

# Solution Methods for Microeconomic Dynamic Stochastic Optimization Problems

2024-04-20

Christopher D. Carroll<sup>1</sup>

Note: The GitHub repo [SolvingMicroDSOPs](https://github.com/llorracc/SolvingMicroDSOPs) associated with this document contains python code that produces all results, from scratch, except for the last section on indirect inference. The numerical results have been confirmed by showing that the answers that the raw python produces correspond to the answers produced by tools available in the [Econ-ARK](#) toolkit, more specifically those in the [HARK](#) which has full [documentation](#). The MSM results at the end have been superseded by tools in the [EstimatingMicroDSOPs](#) repo.

## Abstract

These notes describe tools for solving microeconomic dynamic stochastic optimization problems, and show how to use those tools for efficiently 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 Python code that solves the problems described in the text.

**Keywords**     Dynamic Stochastic Optimization, Method of Simulated Moments, Structural Estimation, Indirect Inference

**JEL codes**    E21, F41

PDF: <https://github.com/llorracc/SolvingMicroDSOPs/blob/master/SolvingMicroDSOPs.pdf>  
Slides: <https://github.com/llorracc/SolvingMicroDSOPs/blob/master/SolvingMicroDSOPs-Slides.pdf>  
Web: <https://llorracc.github.io/SolvingMicroDSOPs>  
Code: <https://github.com/llorracc/SolvingMicroDSOPs/tree/master/Code>  
Archive: <https://github.com/llorracc/SolvingMicroDSOPs>  
(Contains LaTeX code for this document and software producing figures and results)

---

<sup>1</sup>Carroll: Department of Economics, Johns Hopkins University, Baltimore, MD, [ccarroll@jhu.edu](mailto:ccarroll@jhu.edu)

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. Relative to earlier drafts, this version 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://www.econ2.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://www.econ2.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 Tokunaka 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. This document can be cited as Carroll (2023a) in the references.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Problem</b>	<b>1</b>
<b>3</b>	<b>Normalization</b>	<b>2</b>
<b>4</b>	<b>The Usual Theory, and a Bit More Notation</b>	<b>4</b>
4.1	Steps . . . . .	4
4.2	The Usual Theory, Notated . . . . .	4
4.3	Summing Up . . . . .	5
<b>5</b>	<b>Solving the Next-to-Last Period</b>	<b>6</b>
5.1	Discretizing the Distribution . . . . .	6
5.2	The Approximate Consumption and Value Functions . . . . .	8
5.3	An Interpolated Consumption Function . . . . .	8
5.4	Interpolating Expectations . . . . .	10
5.5	Value Function versus First Order Condition . . . . .	10
5.6	Transformation . . . . .	15
5.7	The Natural Borrowing Constraint and the $a_{t-1}$ Lower Bound . . . . .	16
5.8	The Method of Endogenous Gridpoints . . . . .	18
5.9	Improving the $a$ Grid . . . . .	20
5.10	Program Structure . . . . .	22
5.11	Results . . . . .	22
5.12	The Method of Moderation . . . . .	22
5.13	Approximating the Slope Too . . . . .	26
5.14	Value . . . . .	29
5.15	Refinement: A Tighter Upper Bound . . . . .	31
5.16	Extension: A Stochastic Interest Factor . . . . .	33
5.17	Imposing ‘Artificial’ Borrowing Constraints . . . . .	34
<b>6</b>	<b>Recursion</b>	<b>36</b>
6.1	Theory . . . . .	36
<b>7</b>	<b>Multiple Control Variables</b>	<b>37</b>
7.1	Theory . . . . .	37
7.2	Stages Within a Period . . . . .	38
7.3	Application . . . . .	40
7.4	Implementation . . . . .	41
7.5	Results With Multiple Controls . . . . .	41

<b>8</b>	<b>The Infinite Horizon</b>	<b>42</b>
8.1	Convergence . . . . .	43
<b>9</b>	<b>Structural Estimation</b>	<b>44</b>
9.1	Life Cycle Model . . . . .	44
9.2	Estimation . . . . .	46
9.2.1	An Aside to Computing Sensitivity Measures . . . . .	50
9.3	Results . . . . .	50
<b>A</b>	<b>SCF Data</b>	<b>54</b>

# 1 Introduction

{sec:introduction}

These lecture notes provide a gentle introduction to a particular set of solution tools for the canonical consumption-saving/portfolio allocation problem. Specifically, the notes describe and solve optimization problems for a consumer facing uninsurable idiosyncratic risk to nonfinancial income (e.g., labor or transfer income), first without and then with optimal portfolio choice,<sup>1</sup> with detailed intuitive discussion of 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 is 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 shows how to use these methods (via the statistical ‘method of simulated moments’ (MSM for short) to estimate structural parameters like the coefficient of relative risk aversion (*a la* Gourinchas and Parker (2002) and Cagetti (2003))).

## 2 The Problem

{sec:the-problem}

The usual analysis of dynamic stochastic programming problems packs a great many events (intertemporal choice, stochastic shocks, intertemporal returns, income growth, the taking of expectations, and more) into a single tock in which the agent makes an optimal choice taking all these elements into account. For the dissection here, we will be careful to disarticulate everything that happens in the problem explicitly into separate tocks so that each element can be scrutinized and understood in isolation.

We are interested in the behavior a consumer who begins period  $t$  with a certain amount of ‘capital’  $\mathbf{k}_t$ , which is immediately rewarded by a return factor  $R_t$  with the proceeds deposited in a **bankbalance**:

$$\mathbf{b}_t = \mathbf{k}_t R_t. \quad (1) \quad \{\text{eq:bLvl}\}$$

Simultaneously with the realization of the capital return, the consumer also receives noncapital income  $\mathbf{y}_t$ , which is determined by multiplying the consumer’s ‘permanent income’  $\mathbf{p}_t$  by a transitory shock  $\boldsymbol{\theta}_t$ :

$$\mathbf{y}_t = \mathbf{p}_t \boldsymbol{\theta}_t \quad (2) \quad \{\text{eq:yLvl}\}$$

whose whose expectation is 1 (that is, before realization of the transitory shock, the consumer’s expectation is that actual income will on average be equal to permanent income  $\mathbf{p}_t$ ).

---

<sup>1</sup>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.

The combination of bank balances  $\mathbf{b}$  and income  $\mathbf{y}$  define's the consumer's 'market resources' (sometimes called 'cash-on-hand,' following Deaton (1992)):

$$\mathbf{m}_t = \mathbf{b}_t + \mathbf{y}_t, \quad (3) \quad \{\text{eq:mLvl}\}$$

available to be spent on consumption  $\mathbf{c}_t$  even for a consumer subject to a liquidity constraint that requires  $\mathbf{c} \leq \mathbf{m}$ .

The consumer's goal is to maximize discounted utility from consumption over the rest of a lifetime ending at date  $T$ :

$$\max \mathbb{E}_t \left[ \sum_{n=0}^{T-t} \beta^n u(\mathbf{c}_{t+n}) \right]. \quad (4) \quad \{\text{eq:MaxProb}\}$$

Income evolves according to:

$$\begin{aligned} \mathbf{p}_{t+1} &= \mathcal{G}_{t+1} \mathbf{p}_t && \text{-- permanent labor income dynamics} \\ \log \boldsymbol{\theta}_{t+n} &\sim \mathcal{N}(-\sigma_{\boldsymbol{\theta}}^2/2, \sigma_{\boldsymbol{\theta}}^2) && \text{-- lognormal transitory shocks } \forall n > 0. \end{aligned} \quad (5) \quad \{\text{eq:permincgrow}\}$$

Equation (5) indicates that we are allowing for a predictable average profile of income growth over the lifetime  $\{\mathcal{G}\}_0^T$  (to capture typical career wage paths, pension arrangements, etc).<sup>2</sup> Finally, the utility function is of the Constant Relative Risk Aversion (CRRA), form,  $u(\bullet) = \bullet^{1-\rho}/(1-\rho)$ .

It is well known that this problem can be rewritten in recursive (Bellman) form

$$\mathbf{v}_t(\mathbf{m}_t, \mathbf{p}_t) = \max_{\mathbf{c}_t} u(\mathbf{c}_t) + \beta \mathbb{E}_t[\mathbf{v}_{t+1}(\mathbf{m}_{t+1}, \mathbf{p}_{t+1})] \quad (6) \quad \{\text{eq:vrecurse}\}$$

subject to the Dynamic Budget Constraint (DBC) implicitly defined by equations (1)-(3) and to the transition equation that defines next period's initial capital as this period's end-of-period assets:

$$\mathbf{k}_{t+1} = \mathbf{a}_t. \quad (7) \quad \{\text{eq:transitionstate}\}$$

### 3 Normalization

{sec: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 at all possible). In the consumption context, 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  $\mathbf{p}_t$  (henceforth for brevity, 'permanent income.')

---

<sup>2</sup>For expositional and pedagogical purposes, this equation assumes that there are no shocks to permanent income. A large literature finds that, in reality, permanent (or at least extremely highly persistent) 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 are 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 (2023b).

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

$$\mathbf{v}_T(\mathbf{m}_T, \mathbf{p}_T) = \frac{\mathbf{m}_T^{1-\rho}}{1-\rho}. \quad (8) \quad \{\text{eq:levelTm1}\}$$

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  $v_T(m_T) = u(m_T)$ .<sup>3</sup> For our CRRA utility function,  $u(xy) = x^{1-\rho}u(y)$ , so (8) can be rewritten as

$$\begin{aligned} \mathbf{v}_T(\mathbf{m}_T, \mathbf{p}_T) &= \mathbf{p}_T^{1-\rho} \frac{m_T^{1-\rho}}{1-\rho} \\ &= (\mathbf{p}_{T-1}\mathcal{G}_T)^{1-\rho} \frac{m_T^{1-\rho}}{1-\rho} \\ &= \mathbf{p}_{T-1}^{1-\rho} \mathcal{G}_T^{1-\rho} v_T(m_T). \end{aligned} \quad (9) \quad \{\text{eq:vT}\}$$

Now define a new (dissected) optimization problem:

$$\begin{aligned} v_t(m_t) &= \max_{c_t} u(c_t) + \beta \mathbb{E}_t[\mathcal{G}_{t+1}^{1-\rho} v_{t+1}(m_{t+1})] \\ \text{s.t.} \\ a_t &= m_t - c_t \\ k_{t+1} &= a_t \\ b_{t+1} &= \underbrace{(\mathbf{R}/\mathcal{G}_{t+1}) k_{t+1}}_{\equiv \mathcal{R}_{t+1}} \\ m_{t+1} &= b_{t+1} + \boldsymbol{\theta}_{t+1}, \end{aligned} \quad (10) \quad \{\text{eq:vNormed}\}$$

where division by the  $\mathcal{G}$  in second-to-last equation to yield a normalized return factor  $\mathcal{R}$  is the consequence of the fact that we have divided  $t+1$  level variables by  $\mathbf{p}_{t+1} = \mathcal{G}\mathbf{p}_t$ .

Then it is easy to see that for  $T-1$ ,

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

and so on back to all earlier periods. Hence, if we solve the problem (10) which has only a single state variable  $m_{T-1}$ , we can obtain the levels of the value function, consumption, and all other variables from the corresponding permanent-income-normalized solution objects by multiplying each by  $\mathbf{p}_T$ , e.g.  $c_T(\mathbf{m}_T, \mathbf{p}_T) = \mathbf{p}_T c_T(m_T/\mathbf{p}_T)$  (or, for the value function,  $\mathbf{v}_T(\mathbf{m}_T, \mathbf{p}_T) = \mathbf{p}_T^{1-\rho} v_T(m_T)$ ). We have thus reduced the problem from two continuous state variables to one (and thereby enormously simplified its solution).

For future reference it will be useful to write the problem (10) in the traditional way, by substituting  $b_{t+1}$ ,  $k_{t+1}$ , and  $a_t$  into  $m_{t+1}$ :

$$v_t(m_t) = \max_c u(c) + \beta \mathbb{E}_t[\mathcal{G}_{t+1}^{1-\rho} v_{t+1}(\overbrace{(m_t - c)(\mathbf{R}/\mathcal{G}_{t+1}) + \boldsymbol{\theta}_{t+1}}^{m_{t+1}})] \quad (11) \quad \{\text{eq:vusual}\}$$

whose (relative) complexity illustrates one reason that dissection is useful.

---

<sup>3</sup>Nonbold value is bold value divided by  $\mathbf{p}^{1-\rho}$  rather than  $\mathbf{p}$ .

## 4 The Usual Theory, and a Bit More Notation

{sec:the-usual-theor

### 4.1 Steps

Generically, we want to think of the Bellman solution as having three tocks:

1. **Arrival:** Incoming state variables (e.g.,  $k_t$ ) are known, but any shocks associated with the period have not been realized and decision(s) have not yet been made
2. **Decision:** All exogenous variables (like income shocks, rate of return shocks, and predictable income growth  $\mathcal{G}$ ) have been realized (so that, e.g.,  $m_t$ 's value is known) and the agent solves the optimization problem
3. **Continuation:** After all decisions have been made, their consequences are revealed by evaluation of the continuing-value function at the values of the 'outgoing' state variables.

In the standard treatment in the literature, the (implicit) default assumption is that the tock where the agent is solving a decision problem is the unique moment at which the problem is defined. This is what implicitly was done above, when (for example) in (10) we related the value  $v_t$  of the current decision to the explicit expectation of the future value  $v_{t+1}$ . The idea, here, though, is to encapsulate the current period's problem as a standalone object, which is solvable contingent on some exogenously-provided continuation-value function  $v_{t \rightarrow}(a)$ .

When we want to refer to a specific tock within period  $t$  we will do so by modifying its period subscript with an indicator character:

Stp	Indicator	State	Usage	Explanation
Arrival	$\leftarrow$ prefix	$k_t$	$v_{\leftarrow t}(k_t)$	value at entry to $t$ (before shocks)
Decision	(blank/none)	$m_t$	$v_t(m_t)$	value of $t$ -decision (after shocks)
Continuation	$\rightarrow$ suffix	$a_t$	$v_{t \rightarrow}(a_t)$	value at exit (after decision)

Notice that different steps of the problem have distinct state variables.  $k$  is the state at the beginning of the period because the shocks that yield  $m$  from  $k$  have not yet been realized. The state variable for the continuation tock is  $a$  because after the consumption decision has been made the model assumes that all that matters is where you have ended up, not how you got there.

### 4.2 The Usual Theory, Notated

Using this new notation, the first order condition for (10) with respect to  $c_t$  is

$$\begin{aligned} u^c(c_t) &= \mathbb{E}_{t \rightarrow} [\beta \mathcal{R}_{t+1} \mathcal{G}_{t+1}^{1-\rho} v_{t+1}^m(m_{t+1})] \\ &= \mathbb{E}_{t \rightarrow} [\beta \mathcal{R} \quad \mathcal{G}_{t+1}^{-\rho} v_{t+1}^m(m_{t+1})] \end{aligned} \tag{12} \quad \{\text{eq:upceqEvt p1}\}$$

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

$$v_t^m(m_t) = \mathbb{E}_{\leftarrow t} [\beta \mathcal{R} \mathcal{G}_{t+1}^{-\rho} v_{t+1}^m(m_{t+1})] \tag{13} \quad \{\text{eq:envelope}\}$$

we can substitute the LHS of (13) for the RHS of (12) to get

$$u^c(c_t) = v_t^m(m_t) \quad (14) \quad \{\text{eq:upcteqvtp}\}$$

and rolling forward one period,

$$u^c(c_{t+1}) = v_{t+1}^m(a_t \mathcal{R}_{t+1} + \theta_{t+1}) \quad (15) \quad \{\text{eq:upctp1EqVpxtb}\}$$

so that substituting the LHS in equation (12) finally gives us the Euler equation for consumption:

$$u^c(c_t) = \mathbb{E}_{t \rightarrow} [\beta R \mathcal{G}_{t+1}^{-\rho} u^c(c_{t+1})]. \quad (16) \quad \{\text{eq:cEuler}\}$$

From the perspective of the beginning of period  $t + 1$  we can write the ‘arrival value’ function and its first derivative as

$$\begin{aligned} v_{\leftarrow(t+1)}(k_{t+1}) &= \mathbb{E}_{\leftarrow(t+1)} [\mathcal{G}_{t+1}^{1-\rho} v_{t+1}(\overbrace{\mathcal{R}_{t+1} k_{t+1} + \theta_{t+1}}^{m_{t+1}})] \\ v_{\leftarrow(t+1)}^k(k_{t+1}) &= \mathbb{E}_{\leftarrow(t+1)} [R \mathcal{G}_{t+1}^{-\rho} v_{t+1}^m(m_{t+1})] \end{aligned} \quad (17) \quad \{\text{eq:vFuncBegtpdefn}\}$$

because they return the expected  $t + 1$  value and marginal value associated with arriving in period  $t + 1$  with any given amount of  $k$  capital.

Finally, recalling that we obtain  $v_{t \rightarrow}(a_t) = \beta v_{\leftarrow(t+1)}(k_{t+1})$  using  $a_t = k_{t+1}$ , note for future use that we can write the Euler equation (16) more compactly as

$$u^c(c_t) = v_{t \rightarrow}^a(m_t - c_t). \quad (18) \quad \{\text{eq:upEqbetaOp}\}$$

### 4.3 Summing Up

For future reference, it will be useful here to write the full expressions for the distinct value functions at the Arrival ( $\leftarrow$ ) and Decision steps. (Recall that  $v_{t \rightarrow}(a)$  is provided to the solution algorithm as an input).

There is no need to use our tock-identifying notation for the model’s variables;  $k$ , for example, will have only one unique value over the course of the solution and therefore a notation like  $k_{t \rightarrow}$  would be useless; the same is true of all other variables. Given the continuation value function  $v_{t \rightarrow}$ , the problem can be written entirely without period subscripts:

$$v_{\leftarrow t}(k_t) = \mathbb{E}_{\leftarrow t} [v_t(\overbrace{k_t \mathcal{R}_t + \theta_t}^{m_t})] \quad (19) \quad \{\text{eq:vBeg}\}$$

$$v_t(m_t) = \max_{\{c\}} u(c) + \mathbb{E}_t [v_{t \rightarrow}(\overbrace{m_t - c}^{a_t})] \quad (20) \quad \{\text{eq:vMid}\}$$

and

$$v_{t \rightarrow}(a) = \beta v_{\leftarrow(t+1)}(\overbrace{k_{t+1}}^a) \quad (21) \quad \{\text{eq:vEndtdefn}\}$$



## 5 Solving the Next-to-Last Period

To reduce clutter, we now temporarily assume that  $\mathcal{G}_t = 1$  for all  $t$ , so that the  $\mathcal{G}$  terms from the earlier derivations disappear, and setting  $t = T$  the problem in the second-to-last period of life can now be expressed as

$$v_{T-1}(m_{T-1}) = \max_{c_{T-1}} u(c_{T-1}) + \beta \mathbb{E}_{(T-1) \rightarrow} \left[ v_t(\underbrace{(m_{T-1} - c_{T-1})\mathcal{R}_T + \theta_T}_{m_T}) \right].$$

Using (1)  $v_T(m) = u(m)$ ; (2) the definition of  $u(m)$ ; and (3) the definition of the expectations operator, this becomes:

$$v_{T-1}(m_{T-1}) = \max_{c_{T-1}} \frac{c_{T-1}^{1-\rho}}{1-\rho} + \beta \int_0^\infty \frac{((m_{T-1} - c_{T-1})\mathcal{R}_T + \vartheta)^{1-\rho}}{1-\rho} d\mathcal{F}(\vartheta)$$

where  $\mathcal{F}(\theta)$  is the cumulative distribution function for  $\theta$ .

# A ‘raw’ solution to the  $T-1$  solution works, but is very slow:

This maximization problem implicitly defines a function  $c_{T-1}(m)$  that yields optimal consumption in period  $T-1$  for any specific numerical level of resources like  $m = 1.7$ . But because there is no general analytical solution to this problem, for any given  $m$  we must use numerical computational tools to find the  $c_{T-1}$  that maximizes the expression. This is excruciatingly slow because for every potential  $c$  to be considered, a definite integral over the interval  $(0, \infty)$  must be calculated numerically, and numerical integration is *very* slow (especially over an unbounded domain!).

### 5.1 Discretizing the Distribution

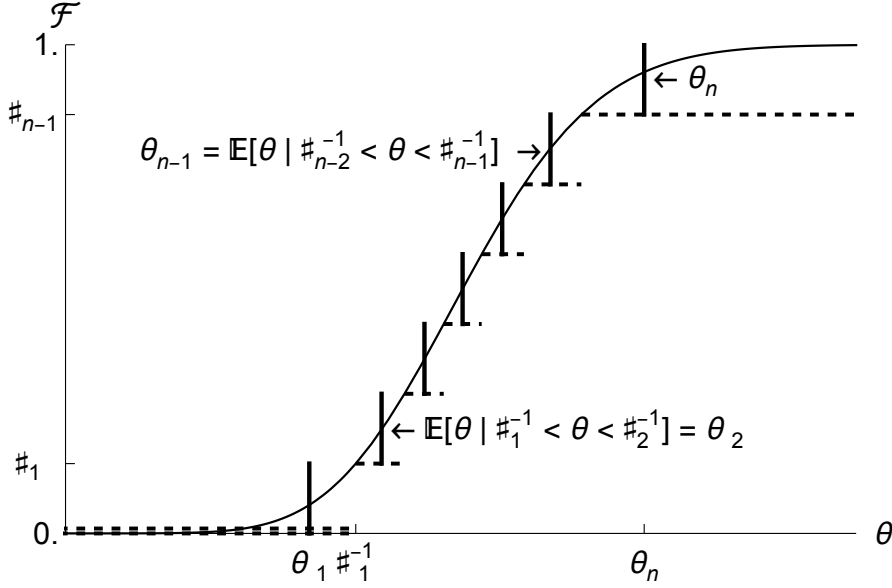
Our first speedup trick is therefore to construct a discrete approximation to the lognormal distribution that can be used in place of numerical integration. That is, we want to approximate the expectation over  $\theta$  of a function  $g(\theta)$  by calculating its value at set of  $n_\theta$  points  $\theta_i$ , each of which has an associated probability weight  $w_i$ :

$$\begin{aligned} \mathbb{E}[g(\theta)] &= \int_{\underline{\theta}}^{\bar{\theta}} g(\vartheta) d\mathcal{F}(\vartheta) \\ &\approx \sum_{\theta=1}^n w_i g(\theta_i) \end{aligned}$$

(because adding  $n$  weighted values to each other is enormously faster than general-purpose numerical integration).

Such a procedure is called a ‘quadrature’ method of integration; Tanaka and Toda (2013) survey a number of options, but for our purposes we choose the one which is easiest to understand: An ‘equiprobable’ approximation (that is, one where each of the values of  $\theta_i$  has an equal probability, equal to  $1/n_\theta$ ).

We calculate such an  $n$ -point approximation as follows.



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

{fig:discreteapprox}

Define a set of points from  $\#_0$  to  $\#_{n_\theta}$  on the  $[0, 1]$  interval as the elements of the set  $\# = \{0, 1/n, 2/n, \dots, 1\}$ .<sup>4</sup> Call the inverse of the  $\theta$  distribution  $\mathcal{F}^{-1}$ , and define the points  $\#_i^{-1} = \mathcal{F}^{-1}(\#_i)$ . Then the conditional mean of  $\theta$  in each of the intervals numbered 1 to  $n$  is:

$$\theta_i \equiv \mathbb{E}[\theta | \#_{i-1}^{-1} \leq \theta < \#_i^{-1}] = \int_{\#_{i-1}^{-1}}^{\#_i^{-1}} \vartheta d\mathcal{F}(\vartheta), \quad (22)$$

and when the integral is evaluated numerically for each  $i$  the result is a set of values of  $\theta$  that correspond to the mean value in each of the  $n$  intervals.

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  $n_\theta$  equiprobable values of  $\theta_i$  which are used to approximate this distribution.<sup>5</sup>

Because one of the purposes of these notes is to connect the math to the code that solves the math, we display here a brief snippet from the notebook that constructs these points.

```
# This is a snippet of code that constructs mass points
# for the equiprobable representation of the problem
```

<sup>4</sup>These points define intervals that constitute a partition of the domain of  $\mathcal{F}$ .

<sup>5</sup>More 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.

We now substitute in our definition of  $v_{(t-1)\rightarrow}$ :

$$v_{(t-1)\rightarrow}(a_{t-1}) = \beta \left( \frac{1}{n_{\theta}} \right) \sum_{i=1}^{n_{\theta}} \frac{(\mathcal{R}_t a_t + \theta_i)^{1-\rho}}{1-\rho} \quad (23) \quad \{\text{eq:vDiscrete}\}$$

so we can rewrite the maximization problem that defines the middle step of period  $t-1$  as

$$v_{t-1}(m_{T-1}) = \max_{c_{T-1}} \left\{ \frac{c_{T-1}^{1-\rho}}{1-\rho} + v_{(t-1)\rightarrow}(m_{T-1} - c_{T-1}) \right\}. \quad (24) \quad \{\text{eq:vEndTm1}\}$$

# The code that corresponds to evaluation of the discretized max problem is

## 5.2 The Approximate Consumption and Value Functions

Given any particular value of  $m_{T-1}$ , a numerical maximization routine can now find the  $c_{T-1}$  that solves (24) in a reasonable amount of time.

The notebook code responsible for computing an estimated consumption function begins in “Solving the Model by Value Function Maximization,” where a vector containing a set of possible values of market resources  $m$  is created (in the code, various  $m$  vectors have names beginning **mVec**; in these notes we will use boldface italics to represent vectors, so we can refer to our collection of  $m$  points as  $\mathbf{m}$  with values indexed by brackets:  $\mathbf{m}[1]$  is the first entry in the vector, up to a last entry  $\mathbf{m}[-1]$ ; we arbitrarily (and suboptimally) pick the first five integers as our five **mVec** gridpoints (in the code, **mVec\_int**= {0., 1., 2., 3., 4.})).

## 5.3 An Interpolated Consumption Function

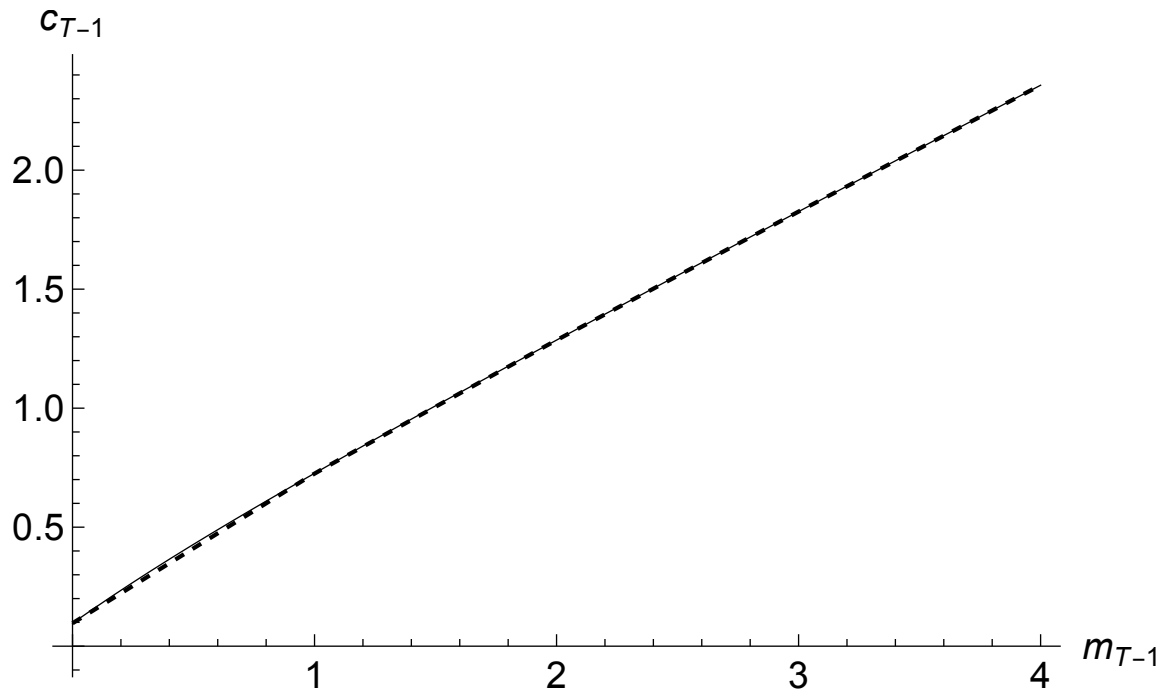
We can now apply our solution to (24) to each of the values in  $\mathbf{m}$ , generating a corresponding optimal  $\mathbf{c}$ . This is called ‘sampling’ the consumption function  $c_{T-1}(m)$ . Using the ordered pairs  $\{\mathbf{m}, \mathbf{c}\}$  we can create a piecewise linear ‘interpolating function’ (a ‘spline’) which when applied to any input  $\mathbf{m}[1] \leq m \leq \mathbf{m}[-1]$  will yield the value of  $c$  that corresponds to a linear ‘connect-the-dots’ interpolation of the value of  $c$  from the values of the two nearest computed  $\{m, c\}$  points.

This is accomplished in “An Interpolated Consumption Function,” which generates an interpolating function that we designate  $\hat{c}_t(m_{T-1})$ .

Figures 2 and 3 show plots of the constructed  $\hat{c}_{T-1}$  and  $\hat{v}_{T-1}$ . 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).<sup>6</sup>

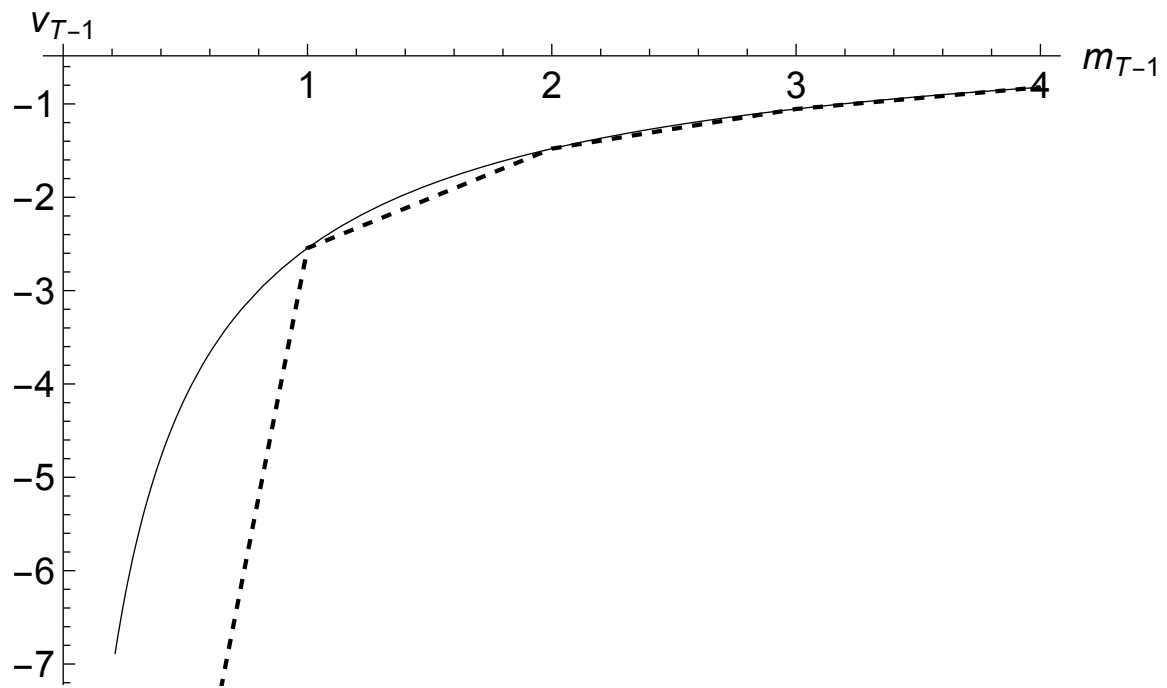
---

<sup>6</sup>For 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  $\dot{c}_{T-1}(m_{T-1})$  (dashed)

{fig:PlotcTm1Simp}



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

{fig:PlotVTm1Simp}

## 5.4 Interpolating Expectations

Piecewise linear ‘spline’ interpolation as described above works well for generating a good approximation to the true optimal consumption function. However, there is a clear inefficiency in the program: Since it uses equation (24), for every value of  $m_{T-1}$  the program must calculate the utility consequences of various possible choices of  $c_{T-1}$  (and therefore  $a_{T-1}$ ) as it searches for the best choice.

For any given index  $j$  in  $\mathbf{m}[j]$ , the algorithm, as it searches for the corresponding optimal  $a$ , will end up calculating  $v_{(t-1)\rightarrow}(\tilde{a}) = \mathbb{E}_{(t-1)\rightarrow}[v_t]$  for many  $\tilde{a}$  values close to the optimal  $a_{T-1}$ . Indeed, even when searching for the optimal  $a$  for a *different*  $m$  (say  $\mathbf{m}[k]$  for  $k \neq j$ ) the search process might compute  $v_{(t-1)\rightarrow}(a)$  for an  $a$  close to the correct optimal  $a$  for  $\mathbf{m}[j]$ . But if that difficult computation does not correspond to the exact solution to the  $\mathbf{m}[k]$  problem, it is discarded.

To avoid solving the problem independently over and over again for multitudes of values of  $a$  that are close to each other, we can employ the same interpolation technique used above to construct a direct numerical approximation to the value function: Define a vector of possible values for end-of-period assets at time  $T-1$ ,  $\mathbf{a}_{T-1}$  (`aVec` in the code). Next, construct  $\mathbf{v}_{t-1} = v_{t-1}(\mathbf{a}_{T-1})$  using equation (23); then construct an approximation  $\hat{v}_{(T-1)\rightarrow}(a_{T-1})$  by passing the vectors `aVec` and `vVec` as arguments to some piecewise-linear interpolator (e.g., `scipy.interpolate`). (These lists contain the points of the  $\mathbf{a}_{T-1}$  and  $\mathbf{v}_{T-1}$  vectors, respectively.)

The notebook section “Interpolating Expectations,” now interpolates the expected value of *ending* the period with a given amount of assets.<sup>7</sup>

Figure 4 compares the true value function to the approximation produced by following the interpolation procedure; the functions are of course identical at the gridpoints of  $a_{T-1}$  and they appear reasonably close except in the region below  $m_{T-1} = 1$ .

Nevertheless, the consumption rule obtained when the approximating  $\hat{v}_{(T-1)\rightarrow}(a_{T-1})$  is used instead of  $v_{(T-1)\rightarrow}(a_{T-1})$  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 concave, which is distressing because Carroll and Kimball (1996) prove that the correct consumption function is strictly concave in a wide class of problems that includes this one.

In all figs,  
replace gothic  
h with notation  
corresponding  
to the lecture  
notes.

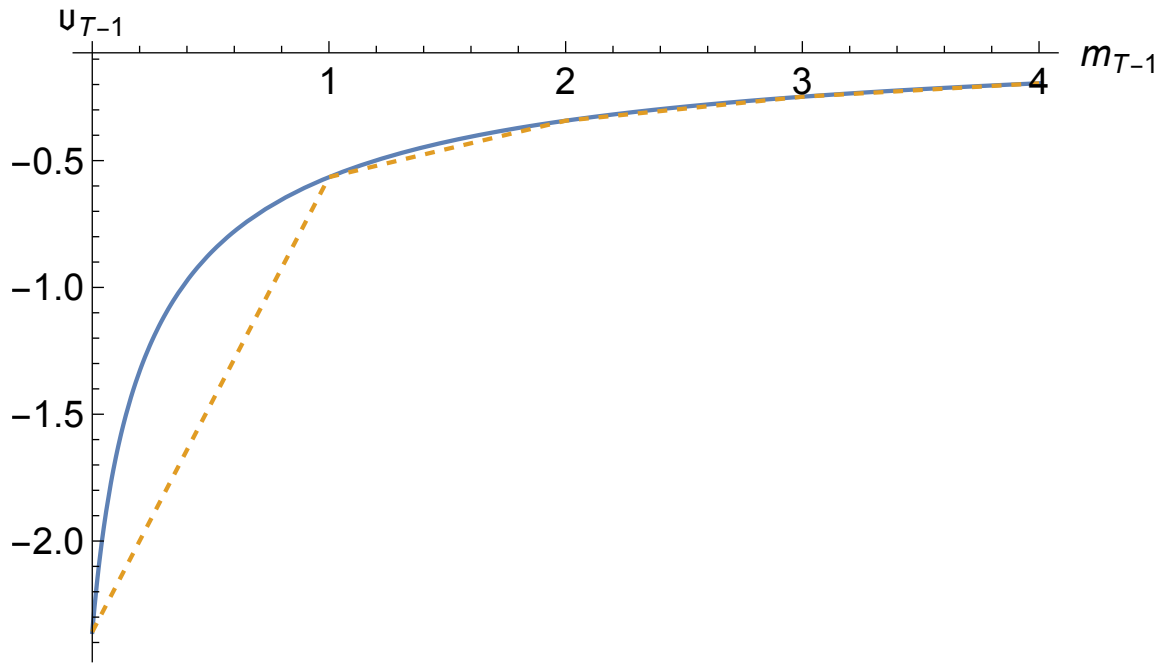
## 5.5 Value Function versus First Order Condition

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

$$\mathbb{E}[-(c - \bar{c})^2] < -(\mathbb{E}[c] - \bar{c})^2 \quad (25)$$

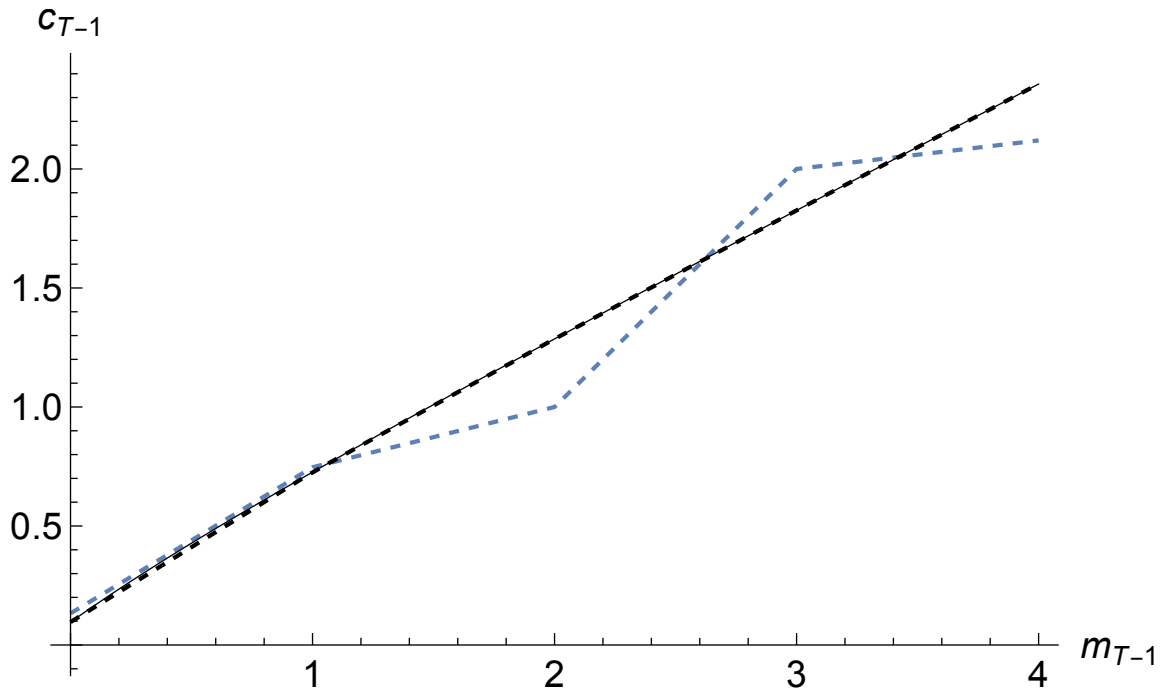
<sup>7</sup>What we are doing here is closely related to ‘the method of parameterized expectations’ of den Haan and Marcet (1990); the only difference is that our method is essentially a nonparametric version.

{subsec:vVsuP}



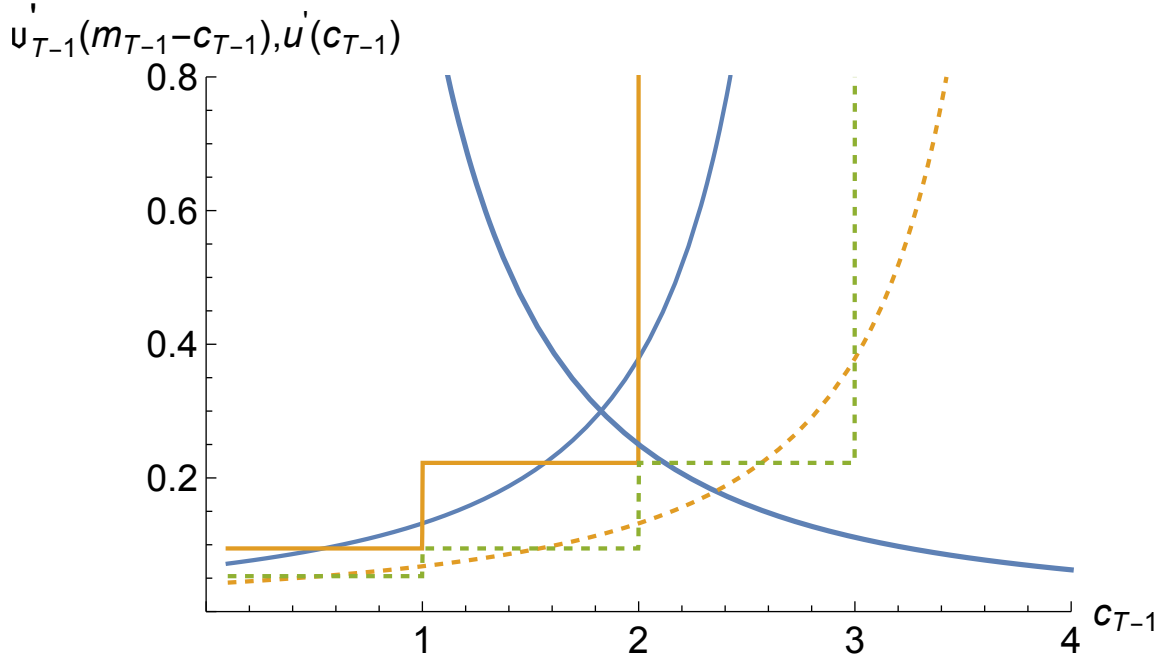
**Figure 4** End-Of-Period Value  $v_{(T-1) \rightarrow}(a_{T-1})$  (solid) versus  $\dot{v}_{(T-1) \rightarrow}(a_{T-1})$  (dashed)

{fig:PlotOTm1Raw}



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

{fig:PlotComparecT}



**Figure 6**  $u^c(c)$  versus  $v_{(T-1) \rightarrow}^a(3-c)$ ,  $v_{(T-1) \rightarrow}^a(4-c)$ ,  $\hat{v}_{(T-1) \rightarrow}^a(3-c)$ ,  $\hat{v}_{(T-1) \rightarrow}^a(4-c)$

{fig:PlotuPrimeVSC}

(where  $\phi$  is the ‘bliss point’ which is assumed always to exceed feasible  $c$ ). 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:

$$\mathbb{E}[-2(c - \phi)] = -2(\mathbb{E}[c] - \phi). \quad (26)$$

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 yield a more accurate approximation to optimal behavior than techniques that approximate the *level* of the value function.

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

$$\begin{aligned} u^c(c_{T-1}) &= \beta \mathbb{E}_{\rightarrow(T-1)}[Ru^c(c_T)] \\ c_{T-1}^{-\rho} &= R\beta \left( \frac{1}{n_{\theta}} \right) \sum_{i=1}^{n_{\theta}} (R(m_{T-1} - c_{T-1}) + \theta_i)^{-\rho}. \end{aligned} \quad (27) \quad \{\text{eq:FOCTm1}\}$$

In the notebook, the “Value Function versus the First Order Condition” section completes the task of finding the values of consumption which satisfy the first order condition in (27) using the `brentq` function from the `scipy` package.

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 (27) as a function of  $c_{T-1}$  under the assumption that  $m_{T-1} = 3$ . Constructing this figure is time-consuming, because for every value of  $c_{T-1}$  plotted we must calculate the RHS of (27). The value of  $c_{T-1}$  for which the RHS and LHS of (27) 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 (approximated) 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 (27) under the assumption that  $m_{T-1} = 4$ , and the intersection of this curve with  $u^c(c_{T-1})$  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  $\dot{v}_{(T-1)\rightarrow}(a_{T-1})$ . Because we have constructed  $\dot{v}_{(T-1)\rightarrow}$  as a linear interpolation, the slope of  $\dot{v}_{(T-1)\rightarrow}(a_{T-1})$  between any two adjacent points  $\{\mathbf{a}[i], \mathbf{a}[i+1]\}$  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  $\dot{v}_{(T-1)\rightarrow}(a_{T-1})$  follows a step function.

The solid-line step function in Figure 6 depicts the actual value of  $\dot{v}_{(T-1)\rightarrow}^a(3 - c_{T-1})$ . When we attempt to find optimal values of  $c_{T-1}$  given  $m_{T-1}$  using  $\dot{v}_{(T-1)\rightarrow}(a_{T-1})$ , the numerical optimization routine will return the  $c_{T-1}$  for which  $u^c(c_{T-1}) = \dot{v}_{(T-1)\rightarrow}^a(m_{T-1} - c_{T-1})$ . Thus, for  $m_{T-1} = 3$  the program will return the value of  $c_{T-1}$  for which the downward-sloping  $u^c(c_{T-1})$  curve intersects with the  $\dot{v}_{(T-1)\rightarrow}^a(3 - c_{T-1})$ ; as the diagram shows, this value is exactly equal to 2. Similarly, if we ask the routine to find the optimal  $c_{T-1}$  for  $m_{T-1} = 4$ , it finds the point of intersection of  $u^c(c_{T-1})$  with  $\dot{v}_{(T-1)\rightarrow}^a(4 - c_{T-1})$ ; 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  $m_{T-1} = 4$ .

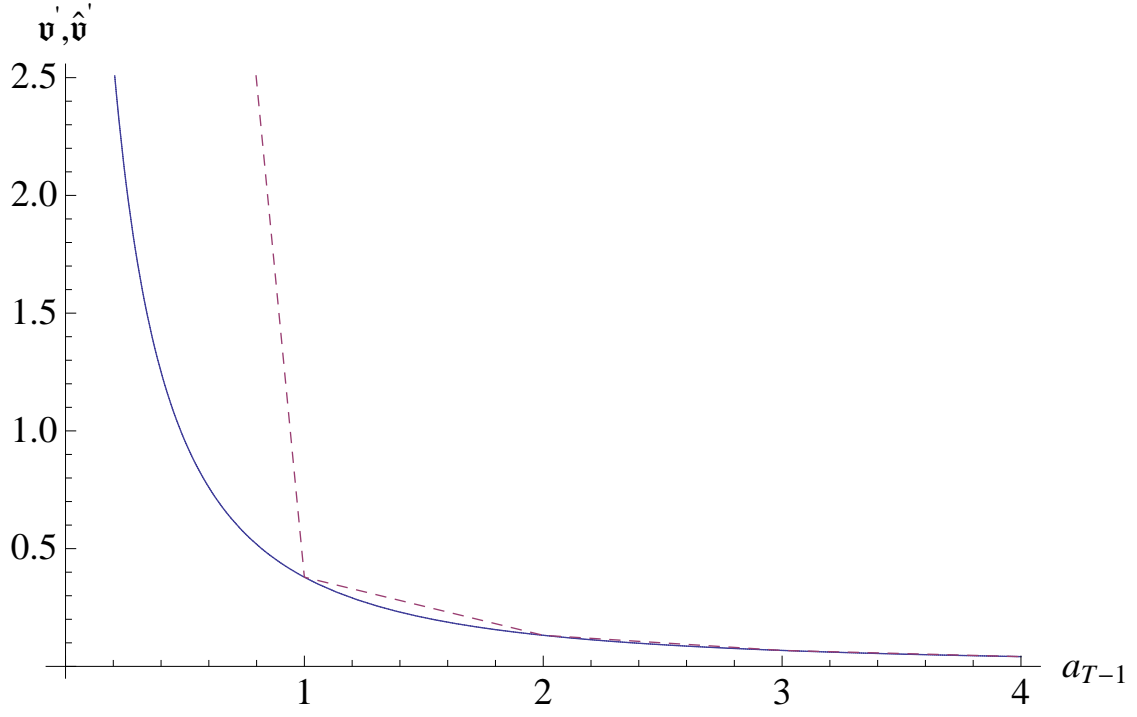
We would obviously obtain much better estimates of the point of intersection between  $u^c(c_{T-1})$  and  $\dot{v}_{(T-1)\rightarrow}^a(m_{T-1} - c_{T-1})$  if our estimate of  $\dot{v}_{(T-1)\rightarrow}^a$  were not a step function. In fact, we already know how to construct linear interpolations to functions, so the obvious next step is to construct a linear interpolating approximation to the *expected marginal value of end-of-period assets function* at the points in  $\mathbf{a}$ :

$$\dot{v}_{(T-1)\rightarrow}^a(\mathbf{a}) = \beta R \left( \frac{1}{n_{\theta}} \right) \sum_{i=1}^{n_{\theta}} (\mathcal{R}_T \mathbf{a} + \boldsymbol{\theta}_i)^{-\rho} \quad (28)$$

yielding  $\mathbf{v}_{(T-1)\rightarrow}^a$  (the vector of expected end-of-period- $(T-1)$  marginal values of assets corresponding to  $\mathbf{aVec}$ ), and construct  $\dot{v}_{(T-1)\rightarrow}^a(a_{T-1})$  as the linear interpolating function that fits this set of points.

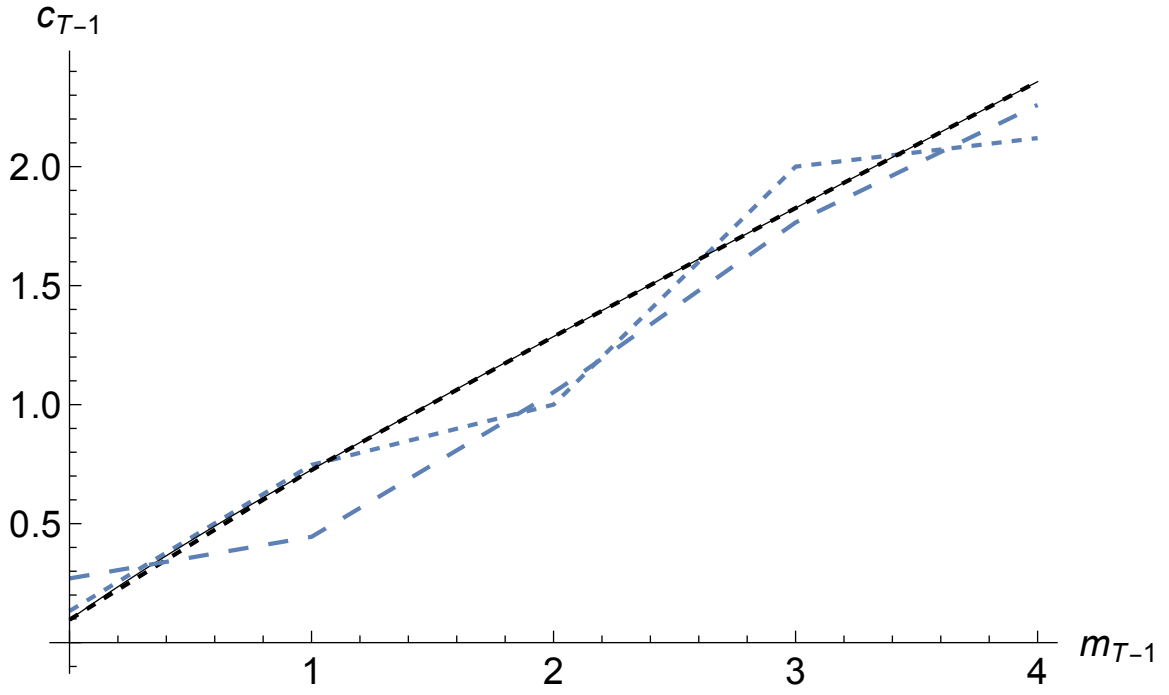
The results are shown in Figure 7. The linear interpolating approximation looks roughly as good (or bad) for the *marginal* value function as it was for the level of the value function. However, Figure 8 shows that the new consumption function (long dashes) is a considerably better approximation of the true consumption function (solid) than was the consumption function obtained by approximating the level of the value function (short dashes).





**Figure 7**  $v_{(T-1) \rightarrow}^a(a_{T-1})$  versus  $\hat{v}_{(T-1) \rightarrow}^a(a_{T-1})$

{fig:PlotOPRawVSI}



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

{fig:PlotcTm1ABC}

## 5.6 Transformation

{subsec:transformat

Even the new-and-improved consumption function diverges notably 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  $v_{(T-1)\rightarrow}^a(a_{T-1})$  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  $\mathcal{R}_T = \beta = \mathcal{G}_T = 1$  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:

$$c_{T-1}^{-\rho} = c_T^{-\rho}. \quad (29)$$

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 its PDV as equivalent to a unit of current wealth. Total resources available at the point when the consumption decision is made is therefore comprised of two types: current market resources  $m_{T-1}$  and ‘human wealth’ (the PDV of future income) of  $h_{(t-1)\rightarrow} = 1$  (because it is the value of human wealth as of the end of the period, and there is only one more period of income of 1 left).

The well-known 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  $m_{T-1} + h_{T-1} = m_{T-1} + 1$ , and since  $v_t^m(m_{T-1}) = u^c(c_{T-1})$ , this implies that

$$v_{t-1}^m(m_{T-1}) = \left( \frac{m_{T-1} + 1}{2} \right)^{-\rho}. \quad (30) \quad \{\text{eq:vPLin}\}$$

Of course, this is a highly nonlinear function. However, if we raise both sides of (30) to the power  $(-1/\rho)$  the result is a linear function:

$$[v_{t-1}^m(m_{T-1})]^{-1/\rho} = \frac{m_{T-1} + 1}{2}. \quad (31)$$

This is a specific example of a general phenomenon: A theoretical literature discussed 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 the key insight in intuitive terms, the nonlinearity we are facing springs in large part from the fact 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

$v_{t-1}^a(a_{T-1})$ , we can define an ‘inverse marginal value’ function

$$\Lambda_{t \rightarrow}^a(a_t) \equiv (v_{t \rightarrow}^a(a_t))^{-1/\rho} \quad (32) \quad \{\text{eq:cGoth}\}$$

which would be linear in the perfect foresight case.<sup>8</sup> We then construct a piecewise-linear interpolating approximation to the  $\Lambda_t^a$  function,  $\hat{\Lambda}_{t \rightarrow}^a(a_t)$ , and for any  $a$  that falls in the range  $\{\mathbf{a}[1], \mathbf{a}[-1]\}$  we obtain our approximation of marginal value from:

$$\hat{v}_t^a(a_t) = [\hat{\Lambda}_t^a(a_t)]^{-\rho} \quad (33)$$

The most interesting thing about all of this, though, is that the  $\Lambda_t^a$  function has another interpretation. Recall our point in (18) that  $u^c(c_t) = v_{t \rightarrow}^a(m_t - c_t)$ . Since with CRRA utility  $u^c(c) = c^{-\rho}$ , this can be rewritten and inverted

$$\begin{aligned} (c_t)^{-\rho} &= v_{t \rightarrow}^a(a_t) \\ c_t &= (v_{t \rightarrow}^a(a_t))^{-1/\rho}. \end{aligned} \quad (34)$$

What this means is that for any given  $a$ , if we can calculate the marginal value associated with ending the period with that  $a$ , then we can learn the level of  $c$  that the consumer must have chosen if they ended up with that  $a$  as the result of an optimal unconstrained choice. This leads us to an alternative interpretation of  $\Lambda^a$ . It is the function that reveals, for any ending  $a$ , how much the agent must have consumed to (optimally) get to that  $a$ . We will therefore henceforth refer to it as the ‘consumed function:’

$$\hat{c}_{t \rightarrow}(a_t) \equiv \hat{\Lambda}_{t \rightarrow}^a(a_t). \quad (35) \quad \{\text{eq:consumedfn}\}$$

Thus, for example, for period  $t - 1$  our procedure is to calculate the vector of  $\mathbf{c}$  points on the consumed function:

$$\mathbf{c} = c_{(t-1) \rightarrow}(\mathbf{a}) \quad (36) \quad \{\text{eq:consumedfnvecs}\}$$

with the idea that we will construct an approximation of the consumed function  $\hat{c}_{(t-1) \rightarrow}$  as the interpolating function connecting these  $\{\mathbf{a}, \mathbf{c}\}$  points.

## 5.7 The Natural Borrowing Constraint and the $a_{t-1}$ Lower Bound

\{\text{subsec:LiqConstrS}\}

This is the appropriate moment to ask an awkward question: How should an interpolated, approximated ‘consumed’ function like  $\hat{c}_{(t-1) \rightarrow}(a_{t-1})$  be extrapolated to return an estimated ‘consumed’ amount when evaluated at an  $a_{t-1}$  outside the range spanned by  $\{\mathbf{a}[1], \dots, \mathbf{a}[n]\}$ ?

For most canned piecewise-linear interpolation tools like `scipy.interpolate`, when the ‘interpolating’ function is evaluated at a point outside the provided range, the algorithm extrapolates under the assumption that the slope of the function remains constant beyond its measured boundaries (that is, the slope is assumed to be equal to the slope of nearest piecewise segment *within* the interpolated range); for example, if the bottommost gridpoint is  $a_1 = \mathbf{a}[1]$  and the corresponding consumed level is  $c_1 = c_{(t-1) \rightarrow}(a_1)$  we could

<sup>8</sup>There is a corresponding inverse for the value function:  $\Lambda_{t \rightarrow}(a_t) = ((1 - \rho)v_{t \rightarrow})^{1/(1-\rho)}$ , and for the marginal marginal value function etc.

calculate the ‘marginal propensity to have consumed’  $\varkappa_1 = \dot{c}_{(t-1)\rightarrow}^a(a_1)$  and construct the approximation as the linear extrapolation below  $\mathbf{a}[1]$  from:

$$\dot{c}_{(t-1)\rightarrow}(a) \equiv c_1 + (a - a_1)\varkappa_1. \quad (37) \quad \{\text{eq:ExtrapLin}\}$$

To see that this will lead us into difficulties, consider what happens to the true (not approximated)  $v_{(t-1)\rightarrow}^a(a_{t-1})$  as  $a_{t-1}$  approaches a quantity we will call the ‘natural borrowing constraint’:  $\underline{a}_{t-1} = -\underline{\theta}\mathcal{R}_t^{-1}$ . From (28) we have

$$\lim_{a_{t-1} \downarrow \underline{a}_{t-1}} v_{(t-1)\rightarrow}^a(a_{t-1}) = \lim_{a_{t-1} \downarrow \underline{a}_{t-1}} \beta R \left( \frac{1}{n_{\theta}} \right) \sum_{i=1}^{n_{\theta}} (a_{t-1}\mathcal{R}_t + \theta_i)^{-\rho}. \quad (38)$$

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  $t$  ( $\mathcal{R}_t \underline{a}_{t-1} = -\theta_1$ ). 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  $t$ , they will have to consume zero in period  $t$ , which yields  $-\infty$  utility and  $+\infty$  marginal utility.

As Zeldes (1989) showed, these reflections show us that the consumer faces a ‘self-imposed’ (or, as above, ‘natural’) borrowing constraint (which springs from the precautionary motive): They 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  $\underline{a}_{t-1}$ ). The constraint is ‘self-imposed’ in the precise sense that if the utility function were different (say, Constant Absolute Risk Aversion), the consumer might 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.<sup>9</sup>

This self-imposed constraint cannot be captured well when the  $v_{(t-1)\rightarrow}^a$  function is approximated by a piecewise linear function like  $\dot{v}_{(t-1)\rightarrow}^m$ , because it is impossible for the linear extrapolation below  $\underline{a}$  to correctly predict  $v_{(t-1)\rightarrow}^a(\underline{a}_{t-1}) = \infty$ .

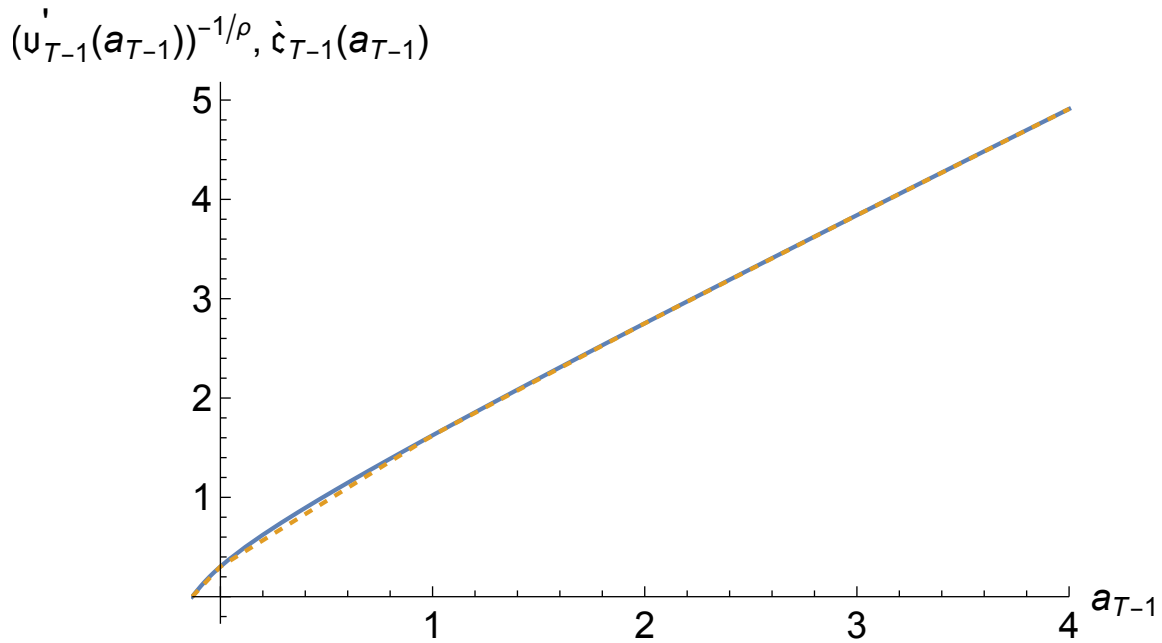
So, the marginal value of saving approaches infinity as  $a_{t-1} \downarrow \underline{a}_{t-1} = -\underline{\theta}\mathcal{R}_t^{-1}$ . But this implies that  $\lim_{a_{t-1} \downarrow \underline{a}_{t-1}} c_{(t-1)\rightarrow}(a_{t-1}) = (v_{(t-1)\rightarrow}^a(a_{t-1}))^{-1/\rho} = 0$ ; that is, as  $a$  approaches its ‘natural borrowing constraint’ minimum possible value, the corresponding amount of worst-case  $c$  must approach *its* lower bound: zero.

The upshot is a realization that all we need to do to address these problems is to prepend each of the  $\mathbf{a}_{t-1}$  and  $\mathbf{c}_{t-1}$  from (36) with an extra point so that the first element in the mapping that produces our interpolation function is  $\{\underline{a}_{t-1}, 0\}$ . This is done in section “The Self-Imposed ‘Natural’ Borrowing Constraint and the  $a_{t-1}$  Lower Bound” of the notebook.

Figure 9 shows the result. The solid line calculates the exact numerical value of the consumed function  $c_{(t-1)\rightarrow}(a_{t-1})$  while the dashed line is the linear interpolating approximation  $\dot{c}_{(t-1)\rightarrow}(a_{t-1})$ . This figure illustrates the value of the transformation: The true

The vertical axis should be relabeled - not gothic c anymore, instead  $\Lambda^a$

<sup>9</sup>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 teaching purposes.



**Figure 9** True  $\Lambda_{(t-1) \rightarrow}^a(a_{t-1})$  vs its approximation  $\hat{\Lambda}_{(t-1) \rightarrow}^a(a_{t-1})$

{fig:GothVInvVSGo}

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 generate  $\hat{v}_{(t-1) \rightarrow}^a(a)$  using our augmented  $[\hat{c}_{(t-1) \rightarrow}(a)]^{-\rho}$  (dashed line) we obtain a *much* closer approximation to the true marginal value function  $v_{(t-1) \rightarrow}^a(a)$  (solid line) than we obtained in the previous exercise which did not do the transformation (Figure 7).<sup>10</sup>

fix the problem articulated in the footnote

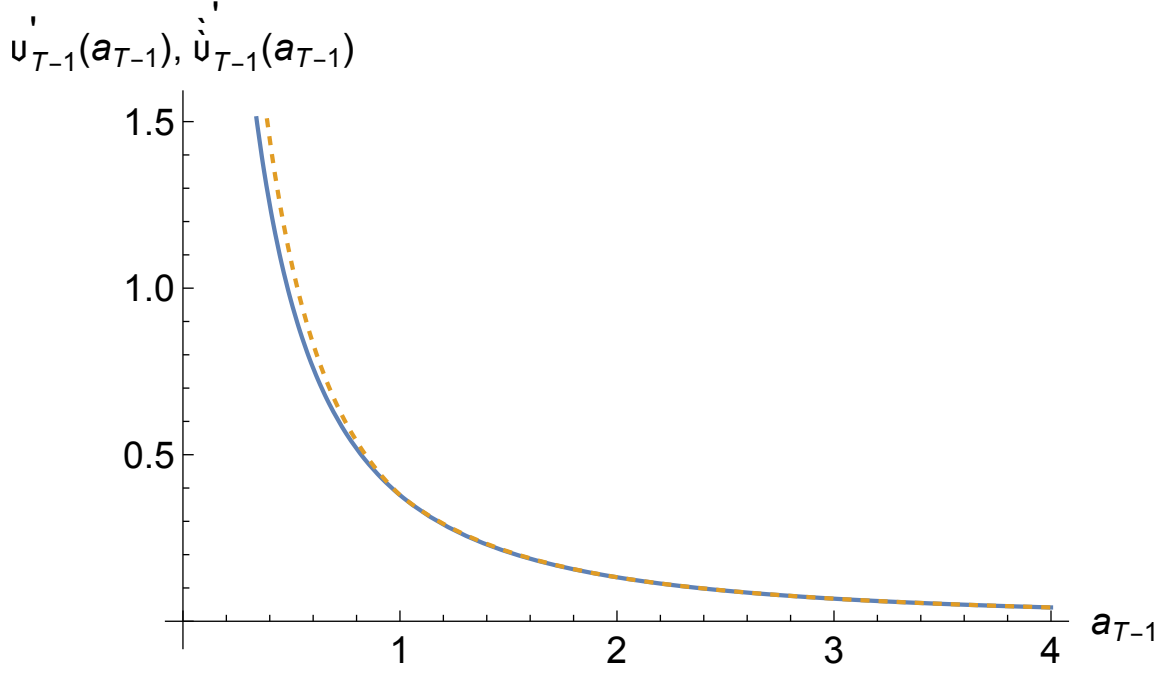
## 5.8 The Method of Endogenous Gridpoints

{subsec:egm}

The solution procedure we articulated above for finding  $c_{t-1}(m)$  still requires us, for each point in  $\mathbf{m}_{t-1}$ , to use a numerical rootfinding algorithm to search for the value of  $c_{t-1}$  that solves  $u^c(c_{t-1}) = v_{(t-1) \rightarrow}^a(m - c_{t-1})$ . Though sections 5.6 and 5.7 developed a highly efficient and accurate procedure to calculate  $\hat{v}_{(t-1) \rightarrow}^a$ , those approximations do nothing to eliminate the need for using a rootfinding operation for calculating, for an arbitrary  $m$ , the optimal  $c$ . And rootfinding is a notoriously computation-intensive (that is, slow!) operation.

Fortunately, it turns out that there is a way to completely skip this slow rootfinding step. The method can be understood by noting that we have already calculated, for a set of arbitrary values of  $\mathbf{a} = \mathbf{a}_{t-1}$ , the corresponding  $\mathbf{c}$  values for which this  $\mathbf{a}$  is optimal.

<sup>10</sup>The vertical axis label uses  $v'$  as an alternative notation for what in these notes we designate as  $v_{(t-1) \rightarrow}^a$ . This will be fixed.



**Figure 10** True  $v_{(t-1)→}^a(a_{t-1})$  vs.  $\hat{v}_{(t-1)→}^a(a_{t-1})$  Constructed Using  $\hat{c}_{(t-1)→}(a_{t-1})$

{fig:GothVVSgothC

But with mutually consistent values of  $\mathbf{c}_{t-1}$  and  $\mathbf{a}_{t-1}$  (consistent, in the sense that they are the unique optimal values that correspond to the solution to the problem), we can obtain the  $\mathbf{m}_{t-1}$  vector that corresponds to both of them from

$$\mathbf{m}_{t-1} = \mathbf{c}_{t-1} + \mathbf{a}_{t-1}. \quad (39)$$

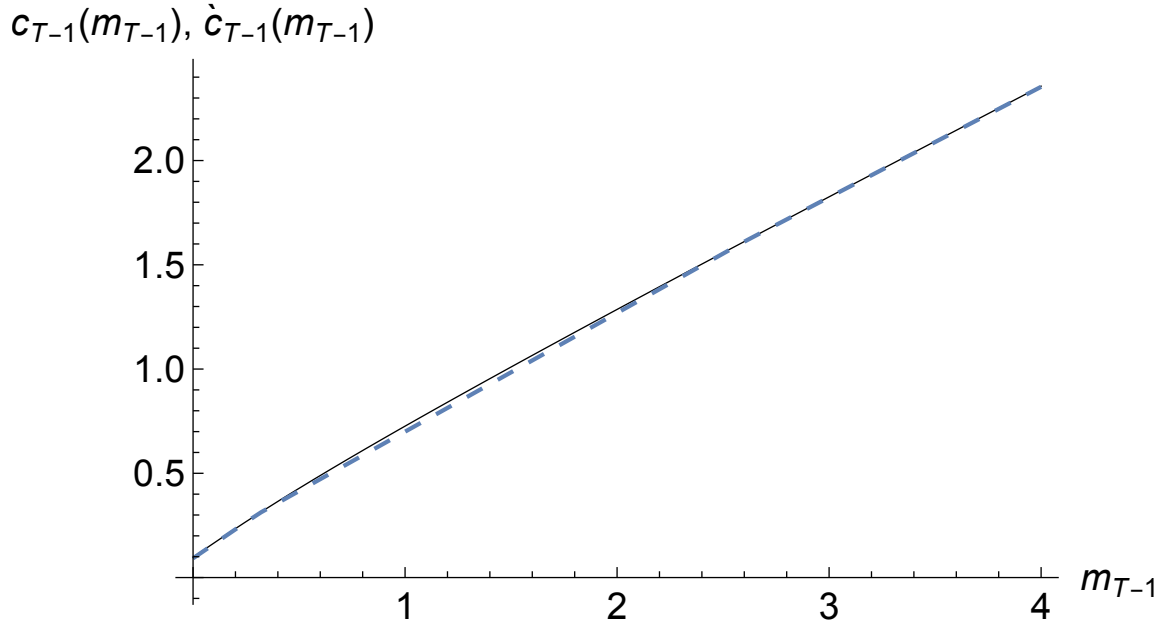
Rename gothic  
class: EndStp

These  $m$  gridpoints are “endogenous” in contrast to the usual solution method of specifying some *ex-ante* (exogenous) grid of values of  $\mathbf{m}$  and then using a rootfinding routine to locate the corresponding optimal consumption vector  $\mathbf{c}$ .

This routine is performed in the “Endogenous Gridpoints” section of the notebook. First, the `gothic.C_Tminus1` function is called for each of the pre-specified values of end-of-period assets stored in `aVec`. These values of consumption and assets are used to produce the list of endogenous gridpoints, stored in the object `mVec_egm`. With the  $\mathbf{c}$  values in hand, the notebook can generate a set of  $\mathbf{m}_{t-1}$  and  $\mathbf{c}_{t-1}$  pairs that can be interpolated between in order to yield  $\hat{c}_{t-1}(m_{t-1})$  at virtually zero computational cost!<sup>11</sup>

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 non-linearity). 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

<sup>11</sup>This is the essential point of Carroll (2006).



**Figure 11**  $c_{t-1}(m)$  (solid) versus  $\hat{c}_{t-1}(m)$  (dashed)

{fig:ComparecTm1}

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

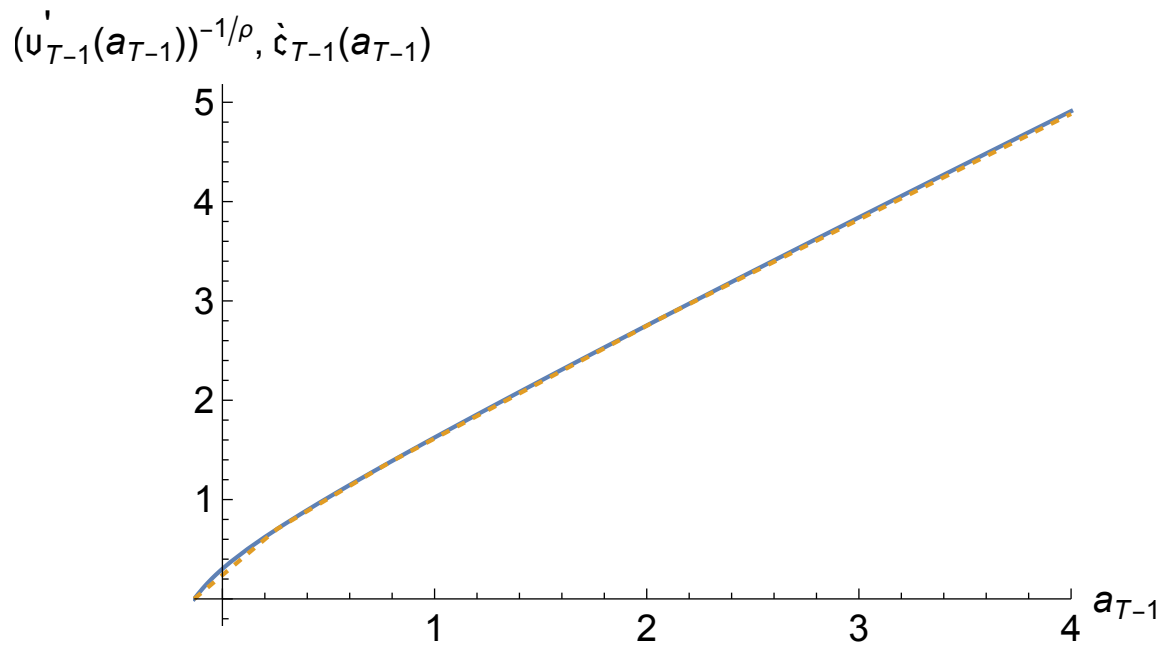
{subsec:improving-t}

Thus far, 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  $\hat{c}_{(t-1) \rightarrow}$  function tends to be farthest from its true value at low values of  $a$ . Combining this with our insight that  $\underline{a}_{t-1}$  is a lower bound, we are now in position to define a more deliberate method for constructing gridpoints for  $a_{t-1}$  – a method that yields values that are more densely spaced at low values of  $a$  where the function is more nonlinear.

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  $n$  – our default is 3) from each point to the next point is constant (instead of, as previously, imposing constancy of the absolute gap between points).

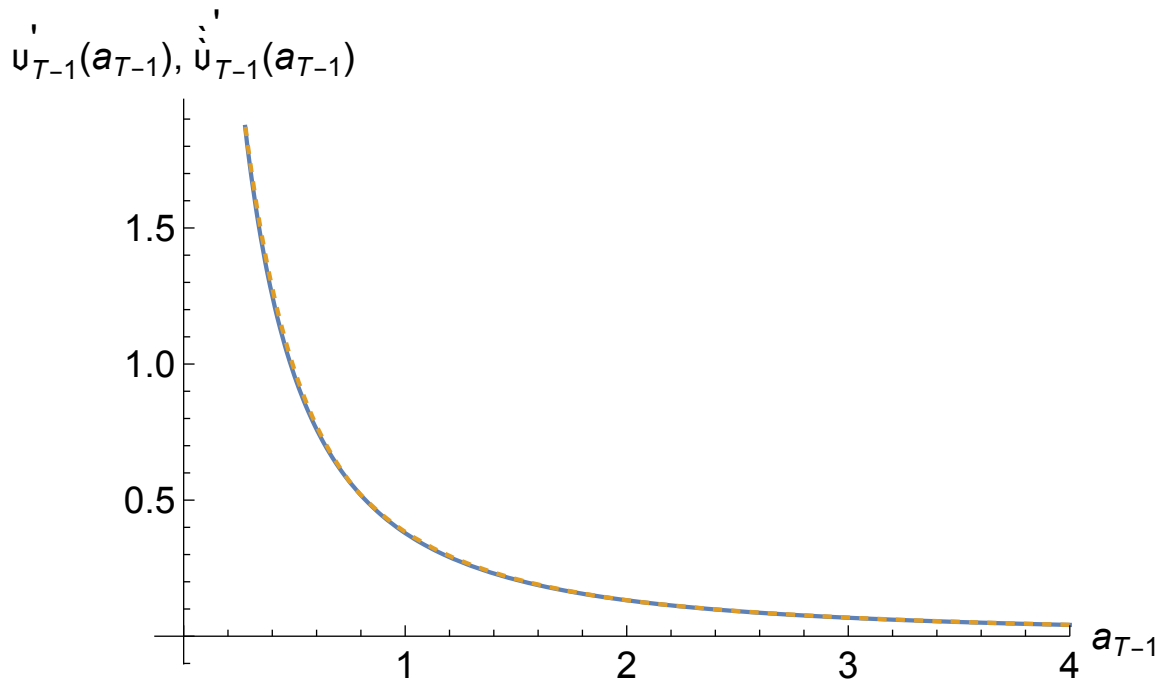
Section “Improve the  $\mathbb{A}_{grid}$ ” begins by defining a function which takes as arguments the specifications of an initial grid of assets and returns the new grid incorporating the multi-exponential approach outlined above.

Notice that the graphs depicted in Figures 12 and 13 are notably closer to their respective truths than the corresponding figures that used the original grid.



**Figure 12**  $c_{(t-1)\rightarrow}(a_{t-1})$  versus  $\dot{c}_{(t-1)\rightarrow}(a_{t-1})$ , Multi-Exponential aVec

{fig:GothVInvVSGo



**Figure 13**  $v^a_{(t-1)\rightarrow}(a_{t-1})$  vs.  $\dot{v}^a_{(t-1)\rightarrow}(a_{t-1})$ , Multi-Exponential aVec

{fig:GothVVSGothC



## 5.10 Program Structure

In section “Solve for  $c_t(m_t)$  in Multiple Periods,” the natural and artificial borrowing constraints are combined with the endogenous gridpoints method to approximate the optimal consumption function for a specific period. Then, this function is used to compute the approximated consumption in the previous period, and this process is repeated for the number of periods specified by  $T$ , as explained earlier.

The essential structure of the program is a loop that iteratively solves for consumption functions by working backward from an assumed final period, using the dictionary `cFunc_life` to store the interpolated consumption functions up to the beginning period. Consumption in a given period is utilized to determine the endogenous gridpoints for the preceding period. This is the sense in which the computation of optimal consumption is done recursively.

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 computational specification of the problem.

## 5.11 Results

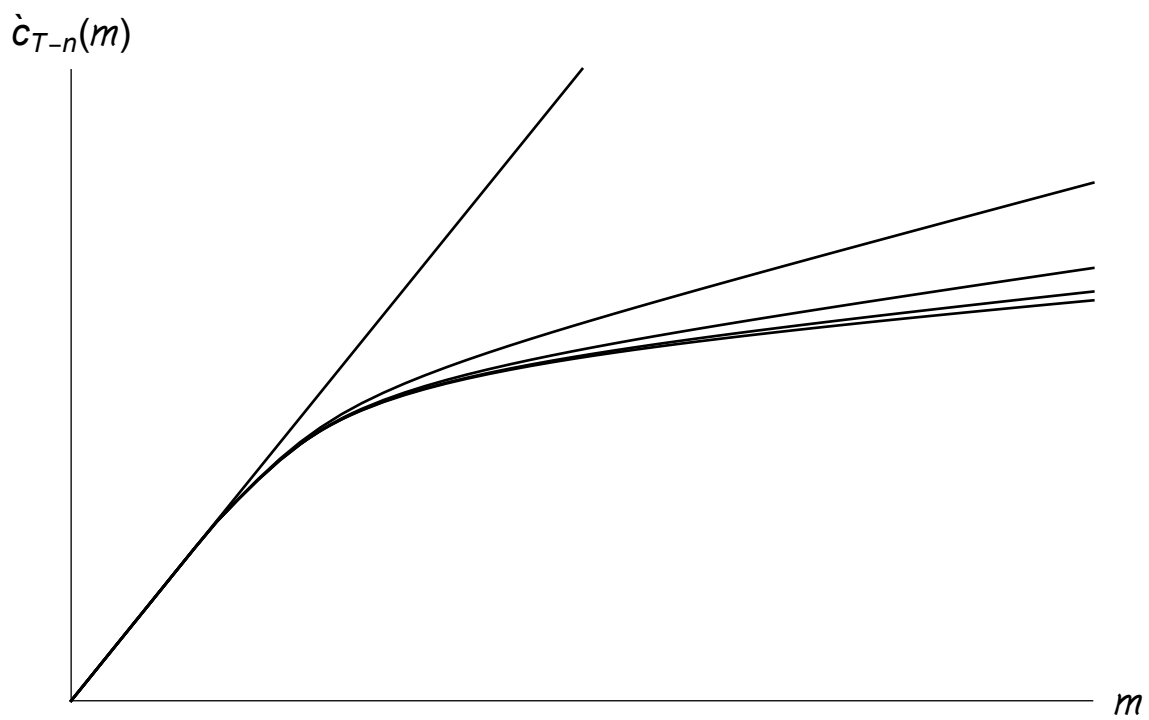
The code creates the relevant  $\hat{c}_t(m_t)$  functions for any period in the horizon specified by the variable  $T$ , at the given values of  $m$ . Figure 14 shows  $\hat{c}_{T-n}(m)$  for  $n = \{20, 15, 10, 5, 1\}$ . At least one feature of this figure is encouraging: the consumption functions converge as the horizon extends, something that Carroll (2023b) shows must be true under certain parametric conditions that are satisfied by the baseline parameter values being used here.

## 5.12 The Method of Moderation

Unfortunately, this endogenous gridpoints solution is not very well-behaved outside the original range of gridpoints targeted by the solution method. (Though other common solution methods are no better outside their own predefined ranges). Figure 15 demonstrates the point by plotting the amount of precautionary saving implied by a linear extrapolation 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, the linearly extrapolated numerical approximation eventually predicts negative precautionary saving (at the point in the figure where the extrapolated locus crosses the horizontal axis).

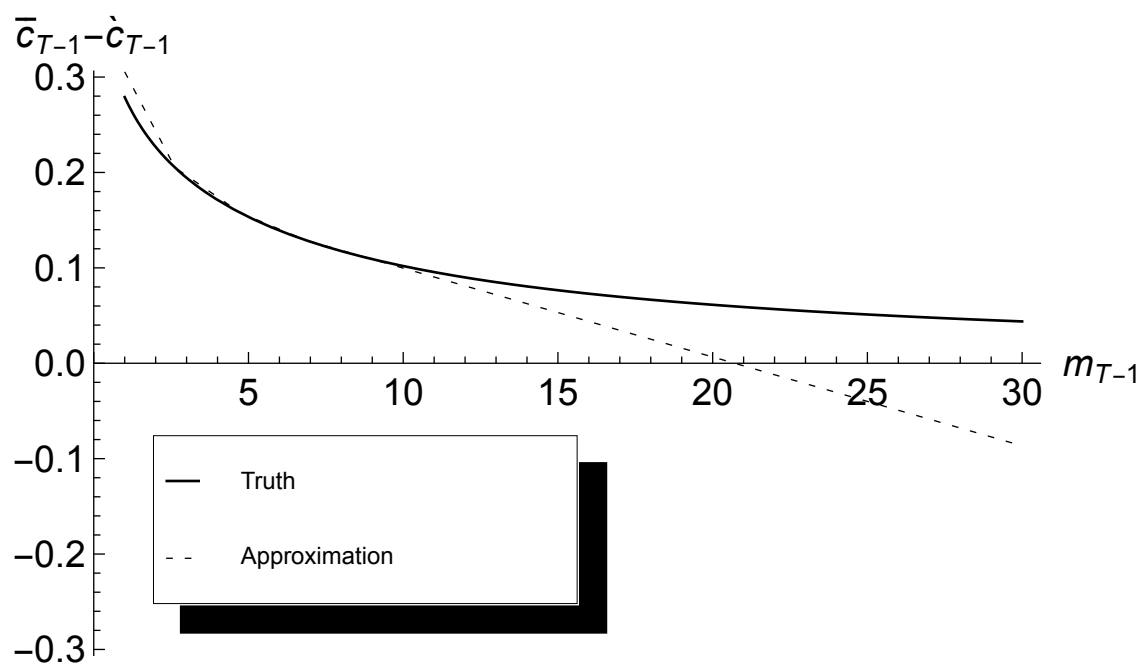
This error cannot be fixed by extending the upper gridpoint; in the presence of serious uncertainty, the consumption rule will need to be evaluated outside of *any* prespecified grid (because starting from the top gridpoint, a large enough realization of the uncertain variable will push next period’s realization of assets above that top; a similar argument applies below the bottom gridpoint). While a judicious extrapolation technique can

{sec:method-of-mod



**Figure 14** Converging  $\hat{c}_{T-n}(m)$  Functions as  $n$  Increases

{fig:PlotCFuncsCon



**Figure 15** For Large Enough  $m_{t-1}$ , Predicted Precautionary Saving is Negative (Oops!)

{fig:ExtrapProblem

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.

As a preliminary to our solution, define  $h_{t \rightarrow}$  as 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 period- $t$  consumer who believes with perfect confidence that the shocks will always take their expected value of 1,  $\theta_{t+n} = \mathbb{E}[\theta] = 1 \forall n > 0$ . The solution to a perfect foresight problem of this kind takes the form<sup>12</sup>

$$\bar{c}_t(m_t) = (m_t + h_{t \rightarrow})\underline{\kappa}_t \quad (40) \quad \{\text{eq:cFuncAbove}\}$$

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

We similarly define  $\underline{h}_{t \rightarrow}$  as ‘minimal human wealth,’ the present discounted value of labor income if the shocks were to take on their worst possible value in every future period  $\theta_{t+n} = \underline{\theta} \forall n > 0$  (which we define as corresponding to the beliefs of 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  $\underline{m}_t = -\underline{h}_{t \rightarrow}$ , 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 explaining the derivation of the ‘natural borrowing constraint’ for period  $T - 1$ ,  $\underline{a}_{t-1}$ ). Since consumption of zero yields negative infinite utility, the solution to realist consumer’s problem is not well defined for values of  $m_t < \underline{m}_t$ , and the limiting value of the realist’s  $c_t$  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:

$$\begin{aligned} \blacktriangle m_t &= m_t + \overbrace{\underline{h}_{t \rightarrow}}^{=-\underline{m}_t} \\ \blacktriangle h_{t \rightarrow} &= h_{t \rightarrow} - \underline{h}_{t \rightarrow}. \end{aligned}$$

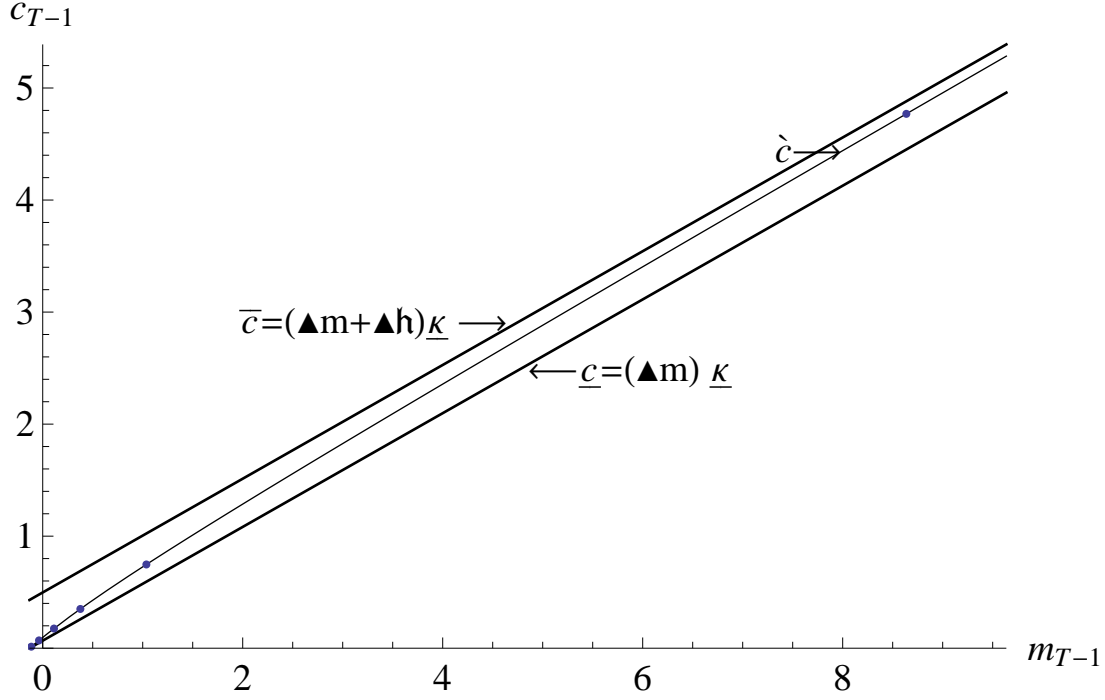
We can now transparently define the optimal consumption rules for the two perfect foresight problems, those of the ‘optimist’ and the ‘pessimist.’ The ‘pessimist’ perceives human wealth to be equal to its minimum feasible value  $\underline{h}_{t \rightarrow}$  with certainty, so consumption is given by the perfect foresight solution

$$\begin{aligned} \underline{c}_t(m_t) &= (m_t + \underline{h}_{t \rightarrow})\underline{\kappa}_t \\ &= \blacktriangle m_t \underline{\kappa}_t. \end{aligned}$$

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

---

<sup>12</sup>For a derivation, see Carroll (2023b);  $\underline{\kappa}_t$  is defined therein as the MPC of the perfect foresight consumer with horizon  $T - t$ .



**Figure 16** Moderation Illustrated:  $\underline{c}_{t-1} < \hat{c}_{t-1} < \bar{c}_{t-1}$

{fig:IntExpFOCInv}

income, and therefore consumes

$$\begin{aligned} \bar{c}_t(m_t) &= (m_t + \underline{h}_{t \rightarrow} - \underline{h}_{t \rightarrow} + h_{t \rightarrow}) \underline{\kappa}_t \\ &= (\blacktriangle m_t + \blacktriangle h_{t \rightarrow}) \underline{\kappa}_t \\ &= \underline{c}_t(m_t) + \blacktriangle h_{t \rightarrow} \underline{\kappa}_t. \end{aligned}$$

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. Figure 16 illustrates the proposition for the consumption rule in period  $T - 1$ .

The proof is more difficult than might be imagined, but the necessary work is done in Carroll (2023b) so we will take the proposition as a fact and proceed by manipulating the inequality:

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

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  $\infty$ ), the object in the middle of the last inequality is

$$\bar{\varphi}_t(\mu_t) \equiv \left( \frac{\bar{c}_t(\underline{m}_t + e^{\mu_t}) - c_t(\underline{m}_t + e^{\mu_t})}{\blacktriangle h_{t \rightarrow \underline{\kappa}_t}} \right), \quad (41) \quad \{\text{eq:koppa}\}$$

and we now define

$$\begin{aligned} {}_t(\mu_t) &= \log \left( \frac{1 - \bar{\varphi}_t(\mu_t)}{\bar{\varphi}_t(\mu_t)} \right) \\ &= \log (1/\bar{\varphi}_t(\mu_t) - 1) \end{aligned} \quad (42) \quad \{\text{eq:chi}\}$$

which has the virtue that it is linear in the limit as  $\mu_t$  approaches  $+\infty$ .

Given , the consumption function can be recovered from

$$\bar{c}_t = \bar{c}_t - \overbrace{\left( \frac{1}{1 + \exp({}_t)} \right)}^{=\bar{\varphi}_t} \blacktriangle h_{t \rightarrow \underline{\kappa}_t}. \quad (43) \quad \{\text{eq:cFuncHi}\}$$

Thus, the procedure is to calculate  ${}_t$  at the points  $\mu_t$  corresponding to the log of the  $\blacktriangle m_t$  points defined above, and then using these to construct an interpolating approximation  $\hat{c}_t$  from which we indirectly obtain our approximated consumption rule  $\hat{\bar{c}}_t$  by substituting  $\hat{c}_t$  for  $\bar{c}_t$  in equation (43).

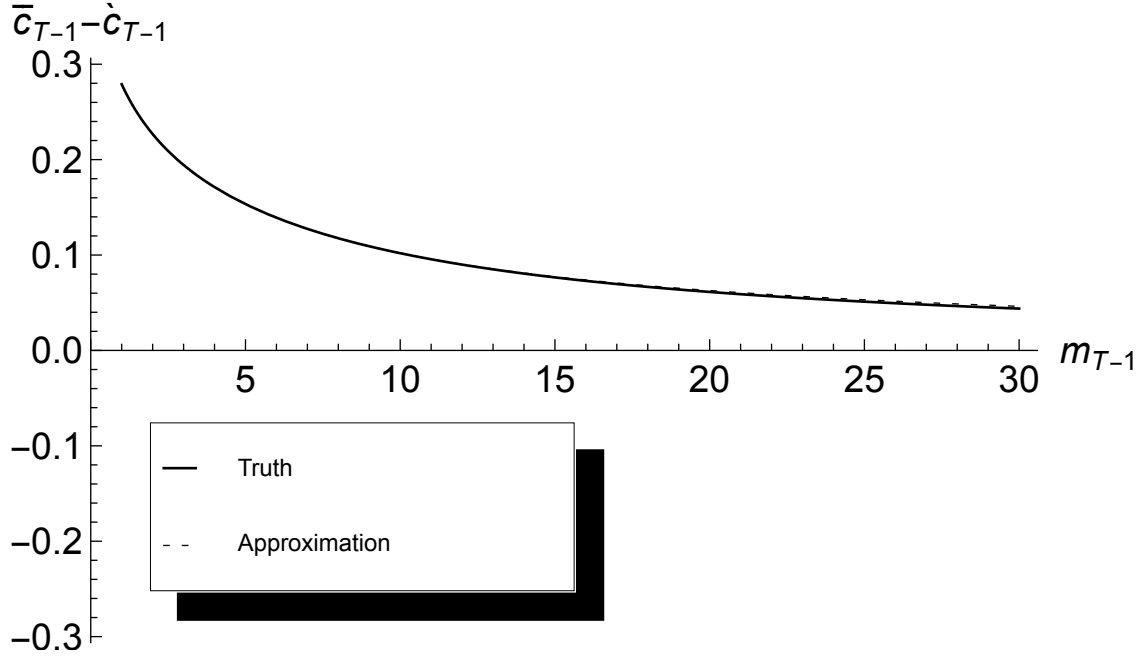
Because this method relies upon the fact that the problem is easy to solve if the decision maker has unreasonable views (either in the optimistic or the pessimistic direction), and because the correct solution is always between these immoderate extremes, we call our solution procedure the ‘method of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

Results are shown in Figure 17; 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 15.

### 5.13 Approximating the Slope Too

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,  $c_{t-1}$ ). 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 (2023b) proves that the true consumption function for this problem is ‘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  $\kappa$ ) at each gridpoint, and then find an interpolating approximation that smoothly matches both the level and the slope at those points.



**Figure 17** Extrapolated  $\dot{\bar{c}}_{t-1}$  Constructed Using the Method of Moderation

{fig:ExtrapProblem}

This requires us to differentiate (41) and (42), yielding

$$\begin{aligned}\bar{\varphi}_t^\mu(\mu_t) &= (\blacktriangle h_{t \rightarrow \underline{\kappa}_t})^{-1} e^{\mu_t} \left( \overbrace{\underline{\kappa}_t - \underline{c}_t^m(\underline{m}_t + e^{\mu_t})}^{\equiv \kappa_t(m_t)} \right) \\ \mu_t &= \left( \frac{-\bar{\varphi}_t^\mu(\mu_t)/\bar{\varphi}_t^2}{1/\bar{\varphi}_t(\mu_t) - 1} \right)\end{aligned}\tag{44}$$

{eq:koppaPrime}

and (dropping arguments) with some algebra these can be combined to yield

$$\mu_t = \left( \frac{\underline{\kappa}_t \blacktriangle m_t \blacktriangle h_{t \rightarrow} (\underline{\kappa}_t - \kappa_t)}{(\bar{c}_t - c_t)(\bar{c}_t - c_t - \underline{\kappa}_t \blacktriangle h_{t \rightarrow})} \right).\tag{45}$$

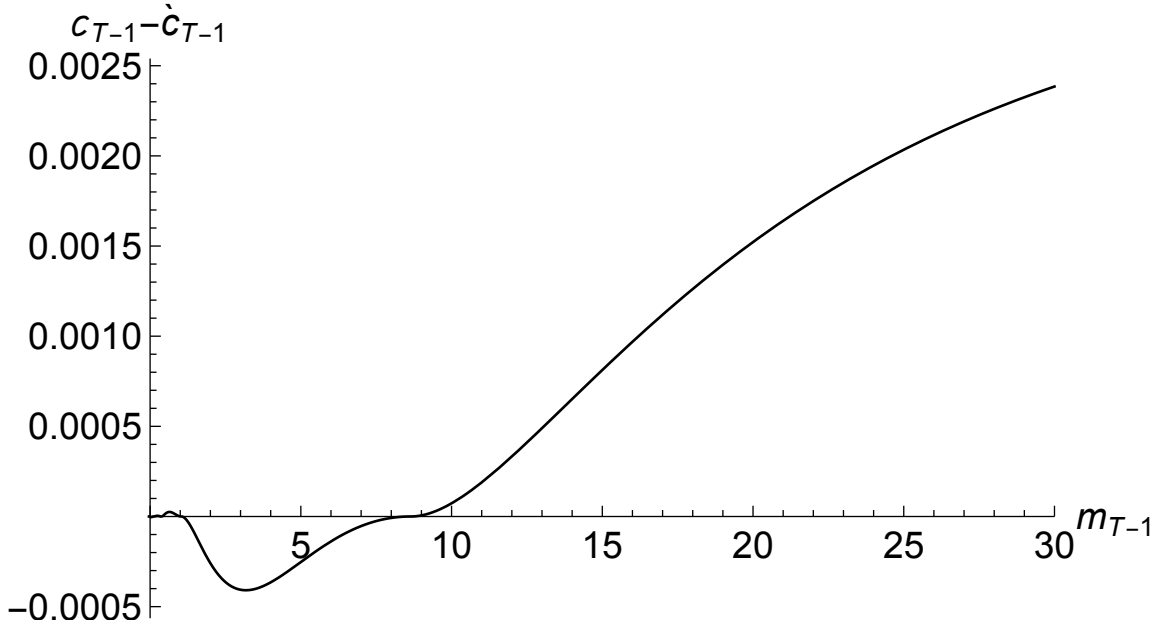
To compute the vector of values of (44) corresponding to the points in  $\mu_t$ , we need the marginal propensities to consume (designated  $\kappa$ ) at each of the gridpoints,  $c_t^m$  (the vector of such values is  $\kappa_t$ ). These can be obtained by differentiating the Euler equation (18) (where we define  $m_{t \rightarrow}(a) \equiv c_{t \rightarrow}(a) + a$ , and drop the (a) arguments to reduce clutter):

$$u^c(c_{t \rightarrow}) = \hat{v}_{t \rightarrow}^a(m_{t \rightarrow} - c_{t \rightarrow}),\tag{46}$$

yielding a marginal propensity to *have consumed*  $c_{t \rightarrow}^a$  at each gridpoint:

$$\begin{aligned}u''(c_{t \rightarrow})c_{t \rightarrow}^a &= \hat{v}_{t \rightarrow}^a(m_{t \rightarrow} - c_{t \rightarrow}) \\ c_{t \rightarrow}^a &= \hat{v}_{t \rightarrow}^a(m_{t \rightarrow} - c_{t \rightarrow})/u''(c_{t \rightarrow})\end{aligned}\tag{47}$$

and the marginal propensity to consume at the beginning of the period is obtained from



**Figure 18** Difference Between True  $c_{t-1}$  and  $\hat{c}_{t-1}$  Is Minuscule

{fig:IntExpFOCInv}

the marginal propensity to have consumed by differentiating the identity with respect to  $a$ :

$$\begin{aligned} c_{t \rightarrow} &= m_{t \rightarrow} - a \\ c_{t \rightarrow}^a + 1 &= m_{t \rightarrow}^a \end{aligned}$$

which, together with the chain rule  $c_{t \rightarrow}^a = c_t^m m_{t \rightarrow}^a$ , yields the MPC from

$$\begin{aligned} c^m(\overbrace{c_{t \rightarrow}^a + 1}^{=m_{t \rightarrow}^a}) &= c_{t \rightarrow}^a \\ c^m &= c_{t \rightarrow}^a / (1 + c_{t \rightarrow}^a). \end{aligned} \tag{48}$$

{eq:MPCfromMPT}

Designating  $\hat{c}_{t-1}$  as the approximated consumption rule obtained using an interpolating polynomial approximation to that matches both the level and the first derivative at the gridpoints, Figure 18 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 exactly 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 plotted in previous figures, because in levels the difference between the approximate and the exact function would not be detectable even to the most eagle-eyed reader.)

## 5.14 Value

Often it is useful to know the value function as well as the consumption rule. Fortunately, many of the tricks used when solving for the consumption rule have a direct analogue in approximation of the value function.

Consider the perfect foresight (or “optimist’s”) problem in period  $T - 1$ . Using the fact that in a perfect foresight model the growth factor for consumption is  $(R\beta)^{1/\rho}$ , we can use the fact that  $c_t = (R\beta)^{1/\rho} c_{t-1}$  to calculate the value function in period  $T - 1$ :

$$\begin{aligned}\bar{v}_{t-1}(m_{t-1}) &\equiv u(c_{t-1}) + \beta u(c_t) \\ &= u(c_{t-1}) (1 + \beta((R\beta)^{1/\rho})^{1-\rho}) \\ &= u(c_{t-1}) (1 + (\beta R)^{1/\rho}/R) \\ &= u(c_{t-1}) \underbrace{\text{PDV}_t^T(c)/c_{t-1}}_{\equiv \mathbb{C}_{t-1}^T}\end{aligned}$$

where  $\mathbb{C}_t^T = \text{PDV}_t^T(c)$  is the present discounted value of consumption, normalized by current consumption. Using the fact demonstrated in Carroll (2023b) that  $\mathbb{C}_t = \kappa_t^{-1}$ , a similar function can be constructed recursively for earlier periods, yielding the general expression

$$\begin{aligned}\bar{v}_t(m_t) &= u(\bar{c}_t) \mathbb{C}_t^T \\ &= u(\bar{c}_t) \underline{\kappa}_t^{-1} \\ &= u(\blacktriangle m_t + \blacktriangle h_{t \rightarrow}) \underline{\kappa}_t \underline{\kappa}_t^{-1} \\ &= u(\blacktriangle m_t + \blacktriangle h_{t \rightarrow}) \underline{\kappa}_t^{1-\rho} \underline{\kappa}_t^{-1} \\ &= u(\blacktriangle m_t + \blacktriangle h_{t \rightarrow}) \underline{\kappa}_t^{-\rho}\end{aligned} \tag{49} \quad \{\text{eq:vFuncPF}\}$$

This can be transformed as

$$\begin{aligned}\bar{\Lambda}_t &\equiv ((1 - \rho)\bar{v}_t)^{1/(1-\rho)} \\ &= c_t (\mathbb{C}_t^T)^{1/(1-\rho)} \\ &= (\blacktriangle m_t + \blacktriangle h_{t \rightarrow}) \underline{\kappa}_t^{-\rho/(1-\rho)}\end{aligned}$$

with derivative

$$\begin{aligned}\bar{\Lambda}_t^m &= (\mathbb{C}_t^T)^{1/(1-\rho)} \underline{\kappa}_t, \\ &= \underline{\kappa}_t^{-\rho/(1-\rho)}\end{aligned}$$

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

$$\begin{aligned}\bar{\Lambda}_t &= ((1 - \rho)\bar{v}_t(m_t))^{1/(1-\rho)} \\ \bar{\Lambda}_t^m &= ((1 - \rho)\bar{v}_t(m_t))^{-1+1/(1-\rho)} \bar{v}_t^m(m_t)\end{aligned}$$

and an excellent approximation to the value function can be obtained by calculating the



values of  $\bar{\lambda}$  at the same gridpoints used by the consumption function approximation, and interpolating among those points.

However, as with the consumption function approximation, we can do even better if we realize that the  $\bar{\lambda}$  function for the optimist's problem is an upper bound for the  $\lambda$  function in the presence of uncertainty, and the value function for the pessimist is a lower bound. Analogously to (41), define an upper-case

$$\hat{Q}_t(\mu_t) = \left( \frac{\bar{\lambda}_t(\underline{m}_t + e^{\mu_t}) - \lambda_t(\underline{m}_t + e^{\mu_t})}{\blacktriangle h_{t \rightarrow \underline{\kappa}_t}(\mathbb{C}_t^T)^{1/(1-\rho)}} \right) \quad (50) \quad \{\text{eq:Koppa}\}$$

with derivative (dropping arguments)

$$\hat{Q}_t^\mu = (\blacktriangle h_{t \rightarrow \underline{\kappa}_t}(\mathbb{C}_t^T)^{1/(1-\rho)})^{-1} e^{\mu_t} (\bar{\lambda}_t^m - \lambda_t^m) \quad (51) \quad \{\text{eq:KoppaPrime}\}$$

and an upper-case version of the equation in (42):

$$\begin{aligned} \hat{X}_t(\mu_t) &= \log \left( \frac{1 - \hat{Q}_t(\mu_t)}{\hat{Q}_t(\mu_t)} \right) \\ &= \log \left( 1/\hat{Q}_t(\mu_t) - 1 \right) \end{aligned} \quad (52) \quad \{\text{eq:Chi}\}$$

with corresponding derivative

$$\hat{X}_t^\mu = \left( \frac{-\hat{Q}_t^\mu / \hat{Q}_t^2}{1/\hat{Q}_t - 1} \right) \quad (53)$$

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

$$\hat{\lambda}_t = \bar{\lambda}_t - \overbrace{\left( \frac{1}{1 + \exp(\hat{X}_t)} \right)}^{=\hat{Q}_t} \blacktriangle h_{t \rightarrow \underline{\kappa}_t}(\mathbb{C}_t^T)^{1/(1-\rho)} \quad (54)$$

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

$$\begin{aligned} \hat{v}_t &= u(\hat{\lambda}_t) \\ \hat{v}_t^m &= u^c(\hat{\lambda}_t) \hat{\lambda}_t^m \\ \hat{v}_t^{mm} &= u^{cc}(\hat{\lambda}_t) (\hat{\lambda}_t^m)^2 + u^c(\hat{\lambda}_t) \hat{\lambda}_t^{mm}. \end{aligned} \quad (55)$$

Although a linear interpolation that matches the level of  $\lambda$  at the gridpoints is simple, a Hermite interpolation that matches both the level and the derivative of the  $\bar{\lambda}_t$  function at the gridpoints has the considerable virtue that the  $\bar{v}_t$  derived from it numerically satisfies the envelope theorem at each of the gridpoints for which the problem has been solved.

If we use the double-derivative calculated above to produce a higher-order Hermite polynomial, our approximation will also match marginal propensity to consume at the gridpoints; this would guarantee that the consumption function generated from the value function would match both the level of consumption 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 within the grid.

### 5.15 Refinement: A Tighter Upper Bound

Carroll (2023b) derives an upper limit  $\bar{\kappa}_t$  for the MPC as  $m_t$  approaches its lower bound. Using this fact plus the strict concavity of the consumption function yields the proposition that

$$c_t(\underline{m}_t + \blacktriangle m_t) < \bar{\kappa}_t \blacktriangle m_t. \quad (56)$$

The solution method described above does not guarantee that approximated consumption will respect this constraint between gridpoints, and a failure to respect the constraint can occasionally cause computational problems in solving or simulating the model. Here, we describe a method for constructing an approximation that always satisfies the constraint.

Defining  $m_t^\#$  as the ‘cusp’ point where the two upper bounds intersect:

$$\begin{aligned} \left( \blacktriangle m_t^\# + \blacktriangle h_{t \rightarrow} \right) \underline{\kappa}_t &= \bar{\kappa}_t \blacktriangle m_t^\# \\ \blacktriangle m_t^\# &= \frac{\underline{\kappa}_t \blacktriangle h_{t \rightarrow}}{(1 - \underline{\kappa}_t) \bar{\kappa}_t} \\ m_t^\# &= \frac{\underline{\kappa}_t h_{t \rightarrow} - \underline{h}_{t \rightarrow}}{(1 - \underline{\kappa}_t) \bar{\kappa}_t}, \end{aligned}$$

we want to construct a consumption function for  $m_t \in (\underline{m}_t, m_t^\#]$  that respects the tighter upper bound:

$$\begin{aligned} \blacktriangle m_t \underline{\kappa}_t &< c_t(\underline{m}_t + \blacktriangle m_t) < \bar{\kappa}_t \blacktriangle m_t \\ \blacktriangle m_t (\bar{\kappa}_t - \underline{\kappa}_t) &> \bar{\kappa}_t \blacktriangle m_t - c_t(\underline{m}_t + \blacktriangle m_t) &> 0 \\ 1 &> \left( \frac{\bar{\kappa}_t \blacktriangle m_t - c_t(\underline{m}_t + \blacktriangle m_t)}{\blacktriangle m_t (\bar{\kappa}_t - \underline{\kappa}_t)} \right) &> 0. \end{aligned}$$

Again defining  $\mu_t = \log \blacktriangle m_t$ , the object in the middle of the inequality is

$$\begin{aligned} \bar{q}_t(\mu_t) &\equiv \frac{\bar{\kappa}_t - c_t(\underline{m}_t + e^{\mu_t}) e^{-\mu_t}}{\bar{\kappa}_t - \underline{\kappa}_t} \\ \bar{q}_t^\mu(\mu_t) &= \frac{c_t(\underline{m}_t + e^{\mu_t}) e^{-\mu_t} - \kappa_t^m(\underline{m}_t + e^{\mu_t})}{\bar{\kappa}_t - \underline{\kappa}_t}. \end{aligned}$$

As  $m_t$  approaches  $-\underline{m}_t$ ,  $\bar{q}_t(\mu_t)$  converges to zero, while as  $m_t$  approaches  $+\infty$ ,  $\bar{q}_t(\mu_t)$  approaches 1.

As before, we can derive an approximated consumption function; call it  $\check{c}_t$ . This function will clearly do a better job approximating the consumption function for low values of  $m_t$  while the previous approximation will perform better for high values of  $m_t$ .

For middling values of  $m$  it is not clear which of these functions will perform better. However, an alternative is available which performs well. Define the highest gridpoint below  $m_t^\#$  as  $\tilde{m}_t^\#$  and the lowest gridpoint above  $m_t^\#$  as  $\hat{m}_t^\#$ . Then there will be a unique

interpolating polynomial that matches the level and slope of the consumption function at these two points. Call this function  $\tilde{c}_t(m)$ .

Using indicator functions that are zero everywhere except for specified intervals,

$$\begin{aligned}\mathbf{1}_{\text{Lo}}(m) &= 1 \text{ if } m \leq \tilde{m}_t^\# \\ \mathbf{1}_{\text{Mid}}(m) &= 1 \text{ if } \tilde{m}_t^\# < m < \hat{m}_t^\# \\ \mathbf{1}_{\text{Hi}}(m) &= 1 \text{ if } \hat{m}_t^\# \leq m\end{aligned}$$

we can define a well-behaved approximating consumption function

$$\dot{c}_t = \mathbf{1}_{\text{Lo}} \dot{\tilde{c}}_t + \mathbf{1}_{\text{Mid}} \dot{\tilde{c}}_t + \mathbf{1}_{\text{Hi}} \dot{\tilde{c}}_t. \quad (57)$$

This just says that, for each interval, we use the approximation that is most appropriate. The function is continuous and once-differentiable everywhere, and is therefore well behaved for computational purposes.

We now construct an upper-bound value function implied for a consumer whose spending behavior is consistent with the refined upper-bound consumption rule.

For  $m_t \geq m_t^\#$ , this consumption rule is the same as before, so the constructed upper-bound value function is also the same. However, for values  $m_t < m_t^\#$  matters are slightly more complicated.

Start with the fact that at the cusp point,

$$\begin{aligned}\bar{v}_t(m_t^\#) &= u(\bar{c}_t(m_t^\#)) \mathbb{C}_t^T \\ &= u(\blacktriangle m_t^\# \bar{\kappa}_t) \mathbb{C}_t^T.\end{aligned}$$

But for *all*  $m_t$ ,

$$\bar{v}_t(m) = u(\bar{c}_t(m)) + v_{t \rightarrow}^-(m - \bar{c}_t(m)),$$

and we assume that for the consumer below the cusp point consumption is given by  $\bar{\kappa} \blacktriangle m_t$  so for  $m_t < m_t^\#$

$$\bar{v}_t(m) = u(\bar{\kappa}_t \blacktriangle m) + v_{t \rightarrow}^-((1 - \bar{\kappa}_t) \blacktriangle m),$$

which is easy to compute because  $v_{t \rightarrow}^-(a_t) = \beta \bar{v}_{t+1}(a_t \mathcal{R} + 1)$  where  $\bar{v}_t$  is as defined above because a consumer who ends the current period with assets exceeding the lower bound will not expect to be constrained next period. (Recall again that we are merely constructing an object that is guaranteed to be an *upper bound* for the value that the ‘realist’ consumer will experience.) At the gridpoints defined by the solution of the consumption problem can then construct

$$\bar{\lambda}_t(m) = ((1 - \rho) \bar{v}_t(m))^{1/(1-\rho)}$$

and its derivatives which yields the appropriate vector for constructing  $\check{X}$  and  $\check{Q}$ . The rest of the procedure is analogous to that performed for the consumption rule and is thus omitted for brevity.

## 5.16 Extension: A Stochastic Interest Factor

Thus far we have assumed that the interest factor is constant at  $R$ . Extending the previous derivations to allow for a perfectly forecastable time-varying interest factor  $R_t$  would be trivial. Allowing for a stochastic interest factor is less trivial.

The easiest case is where the interest factor is i.i.d.,

$$\log \mathbf{R}_{t+n} \sim \mathcal{N}(r + \phi - \sigma_{\mathbf{r}}^2/2, \sigma_{\mathbf{r}}^2) \quad \forall n > 0 \quad (58) \quad \{\text{eq:distRisky}\}$$

where  $\phi$  is the risk premium and the  $\sigma_{\mathbf{r}}^2/2$  adjustment to the mean log return guarantees that an increase in  $\sigma_{\mathbf{r}}^2$  constitutes a mean-preserving spread in the level of the return.

This case is reasonably straightforward because Merton (1969) and Samuelson (1969) showed that for a consumer without labor income (or with perfectly forecastable labor income) the consumption function is linear, with an infinite-horizon MPC<sup>13</sup>

$$\kappa = 1 - \left( \beta \mathbb{E}_{\leftarrow t}[\mathbf{R}_{t+1}^{1-\rho}] \right)^{1/\rho} \quad (59) \quad \{\text{eq:MPCExact}\}$$

and in this case the previous analysis applies once we substitute this MPC for the one that characterizes the perfect foresight problem without rate-of-return risk.

The more realistic case where the interest factor has some serial correlation is more complex. We consider the simplest case that captures the main features of empirical interest rate dynamics: An AR(1) process. Thus the specification is

$$\mathbf{r}_{t+1} - \mathbf{r} = (\mathbf{r}_t - \mathbf{r})\gamma + \epsilon_{t+1} \quad (60)$$

where  $\mathbf{r}$  is the long-run mean log interest factor,  $0 < \gamma < 1$  is the AR(1) serial correlation coefficient, and  $\epsilon_{t+1}$  is the stochastic shock.

The consumer's problem in this case now has two state variables,  $m_t$  and  $\mathbf{r}_t$ , and is described by

$$\begin{aligned} v_t(m_t, \mathbf{r}_t) &= \max_{c_t} u(c_t) + \mathbb{E}_{\leftarrow t}[\beta_{t+1} v_{t+1}(m_{t+1}, \mathbf{r}_{t+1})] \\ \text{s.t.} \\ a_t &= m_t - c_t \\ \mathbf{r}_{t+1} - \mathbf{r} &= (\mathbf{r}_t - \mathbf{r})\gamma + \epsilon_{t+1} \\ \mathbf{R}_{t+1} &= \exp(\mathbf{r}_{t+1}) \\ m_{t+1} &= \underbrace{(\mathbf{R}_{t+1}/\mathcal{G}_{t+1})}_{\equiv \mathcal{R}_{t+1}} a_t + \boldsymbol{\theta}_{t+1}. \end{aligned}$$

We approximate the AR(1) process by a Markov transition matrix using standard techniques. The stochastic interest factor is allowed to take on 11 values centered around the steady-state value  $\mathbf{r}$ . Given this Markov transition matrix, *conditional* on the Markov AR(1) state the consumption functions for the ‘optimist’ and the ‘pessimist’ will still be linear, with identical MPC’s that are computed numerically. Given these MPC’s, the (conditional) realist’s consumption function can be computed for each Markov state, and the converged consumption rules constitute the solution contingent on the dynamics of the stochastic interest rate process.

---

<sup>13</sup>See **CRRA-RateRisk** for a derivation.

In principle, this refinement should be combined with the previous one; further exposition of this combination is omitted here because no new insights spring from the combination of the two techniques.

### 5.17 Imposing ‘Artificial’ Borrowing Constraints

Optimization problems often come with additional constraints that must be satisfied. Particularly common is an ‘artificial’ liquidity constraint that prevents the consumer’s net worth from falling below some value, often zero.<sup>14</sup> The problem then becomes

$$\begin{aligned} v_{t-1}(m_{t-1}) &= \max_{c_{t-1}} u(c_{t-1}) + \mathbb{E}_{t-1}[\beta v_{\rightarrow(T)}(m_t)] \\ \text{s.t.} \\ a_{t-1} &= m_{t-1} - c_{t-1} \\ m_t &= \mathcal{R}_t a_{t-1} + \theta_t \\ a_{t-1} &\geq 0. \end{aligned}$$

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

$$c_{t-1}^{-\rho} = v_{(t-1)\rightarrow}^a(m_{t-1} - c_{t-1}) \tag{61} \quad \{\text{eq:cUnc}\}$$

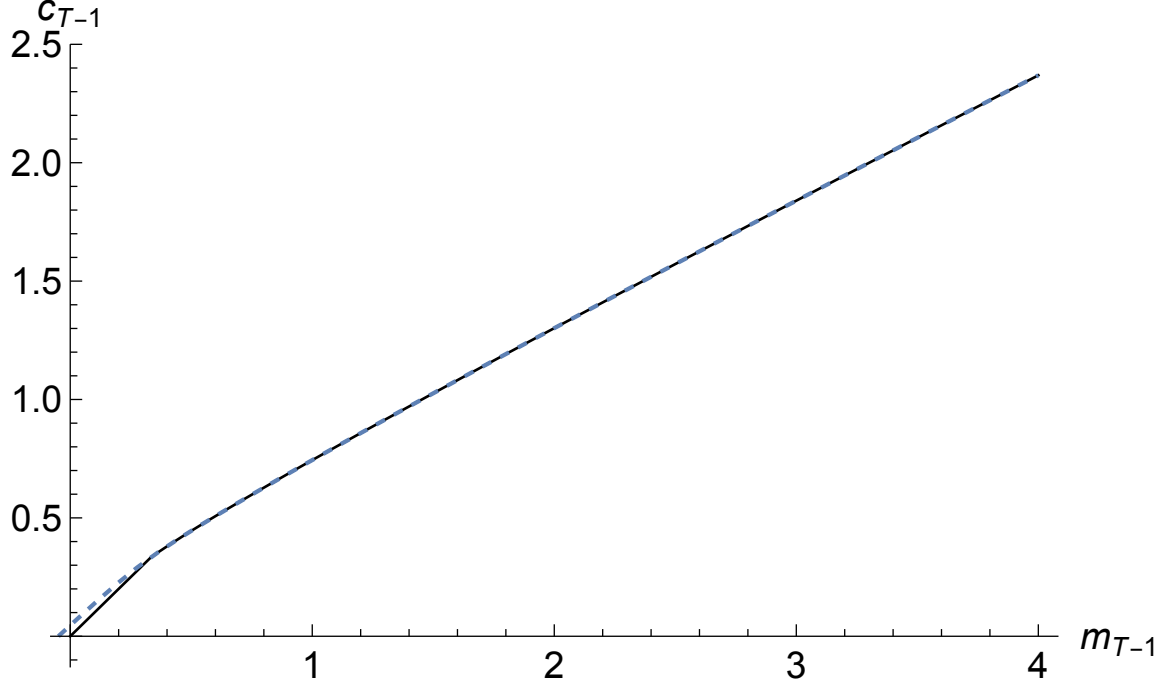
is greater than  $m_{t-1}$ . Call  $\hat{c}_{t-1}^*$  the approximated function returning the level of  $c_{t-1}$  that satisfies (61). Then the approximated constrained optimal consumption function will be

$$\hat{c}_{t-1}(m_{t-1}) = \min[m_{t-1}, \hat{c}_{t-1}^*(m_{t-1})]. \tag{62} \quad \{\text{eq:LiqCons}\}$$

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 ceases to bind. Fortunately, this is easy to calculate. We know that when the constraint is binding the consumer is saving nothing, which yields marginal value of  $v_{(t-1)\rightarrow}^a(0)$ . Further, 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  $\epsilon$  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

---

<sup>14</sup>The 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.



**Figure 19** Constrained (solid) and Unconstrained (dashed) Consumption

{fig:cVScCon}

the level of  $m_{t-1}$  at which the constraint stops binding is:<sup>15</sup>

$$\begin{aligned} u^c(m_{t-1}) &= v_{(t-1) \rightarrow}^a(0) \\ m_{t-1} &= (v_{(t-1) \rightarrow}^a(0))^{(-1/\rho)} \\ &= c_{(t-1) \rightarrow}(0). \end{aligned}$$

The constrained problem is solved in section “Artificial Borrowing Constraint” of the notebook, where the variable `constrained` is set to be a boolean type object. If the value of `constrained` is true, then the constraint is binding and their consumption behavior is computed to match (62). The resulting consumption rule is shown in Figure 19. For comparison purposes, the approximate consumption rule from Figure 19 is reproduced here as the solid line; this is accomplished by setting the boolean value of `constrained` to false.

The presence of the liquidity constraint requires three changes to the procedures outlined above:

1. We redefine  $\underline{h}_{t \rightarrow}$ , which now is the PDV of receiving  $\theta_{t+1} = \underline{\theta}$  next period and  $\theta_{t+n} = 0 \forall n > 1$  – that is, the pessimist believes he will receive nothing beyond period  $t + 1$
2. We augment the end-of-period `aVec` with zero and with a point with a small

---

<sup>15</sup>The logic here repeats an insight from Deaton (1991).

positive value so that the generated  $\mathbf{mVec}$  will the binding point  $m^\#$  and a point just above it (so that we can better capture the curvature around that point)

3. We redefine the optimal consumption rule as in equation (62). This ensures that the liquidity-constrained ‘realist’ will consume more than the redefined ‘pessimist,’ so that we will have  $\varphi$  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.

## 6 Recursion

{sec: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 (28) and (18):

$$\begin{aligned} v_{t \rightarrow}^a(a_t) &= \mathbb{E}_{\leftarrow t}[\beta \mathcal{RG}_{t+1}^{-\rho} u^c(c_{t+1}(\mathcal{R}_{t+1}a_t + \boldsymbol{\theta}_{t+1}))] \\ u^c(c_t) &= v_{t \rightarrow}^a(m_t - c_t). \end{aligned}$$

Assuming that the problem has been solved up to period  $t+1$  (and thus assuming that we have an approximated  $\check{c}_{t+1}(m_{t+1})$ ), 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. (Here, we use the original, rather than the “refined,” method for constructing consumption functions; the generalization of the algorithm below to use the refined method presents no difficulties.)

1. For the grid of values  $a_{t,i}$  in  $\mathbf{aVec\_eee}$ , numerically calculate the values of  $c_t^-(a_{t,i})$  and  $c_t^a(a_{t,i})$ ,

$$\begin{aligned} c_{t,i}^- &= (v_{t \rightarrow}^a(a_{t,i}))^{-1/\rho}, \\ &= (\beta \mathbb{E}_{\leftarrow t} [\mathcal{RG}_{t+1}^{-\rho} (\check{c}_{t+1}(\mathcal{R}_{t+1}a_{t,i} + \boldsymbol{\theta}_{t+1}))^{-\rho}])^{-1/\rho}, \\ c_{t,i}^a &= -(1/\rho) (v_{t \rightarrow}^a(a_{t,i}))^{-1-1/\rho} v_{t \rightarrow}^{aa}(a_{t,i}), \end{aligned} \tag{63} \quad \{\text{eq:vEndeq}\}$$

generating vectors of values  $\mathbf{c}_t$  and  $\mathbf{c}_t^a$ .

2. Construct a corresponding vector of values of  $\mathbf{m}_t = \mathbf{c}_t + \mathbf{a}_t$ ; similarly construct a corresponding list of MPC’s  $\boldsymbol{\kappa}_t$  using equation (48).
3. Construct a corresponding vector  $\boldsymbol{\mu}_t$ , the levels and first derivatives of  $\boldsymbol{\varphi}_t$ , and the levels and first derivatives of  $\boldsymbol{\chi}_t$ .

4. Construct an interpolating approximation  $\hat{\lambda}_t$  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  $\mathbf{v}_t$ , the levels and first derivatives of  $\mathbf{Q}_t$ , and the levels and first derivatives of  $\hat{\mathbf{X}}_t$ ; and construct an interpolating approximation function  $\hat{X}_t$  that matches those points.

With  $\hat{\lambda}_t$  in hand, our approximate consumption function is computed directly from the appropriate substitutions in (43) and related equations. 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 an item for constructing  $\hat{v}_t^a(m_t)$ . This is because with  $\hat{c}_t(m_t)$  in hand, we simply *define*  $\hat{v}_t^a(m_t) = u^c(\hat{c}_t(m_t))$  so there is no need to construct interpolating approximations - the function arises ‘free’ (or nearly so) from our constructed  $\hat{c}_t(m_t)$  via the usual envelope result (cf. (13)).

## 7 Multiple Control Variables

We now consider how to solve problems with multiple control variables. (To reduce notational complexity, in this section we set  $\mathcal{G}_t = 1 \forall t$ ; the time- or age-varying growth factor will return when we consider life cycle problems below).

### 7.1 Theory

The new control variable that the consumer can now choose is captured by the Greek character called ‘sigma,’ which represents the share  $\varsigma$  of their disposable assets to invest in the risky asset (conventionally interpreted as the stock market). Designating the return factor for the risky asset as  $\mathbf{R}$  and the share of the portfolio invested in  $\mathbf{R}$  as  $\varsigma$ , the realized portfolio rate of return  $\mathfrak{R}$  as a function of the share  $\varsigma$  is:

$$\mathfrak{R}(\varsigma) = R + (\mathbf{R} - R)\varsigma. \quad (64)$$

If we imagine the portfolio share decision as being made simultaneously with the  $c_t$  decision, a traditional way of writing the problem is (substituting the budget constraint):

$$v_t(m_t) = \max_{\{c_t, \varsigma\}} u(c_t) + \mathbb{E}_t[\beta v_{t+1}((m_t - c_t)(R + (\mathbf{R} - R)\varsigma) + \theta_{t+1})] \quad (65)$$

where I have deliberately omitted the period-designating subscripts for  $\varsigma$  and the return factors to highlight the point that, once the consumption and  $\varsigma$  decisions have been made, it makes no difference to this equation whether we suppose that the risky return factor  $\mathbf{R}$  is revealed a nanosecond before the end of period  $t$  or a nanosecond after the beginning of  $t + 1$ .

But as a notational choice, there is good reason to designate the realization as happening in  $t + 1$ . A standard way of motivating stochastic returns and wages is to attribute



them to “productivity shocks” and to assume that the productivity shock associated with a date is the one that affects the production function for that date.

## 7.2 Stages Within a Period

Solving simultaneously for the two variables  $\varsigma$  and  $c$  can be computationally challenging. Fortunately, there is a simple solution: Break the problem into two ‘stages.’<sup>16</sup>

As demonstrated in (65), the mathematical solution for the optimal portfolio share is the same whether we conceive the shocks as occurring at the end of  $t$  or the beginning of  $t + 1$ . But our tripartite dissection of the problem above into  $\{t_{\leftarrow}, t, t_{\rightarrow}\}$  steps was motivated partly by a desire to invent a notation that can allow for construction of standalone ‘stages’ whose only connection to the subsequent problem is through the continuation-value function.

To illustrate this point, consider the standalone problem of an ‘investor’ whose continuation value function  $v_{[\varsigma]\rightarrow}$  depends only on the amount of liquid assets  $\ell_{\rightarrow}$  with which they end up after the realization of the stochastic  $\mathbf{R}$  return. The expected value that the investor will obtain from any combination of initial  $\ell_{\leftarrow}$  and  $\varsigma$  is the expectation of the continuation value function over the liquid assets that result from the portfolio choice:

$$v_{[\varsigma]}(\ell_{\leftarrow}) = \max_{\varsigma} \mathbb{E} [v_{[\varsigma]\rightarrow} (\mathfrak{R}(\varsigma)\ell_{\leftarrow})] \quad (66)$$

where we have omitted any designator like  $t$  for the period in which this problem is solved because, with the continuation value function defined already as  $v_{t_{\rightarrow}}(\ell_{\rightarrow})$ , the problem is self-contained – it need make no reference to events later or earlier than its own moment, defined as the interval within which risky returns are realized. The solution to this problem will yield an optimal  $\varsigma$  decision rule  $\hat{\varsigma}(\ell_{\leftarrow})$ . Finally, we can specify the value of an investor ‘arriving’ with  $\ell_{\leftarrow}$  as the expected value that will be obtained when the investor invests optimally, generating the (optimally) stochastic portfolio return factor  $\hat{\mathfrak{R}}(\ell_{\leftarrow}) = \mathfrak{R}(\hat{\varsigma}(\ell_{\leftarrow}))$ :

$$v_{\leftarrow}(\ell_{\leftarrow}) = \mathbb{E}[v_{[\varsigma]\rightarrow}(\overbrace{\hat{\mathfrak{R}}(\ell_{\leftarrow})\ell_{\leftarrow}}^{\ell_{\rightarrow}})]. \quad (67)$$

The reward for all our painful notational investment is that it is now clear that *exactly the same code* for solving the portfolio share problem can be used in two distinct models: One in which the  $\mathbf{R}$  shocks are realized just after the beginning of  $t$  and one in which they are realized just before the end. In the first case, the ‘incoming’ amount of liquid assets  $\ell_{\leftarrow}$  corresponds to the capital  $k_t$  with which the agent enters the period while the exiting amount of assets corresponds to  $b_t$  (before-labor-income resources) as in (10) (with the substitution of the  $\hat{\mathfrak{R}}$  for  $\mathcal{R}$ ). The second case is only a tiny bit trickier: Our assumption that the return shocks happen at the end of  $t$  means that the problem needs to be altered slightly to bring the steps involving the realization of risky returns fully into period  $t$ ; the variable with which the agent ends the period is now  $b_t$  and to avoid confusion with the prior model in which we assumed  $k_{t+1} = a_t$  we will now define

---

<sup>16</sup>cite mnw and ael papers.

$\kappa_{t+1} = b_t$ . The continuation value function now becomes

$$v_{t \rightarrow}(b_t) = \beta v_{\leftarrow(t+1)}(\kappa_{t+1}) \quad (68)$$

while the dynamic budget constraint for  $m$  changes to

$$m_t = \kappa_t + \theta_t \quad (69)$$

and the problem in the decision step is now

$$v_t(m_t) = \max_c u(c) + \mathbb{E}_t[v_{t \rightarrow}(m_t - c)] \quad (70)$$

while value in the arrival step is now

$$v_{\leftarrow t}(\kappa_t) = \mathbb{E}_{\leftarrow t}[v_t(m_t)] \quad (71)$$

which, *mutatis mutandis*, is the same as in (10).

The upshot is that all we need to do is change some of the transition equations and we can use the same code (both for the  $\varsigma$ -stage and the  $c$ -stage) to solve the problem with either assumption about the timing of portfolio choice. There is even an obvious notation for the two problems:  $v_{\leftarrow t[\varsigma c]}$  can be the arrival (beginning-of-period) value function for the version where the portfolio share is chosen at the beginning of the period, and  $v_{\leftarrow t[c\varsigma]}$  is initial value for the the problem where the share choice is at the end.

While the investor's problem cannot be solved using the endogenous gridpoints method,<sup>17</sup> we can solve it numerically for the optimal  $\varsigma$  at a vector of  $\mathbf{a}$  (`aVec` in the code) and then construct an approximated optimal portfolio share function  $\hat{\varsigma}(\mathbf{a})$  as the interpolating function among the members of the  $\{\mathbf{a}, \boldsymbol{\varsigma}\}$  mapping. Having done this, we can now calculate a vector of values that correspond to `aVec`

$$\mathbf{v}_{\rightarrow} = v_{t \rightarrow}(\mathbf{a}, \hat{\varsigma}(\mathbf{a})), \quad (72) \quad \{\text{eq:vShrEnd}\}$$

which can be evaluated at the same vector of points  $\mathbf{a}$  at which  $\boldsymbol{\varsigma}$  was calculated, generating the vector  $\mathbf{v}_{[\varsigma] \rightarrow}$  whose inverse can again be approximated as above (along with approximations of marginal value, marginal marginal value, etc).

The beauty of this procedure is that, with the approximation to  $v_{\text{arvlt}[\varsigma]}^a(a)$  in hand, we can construct our approximation to the consumption function using *exactly the same EGM procedure* that we used in solving the problem *without* a portfolio choice (see (32)):

$$\mathbf{c} \equiv (v_{t \rightarrow}^a(\mathbf{a}))^{-1/\rho}, \quad (73) \quad \{\text{eq:cVecPort}\}$$

which, following a procedure identical to that in the EGM subsection 5.8, yields an approximated consumption function  $\hat{c}_t(m)$ . Thus, again, we can construct the consumption function at nearly zero cost (once we have calculated  $\mathbf{v}^a$ ).

---

<sup>17</sup>Because  $v_{t[\varsigma] \rightarrow}^{\varsigma}$  is not invertible with respect to  $\varsigma$ , see [references to MNW and AEL's work].

### 7.3 Application

{subsec:MCApplca}

In specifying the stochastic process for  $\mathbf{R}_{t+1}$ , we follow the common practice of assuming that returns are lognormally distributed,  $\log \mathbf{R} \sim \mathcal{N}(\phi + \mathbf{r} - \sigma_{\mathbf{r}}^2/2, \sigma_{\mathbf{r}}^2)$  where  $\phi$  is the equity premium over the thin returns  $\mathbf{r}$  available on the riskless asset.<sup>18</sup>

As with labor income uncertainty, it is necessary to discretize the rate-of-return risk in order to have a problem that is soluble in a reasonable amount of time. We follow the same procedure as for labor income uncertainty, generating a set of  $n_{\mathbf{r}}$  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  $\mathfrak{R}_{i,j}$  (where dependence on  $i$  is allowed to permit the possibility of nonzero correlation between the return on the risky asset and the  $\boldsymbol{\theta}$  shock to labor income (for example, in recessions the stock market falls and labor income also declines)).

The direct expressions for the derivatives of  $v_{t \rightarrow}$  are

$$\begin{aligned} v_{t \rightarrow}^a(a_t, \varsigma_t) &= \beta \left( \frac{1}{n_{\mathbf{r}} n_{\boldsymbol{\theta}}} \right) \sum_{i=1}^{n_{\boldsymbol{\theta}}} \sum_{j=1}^{n_{\mathbf{r}}} \mathfrak{R}_{i,j} (c_{t+1}(\mathfrak{R}_{i,j} a_t + \boldsymbol{\theta}_i))^{-\rho} \\ v_{t \rightarrow}^{\varsigma}(a_t, \varsigma_t) &= \beta \left( \frac{1}{n_{\mathbf{r}} n_{\boldsymbol{\theta}}} \right) \sum_{i=1}^{n_{\boldsymbol{\theta}}} \sum_{j=1}^{n_{\mathbf{r}}} (\mathbf{R}_{i,j} - \mathbf{R}) (c_{t+1}(\mathfrak{R}_{i,j} a_t + \boldsymbol{\theta}_i))^{-\rho}. \end{aligned} \quad (74)$$

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  $n_{\mathbf{r}} \times n_{\boldsymbol{\theta}}$  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 (inverses of the) derivatives of the  $v_{t \rightarrow}$  function.

Details of the construction of the interpolating approximation are given below; assume for the moment that we have the approximations  $\hat{v}_{t \rightarrow}^a$  and  $\hat{v}_{t \rightarrow}^{\varsigma}$  in hand and we want to proceed. As noted above in the discussion of (65), nonlinear equation solvers can find the solution to a set of simultaneous equations. Thus we could ask one to solve

$$\begin{aligned} c_t^{-\rho} &= \hat{v}_{t \rightarrow}^a(m_t - c_t, \varsigma_t) \\ 0 &= \hat{v}_{t \rightarrow}^{\varsigma}(m_t - c_t, \varsigma_t) \end{aligned} \quad (75) \quad \{\text{eq:FOCwrtw}\}$$

simultaneously for  $c$  and  $\varsigma$  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

$$\begin{aligned} \tilde{v}_{t \rightarrow}(a_t) &= \max_{\varsigma_t} v_{t \rightarrow}(a_t, \varsigma_t) \\ &\text{s.t.} \\ 0 &\leq \varsigma_t \leq 1 \end{aligned}$$

<sup>18</sup>This guarantees that  $\mathbb{E}[\mathbf{R}] = \boldsymbol{\varphi}$  is invariant to the choice of  $\sigma_{\boldsymbol{\phi}}^2$ ; see `LogELogNorm`.

where the tilde over  $\tilde{v}(a)$  indicates that this is the  $v$  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  $\tilde{\tilde{v}}_t^a(a_t)$ .<sup>19</sup> With this function in hand, we can use the first order condition from the single-control problem

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

to solve for the optimal level of consumption as a function of  $m_t$  using the endogenous gridpoints method described above. Thus we have transformed the multidimensional optimization problem into a sequence of two simple optimization problems.

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.

## 7.4 Implementation

Following the discussion from section 7.1, to provide a numerical solution to the problem with multiple control variables, we must define expressions that capture the expected marginal value of end-of-period assets with respect to the level of assets and the share invested in risky assets. This is addressed in “Multiple Control Variables.”

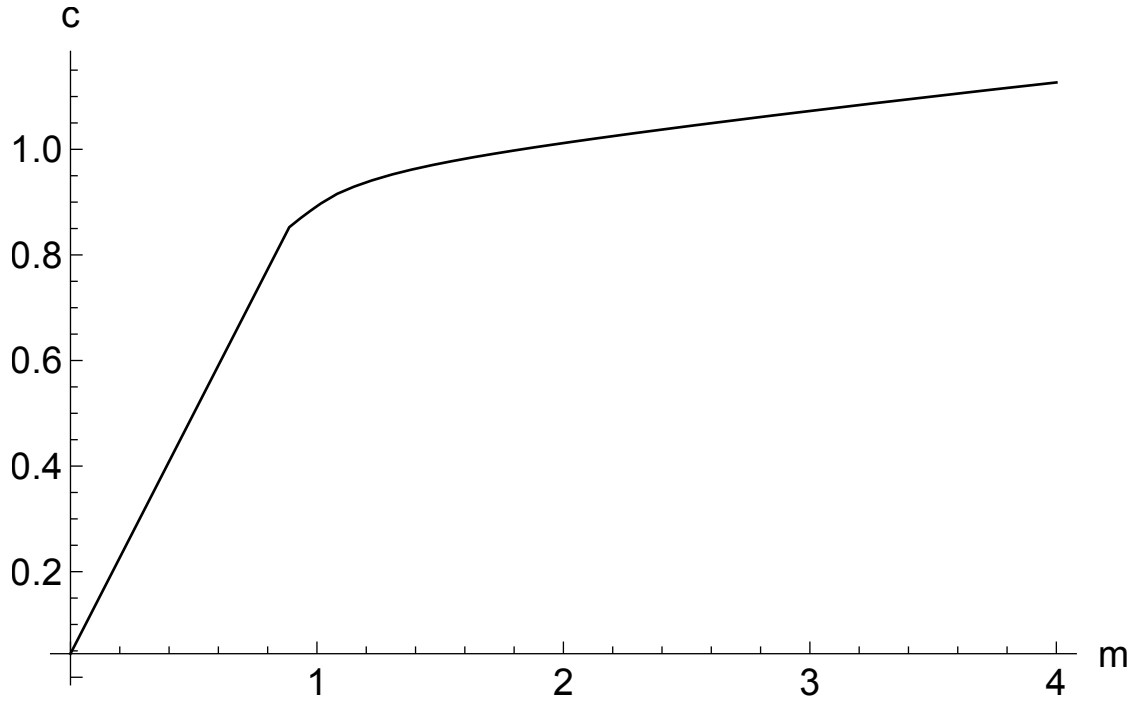
## 7.5 Results With Multiple Controls

Figure 20 plots the  $t - 1$  consumption function generated by the program; qualitatively it does not look much different from the consumption functions generated by the program without portfolio choice.

But Figure 21 which plots the optimal portfolio share as a function of the level of assets, 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 - see Carroll (2002)) prediction reflects the nature of the risks the consumer faces. Those consumers who are poor in measured financial wealth will likely derive a high proportion of future consumption from their labor income. Since by assumption labor income risk is

<sup>19</sup>A faster solution could be obtained by, for each element in `aVec`, computing  $v_{t \rightarrow}^\varsigma(m_t - c_t, \varsigma)$  of a grid of values of  $\varsigma$ , 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.

{subsec:results-with



**Figure 20**  $c(m_1)$  With Portfolio Choice

{fig:PlotctMultCont}

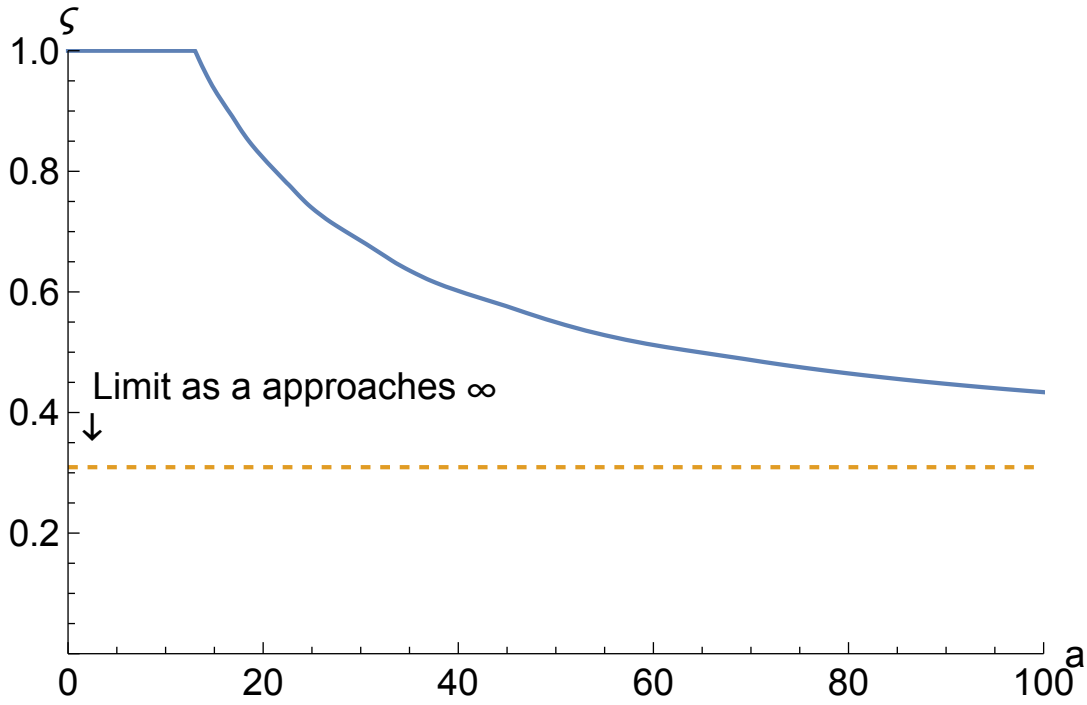
uncorrelated with rate-of-return risk, the covariance between their future consumption and future stock returns is relatively low. By contrast, persons with relatively large wealth will be paying for a large proportion of future consumption out of that wealth, and hence if they invest too much of it 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

{sec:the-infinite-hor}

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 (2023b)), 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 14. 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.

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.



**Figure 21** Portfolio Share in Risky Assets in First Period  $\varsigma(a)$

{fig:PlotRiskyShare}

## 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 (2023b) proves will exist in problems of this kind **under certain conditions**: The  $\hat{m}$  such that

$$\mathbb{E}_t[m_{t+1}/m_t] = 1 \text{ if } m_t = \hat{m} \quad (76) \quad \{\text{eq:mTrgNrmet}\}$$

where the 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  $\hat{m}$ . So for our problem, a solution is declared to have converged if the following criterion is met:  $|\hat{m}_{t+1} - \hat{m}_t| < \epsilon$ , where  $\epsilon$  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.

## 9 Structural Estimation

{sec:structural-estim}

This section describes how to use the methods developed above to structurally estimate a life-cycle consumption model, following closely the work of [Cagetti \(2003\)](#).<sup>20</sup> 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.

### 9.1 Life Cycle Model

{subsec:life-cycle-m}

Realistic calibration of a life cycle model needs to take into account a few things that we omitted from the bare-bones model described above. For example, the whole point of the life cycle model is that life is finite, so we need to include a realistic treatment of life expectancy; this is done easily enough, by assuming that utility accrues only if you live, so effectively the rising mortality rate with age is treated as an extra reason for discounting the future. Similarly, we may want to capture the demographic evolution of the household (e.g., arrival and departure of kids). A common way to handle that, too, is by modifying the discount factor (arrival of a kid might increase the total utility of the household by, say, 0.2, so if the ‘pure’ rate of time preference were 1.0 the ‘household-size-adjusted’ discount factor might be 1.2. We therefore modify the model presented above to allow age-varying discount factors that capture both mortality and family-size changes (we just adopt the factors used by [Cagetti \(2003\)](#) directly), with the probability of remaining alive between  $t$  and  $t + n$  captured by  $\mathcal{L}$  and with  $\hat{\beta}$  now reflecting all the age-varying discount factor adjustments (mortality, family-size, etc). Using  $\beth$  (the Hebrew cognate of  $\beta$ ) for the ‘pure’ time preference factor, the value function for the revised problem is

$$v_t(\mathbf{p}_t, \mathbf{m}_t) = \max_{\{\mathbf{c}\}_t^T} u(\mathbf{c}_t) + \mathbb{E}_{t \rightarrow} \left[ \sum_{n=1}^{T-t} \beth^n \mathcal{L}_t^{t+n} \hat{\beta}_t^{t+n} u(\mathbf{c}_{t+n}) \right] \quad (77) \quad \{\text{eq:lifecyclemax}\}$$

subject to the constraints

$$\begin{aligned} \mathbf{a}_t &= \mathbf{m}_t - \mathbf{c}_t \\ \mathbf{p}_{t+1} &= \mathcal{G}_{t+1} \mathbf{p}_t \Psi_{t+1} \\ \mathbf{y}_{t+1} &= \mathbf{p}_{t+1} \boldsymbol{\theta}_{t+1} \\ \mathbf{m}_{t+1} &= \mathbf{R} \mathbf{a}_t + \mathbf{y}_{t+1} \end{aligned}$$

where

$$\begin{aligned} \mathcal{L}_t^{t+n} &: \text{probability to } \mathcal{L} \text{ive until age } t+n \text{ given alive at age } t \\ \hat{\beta}_t^{t+n} &: \text{age-varying discount factor between ages } t \text{ and } t+n \\ \Psi_t &: \text{mean-one shock to permanent income} \\ \beth &: \text{time-invariant ‘pure’ discount factor} \end{aligned}$$

---

<sup>20</sup>Similar structural estimation exercises have been also performed by [Palumbo \(1999\)](#) and [Gourinchas and Parker \(2002\)](#).

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.

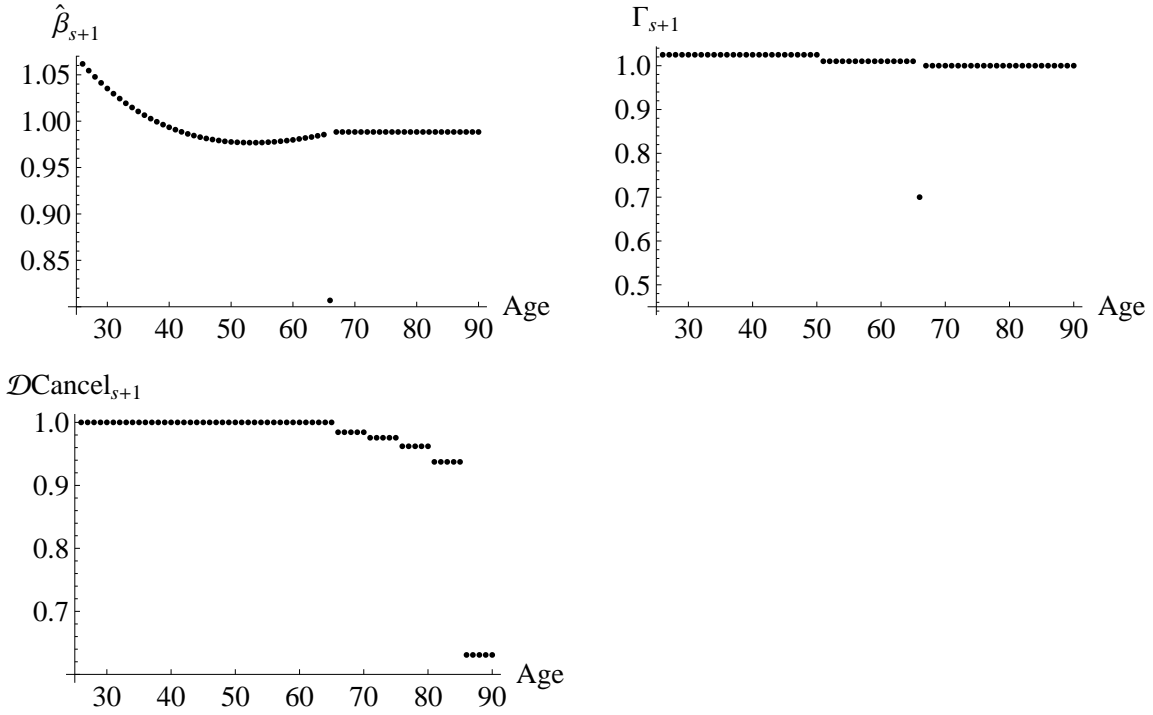
Transitory and permanent shocks are distributed as follows:

$$\Xi_s = \begin{cases} 0 & \text{with probability } \wp > 0 \\ \theta_s/\wp & \text{with probability } (1 - \wp), \text{ where } \log \theta_s \sim \mathcal{N}(-\sigma_\theta^2/2, \sigma_\theta^2) \end{cases} \quad (78)$$

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

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  $\sigma_\Psi = 0.1$ .<sup>21</sup> The income growth profile  $\mathcal{G}_t$  is from Carroll (1997) and the values of  $\mathcal{L}_t$  and  $\hat{\beta}_t$  are obtained from Cagetti (2003) (Figure 22).<sup>22</sup> The interest rate is assumed to equal 1.03. The model parameters are included in Table 1.



**Figure 22** Time Varying Parameters

{fig:TimeVaryingPa

<sup>21</sup>Note 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  $\sigma_\theta$  is almost certainly increased by measurement error.

<sup>22</sup>The 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).



**Table 1** Parameter Values

{table:StrEstParam

$\sigma_{\theta}$	0.1	Carroll (1992)
$\sigma_{\Psi}$	0.1	Carroll (1992)
$\wp$	0.005	Carroll (1992)
$\mathcal{G}_s$	figure 22	Carroll (1997)
$\hat{\beta}_s, \mathcal{L}_s$	figure 22	Cagetti (2003)
R	1.03	Cagetti (2003)

The structural estimation of the parameters  $\beth$  and  $\rho$  is carried out using the procedure specified in the following section, which is then implemented in the `StructEstimation.py` file. This file consists of two main components. The first section defines the objects required to execute the structural estimation procedure, while the second section executes the procedure and various optional experiments with their corresponding commands. The next section elaborates on the procedure and its accompanying code implementation in greater detail.

## 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  $\beth$  and  $\rho$  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.<sup>23</sup> 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 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 of the

---

<sup>23</sup>Cagetti (2003) matches wealth levels rather than wealth to income ratios. We believe it is more appropriate to match ratios both because the ratios are the state variable in the theory and because empirical moments for ratios of wealth to income are not influenced by the method used to remove the effects of inflation and productivity growth.

model which can be written in Bellman form as:

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

with the first order condition:

$$u^c(c_t) = \beta \mathcal{L}_{t+1} \hat{\beta}_{t+1} R \mathbb{E}_t [u^c(\Psi_{t+1} \mathcal{G}_{t+1} c_{t+1} (a_t \mathcal{R}_{t+1} + \theta_{t+1}))]. \quad (79)$$

{eq:FOCLifeCycle}

The first substantive task in this estimation procedure is to solve for the consumption functions at each age. We need to discretize the shock distribution and solve for the policy functions by backward induction using equation (79) following the procedure in sections 5 and ‘Recursion.’ 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.

Correspondingly, the first part of the `StructEstimation.py` file begins by defining the agent type by inheriting from the baseline agent type `IndShockConsumerType`, with the modification to include time-varying discount factors. Next, an instance of this “life-cycle” consumer is created for the estimation procedure. The number of periods for the life cycle of a given agent is set and, 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. 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.

To complete the creation of the consumer type needed for the simulation, a history of shocks is drawn for each agent across all periods by invoking the `make_shock_history` function. This involves discretizing the shock distribution for as many points as the number of agents we want to simulate and then randomly permuting this shock vector as many times as we need to simulate the model for. In this way, we obtain a time varying shock for each agent. 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.)

With the age-varying consumption functions derived from the life-cycle agent, we can proceed to generate simulated data and compute the corresponding medians. Estimating the model involves comparing these simulated medians with empirical medians,

measuring the model's success by calculating the difference between the two. However, before performing the necessary steps of solving and simulating the model to generate simulated moments, it's important to note a difficulty in producing the target moments using the available data.

Specifically, defining  $\xi$  as the set of parameters to be estimated (in the current case  $\xi = \{\rho, \mathfrak{N}\}$ ), we could search for the parameter values which solve

$$\min_{\xi} \sum_{\tau=1}^7 |\varsigma^{\tau} - \mathbf{s}^{\tau}(\xi)| \quad (80) \quad \{\text{eq:naivePowell}\}$$

where  $\varsigma^{\tau}$  and  $\mathbf{s}^{\tau}$  are respectively the empirical and simulated medians of the wealth to permanent income ratio for age group  $\tau$ . 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 (80) 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:

$$\min_{\xi} \sum_i^N \omega_i |\varsigma_i^{\tau} - \mathbf{s}^{\tau}(\xi)| \quad (81) \quad \{\text{eq:StructEstim}\}$$

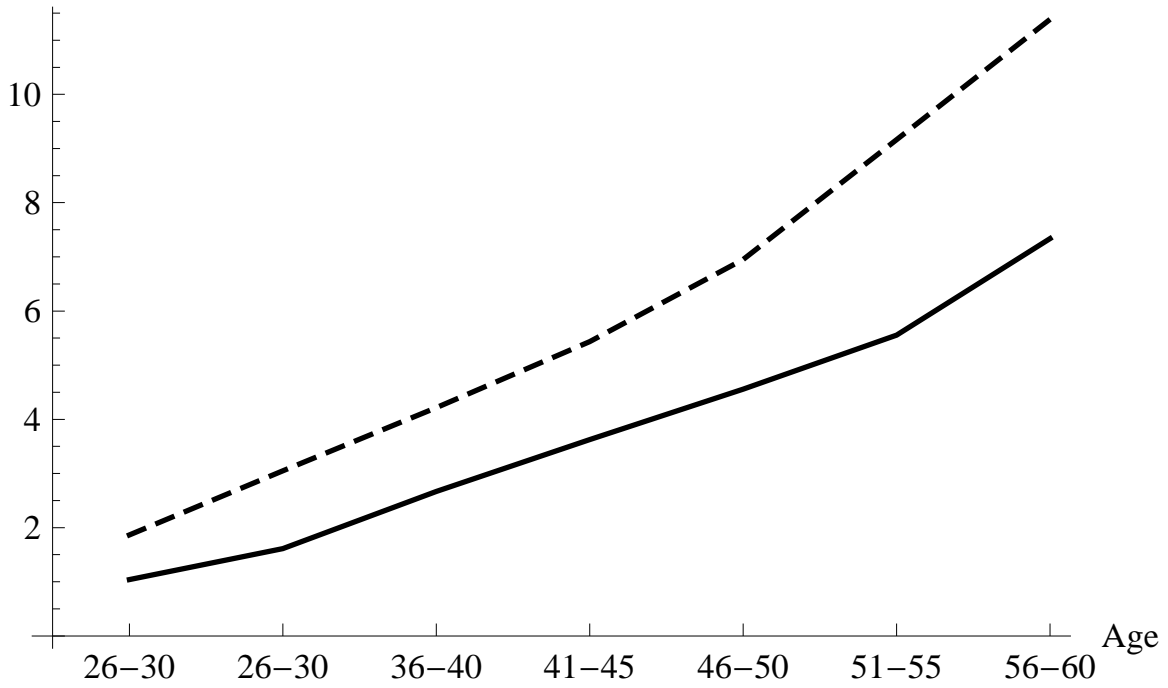
where  $\omega_i$  is the weight of household  $i$  in the entire population,<sup>24</sup> and  $\varsigma_i^{\tau}$  is the empirical wealth to permanent income ratio of household  $i$  whose head belongs to age group  $\tau$ .  $\omega_i$  is needed because unequal weight is assigned to each observation 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.

With this in mind, we turn our attention to the computation of the weighted median wealth target moments for each age cohort using this data from the 2004 Survey of Consumer Finances on household wealth. The objects necessary to accomplish this task are `weighted_median` and `get_targeted_moments`. The actual data are taken from several waves of the SCF and the medians and means for each age category are plotted in figure 23. More details on the SCF data are included in appendix A.

We now turn our attention to the the two key functions in this section of the code file. The first, `simulate_moments`, executes the solving (`solve`) and simulation (`simulation`) steps for the defined life-cycle agent. Subsequently, the function uses the agents' tracked levels of wealth based on their optimal consumption behavior to compute and store the simulated median wealth to income ratio for each age cohort. The second function,

---

<sup>24</sup>The Survey of Consumer Finances includes many more high-wealth households than exist in the population as a whole; therefore if one wants to produce population-representative statistics, one must be careful to weight each observation by the factor that reflects its "true" weight in the population.



**Figure 23** Wealth to Permanent Income Ratios from SCF (means (dashed) and medians (solid))

{fig:MeanMedianSCF}

`smmObjectiveFxn`, calls the `simulate_moments` function to create the objective function described in (81), which is necessary to perform the SMM estimation.

Thus, for a given pair of the parameters to be estimated, the single call to the function `smmObjectiveFxn` executes the following:

1. solves for the consumption functions for the life-cycle agent
2. simulates the data and computes the simulated medians
3. returns the value of equation (81)

We delegate the task of finding the coefficients that minimize the `smmObjectiveFxn` function to the `minimize_nelder_mead` function, which is defined elsewhere and called in the second part of this file. This task can be quite slow and rather problematic if the `smmObjectiveFxn` 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.

The final object defined in this first part of the `StructEstimation.py` file is `calculateStandardErrorsByBootstrap`. As the name suggests, the purpose of this function is to compute the standard errors by bootstrap.<sup>25</sup> This involves:

1. drawing new shocks for the simulation

---

<sup>25</sup>For a treatment of the advantages of the bootstrap see Horowitz (2001)

2. drawing a random sample (with replacement) of actual data from the SCF
3. obtaining new estimates for  $\rho$  and  $\gamma$

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

### 9.2.1 An Aside to Computing Sensitivity Measures

A common drawback in commonly used structural estimation procedures is a lack of transparency in its estimates. As Andrews et al. (2017) notes, a researcher employing such structural empirical methods may be interested in how alternative assumptions (such as misspecification or measurement bias in the data) would “change the moments of the data that the estimator uses as inputs, and how changes in these moments affect the estimates.” The authors provide a measure of sensitivity for given estimator that makes it easy to map the effects of different assumptions on the moments into predictable bias in the estimates for non-linear models.

In the language of Andrews et al. (2017), section 9 is aimed at providing an estimator  $\xi = \{\rho, \gamma\}$  that has some true value  $\xi_0$  by assumption. Under the assumption  $a_0$  of the researcher, the empirical targets computed from the SCF is measured accurately. These moments of the data are precisely what determine our estimate  $\hat{\xi}$ , which minimizes (81). Under alternative assumptions  $a$ , such that a given cohort is mismeasured in the survey, a different estimate is computed. Using the plug-in estimate provided by the authors, we can see quantitatively how our estimate changes under these alternative assumptions  $a$  which correspond to mismeasurement in the median wealth to income ratio for a given age cohort.

## 9.3 Results

The second part of the file `StructEstimation.py` defines a function `main` which produces our  $\rho$  and  $\gamma$  estimates with standard errors using 10,000 simulated agents by setting the positional arguments `estimate_model` and `compute_standard_errors` to `true`.<sup>26</sup> Results are reported in Table 2.<sup>27</sup>

The literature on consumption and saving behavior over the lifecycle in the presenece of labor income uncertainty<sup>28</sup> warns us to be careful in disentangling the effect of time preference and risk aversion when describing the optimal behavior of households in this setting. Since the precautionary saving motive dominates in the early stages of life, the coefficient of relative risk aversion (as well as expected labor income growth) has

---

<sup>26</sup>The procedure is: First we calculate the  $\rho$  and  $\gamma$  estimates as the minimizer of equation (81) using the actual SCF data. Then, we apply the **Bootstrap** function several times to obtain the standard error of our estimates.

<sup>27</sup>Differently from Cagetti (2003) who estimates a different set of parameters for college graduates, high school graduates and high school dropouts graduates, we perform the structural estimation on the full population.

<sup>28</sup>For example, see Gourinchas and Parker (2002) for an exposition of this.

**Table 2** Estimation Results

{tab:EstResults}

$\rho$	$\beta$
3.69	0.88
(0.047)	(0.002)

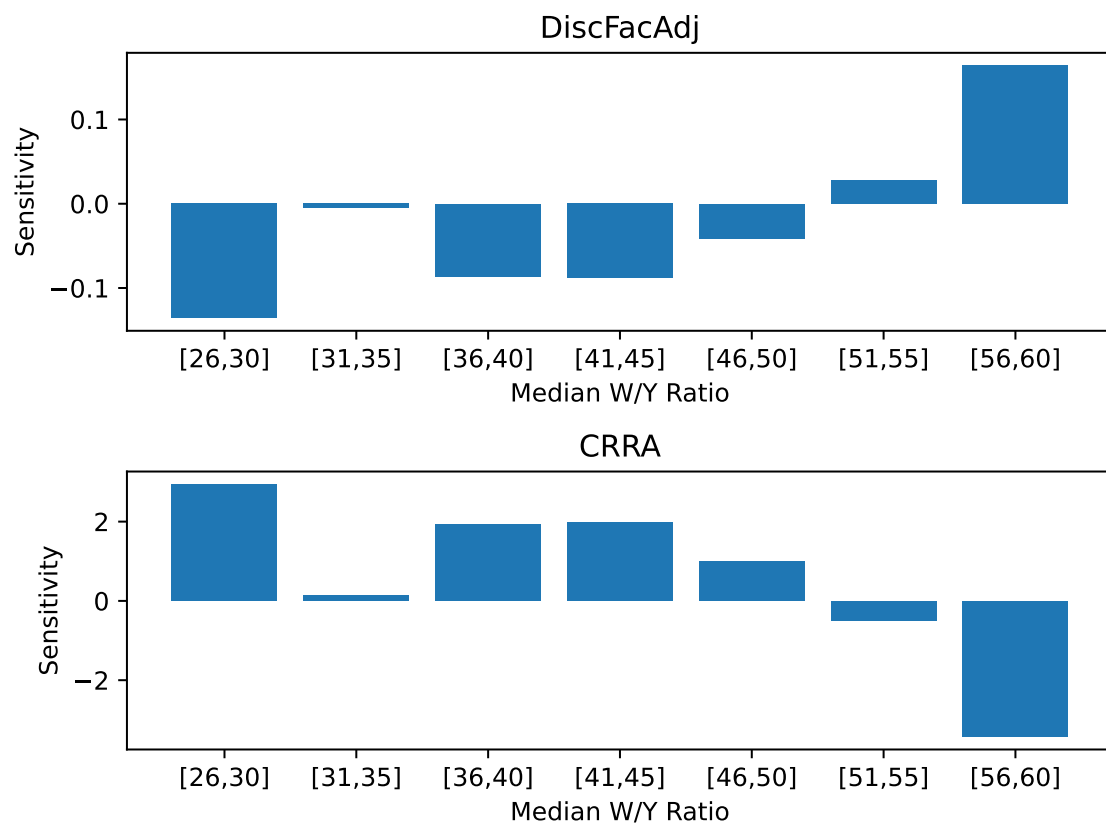
a larger effect on optimal consumption and saving behavior through their magnitude relative to the interest rate. Over time, life-cycle considerations (such as saving for retirement) become more important and the time preference factor plays a larger role in determining optimal behavior for this cohort.

Using the positional argument `compute_sensitivity`, Figure 24 provides a plot of the plug-in estimate of the sensitivity measure described in 9.2.1. As you can see from the figure the inverse relationship between  $\rho$  and  $\beta$  over the life-cycle is retained by the sensitivity measure. Specifically, under the alternative assumption that *a particular cohort is mismeasured in the SCF dataset*, we see that the y-axis suggests that our estimate of  $\rho$  and  $\beta$  change in a predictable way.

Suppose that there are not enough observations of the oldest cohort of households in the sample. Suppose further that the researcher predicts that adding more observations of these households to correct this mismeasurement would correspond to a higher median wealth to income ratio for this cohort. In this case, our estimate of the time preference factor should increase: the behavior of these older households is driven by their time preference, so a higher value of  $\beta$  is required to match the affected wealth to income targets under this alternative assumption. Since risk aversion is less important in explaining the behavior of this cohort, a lower value of  $\rho$  is required to match the affected empirical moments.

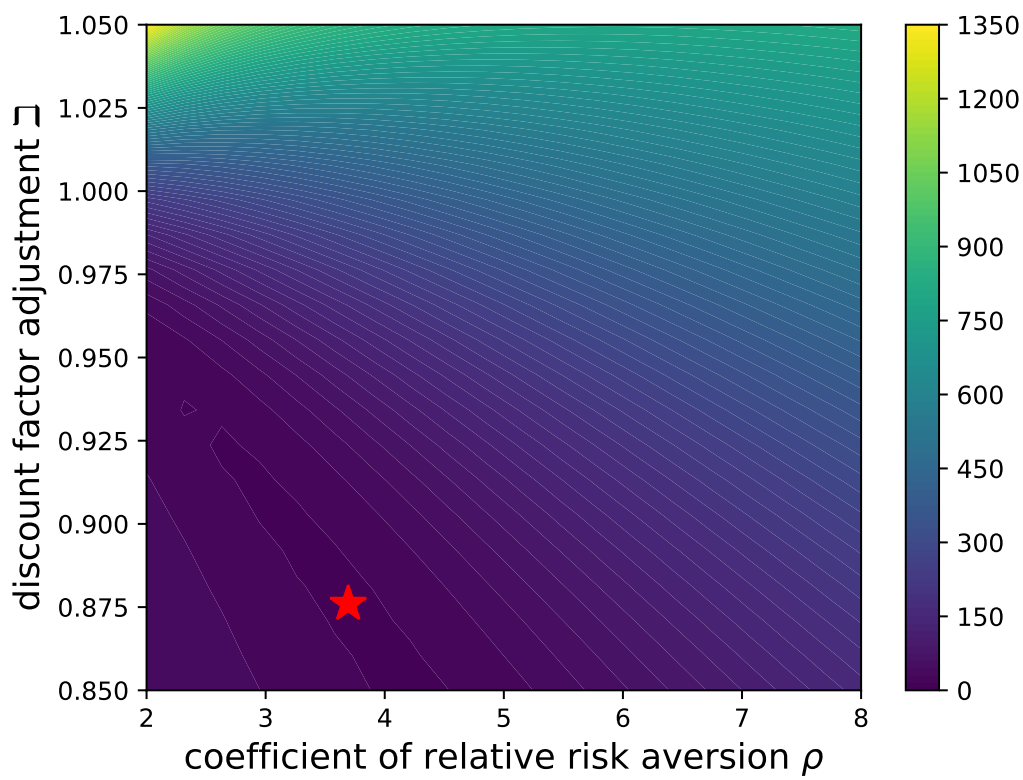
To recap, the sensitivity measure not only matches our intuition about the inverse relationship between  $\rho$  and  $\beta$  over the life-cycle, but provides a quantitative estimate of what would happen to our estimates of these parameters under the alternative assumption that the data is mismeasured in some way.

By setting the positional argument `make_contour_plot` to true, Figure 25 shows the contour plot of the `smmObjectiveFxn` function and the parameter estimates. The contour plot shows equally spaced isoquants of the `smmObjectiveFxn` function, i.e. the pairs of  $\rho$  and  $\beta$  which lead to the same deviations between simulated and empirical medians (equivalent values of equation (81)). Interestingly, 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.



**Figure 24** Sensitivity of Estimates  $\{\rho, \gamma\}$  regarding Alternative Mismeasurement Assumptions.

{fig:PlotSensitivity}



**Figure 25** Contour Plot (larger values are shown lighter) with  $\{\rho, \beta\}$  Estimates (red dot).

{fig:PlotContourMe



# Appendices

## A SCF Data

{app: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).<sup>29</sup> 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. 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.<sup>30</sup>

---

<sup>29</sup>Cagetti (2003) argues that younger households should be dropped since educational choice is not modeled. Also, he drops singles, since they include a large number of single mothers whose saving behavior is influenced by welfare.

<sup>30</sup>Please refer to the archive code for details of how these after-tax measures of  $P$  are constructed.

## References

- ANDREWS, I., M. GENTZKOW, AND J. M. SHAPIRO (2017): “Measuring the sensitivity of parameter estimates to estimation moments,” *The Quarterly Journal of Economics*, 132, 1553–1592.
- ATTANASIO, O., J. BANKS, C. MEGHIR, AND G. WEBER (1999): “Humps and Bumps in Lifetime Consumption,” *Journal of Business and Economic Statistics*, 17, 22–35.
- CAGETTI, M. (2003): “Wealth Accumulation Over the Life Cycle and Precautionary Savings,” *Journal of Business and Economic Statistics*, 21, 339–353.
- CARROLL, C. D. (1992): “The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence,” *Brookings Papers on Economic Activity*, 1992, 61–156, <https://www.econ2.jhu.edu/people/ccarroll/BufferStockBPEA.pdf>, <https://www.econ2.jhu.edu/people/ccarroll/BufferStockBPEA.pdf>.
- (1997): “Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis,” *Quarterly Journal of Economics*, CXII, 1–56.
- (2002): “Portfolios of the Rich,” in *Household Portfolios: Theory and Evidence*, Cambridge, MA: MIT Press, <https://www.econ2.jhu.edu/people/ccarroll/richportfolios.pdf>, <https://www.econ2.jhu.edu/people/ccarroll/richportfolios.pdf>.
- (2006): “The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems,” *Economics Letters*, 91, 312–320, 10.1016/j.econlet.2005.09.013, <https://www.econ2.jhu.edu/people/ccarroll/EndogenousGridpoints.pdf>.
- (2023a): “Solving Microeconomic Dynamic Stochastic Optimization Problems,” *Econ-ARK REMARK*, <https://llorracc.github.io/SolvingMicroDSOPs>.
- (2023b): “Theoretical Foundations of Buffer Stock Saving,” *Revise and Resubmit, Quantitative Economics*, <https://zenodo.org/badge/latestdoi/304124725>.
- CARROLL, C. D., AND M. S. KIMBALL (1996): “On the Concavity of the Consumption Function,” *Econometrica*, 64, 981–992, <https://www.econ2.jhu.edu/people/ccarroll/concavity.pdf>, <https://www.econ2.jhu.edu/people/ccarroll/concavity.pdf>.
- CARROLL, C. D., AND A. A. SAMWICK (1997): “The Nature of Precautionary Wealth,” *Journal of Monetary Economics*, 40, 41–71, <https://www.econ2.jhu.edu/people/ccarroll/papers/nature.pdf>.
- DEATON, A. S. (1991): “Saving and Liquidity Constraints,” *Econometrica*, 59, 1221–1248, <http://www.jstor.org/stable/2938366>, <https://www.jstor.org/stable/2938366>.

- (1992): *Understanding Consumption*, New York: Oxford University Press.
- GOURINCHAS, P.-O., AND J. PARKER (2002): “Consumption Over the Life Cycle,” *Econometrica*, 70, 47–89.
- DEN HAAN, W. J., AND A. MARCET (1990): “Solving the Stochastic Growth Model by Parameterizing Expectations,” *Journal of Business and Economic Statistics*, 8, 31–34, Available at <http://ideas.repec.org/a/bes/jnlbes/v8y1990i1p31-34.html>.
- HOROWITZ, J. L. (2001): “The Bootstrap,” in *Handbook of Econometrics* ed. by Heckman, J. J., and Leamer, E. Volume 5: Elsevier/North Holland.
- KOPECKY, K. A., AND R. M. SUEN (2010): “Finite State Markov-Chain Approximations To Highly Persistent Processes,” *Review of Economic Dynamics*, 13, 701–714, <http://www.karenkopecky.net/RouwenhorstPaper.pdf>, <http://www.karenkopecky.net/RouwenhorstPaper.pdf>.
- MERTON, R. C. (1969): “Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case,” *Review of Economics and Statistics*, 51, 247–257.
- PALUMBO, M. G. (1999): “Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle,” *Review of Economic Studies*, 66, 395–421, Available at <http://ideas.repec.org/a/bla/restud/v66y1999i2p395-421.html>.
- SAMUELSON, P. A. (1969): “Lifetime Portfolio Selection by Dynamic Stochastic Programming,” *Review of Economics and Statistics*, 51, 239–46.
- TANAKA, K., AND A. A. TODA (2013): “Discrete approximations of continuous distributions by maximum entropy,” *Economics letters*, 118, 445–450, [10.1016/j.econlet.2012.12.020](https://doi.org/10.1016/j.econlet.2012.12.020).
- ZELDES, S. P. (1989): “Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence,” *Quarterly Journal of Economics*, 104, 275–298.