endogenous” in contrast to the usual solution method of specifying some ex-ante grid of values of m_{T-1} and then using a rootfinding routine to locate the corresponding optimal c_{T-1} .

Thus, we can generate a set of $m_{T-1,i}$ and $c_{T-1,i}$ pairs that can be interpolated between in order to yield $\hat{c}(m_{T-1})$ at virtually zero computational cost once we have the $\hat{\mathbf{c}}_{T-1,i}$ values in hand!¹⁷ One might worry about whether the $\{m, c\}$ points obtained in this way will provide a good representation of the consumption function as a whole, but in practice there are good reasons why they work well (basically, this procedure generates a set of gridpoints that is naturally dense right around the parts of the function with the greatest nonlinearity). Figure 11 plots the actual consumption function c_{T-1} and the approximated consumption function \hat{c}_{T-1} derived by the method of endogenous grid points. Compared to the approximate consumption functions illustrated in Figure 8 \hat{c}_{T-1} is quite close to the actual consumption function.

5.9 Improving the a Grid

Until now we have arbitrarily used a gridpoints of $\{0., 1., 2., 3., 4.\}$ (augmented in the last subsection by \underline{a}_{T-1}). But it has been obvious from the figures that the approximated \hat{c}_{T-1} function is always farthest from its true value \mathbf{c}_{T-1} at low values of a . Combining this with our insight that \underline{a}_{T-1} is a lower bound, we are now in

¹⁷This is the essential point of Carroll (2006).

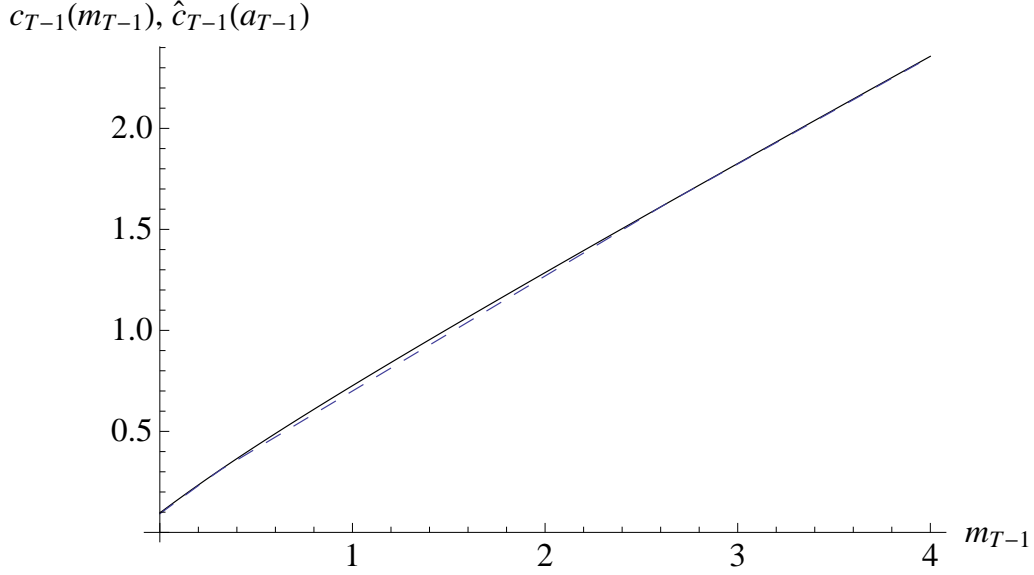


Figure 11 $c_{T-1}(m_{T-1})$ (solid) versus $\hat{c}_{T-1}(m_{T-1})$ (dashed)

position to define a more deliberate method for constructing gridpoints for a_{T-1} – a method that yields values that are more densely spaced than the uniform grid at low values of a . A pragmatic choice that works well is to find the values such that (1) the last value *exceeds the lower bound* by the same amount \bar{a}_{T-1} as our original maximum gridpoint (in our case, 4.); (2) we have the same number of gridpoints as before; and (3) the *multi-exponential growth rate* (that is, $e^{e^{e^{\dots}}}$ for some number of exponentiations n) from each point to the next point is constant (instead of, as previously, imposing constancy of the absolute gap between points).

The results (generated by the program `2periodIntExpFOCInvEEE.m`) are depicted in Figures 12 and 13, which are notably closer to their respective truths than the corresponding figures that used the original grid.

5.10 Approximating Precautionary Saving

One difficulty remains: The solution we have obtained is not very well-behaved above the range of the gridpoints for which the problem has been numerically solved. Figure 14 illustrates the point by plotting the amount of precautionary saving implied by a linear extrapolation *a la* (37) of our approximated consumption rule (the consumption of the perfect foresight consumer \bar{c}_{T-1} minus our approximation to optimal consumption under uncertainty, \hat{c}_{T-1}). Although theory proves that precautionary saving is always positive (though declining in m), the linearly extrapolated numerical approximation eventually predicts negative precautionary saving (at the point in the figure where the extrapolated locus crosses the horizontal axis).

This problem cannot be solved by extending the upper gridpoint; in the presence of serious uncertainty, the consumption rule will need to be evaluated outside of *any*

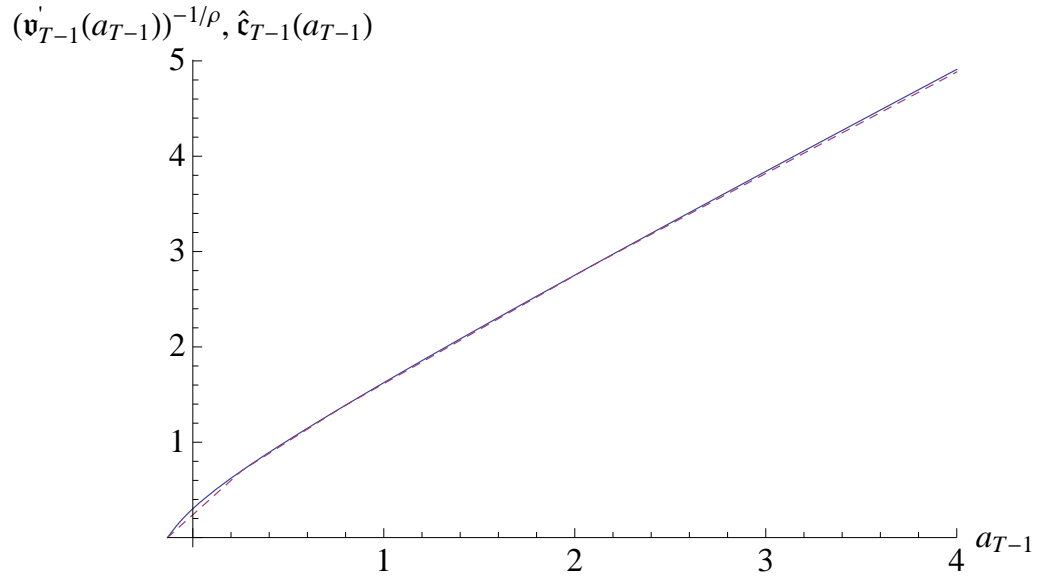


Figure 12 $\mathbf{c}_{T-1}(a_{T-1})$ versus $\hat{\mathbf{c}}_{T-1}(a_{T-1})$, Multi-Exponential aVec

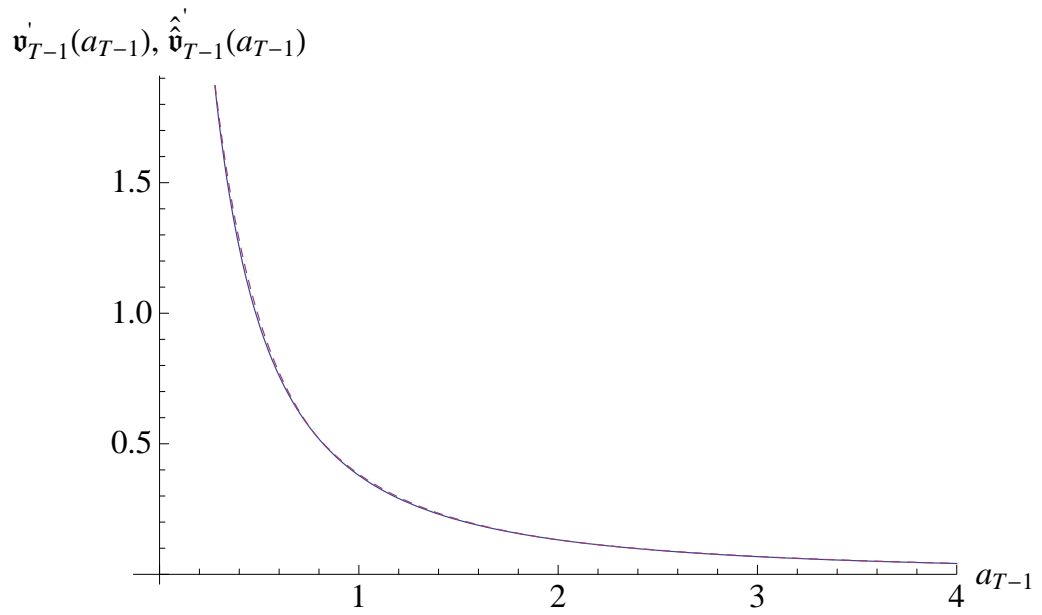


Figure 13 $\mathbf{v}'_{T-1}(a_{T-1})$ vs. $\hat{\mathbf{v}}'_{T-1}(a_{T-1})$, Multi-Exponential aVec

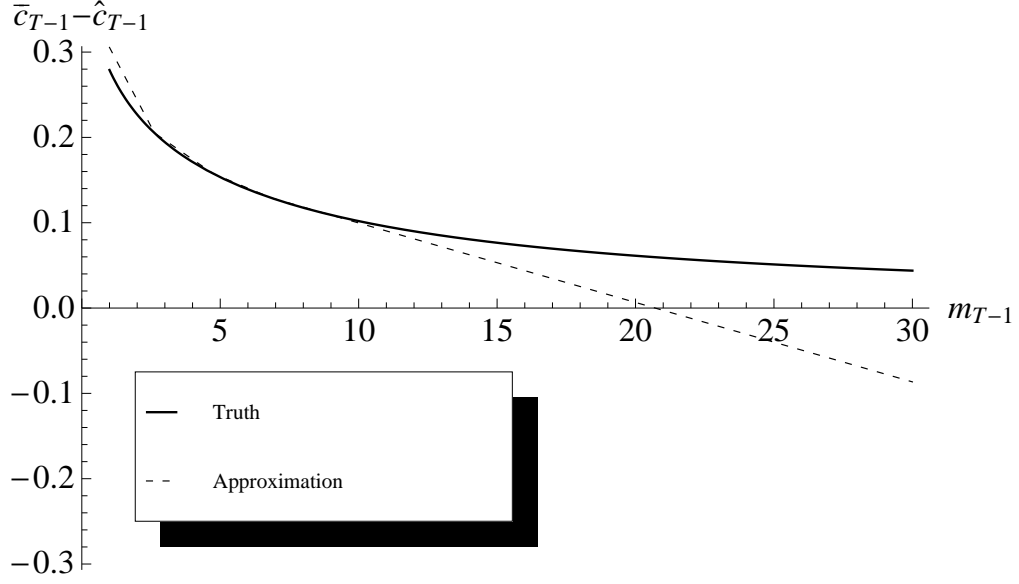


Figure 14 For Large Enough m_{T-1} , Predicted Precautionary Saving is Negative (Oops!)

prespecified grid (because starting from the top gridpoint, a large enough realization of the uncertain variable will push next period's realization of assets above that top). While a judicious extrapolation technique can prevent this problem from being fatal (for example by carefully excluding negative precautionary saving), the problem is often dealt with using inelegant methods whose implications for the accuracy of the solution are difficult to gauge.

5.11 The Method of Moderation

An elegant solution flows directly from the definition of precautionary saving as the amount by which consumption is reduced by uncertainty.

As a preliminary, define \mathfrak{h}_t as the level of end-of-period human wealth (the present discounted value of future labor income) for a perfect foresight version of the problem of a ‘risk optimist:’ a consumer who believes with perfect confidence that the transitory ‘shock’ will always take the value 1, $\theta_{t+n} = \mathbb{E}[\theta] = 1 \ \forall \ n > 0$). It has long been known that the solution to a perfect foresight problem of this kind takes the form¹⁸

$$\bar{c}_t(m_t) = (m_t + \mathfrak{h}_t)\underline{\kappa}_t \quad (45)$$

for a constant marginal propensity to consume $\underline{\kappa}_t$ given below.

We similarly define $\underline{\mathfrak{h}}_t$ as ‘minimal human wealth,’ the present discounted value of labor income if the transitory shock were to take on its worst possible value in every future period $\theta_{t+n} = \underline{\theta} \ \forall \ n > 0$ (which we define as corresponding to the beliefs of

¹⁸For a derivation, see Carroll (2011); $\underline{\kappa}_t$ is defined therein as the MPC of the perfect foresight consumer with horizon $T - t$.

a ‘pessimist’). We will call a ‘realist’ the consumer who correctly perceives the true probabilities of the future risks and optimizes accordingly.

A first useful point is that, for the realist, a lower bound for the level of market resources is $\underline{m}_t = -\underline{h}_t$, because if m_t equalled this value then there would be a positive finite chance (however small) of receiving $\theta_{t+n} = \underline{\theta}$ in every future period, which would require the consumer to set c_t to zero in order to guarantee that the intertemporal budget constraint holds (this is the multiperiod generalization of the discussion in section 5.7 about \underline{a}_{T-1}). Since consumption of zero yields negative infinite utility, the solution to consumer’s problem is not well defined for values of $m_t < \underline{m}_t$, and the limiting value of c_t is zero as $m_t \downarrow \underline{m}_t$.

Given this result, it will be convenient to define ‘excess’ market resources as the amount by which actual resources exceed the lower bound, and ‘excess’ human wealth as the amount by which mean expected human wealth exceeds guaranteed minimum human wealth:

$$\blacktriangle m_t = m_t + \overbrace{\underline{h}_t}^{=-\underline{m}_t} \quad (46)$$

$$\blacktriangle h_t = h_t - \underline{h}_t. \quad (47)$$

Using these definitions, we can now transparently define the optimal consumption rules for the two perfect foresight problems, of the ‘optimist’ and the ‘pessimist.’ The ‘pessimist’ perceives his human wealth to be equal to its minimum feasible value \underline{h}_t with certainty, so his consumption is given by the perfect foresight solution

$$\underline{c}_t(m_t) = (m_t + \underline{h}_t)\underline{\kappa}_t \quad (48)$$

$$= \blacktriangle m_t \underline{\kappa}_t. \quad (49)$$

The ‘optimist,’ on the other hand, pretends that there is no uncertainty about future income, and therefore consumes

$$\bar{c}_t(m_t) = (m_t + \underline{h}_t - \underline{h}_t + h_t)\underline{\kappa}_t \quad (50)$$

$$= (\blacktriangle m_t + \blacktriangle h_t)\underline{\kappa}_t \quad (51)$$

$$= \underline{c}_t(m_t) + \blacktriangle h_t \underline{\kappa}_t. \quad (52)$$

It seems obvious that the spending of the ‘realist’ will be strictly greater than that of the pessimist and strictly less than that of the optimist. This is more difficult to prove than might be imagined, but the necessary work is done in [Carroll \(2011\)](#) so we will take it as a fact and proceed by manipulating the inequality:

$$\begin{aligned} \blacktriangle m_t \underline{\kappa}_t &< c_t(\underline{m}_t + \blacktriangle m_t) &< (\blacktriangle m_t + \blacktriangle h_t)\underline{\kappa}_t \\ -\blacktriangle m_t \underline{\kappa}_t &> -c_t(\underline{m}_t + \blacktriangle m_t) &> -(\blacktriangle m_t + \blacktriangle h_t)\underline{\kappa}_t \\ \blacktriangle h_t \underline{\kappa}_t &> \frac{\bar{c}_t(\underline{m}_t + \blacktriangle m_t) - c_t(\underline{m}_t + \blacktriangle m_t)}{\underline{\kappa}_t} &> 0 \\ 1 &> \underbrace{\left(\frac{\bar{c}_t(\underline{m}_t + \blacktriangle m_t) - c_t(\underline{m}_t + \blacktriangle m_t)}{\blacktriangle h_t \underline{\kappa}_t} \right)}_{\equiv q_t} &> 0 \end{aligned}$$

where the fraction in the middle of the last inequality is the ratio of actual precautionary saving (the numerator is the difference between perfect-foresight consumption and optimal consumption in the presence of uncertainty) to the maximum conceivable amount of precautionary saving (the amount that would be undertaken

by the pessimist who consumes nothing out of any future income beyond the perfectly certain component).

Defining $\mu_t = \log \blacktriangle m_t$ (which can range from $-\infty$ to ∞), so that the object in the middle of the inequality above is

$$\varrho_t(\mu_t) = \left(\frac{\bar{c}_t(\underline{m}_t + e^{\mu_t}) - c_t(\underline{m}_t + e^{\mu_t})}{\blacktriangle \mathfrak{h}_t \underline{k}_t} \right), \quad (53)$$

we now define

$$\chi_t(\mu_t) = \log \left(\frac{1 - \varrho_t(\mu_t)}{\varrho_t(\mu_t)} \right) \quad (54)$$

$$= \log (1/\varrho_t(\mu_t) - 1) \quad (55)$$

which has the virtue that it is linear in the limit as μ_t approaches either $-\infty$ or $+\infty$.

If we approximate $\chi_t(\mu_t)$ at the points $\vec{\mu}_t$ corresponding to the log of the \vec{m}_t points defined above, then we can construct an interpolating $\hat{\chi}_t(\mu_t)$ from which we indirectly obtain our new-and-improved approximating function:

$$\hat{c}_t = \bar{c}_t - \overbrace{\left(\frac{1}{1 + \exp(\hat{\chi}_t)} \right)}^{=\varrho_t} \blacktriangle \mathfrak{h}_t \underline{k}_t. \quad (56)$$

Because this method relies upon the fact that the problem is easy to solve if the decision maker has unreasonable views (either in the optimistic or the pessimistic direction), and because the correct solution is always between these immoderate extremes, we call the solution method used here the ‘method of moderation.’

The results are shown in Figure 15; a reader with very good eyesight might be able to detect the barest hint of a discrepancy between the Truth and the Approximation at the far righthand edge of the figure – a stark contrast with the calamitous divergence evident in Figure 14.

5.12 Approximating the Slope As Well As the Level

A final insight leads to a final major improvement in the accuracy of the solution. Until now, we have calculated the level of consumption at various different gridpoints and used linear interpolation (either directly for c_{T-1} or indirectly for, say, χ_{T-1}). But the resulting piecewise linear approximations have the unattractive feature that they are not differentiable at the ‘kink points’ that correspond to the gridpoints where the slope of the function changes discretely.

Carroll (2011) shows that the true consumption function for problems of the kind being considered here is in fact ‘smooth.’ It exhibits a well-defined unique marginal propensity to consume at every positive value of m . This suggests that we should calculate, not just the level of consumption, but also the marginal propensity to consume (henceforth κ) at each gridpoint, and then find an interpolating approximation that smoothly matches both the level and the slope at those points.

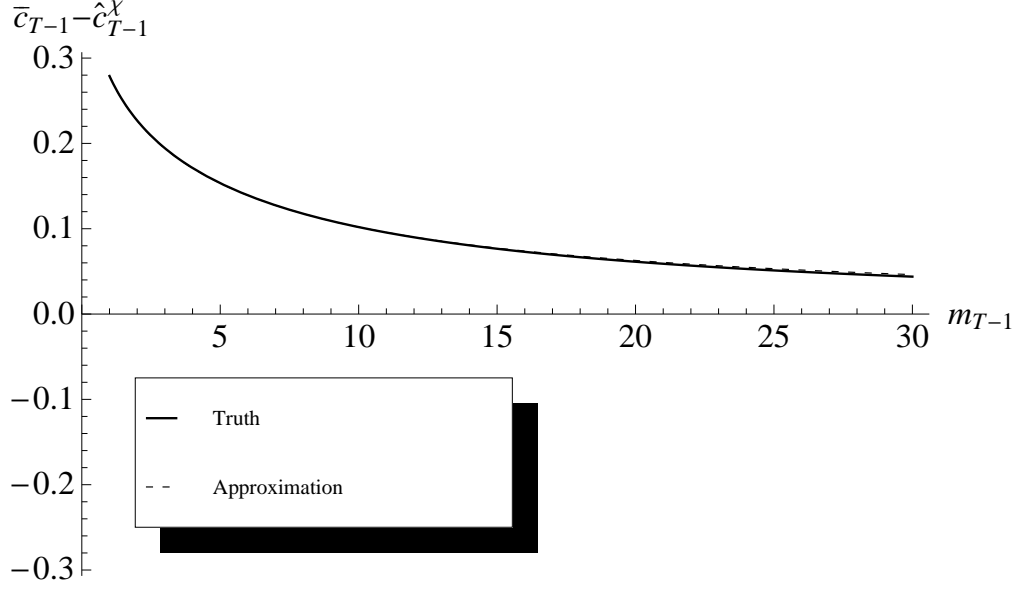


Figure 15 Extrapolated \hat{c}_{T-1}^χ Constructed Using the Method of Moderation

This requires us to differentiate (53) and (55), yielding

$$\varphi_t^\mu(\mu_t) = (\mathbf{\Delta} \mathbf{h}_t \underline{\kappa}_t)^{-1} e^{\mu_t} \left(\underline{\kappa}_t - \overbrace{\mathbf{c}_t^m(\underline{m}_t + e^{\mu_t})}^{\equiv \kappa_t(m_t)} \right) \quad (57)$$

$$\chi_t^\mu(\mu_t) = \left(\frac{-\varphi_t^\mu(\mu_t)/\varphi_t^2}{1/\varphi_t(\mu_t) - 1} \right) \quad (58)$$

and with some algebra these can be combined to yield

$$\chi_t^\mu = \left(\frac{\underline{\kappa}_t \mathbf{\Delta} m_t \mathbf{\Delta} \mathbf{h}_t (\underline{\kappa}_t - \kappa_t)}{(\bar{c}_t - c_t)(\bar{c}_t - c_t - \underline{\kappa}_t \mathbf{\Delta} \mathbf{h}_t)} \right). \quad (59)$$

To compute the vector of values of (57) corresponding to the points in $\vec{\mu}_t$, we need the marginal propensities to consume (designated κ) at each of the gridpoints, \mathbf{c}_t^m (the vector of such values is written $\vec{\kappa}_t$). This can be obtained by differentiating the Euler equation (22) (where we define $\mathbf{m}_t(a) \equiv \mathbf{c}_t(a) + a$):

$$\mathbf{u}'(\mathbf{c}_t) = \hat{\mathbf{v}}_t^a(\mathbf{m}_t - \mathbf{c}_t) \quad (60)$$

with respect to a , yielding a marginal propensity to *have consumed* \mathbf{c}^a at each gridpoint:

$$\mathbf{u}''(\mathbf{c}_t) \mathbf{c}_t^a = \hat{\mathbf{v}}_t^{aa}(\mathbf{m}_t - \mathbf{c}_t) \quad (61)$$

$$\mathbf{c}_t^a = \hat{\mathbf{v}}_t^{aa}(\mathbf{m}_t - \mathbf{c}_t) / \mathbf{u}''(\mathbf{c}_t) \quad (62)$$

and the marginal propensity to consume at the beginning of the period is obtained from the marginal propensity to have consumed by noting that

$$\begin{aligned} \mathbf{c} &= \mathbf{m} - a \\ \mathbf{c}^a + 1 &= \mathbf{m}^a \end{aligned}$$

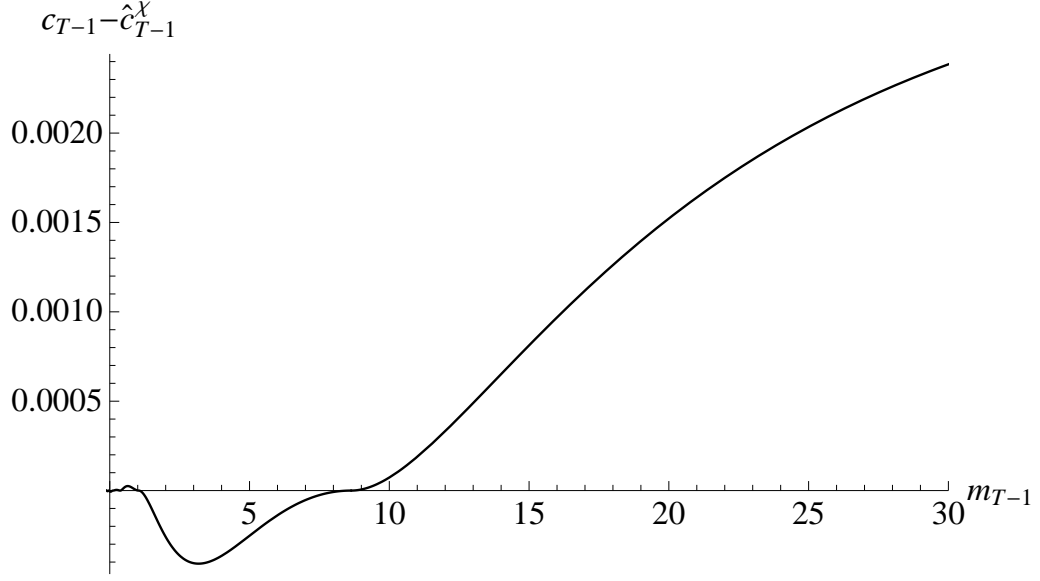


Figure 16 Difference Between True c_{T-1} and \hat{c}_{T-1}^χ Is Minuscule

which, together with the chain rule $\mathfrak{c}^a = c^m \mathfrak{m}^a$, yields the MPC from

$$\begin{aligned} c^m(\overbrace{\mathfrak{c}^a + 1}^{=\mathfrak{m}^a}) &= \mathfrak{c}^a \\ c^m &= \mathfrak{c}^a / (1 + \mathfrak{c}^a). \end{aligned}$$

Designating \hat{c}_{T-1}^χ as the approximated consumption rule obtained using an interpolating polynomial approximation to χ that matches both the level and the first derivative at the gridpoints, Figure 16 plots the difference between this latest approximation and the true consumption rule for period $T - 1$ up to the same large value (far beyond the largest gridpoint) used in prior figures. Of course, at the gridpoints the approximation will match the true function; but this figure illustrates that the approximation is quite accurate far beyond the last gridpoint (which is the last point at which the difference touches the horizontal axis). (We plot here the difference between the two functions rather than the level, as was done in previous figures, because in levels the approximation error would not be detectable even to the most eagle-eyed reader.)

5.13 Imposing ‘Artificial’ Borrowing Constraints

Optimization problems often come with additional constraints that must be satisfied. Particularly common is what I will call an ‘artificial’ liquidity constraint that prevents the consumer’s net worth from falling below some value, often zero.¹⁹

¹⁹The word ‘artificial’ is chosen only because of its clarity in distinguishing this from the case of the ‘natural’ borrowing constraint examined above; no derogation is intended – constraints of this kind certainly exist in the real world.

With such an additional constraint, the problem is

$$\begin{aligned}
v_{T-1}(m_{T-1}) &= \max_{c_{T-1}} u(c_{T-1}) + \mathbb{E}_{T-1}[\beta_T \mathcal{G}_T^{1-\rho} v_T(m_T)] \\
&\text{s.t.} \\
a_{T-1} &= m_{T-1} - c_{T-1} \\
m_T &= \mathcal{R}_T a_{T-1} + \theta_T \\
a_{T-1} &\geq 0.
\end{aligned}$$

By definition, the constraint will bind if the unconstrained consumer would choose a level of spending that would violate the constraint. In this case, that means that the constraint binds if the c_{T-1} that satisfies the unconstrained FOC

$$c_{T-1}^{-\rho} = \mathbf{v}'_{T-1}(m_{T-1} - c_{T-1}) \quad (63)$$

is greater than m_{T-1} . Call \check{c}_{T-1} the function returning the level of c_{T-1} that satisfies (63). Then the constrained optimal level of consumption will be

$$c_{T-1}(m_{T-1}) = \min[m_{T-1}, \check{c}_{T-1}(m_{T-1})]. \quad (64)$$

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, this is easy to calculate. We know that when the constraint is binding the consumer is saving nothing, which yields marginal value of $\mathbf{v}'_{T-1}(0)$. Further, 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 stops binding is:

$$\begin{aligned}
u'(m_{T-1}) &= \mathbf{v}'_{T-1}(0) \\
m_{T-1} &= (\mathbf{v}'_{T-1}(0))^{(-1/\rho)} \\
&= \mathbf{c}_{T-1}(0).
\end{aligned} \quad (65)$$

The constrained problem is solved by `2periodIntExpFOCInvPesReaOptCon.m`; the resulting consumption rule is shown in Figure 17. For comparison purposes, the approximate consumption rule from Figure 17 is reproduced here as the solid line. The presence of the liquidity constraint has induced three changes: the first is to redefine \mathbf{h}_t , which now is the PDV of receiving $\underline{\theta}$ next period and nothing beyond period $t + 1$; the second is to augment the end-of-period asset $a\mathbf{Vec}$ by a point near next-period's binding point m^\sharp (so that we can better capture the curvature around that point), and to compute this period's binding point for future use; the third is to redefine the optimal consumption rule as in equation (64). This ensures that the liquidity-constrained 'realist' will consume more than the redefined 'pessimist,' so

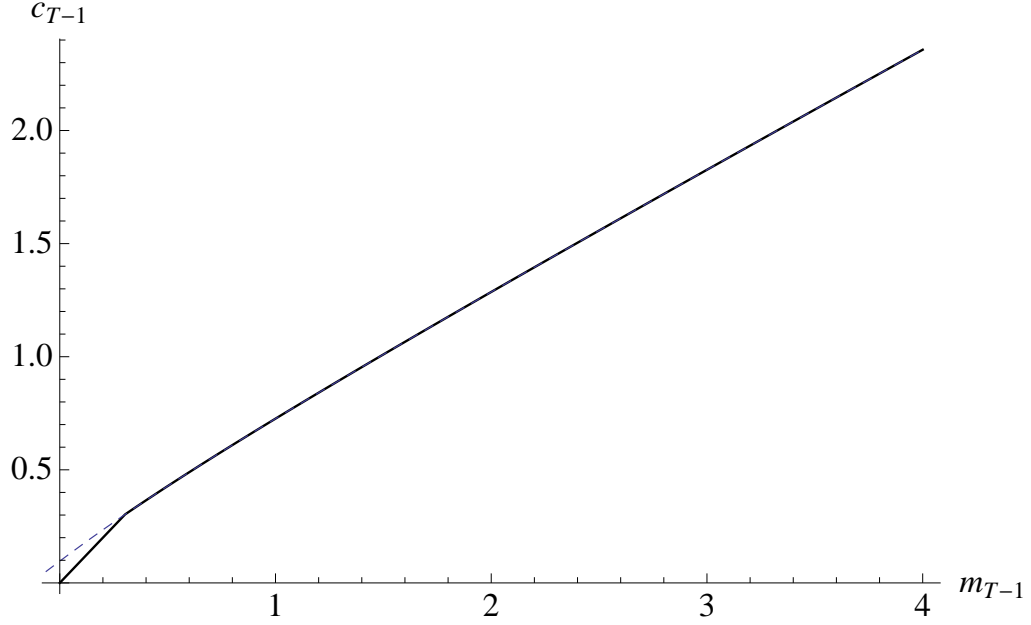


Figure 17 Constrained (solid) and Unconstrained (dashed) Consumption

that we will have φ still between 0 and 1 and the ‘method of moderation’ will proceed smoothly. 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.

5.14 Value

Our focus thus far has been on approximating the optimal consumption rule. Often it is also useful to know the value function associated with a problem. Fortunately, many of the tricks used when solving the consumption problem have a direct analogue in approximation of the value function.

Consider the perfect foresight (or ‘optimist’) case in period $T - 1$:

$$\begin{aligned}
 v_{T-1}(m_{T-1}) &= u(c_{T-1}) + \beta_T u(c_T) \\
 &= u(c_{T-1}) (1 + \beta_T ((\beta_T R)^{1/\rho})^{1-\rho}) \\
 &= u(c_{T-1}) (1 + \beta_T (\beta_T R)^{1/\rho-1}) \\
 &= u(c_{T-1}) (1 + (\beta_T R)^{1/\rho} / R) \\
 &= u(c_{T-1}) \underbrace{\text{PDV}_t^T(c) / c_{T-1}}_{\equiv \mathbb{C}_t^T}
 \end{aligned}$$

where $\text{PDV}_t^T(c)$ is the present discounted value through the end of the horizon of the stream of consumption that the consumer foresees experiencing. A similar function can be constructed recursively for earlier periods, yielding the general expression

$$\bar{v}_t(m_t) = u(c_t) \mathbb{C}_t^T \quad (66)$$

which can be transformed as

$$\bar{\Lambda}_t(m_t) = ((1 - \rho)\bar{v}_t)^{1/(1-\rho)} \quad (67)$$

$$= \bar{c}_t(m_t)(\mathbb{C}_t^T)^{1/(1-\rho)} \quad (68)$$

with derivative

$$\bar{\Lambda}_t^m(m_t) = (\mathbb{C}_t^T)^{1/(1-\rho)} \underline{\kappa}_t \quad (69)$$

and since \mathbb{C}_t^T is a constant while the consumption function is linear, $\bar{\Lambda}_t$ will also be linear.

We apply the same transformation to the value function for the problem with uncertainty (the realist's problem) and differentiate:

$$\Lambda_t = ((1 - \rho)v_t(m_t))^{1/(1-\rho)} \quad (70)$$

$$\Lambda_t^m = ((1 - \rho)v_t(m_t))^{-1+1/(1-\rho)} v_t^m(m_t) \quad (71)$$

and an excellent approximation to the value function can be obtained by calculating the values of Λ at the same gridpoints used by the consumption function approximation, and interpolating among those points.

However, as with the consumption approximation, we can do even better if we realize that the $\bar{\Lambda}$ function for the perfect foresight (or optimist's) problem is an upper bound for the Λ function in the presence of uncertainty, and the value function for the pessimist is a lower bound. Analogously to (53), define an upper-case

$$\Omega_t(\mu_t) = \left(\frac{\bar{\Lambda}_t(\underline{m}_t + e^{\mu_t}) - \Lambda_t(\underline{m}_t + e^{\mu_t})}{\mathbf{\Delta} \mathbf{h}_t \underline{\kappa}_t (\mathbb{C}_t^T)^{1/(1-\rho)}} \right) \quad (72)$$

with derivative

$$\Omega_t^\mu(\mu_t) = (\mathbf{\Delta} \mathbf{h}_t \underline{\kappa}_t (\mathbb{C}_t^T)^{1/(1-\rho)})^{-1} e^{\mu_t} (\bar{\Lambda}_t^m - \Lambda_t^m) \quad (73)$$

and an upper-case version of the χ equation in (55):

$$X_t(\mu_t) = \log \left(\frac{1 - \Omega_t(\mu_t)}{\Omega_t(\mu_t)} \right) \quad (74)$$

$$= \log(1/\Omega_t(\mu_t) - 1) \quad (75)$$

with corresponding derivative

$$X_t^\mu(\mu_t) = \left(\frac{-\Omega_t^\mu(\mu_t)/\Omega_t^2}{1/\Omega_t(\mu_t) - 1} \right) \quad (76)$$

and if we approximate these objects then invert them (as above with the φ and χ functions) we obtain a very high-quality approximation to our inverted value function at the same points for which we have our approximated consumption function:

$$\hat{\Lambda}_t = \bar{\Lambda}_t - \overbrace{\left(\frac{1}{1 + \exp(\hat{X}_t)} \right)}^{=\Omega_t} \mathbf{\Delta} \mathbf{h}_t \underline{\kappa}_t (\mathbb{C}_t^T)^{1/(1-\rho)} \quad (77)$$

from which we obtain our approximation to the value function and its derivative as

$$\hat{v}_t = u(\hat{\Lambda}_t) \quad (78)$$

$$\hat{v}_t^m = u'(\hat{\Lambda}_t) \hat{\Lambda}_t^m. \quad (79)$$

This approximation to the value function has the considerable virtue that it numerically satisfies the envelope theorem at each of the gridpoints for which the problem has been solved. (If even greater accuracy were desired, we could differentiate one more time to obtain terms involving the second derivative of the value function and the marginal propensity to consume at the gridpoints, which could be matched to the corresponding interpolating function; this would guarantee that the consumption function generated from the value function would match both the level of consumption (as now) and the marginal propensity to consume at the gridpoints; the numerical differences between the newly constructed consumption function and the highly accurate one constructed earlier would be negligible).

6 Recursion

6.1 Theory

Before we solve for periods earlier than $T - 1$, we assume for convenience that in each such period a liquidity constraint exists of the kind discussed above, preventing c from exceeding m . This simplifies things a bit because now we can always consider an \mathbf{aVec} that starts with zero as its smallest element.

Recall now equations (21) and (22):

$$\begin{aligned}\mathbf{v}'_t(a_t) &= \mathbb{E}_t[\beta \mathcal{RG}_{t+1}^{-\rho} \mathbf{u}'(c_{t+1}(\mathcal{R}_{t+1}a_t + \theta_{t+1}))] \\ \mathbf{u}'(c_t) &= \mathbf{v}'_t(m_t - c_t).\end{aligned}$$

Assuming that the problem has been solved up to period $t + 1$ (and thus assuming that we have a numerical function $\hat{\chi}_{t+1}(\mu_{t+1})$ and hence the derived $\hat{c}_{t+1}(m_{t+1})$), the first of these equations tells us how to calculate $\mathbf{v}'_t(a_t)$, and the second tells us how to calculate $\hat{c}_t(m_t)$ given $\mathbf{v}_t(a_t)$. 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 $a_{t,i}$ in \mathbf{aVec}_t , numerically calculate the values of $\mathbf{c}_t(a_{t,i})$ and $\mathbf{c}'_t(a_{t,i})$,

$$\mathbf{c}_{t,i} = (\mathbf{v}'_t(a_{t,i}))^{-1/\rho}, \quad (80)$$

$$= (\beta \mathbb{E}_t [\mathcal{RG}_{t+1}^{-\rho} (\hat{c}_{t+1}(\mathcal{R}_{t+1}a_{t,i} + \theta_{t+1}))^{-\rho}])^{-1/\rho}, \quad (81)$$

generating vectors of values $\vec{\mathbf{c}}_t$ and $\vec{\mathbf{c}}'_t$ (where the letter \varkappa is a variant of κ ; we need a variant because κ itself is reserved for the marginal propensity to consume as of the beginning of the period, and here we are calculating the marginal propensity to *have consumed*).

2. Construct a corresponding list of values of $c_{t,i}$ and $m_{t,i}$ from $c_{t,i} = \mathbf{c}_{t,i}$ and $m_{t,i} = c_{t,i} + a_{t,i}$; similarly construct a corresponding list of $\kappa_{t,i}$.
3. Construct a corresponding list of $\mu_{t,i}$, the levels and first derivatives of $\varphi_{t,i}$, and the levels and first derivatives of $\chi_{t,i}$.

4. Construct an interpolating approximation $\hat{\chi}_t$ that smoothly matches both the level and the slope at those points.
5. If we are to approximate the value function, construct a corresponding list of values of $v_{t,i}$, the levels and first derivatives of $\Omega_{t,i}$, and the levels and first derivatives of $X_{t,i}$; and construct an interpolating approximation \hat{X}_t that smoothly matches both the level and the slope at those points.

With $\hat{\chi}_t$ in hand, our approximate consumption function is computed directly from \hat{c}_t^x . With this consumption rule in hand, 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 $\hat{v}'_t(a_t)$ or $\hat{v}'_t(m_t)$. This is because with $\hat{c}_t(a_t)$ and $\hat{c}_t(m_t)$ in hand, we simply *define* $\hat{v}'_t(a_t) = [\hat{c}_t(a_t)]^{-\rho}$ and $\hat{v}'_t(m_t) = u'(\hat{c}_t(m_t))$ so there is no need to construct interpolating approximations to these functions - they arise ‘free’ (or nearly so) from our constructed $\hat{c}_t(a_t)$ and $\hat{c}_t(m_t)$.

The program `multiPeriodCon.m`²⁰ presents a fairly general and flexible approach to solving problems of this kind. The essential structure of the program is a loop that simply works its way back from an assumed last period of life, using the command `AppendTo` to record the interpolated $\hat{\chi}_t$ 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 calibrate a nonconstant path of expected income growth over the lifetime that matches the empirical profile; allowing for such a calibration is the reason we have included the \mathcal{G}_t^T vector in our specification of the problem.

6.2 *Mathematica* Background

Mathematica has several features that are useful in solving the multiperiod problem.

- It can treat a user-created function as an object just like a number or a character.
- *Mathematica* uses the ‘list’ as its basic data structure. A *Mathematica* ‘list’ is a very powerful and flexible data construct. A list of length N in *Mathematica* can hold essentially anything in each of its N positions - a function, a number, another list, a symbolic expression, or any other object that *Mathematica* can recognize. The items at position i in a list named `ExampleList` are retrieved or addressed using the syntax `ExampleList[[i]]`.
- The function `Apply[FuncName_, DataListName_]` takes the function whose name is `FuncName` (for example, `Vt`) and the data in `DataListName` (for example, `{1, 19}`) and returns the result that would have been returned by calling the function `Vt[1, 19]`.
- The function `Map[FuncToApply_, DataToApplyItTo_]` takes a list of possible arguments to the function `FuncToApply` and applies that function to each of the elements of that list sequentially. For example, `Map[Sin, {1, 2, 3}]` would return a list `{Sin[1], Sin[2], Sin[3]}`.

²⁰There is also a parallel `multiPeriod.m` file that solves the unconstrained multi-period problem.

6.3 Program Structure

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

6.3.1 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, 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. Add various period- t parameters to their respective lifecycle lists, which is accomplished by calling the `AddNewPeriodToParamLifeDates` function.
2. For each a in `aVec`, construct \mathbf{c} as follows:

$$\mathbf{c}_t(a_t) = \left(\beta \mathbb{E}_t \left[R \mathcal{G}_{t+1}^{-\rho} (\hat{\mathbf{c}}_{t+1}(\mathcal{R}_{t+1}a_t + \theta_{t+1}))^{-\rho} \right] \right)^{-1/\rho} \quad (82)$$

$$= \left(\beta \frac{1}{n} \sum_{i=1}^n R \left[\mathcal{G}_{t+1}^{-\rho} (\hat{\mathbf{c}}_{t+1}(\mathcal{R}_{t+1}a_t + \theta_i))^{-\rho} \right] \right)^{-1/\rho}. \quad (83)$$

Obviously, $\mathbf{c}(a)$ depends on the constructed consumption function from one period later in life. We also construct the corresponding `mVec`, `κVec`, etc. by calling the `AddNewPeriodToSolvedLifeDates` function.

3. For each m in `mVec`, we can define `▲mVec`, find the corresponding optimal consumption vector for a pessimist and an optimist, construct the φ and χ vectors, and finally an interpolation function $\hat{\chi}_t$. Similarly we can construct an interpolation function \hat{X}_t that approximates the value function. The whole process is done by calling the `AddNewPeriodToSolvedLifeDatesPesReaOpt` function.
4. Various period- t functions are derived from $\hat{\chi}_t$ and \hat{X}_t (in `functions_ConsNVal.m`). Note that the liquidity constraint is dealt with by comparing the unconstrained solution `cFromχ` with the 45 degree line.

6.4 Results

As written, the program creates $\hat{\chi}_t(\mu_t)$ functions and hence we can derive $\hat{\mathbf{c}}_t(m_t)$, etc., which can be evaluated in any period for any value of m .

As an illustration, Figure 18 shows $\hat{\mathbf{c}}_t(m_t)$ for $t = \{20, 15, 10, 5, 1\}$. At least one feature of this figure is encouraging: the consumption functions converge as the horizon extends, something that Carroll (2011) shows must be true under certain parametric conditions that are satisfied by the baseline parameter values being used here.

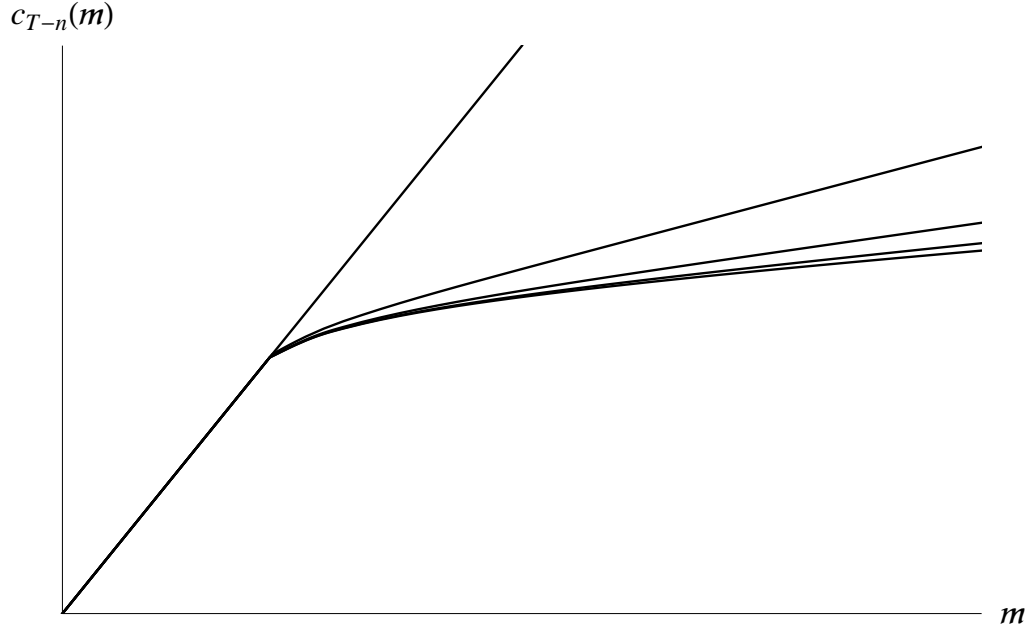


Figure 18 Converging $c_{T-n}(m_{T-n})$ Functions as n Increases

7 Multiple Control Variables

We now consider how to solve problems with multiple control variables. (To reduce notational complexity, in this section we set $\mathcal{G}_t = 1 \forall t$.)

7.1 Theory

The new control variable that the consumer can now choose is the portion of the portfolio to invest in risky assets. Designating the gross return on the risky asset as \mathfrak{R}_{t+1} , and using ς_t to represent the proportion of the portfolio invested in this asset between t and $t+1$ (restricted here, as often in the literature, to values between 0 and 1, corresponding to an assumption that the consumer cannot be ‘net short’ and cannot issue net equity), the overall return on the consumer’s portfolio between t and $t+1$ will be:

$$\mathbb{R}_{t+1} = R(1 - \varsigma_t) + \mathfrak{R}_{t+1}\varsigma_t \quad (84)$$

$$= R + (\mathfrak{R}_{t+1} - R)\varsigma_t \quad (85)$$

and the maximization problem is

$$\begin{aligned} v_t(m_t) &= \max_{\{c_t, \varsigma_t\}} u(c_t) + \beta \mathbb{E}_t[v_{t+1}(m_{t+1})] \\ &\text{s.t.} \\ \mathbb{R}_{t+1} &= R + (\mathfrak{R}_{t+1} - R)\varsigma_t \\ m_{t+1} &= (m_t - c_t)\mathbb{R}_{t+1} + \theta_{t+1} \\ 0 &\leq \varsigma_t \leq 1, \end{aligned}$$

or

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

The first order condition with respect to c_t is almost identical to that in the single-control problem, equation (14), with the only difference being that the nonstochastic interest factor \mathbf{R} is now replaced by \mathbb{R}_{t+1} ,

$$u'(c_t) = \beta \mathbb{E}_t[\mathbb{R}_{t+1} v'_{t+1}(m_{t+1})], \quad (86)$$

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

$$u'(c_t) = v'_t(m_t)$$

yielding the Euler equation for consumption

$$u'(c_t) = \mathbb{E}_t[\beta \mathbb{R}_{t+1} u'(c_{t+1})]. \quad (87)$$

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

$$\begin{aligned} 0 &= \mathbb{E}_t[v'_{t+1}(m_{t+1})(\mathfrak{R}_{t+1} - \mathbf{R})a_t] \\ &= a_t \mathbb{E}_t[u'(c_{t+1}(m_{t+1}))(\mathfrak{R}_{t+1} - \mathbf{R})]. \end{aligned} \quad (88)$$

As before, it will be useful to define \mathbf{v}_t as a function that yields the expected $t+1$ value 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 chosen values of both of those variables, because expected end-of-period value will depend not just on how much the agent saves, but also on how the saved assets are allocated between the risky and riskless assets. Thus we define

$$\mathbf{v}_t(a_t, \varsigma_t) = \mathbb{E}_t[\beta v_{t+1}(m_{t+1})]$$

which has derivatives

$$\begin{aligned} \mathbf{v}_t^a &= \mathbb{E}_t[\beta \mathbb{R}_{t+1} v_{t+1}^m(m_{t+1})] \\ \mathbf{v}_t^\varsigma &= \mathbb{E}_t[\beta (\mathfrak{R}_{t+1} - \mathbf{R}) v_{t+1}^m(m_{t+1})] a_t \end{aligned}$$

implying that the first order conditions (87) and (88) and can be rewritten

$$u'(c_t) = \mathbf{v}_t^a(m_t - c_t, \varsigma_t) \quad (89)$$

$$0 = \mathbf{v}_t^\varsigma(a_t, \varsigma_t). \quad (90)$$

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, $\log \mathfrak{R} \sim \mathcal{N}(\phi + \mathbf{r} - \sigma_\phi^2/2, \sigma_\phi^2)$ where ϕ is the equity premium over the returns \mathbf{r} available on the riskless asset.²¹

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

²¹This guarantees that $\mathbb{E}[\mathfrak{R}] = \Phi$ is invariant to the choice of σ_ϕ^2 ; see **LogELogNorm**.

follow the same procedure as for labor income uncertainty, generating a set of m equiprobable shocks to the rate of return; in a slight abuse of notation, we will designate the portfolio-weighted return (contingent on the chosen portfolio share in equity, and potentially contingent on any other aspect of the consumer's problem) simply as $\mathbb{R}_{i,j}$ (where dependence on i is allowed to permit the possibility of nonzero correlation between the return on the risky asset and the shock to labor income (for example, in recessions the stock market falls and labor income also declines)).

The direct expressions for the derivatives of \mathbf{v}_t are

$$\mathbf{v}_t^a(a_t, \varsigma_t) = \beta \left(\frac{1}{mn} \right) \sum_{i=1}^n \sum_{j=1}^m [\mathbb{R}_{i,j} (c_{t+1}(\mathbb{R}_{i,j}a_t + \theta_i))^{-\rho}] \quad (91)$$

$$\mathbf{v}_t^\varsigma(a_t, \varsigma_t) = \beta \left(\frac{1}{mn} \right) \sum_{i=1}^n \sum_{j=1}^m [(\mathfrak{R}_{i,j} - R) (c_{t+1}(\mathbb{R}_{i,j}a_t + \theta_i))^{-\rho}]. \quad (92)$$

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 \mathbf{v} function.

Details of the construction of the interpolating approximation are given below; assume for the moment that we have the approximations $\hat{\mathbf{v}}_t^a$ and $\hat{\mathbf{v}}_t^\varsigma$ 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

$$c_t^{-\rho} = \hat{\mathbf{v}}_t^a(m_t - c_t, \varsigma_t) \quad (93)$$

$$0 = \hat{\mathbf{v}}_t^\varsigma(m_t - c_t, \varsigma_t) \quad (94)$$

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

$$\bar{\mathbf{v}}_t(a_t) = \max_{\{\varsigma_t\}} \mathbf{v}_t(a_t, \varsigma_t) \quad (95)$$

s.t.

$$0 \leq \varsigma_t \leq 1 \quad (96)$$

where the bar accent on \mathbf{v} indicates that this is the \mathbf{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 `aVec` and use the results to construct the interpolating function $\hat{\mathbf{v}}_t^a(a_t)$.²² With this function in hand, we can use the first order condition from the single-control problem

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

to solve for the optimal level of consumption as a function of m_t . Thus we have

²²A faster solution could be obtained by, for each element in `aVec`, computing $\hat{\mathbf{v}}_t^\varsigma(m_t - c_t, \varsigma)$ of a grid of values of ς , and then using an approximating interpolating function (rather than the full expectation) in the `FindRoot` command. The associated speed improvement is fairly modest, however, so this route was not pursued.

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 constrained problem with multiple control variables is `multicontrolCon.m`.

Some of the functions defined in `multicontrolCon.m` correspond to the derivatives of $\mathbf{v}_t(a_t, \varsigma_t)$.

The first function definition that does not resemble anything in `multiperiod.m` is `varsRaw[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 (94), 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 $\bar{\mathbf{v}}_t^a(a_t)$ and $\hat{\varsigma}_t(a_t)$ will be constructed.

As $\hat{\varsigma}_t(a_t)$ can be constructed by `varsRaw[at_]`, $\bar{\mathbf{v}}_t^a(a_t)$ is constructed by another newly defined function `varsOpt[at_]`, where the naming convention is obviously that ‘Opt’ stands for ‘Optimized.’ With $\bar{\mathbf{v}}_t^a(a_t)$ in hand (as well as the appropriately redefined $\bar{\mathbf{v}}_t(a_t)$ and $\bar{\mathbf{v}}_t^{aa}(a_t)$) the analysis is essentially identical to that for the standard multi-period problem with a single control variable.

The structure of the program in detail is as follows. First, perform the usual initializations. Then initialize `varsVec` and the other variables specific to the multiple control problem.²³ 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 $\mathbf{v}_t^\varsigma(a_t, \varsigma_t)$ using the usual calculation over the tensor defined by the combinations of the elements of `varsVec` and `varsVec`.
2. For any level of saving `at`, the function `varsRaw[at_]` performs a rootfinding operation²⁴

$$0 = \mathbf{v}_t^\varsigma(a_t, \varsigma_t) \quad (97)$$

s.t.

$$0 \leq \varsigma_t \leq 1 \quad (98)$$

²³Note 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.

²⁴Alternatively, the rootfinding operation would be $0 = \hat{\mathbf{v}}_t^\varsigma(a_t, \varsigma_t)$, where the interpolation function of $\mathbf{v}_t^\varsigma(a_t, \varsigma_t)$ is used instead. However, the results obtained (especially $\hat{\varsigma}_t(a_t)$) are much less satisfactory.

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

3. Construct the function `vaOpt[at_]`

$$\bar{\mathbf{v}}_t^a(a_t) \equiv \mathbf{v}_t^a(a_t, \hat{\varsigma}_t(a_t)) \quad (99)$$

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

4. Using $\bar{\mathbf{v}}_t^a(a_t) \equiv \mathbf{vaOpt[at_]$ and the redefined $\bar{\mathbf{v}}_t(a_t)$ and $\bar{\mathbf{v}}_t^{aa}(a_t)$ (in place of $\mathbf{v}_t^a(a_t) \equiv \mathbf{va[at_]$ in `multiperiod.m`), follow the same procedures as in `multiperiod.m` to generate $\hat{c}_t(m_t)$.

7.4 Results

Figure 19 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 ?? 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 optimal choice is for the agent to invest a proportion 1 (100 percent) of the portfolio in stocks (instead of the safe bank account with riskless return R) is 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 assets in stocks, while the rich should be cautious, holding more of their wealth in safe bank deposits and less in stocks. This seemingly bizarre (and highly counterfactual) prediction reflects the nature of the risks the consumer faces. 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 relatively large wealth will be paying for a large proportion of future consumption out of that wealth, and hence if they invest too much of it in stocks their consumption will have a high covariance with stock returns. Consequently, they reduce that correlation by holding some of their wealth in the riskless form.

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 18. Furthermore, Deaton (1991), Carroll (1992; 1997) and others have argued that the ‘buffer-stock’ saving behavior that emerges under some further restrictions on 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.

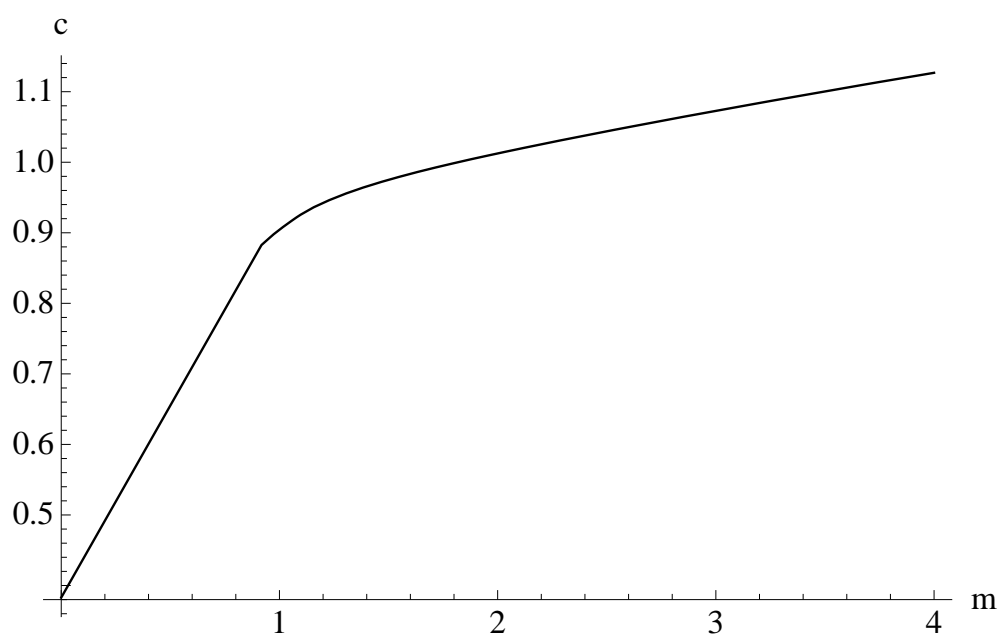


Figure 19 $c(m_1)$ With Portfolio Choice

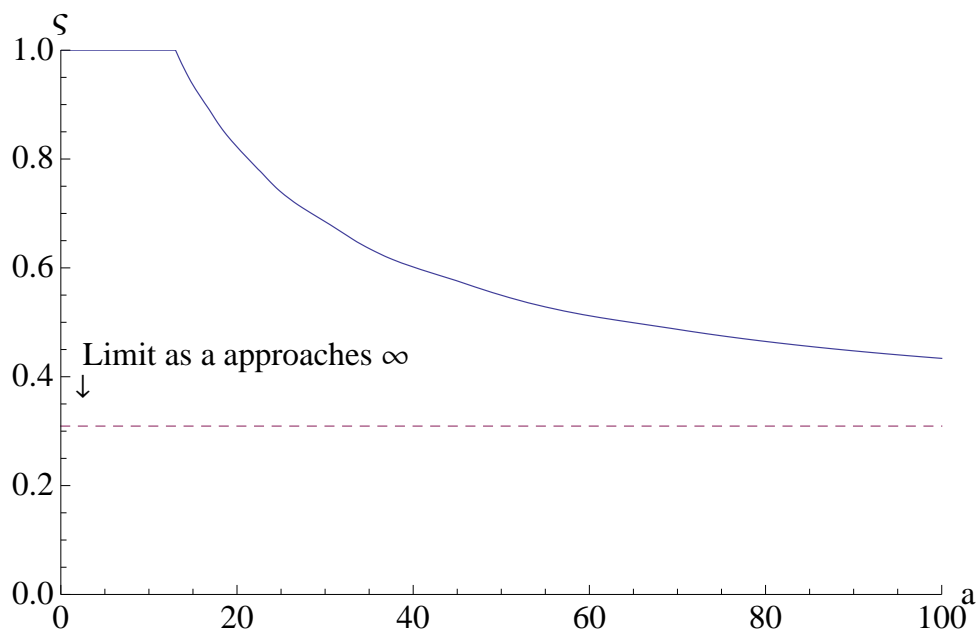


Figure 20 Portfolio Share in Risky Assets in First Period $\varsigma(a)$

Of course, the simplest such method is to solve the problem as specified above for a large number of periods. This is feasible, but there are much faster methods.

8.1 Convergence

In solving an infinite-horizon problem, it is necessary to have some metric that determines when to stop because a solution that is ‘good enough’ has been found.

A natural metric is defined by the unique ‘target’ level of wealth that Carroll (2011) proves will exist in problems of this kind: The \tilde{m} such that

$$\mathbb{E}_t[m_{t+1}/m_t] = 1 \text{ if } m_t = \tilde{m} \quad (100)$$

where the \vee accent is meant to signify that this is the value that other m ’s ‘point to.’

Given a consumption rule $c(m)$ it is straightforward to find the corresponding \tilde{m} . So for our problem, a solution is declared to have converged if the following criterion is met: $|\tilde{m}_{t+1} - \tilde{m}_t| < \epsilon$, where ϵ is a very small number and measures our degree of convergence tolerance.

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 Coarse then Fine θVec

The speed of solution is roughly proportionate²⁵ 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 solution. The program `multiperiodCon_infhor.m` begins with three gridpoints, and then solves for successively finer θVec .

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).²⁶ 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.)

²⁵It is also true that the speed of each iteration is directly proportional to the number of gridpoints in `aVec`, at which the problem must be solved. However given our method of moderation, now the problem could be solved very precisely based on five gridpoints only. Hence we do not pursue the process of “Coarse then Fine `aVec`”.

²⁶Similar structural estimation exercises have been also performed by Palumbo (1999) and Gourinchas and Parker (2002).

9.1 Life Cycle Model

The decision problem for the household at age t is:

$$\max \left\{ u(\mathbf{c}_t) + \mathbb{E}_t \left[\sum_{s=t+1}^T \beth^{s-t} \left(\Pi_{i=t+1}^s \hat{\beta}_i \mathfrak{D}_i \right) u(\mathbf{c}_s) \right] \right\} \quad (101)$$

subject to the constraints

$$\begin{aligned} \mathbf{a}_s &= \mathbf{m}_s - \mathbf{c}_s \\ \mathbf{m}_{s+1} &= \mathbf{R}\mathbf{a}_s + Y_{s+1} \\ Y_{s+1} &= \mathbf{p}_{s+1}\theta_{s+1} \\ \mathbf{p}_{s+1} &= \mathcal{G}_{s+1}\mathbf{p}_s\Psi_{s+1} \end{aligned}$$

where

- \mathfrak{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 $s - 1$ and s
- Ψ_s : mean-one shock to permanent income
- \beth : time-invariant discount factor

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 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 (101) which captures the effect of time-varying demographic variables (e.g. changes in family size).

Transitory and permanent shocks are distributed as follows:

$$\Xi_s = \begin{cases} 0 & \text{with probability } \wp > 0 \\ \theta_s/\wp & \text{with probability } (1 - \wp), \text{ where } \log \theta_s \sim \mathcal{N}(-\sigma_\theta^2/2, \sigma_\theta^2) \end{cases} \quad (102)$$

$$\log \psi_s \sim \mathcal{N}(-\sigma_\psi^2/2, \sigma_\psi^2) \quad (103)$$

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 $\sigma_\psi = 0.1$.²⁷ The income growth profile \mathcal{G}_s is from Carroll (1997) and the values of \mathfrak{D}_s and $\hat{\beta}_s$ are obtained from Cagetti (2003) (Figure 21).²⁸ The interest rate is assumed to equal 1.03. The model parameters are included in Table 1.

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

²⁷Note that $\sigma_\theta = 0.1$ 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.

²⁸The income growth profile is the one used by Carroll 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).

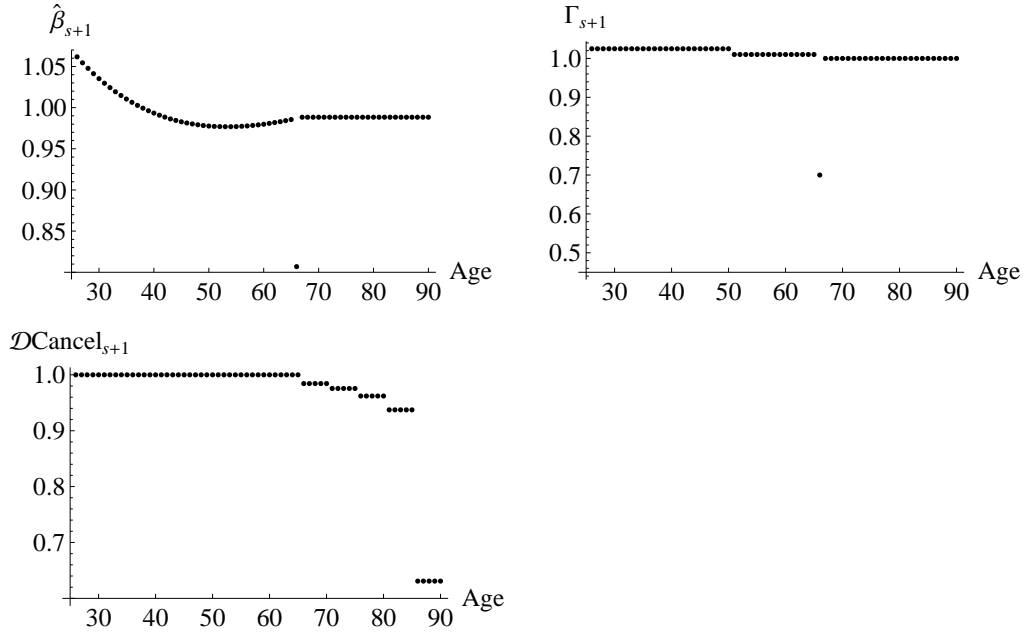


Figure 21 Time Varying Parameters

Table 1 Parameter Values

σ_θ	0.1	Carroll (1992)
σ_ψ	0.1	Carroll (1992)
\wp	0.005	Carroll (1992)
\mathcal{G}_s	figure 21	Carroll (1997)
$\hat{\beta}_s, \mathcal{D}_s$	figure 21	Cagetti (2003)
R	1.03	Cagetti (2003)

9.2 Estimation

When economists say that they are performing “structural estimation” of a model like this, they mean that they have devised a formal procedure for searching for values for the parameters \mathfrak{A} 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. Here, we choose to match the median of the wealth to permanent income ratio across 7 age groups, from age 26 – 30 up to 56 – 60.²⁹ The choice of matching the medians rather the means is motivated by the fact that the wealth distribution is much more concentrated at the top 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

²⁹Cagetti (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.

match the behavior of Homer Simpson, we choose 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:

$$\begin{aligned}
v_t(m_t) &= \max_{c_t} \left\{ u(c_t) + \beta_{t+1} \mathbb{E}_t[(\psi_{t+1} \mathcal{G}_{t+1})^{1-\rho} v_{t+1}(m_{t+1})] \right\} \\
&\text{s.t.} \\
a_t &= m_t - c_t \\
m_{t+1} &= a_t \underbrace{\left(\frac{R}{\psi_{t+1} \mathcal{G}_{t+1}} \right)}_{\equiv \mathcal{R}_{t+1}} + \theta_{t+1}
\end{aligned}$$

with the first order condition:

$$u'(c_t) = \beta_{t+1} R \mathbb{E}_t[u'(\psi_{t+1} \mathcal{G}_{t+1} c_{t+1} (a_t \mathcal{R}_{t+1} + \theta_{t+1}))]. \quad (104)$$

The first step is to solve for the consumption functions at each age using the routines included in the `setup_ConsFn.m` file. We need to discretize the shock distribution and solve for the policy functions by backward induction using equation (104) following the procedure in sections 5 and 6 (`ConstructcFuncLife`). 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 must ensure that at each backward iteration the right parameter values are used.

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.m` file. We first have to draw the shocks for each agent and period. This involves discretizing the shock distribution for as many points as the number of 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`). 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. (Similarly, in order to speed up the process, at each backward iteration we compute the consumption function and other variables as a vector at once.) Then, following Cagetti (2003), we initialize the wealth-to-income ratio of agents at age 25 by randomly assigning the equal probability values to 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, using the appropriate 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 can estimate the model by calculating empirical medians and measure 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 $\xi = [\rho, \Xi]$), we could search for the parameter values which solve

$$\min_{\xi} \sum_{\tau=1}^7 |\varsigma^{\tau} - \mathbf{s}^{\tau}(\xi)| \quad (105)$$

where ς^{τ} and \mathbf{s}^{τ} 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 as though they reflected perfect measurements of the truth. Imagine, however, that one of the age groups happened to have (in the consumer survey) 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 (105) assigns equal importance to a deviation between the model and the data for all age groups.

We can get around this problem (and a variety of others) by instead minimizing a slightly more complex object:

$$\min_{\xi} \sum_i^N \omega_i |\varsigma_i^{\tau} - \mathbf{s}^{\tau}(\xi)| \quad (106)$$

where ω_i is the weight of household i in the entire population,³⁰ 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 that 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 22. More details on the SCF data are included in appendix A.

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

```
GapEmpiricalSimulatedMedians[ $\rho, \Xi$ ] := [ ConstructcFuncLife[ $\rho, \Xi$ ];
                                         Simulate;
                                          $\sum_i^N \omega_i |\varsigma_i^{\tau} - \mathbf{s}^{\tau}(\xi)|$ 
                                         ];
```

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

1. solves for the consumption functions by calling `ConstructcFuncLife`
2. simulates the data and computes the simulated medians by calling `Simulate`

³⁰The 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.

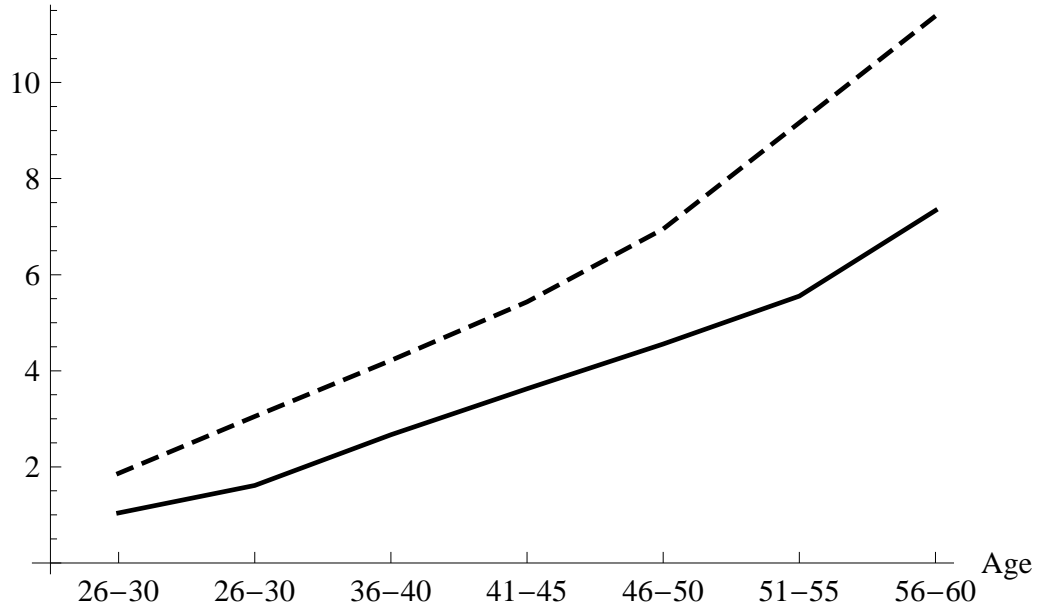


Figure 22 Wealth to Permanent Income Ratios from SCF (means (dashed) and medians (solid))

3. returns the value of equation (106)

We delegate the task of finding the coefficients that minimize the `GapEmpiricalSimulatedMedians` function to the *Mathematica* built-in numerical minimizer `FindMinimum`. 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 by experimenting with a variety of alternative starting values for the parameter search.

Finally the standard errors are computed by bootstrap using the routines in the `setup_Bootstrap.m` file.³¹ 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 λ

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.m` produces our ρ and λ estimates with standard errors using 10,000 simulated agents.³² Results are reported in Table 2.³³ Figure 23 shows the contour plot of the `GapEmpiricalSimulatedMedians` function and the parameter estimates. The contour plot shows equally spaced isoquants of the

³¹For a treatment of the advantages of the bootstrap see Horowitz (2001)

³²The procedure is: First we calculate the ρ and λ estimates as the minimizer of equation (106) using the actual SCF data. Then, we apply the `Bootstrap` function several times to obtain the standard error of our estimates.

³³Differently 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.

`GapEmpiricalSimulatedMedians` function, i.e. the pairs of ρ and γ which lead to the same deviations between simulated and empirical medians (equivalent values of equation (106)). 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.

Table 2 Estimation Results

ρ	γ
4.68	1.00
(0.13)	(0.00)

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, 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).

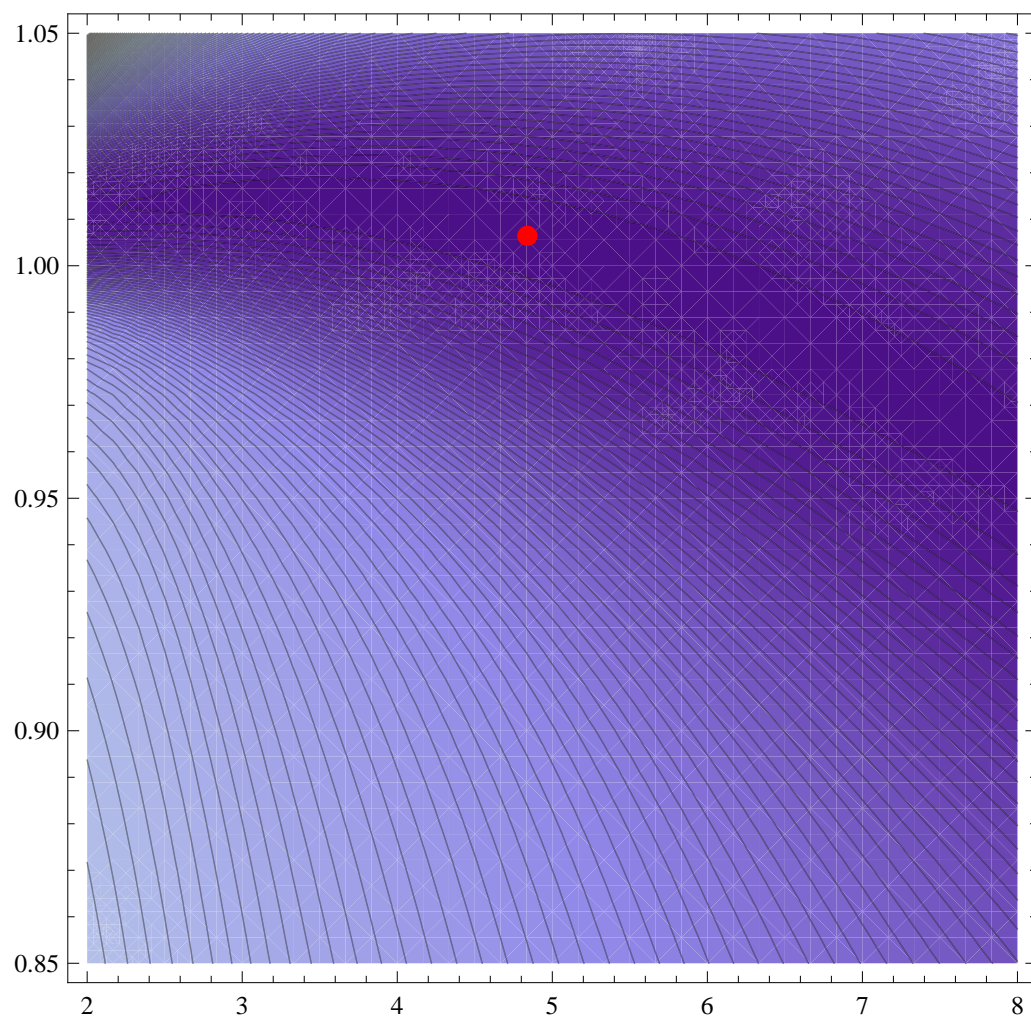


Figure 23 Contour Plot (larger values are shown lighter) with $\{\rho, \gamma\}$ Estimates (red dot).

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).³⁴ 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.³⁵

³⁴Cagetti (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.

³⁵Please 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, <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, <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, <http://econ.jhu.edu/people/ccarroll/EndogenousGridpoints.pdf>.
- (2011): “Theoretical Foundations of Buffer Stock Saving,” *Manuscript, Department of Economics, Johns Hopkins University*, <http://econ.jhu.edu/people/ccarroll/papers/BufferStockTheory>.
- (Current): “Math Facts Useful for Graduate Macroeconomics,” Online Lecture Notes.
- CARROLL, CHRISTOPHER D., AND MILES S. KIMBALL (1996): “On the Concavity of the Consumption Function,” *Econometrica*, 64(4), 981–992, <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, <http://econ.jhu.edu/people/ccarroll/nature.pdf>.
- DEATON, ANGUS S. (1991): “Saving and Liquidity Constraints,” *Econometrica*, 59, 1221–1248.
- DEN HAAN, WOUTER J, AND ALBERT MARCET (1990): “Solving the Stochastic Growth Model by Parameterizing Expectations,” *Journal of Business and Economic Statistics*, 8(1), 31–34, Available at <http://ideas.repec.org/a/bes/jnlbes/v8y1990i1p31-34.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.
- KOPECKY, KAREN A., AND RICHARD M.H. SUEN (2010): “Finite State Markov-Chain Approximations To Highly Persistent Processes,” *Review of Economic Dynamics*, 13(3), 701–714, <http://www.karenkopecky.net/RouwenhorstPaper.pdf>.
- MERTON, ROBERT C. (1969): “Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case,” *Review of Economics and Statistics*, 50, 247–257.
- 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>.
- SAMUELSON, PAUL A. (1969): “Lifetime Portfolio Selection by Dynamic Stochastic Programming,” *Review of Economics and Statistics*, 51, 239–46.
- VALENCIA, FABIAN (2006): “Banks’ Financial Structure and Business Cycles,” Ph.D. thesis, Johns Hopkins University.