Solution Methods for Microeconomic Dynamic Stochastic Optimization Problems

2024-03-27

Christopher D. Carroll[1]

Note: The GitHub repo SolvingMicroDSOPs repo 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 The MSM results at the end have have been superseded by tools in the EstimatingMicroDSOPs repo.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 *Mathematica*, Matlab, and Python software 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/econ-ark/SolvingMicroDSOPs/blob/master/SolvingMicroDSOPs-Slides.pdf
Web: https://econ-ark.github.io/SolvingMicroDSOPs
Code: https://github.com/econ-ark/SolvingMicroDSOPs/tree/master/Code
Archive: https://github.com/econ-ark/SolvingMicroDSOPs
(Contains LaTeX code for this document and software producing figures and results)

[1]Carroll: Department of Economics, Johns Hopkins University, Baltimore, MD, 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 Tokuoka for drafting the section on structural estimation, to Damiano Sandri for exceptionally insightful help in revising and updating the method of simulated moments estimation section, and to Weifeng Wu and Metin Uyanik for revising to be consistent with the 'method of moderation' and other improvements. All errors are my own. This document can be cited as Carroll (2023a) in the references.

Contents

1 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,[1] with detailed intuitive discussion of the various mathematical and computational techniques that, together, speed the solution by many orders of magnitude compared to "brute force" methods. The problem is solved with and without liquidity constraints, and the infinite horizon solution 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') 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 step in which the agent makes an optimal choice taking account of all of these elements. For the detailed analysis here, we will be careful to disarticulate everything that happens in the problem explicitly into separate steps 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 **b**ank account **b**alance:

$$\mathbf{b}_t = \mathbf{k}_t \mathsf{R}_t. \tag{1}$$ {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 θ_t:

$$\mathbf{y}_t = \mathbf{p}_t \theta_t \tag{2}$$ {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 $\mathbf{p}_t$).

The combination of bank balances $\mathbf{b}$ and income $\mathbf{y}$ define's the consumer's 'market

[1]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.

resources' (sometimes called 'cash-on-hand,' following Deaton (1992)):

$$\mathbf{m}_t = \mathbf{b}_t + \mathbf{y}_t, \tag{3}$$ {eq:mLvl}

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

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

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

Income evolves according to:

$$\begin{aligned} \mathbf{p}_{t+1} &= \mathcal{G}_{t+1}\mathbf{p}_t && \text{- permanent labor income dynamics} \\ \log\ \theta_{t+n} &\sim\ \mathcal{N}(-\sigma_\theta^2/2, \sigma_\theta^2) && \text{- lognormal transitory shocks } \forall\ n > 0. \end{aligned} \tag{5}$$ {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).[2] Finally, the utility function is of the Constant Relative Risk Aversion (CRRA), form, $\mathrm{u}(\bullet) = \bullet^{1-\rho}/(1-\rho)$.

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

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

$$\mathbf{k}_{t+1} = \mathbf{a}_t. \tag{7}$$ {eq:transition.state}

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). Here, 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.')

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

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

[2] 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).

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 $\mathrm{v}_T(m_T) = \mathrm{u}(m_T)$.[3] For our CRRA utility function, $\mathrm{u}(xy) = x^{1-\rho}\mathrm{u}(y)$, so (8) can be rewritten as

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

Now define a new optimization problem:

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

where the last equation is the normalized version of the transition equation for $\mathbf{m}_{t+1}$.[4] Then it is easy to see that for $t = T-1$,

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

and so on back to all earlier periods. Hence, if we solve the problem (10) which (when optimal consumption is being chosen) has only a single state variable m_t, 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. $\mathbf{c}_t(\mathbf{m}_t, \mathbf{p}_t) = \mathbf{p}_t\mathrm{c}_t(\mathbf{m}_t/\mathbf{p}_t)$ (or, for the value function, $\mathbf{v}_t(\mathbf{m}_t, \mathbf{p}_t) = \mathbf{p}_t^{1-\rho}\mathrm{v}_t(m_t)$). We have thus reduced the problem from two continuous state variables to one (and thereby enormously simplified its solution).

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

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

[4] Derivation:

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

4 The Usual Theory, and A Bit More Notation

{sec:usualtheory}

4.1 Steps

Generically, we want to think of the Bellman problem itself as having three steps:

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, 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, it is possible to calculate the consequences of the decision, taking as given the 'outgoing' state variables (e.g., a) – sometimes called 'post-state' variables.

The (implicit) default assumption is often to think of the step of the problem where the agent is solving a decision problem as defining the unique moment at which the problem is defined. This is what implicitly was done above, when (for example) in (10) we related current value v_t to the expectation of future value v_{t+1}.

When we want to refer to a specific step within period t we will do so by preceding it by an indicator character:

Step	Indicator	Usage	Explanation
Arrival	$\leftarrow$	$\mathrm{v}_{\leftarrow t}(k_t)$	value of entering t (before shocks)
Decision	(blank/none)	$\mathrm{v}_t(m_t)$	value of t-decision (after shocks)
Continuation	$\rightarrow$	$\mathrm{v}_{t\rightarrow}(a_t)$	value of outgoing state (after decision)

This notation allows us to capture the fact that the value of the consumer's circumstances can be computed at any of the three steps (and is a function of a different state variable at each step).

Note that there is no need to use these subscripts for the model's variables; while a variable like $a_t = m_t - c_t$ takes on its value in the transition from the Decision to the Continuation step, a will have only one unique value over the course of the period and therefore a notation like $a_{\leftarrow t}$ would be useless because the variable does not have a value until the continuation step is reached. Each variable in our problem has a unique value defined at some point during the period, so there is no ambiguity in referring to them with normal notation like a_t.

4.2 The Usual Theory

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

$$\begin{aligned} \mathrm{u}^c(c_t) &= \mathbb{E}_{t\rightarrow}[\beta\mathcal{R}_{t+1}\mathcal{G}_{t+1}^{1-\rho}\mathrm{v}^m_{t+1}(m_{t+1})] \\ &= \mathbb{E}_{t\rightarrow}[\beta\mathsf{R} \quad \mathcal{G}_{t+1}^{-\rho}\mathrm{v}^m_{t+1}(m_{t+1})] \end{aligned} \tag{11}$$

{eq:upceqEvtp1}

and because the Envelope theorem tells us that

$$\mathrm{v}_t^m(m_t) = \mathbb{E}_{\leftarrow t}[\beta \mathsf{R} \mathcal{G}_{t+1}^{-\rho} \mathrm{v}_{t+1}^m(m_{t+1})] \tag{12}$$ {eq:envelope}

we can substitute the LHS of (12) for the RHS of (11) to get

$$\mathrm{u}^c(c_t) = \mathrm{v}_t^m(m_t) \tag{13}$$ {eq:Envelope}

and rolling forward one period,

$$\mathrm{u}^c(c_{t+1}) = \mathrm{v}_{t+1}^m(a_t \mathcal{R}_{t+1} + \theta_{t+1}) \tag{14}$$ {eq:upctp1EqVpxtp

and substituting the LHS in equation (11) finally gives us the Euler equation for consumption:

$$\mathrm{u}^c(c_t) = \mathbb{E}_{t\rightarrow}[\beta \mathsf{R} \mathcal{G}_{t+1}^{-\rho} \mathrm{u}^c(c_{t+1})]. \tag{15}$$ {eq:cEuler}

For future reference, it may be useful here to write the distinct value functions at each step (using the transition equations from (10)):

$$\mathrm{v}_{\leftarrow t}(k_t) = \mathbb{E}_{\leftarrow t}[\mathrm{v}_t(\overbrace{k_t \mathcal{R}_t + \theta_t}^{=m_t})] \tag{16}$$ {eq:vBegStep}

$$\mathrm{v}_t(m_t) = \mathrm{u}(\mathrm{c}_t(m_t)) + \mathrm{v}_{t\rightarrow}(\overbrace{m_t - c_t}^{a_t}) \tag{17}$$ {eq:vMidStep}

and

$$\mathrm{v}_{t\rightarrow}(a_t) = \beta \mathrm{v}_{\leftarrow t+1}(\underbrace{k_{t+1}}_{=a_t}) \tag{18}$$ {eq:vEndtdefn}

where the last line illustrates the notation for addressing the beginning step of the successor period.

Putting all this together, from the perspective of the beginning of period $t+1$ we can write the 'arrival value' function and its first derivative as

$$\begin{aligned} \mathrm{v}_{\leftarrow t+1}(k_{t+1}) &= \mathbb{E}_{\leftarrow t+1}[\ \ \mathcal{G}_{t+1}^{1-\rho} \mathrm{v}_{t+1}(\mathcal{R}_{t+1} k_t + \theta_{t+1})] \\ \mathrm{v}_{\leftarrow t+1}^k(k_{t+1}) &= \mathbb{E}_{\leftarrow t+1}[\mathsf{R} \mathcal{G}_{t+1}^{-\rho} \mathrm{v}_{t+1}^m(m_{t+1})] \end{aligned} \tag{19}$$ {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**apital. Finally, note for future use that since $k_{t+1} = a_t$ and $\mathrm{v}_{t\rightarrow}(a_t) = \beta \mathrm{v}_{\leftarrow t+1}(k_{t+1})$, the first order condition (11) can now be rewritten compactly as

$$\mathrm{u}^c(c_t) = \mathrm{v}_{t\rightarrow}^a(m_t - c_t). \tag{20}$$ {eq:upEqbetaOp}

5 Solving the Next-to-Last Period

{sec:NextToLast}

For convenience assuming that $\mathcal{G} = 1$ so that the $\mathcal{G}$ terms disappear from the earlier derivations, the problem in the second-to-last period of life can now be expressed as

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

where $\mathbb{E}_{T-1_{\rightarrow}}$ indicates that the expectation is taken as of the end of period $T-1$.

Using (1) the fact that $\mathrm{v}_T = \mathrm{u}(c_T)$; (2) the definition of $\mathrm{u}(c_T)$; (3) the definition of the expectations operator, this becomes:

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

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

The maximization implicitly defines a function $\mathrm{c}_{T-1}(m_{T-1})$ that yields optimal consumption in period $T-1$ for any specific numerical level of resources like $m_{T-1} = 1.7$. But because there is no general analytical solution to this problem, for any given m_{T-1} we must use numerical computational tools to find the c_{T-1} that maximizes the expression. This is excruciatingly slow because for every potential c_{T-1} to be considered, 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 θ of a function $g(\theta)$ by calculating its value at set of n_θ points θ_i, each of which has an associated probability weight w_i:

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

(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 θ_i has an equal probability, equal to $1/n_\theta$).

We calculate such an n-point approximation as follows.

Define a set of points from $\sharp_0$ to $\sharp_{n_\theta}$ on the $[0, 1]$ interval as the elements of the set

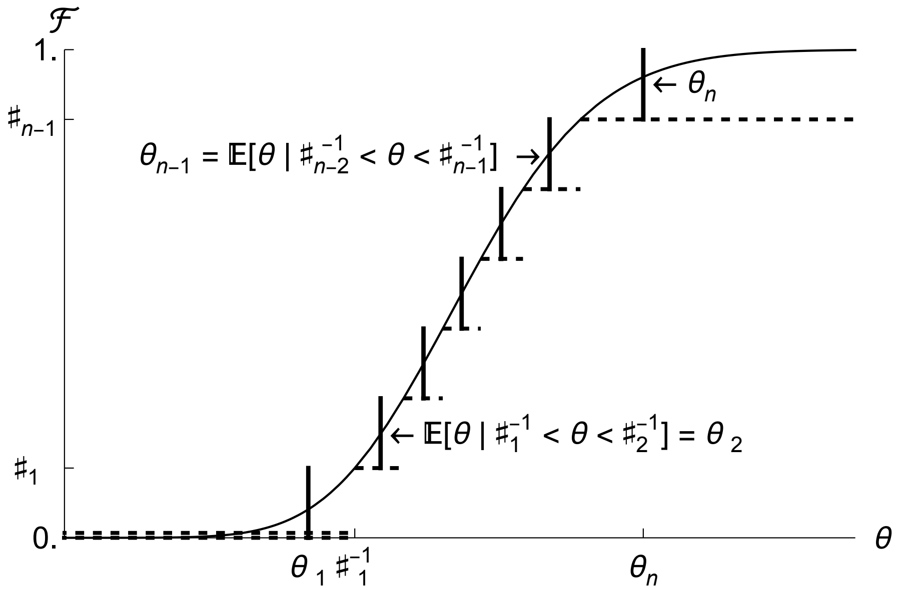

Figure 1 Equiprobable Discrete Approximation to Lognormal Distribution $\mathcal{F}$

{fig:discreteapprox}

$\natural = \{0, 1/n, 2/n, \ldots, 1\}$.[5] Call the inverse of the θ distribution $\mathcal{F}^{-1}$, and define the points $\natural_i^{-1} = \mathcal{F}^{-1}(\natural_i)$. Then the conditional mean of θ in each of the intervals numbered 1 to n is:

$$\theta_i \equiv \mathbb{E}[\theta|\natural_{i-1}^{-1} \leq \theta < \natural_i^{-1}] = \int_{\natural_{i-1}^{-1}}^{\natural_i^{-1}} \vartheta \; d\mathcal{F}\,(\vartheta), \tag{21}$$

and when the integral is evaluated numerically for each i the result is a set of values of θ 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_θ equiprobable values of θ_i which are used to approximate this distribution.[6]

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

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:

Substituting into our definition of $\mathrm{v}_{T-1_\rightarrow}$,

$$\mathrm{v}_{T-1_\rightarrow}(a_{T-1}) = \beta \left(\frac{1}{n_\theta}\right) \sum_{i=1}^{n_\theta} \frac{(\mathcal{R}_T a_T + \theta_i)^{1-\rho}}{1-\rho} \tag{22}$$

{eq:vDiscrete}

[5] These points define intervals that constitute a partition of the domain of $\mathcal{F}$.

[6] 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.

so we can rewrite the maximization problem that defines the middle stage of the period as

$$\mathrm{v}_{T-1}(m_{T-1}) = \max_{c_{T-1}} \left\{ \frac{c_{T-1}^{1-\rho}}{1-\rho} + \mathrm{v}_{\leftarrow T-1}(m_{T-1} - c_{T-1}) \right\}. \tag{23}$$ {eq:vEndTm1}

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

5.2 The Approximate Consumption and Value Functions

Given a particular value of m_{T-1}, a numerical maximization routine can now find the c_{T-1} that maximizes (23) in a reasonable amount of time.

The heart of the program responsible for computing an estimated consumption function begins in “Solving the Model by Value Function Maximization,” where a grid characterizing the possible values of market resources (m) is initialized (in the code, various m vectors have names beginning `mVec`), and for each of the `mVec` values, the consumption values c that solve the minimization problem equivalent to (23) are computed. We arbitrarily pick the first five integers as our five `mVec` gridpoints. (That is, `mVec_int`$=\{0., 1., 2., 3., 4.\}$.

5.3 An Interpolated Consumption Function

{subsec:LinInterp}

Given a set of points on a function (in this case, the consumption function $\mathrm{c}_{T-1}(m)$), we can create a piecewise linear ‘interpolating function’ (a ‘spline’) which when applied to an input m will yield the value of c that corresponds to a linear ‘connect-the-dots’ interpolation of the value of c from the points, creating a function that is an approximation of the function whose points have been sampled.

This is accomplished in “An Interpolated Consumption Function,” which defines an approximation to the consumption function $\grave{\mathrm{c}}_T(m_{T-1})$. That is, when called with an m_{T-1} that is equal to one of the points in mVec_int, $\grave{\mathrm{c}}_{T-1}$ returns the associated value of $\boldsymbol{c}_{\mathtt{T-1}}$, and when called with a value of m_{T-1} that is not exactly equal to one of the `mVec_int`, returns the value of c that reflects a linear interpolation between the $\boldsymbol{c}_{\mathtt{T-1}}$ points associated with the two `mVec_int` points immediately above and below m_{T-1}.

Figures 2 and 3 show plots of the constructed $\grave{\mathrm{c}}_{T-1}$ and $\grave{\mathrm{v}}_{T-1}$. While the $\grave{\mathrm{c}}_{T-1}$ function looks very smooth, the fact that the $\grave{\mathrm{v}}_{T-1}$ 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).[7]

5.4 Interpolating Expectations

Picewise linear ‘spline’ interpolation as described above works well for generating a good approximation to the true optimal consumption function. However, there is a

[7]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.

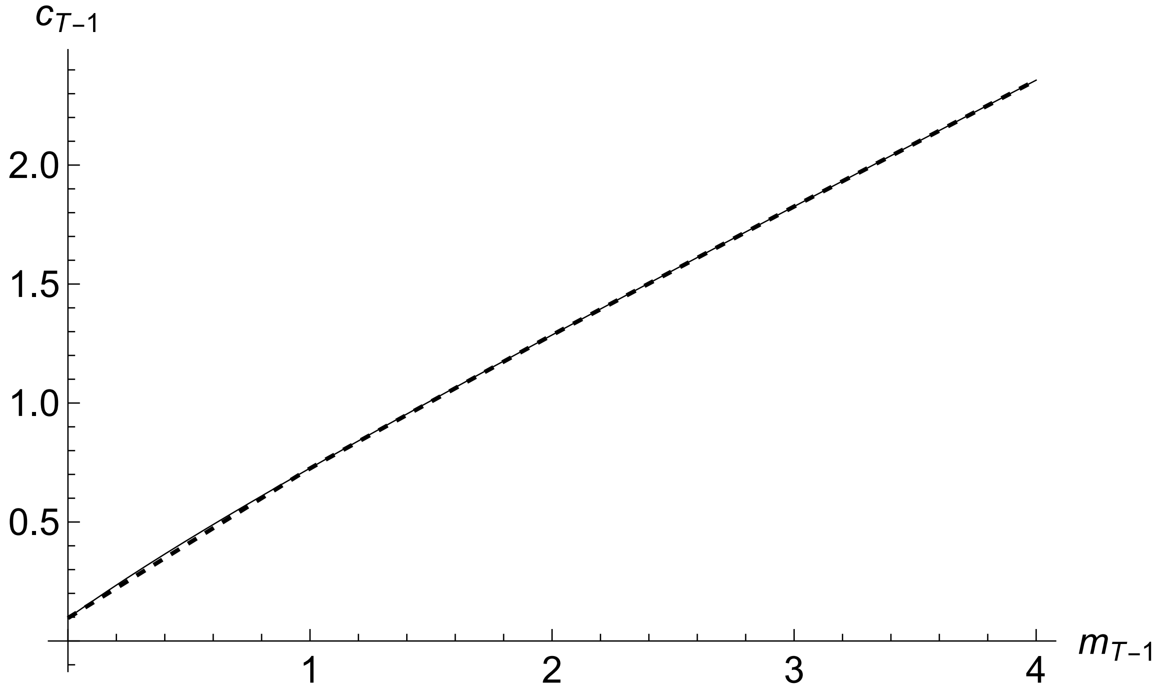

Figure 2 $\mathrm{c}_{T-1}(m_{T-1})$ (solid) versus $\grave{\mathrm{c}}_{T-1}(m_{T-1})$ (dashed)

{fig:PlotcTm1Simpl

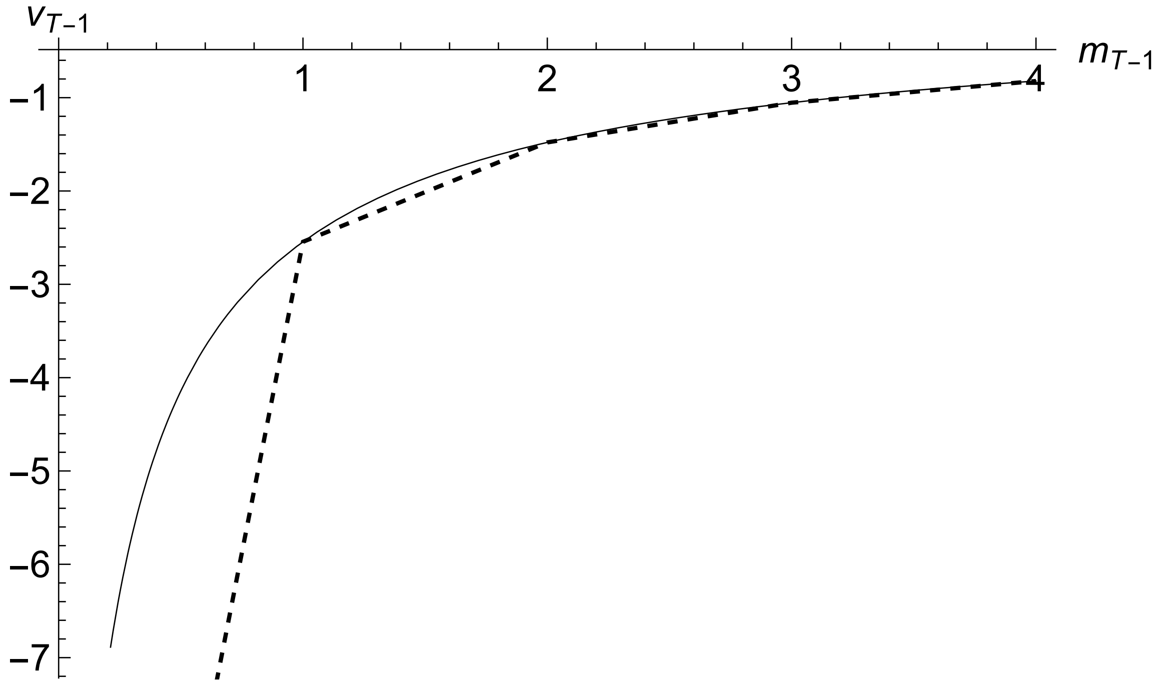

Figure 3 v_{T-1} (solid) versus $\grave{\mathrm{v}}_{T-1}(m_{T-1})$ (dashed)

{fig:PlotVTm1Simp

clear inefficiency in the program: Since it uses equation (23), for every value of m_{T-1} the program must calculate the utility consequences of various possible choices of c_{T-1} as it searches for the best choice.

For any given value of a_{T-1}, notice that there is a good chance that the program may end up calculating the corresponding v_T many times while maximizing utility from different m_{T-1}'s. For example, it is possible that the program will calculate the value of ending the period with $a_{T-1} = 0$ dozens of times. It would be much more efficient if the program could make the calculation of $\mathrm{v}_T(0)$ once and then merely recall the value when it is needed again.

Something like this can be achieved using 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 $T-1$, $\boldsymbol{a}_{\mathtt{T-1}}$ (`aVec` in the code). Next, construct $\boldsymbol{v}_{\mathtt{T-1}} = \mathrm{v}_{T-1}(\boldsymbol{a}_{\mathtt{T-1}})$ using equation (23); then construct an approximation $\grave{\mathrm{v}}_{T-1_\rightarrow}(a_{T-1})$ by passing the lists `aVec` and `vVec` as arguments. These lists contain the points of the $\boldsymbol{a}_{\mathtt{T-1}}$ and $\mathbf{v}_{\mathtt{T-1}}$ vectors, respectively.

As seen in the section "Interpolating Expectations," we are now interpolating for the function that reveals the expected value of *ending* the period with a given amount of assets.[8]

Figure 4 compares the true value function to the approximation produced 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 $m_{T-1} = 1$.

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

Nevertheless, the consumption rule obtained when the approximating $\grave{\mathrm{v}}_{T-1_\rightarrow}(a_{T-1})$ is used instead of $\mathrm{v}_{T-1_\rightarrow}(a_{T-1})$ is surprisingly bad, as shown in figure 5. For example, when m_{T-1} goes from 2 to 3, $\grave{c}_{T-1}$ goes from about 1 to about 2, yet when m_{T-1} goes from 3 to 4, $\grave{c}_{T-1}$ goes from about 2 to about 2.05. The function fails even to be strictly concave, which is distressing because Carroll and Kimball (1996) prove that the correct consumption function is strictly concave in a wide class of problems that includes this problem.

5.5 Value Function versus First Order Condition

{subsec:vVsuP}

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,

$$\mathbb{E}[-(c-\cancel{c})^2] < -(\mathbb{E}[c]-\cancel{c})^2 \tag{24}$$

(where $\cancel{c}$ 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

[8]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.

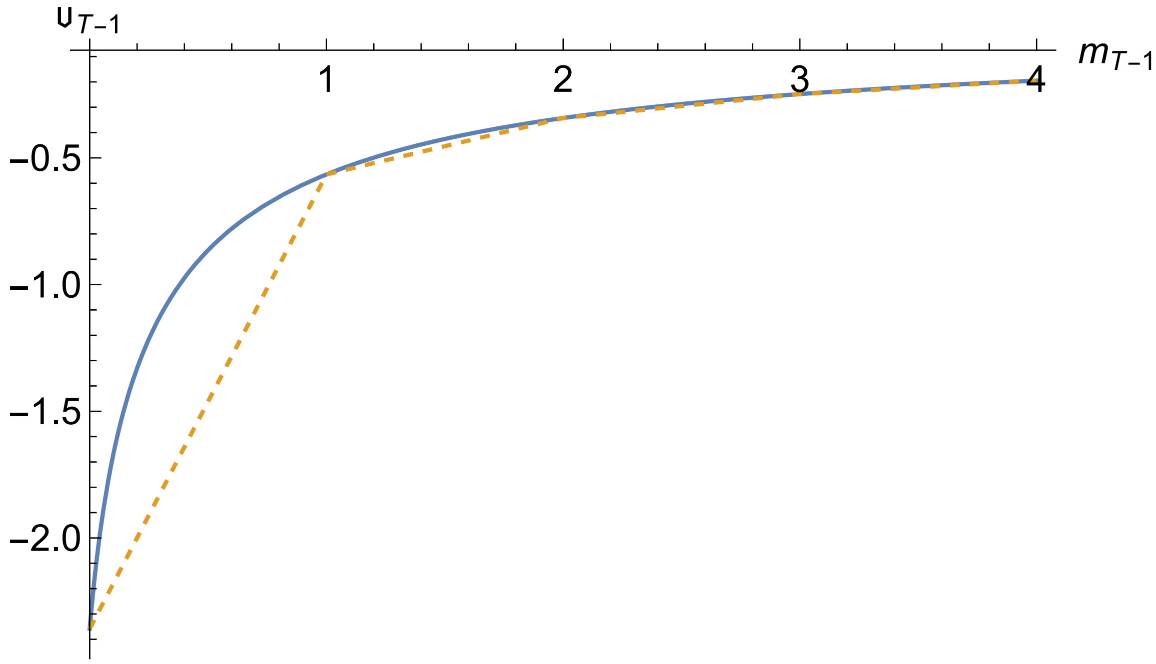

Figure 4 End-Of-Period Value $\mathrm{v}_{T-1_{\rightarrow}}(a_{T-1})$ (solid) versus $\grave{\mathrm{v}}_{T-1_{\rightarrow}}(a_{T-1})$ (dashed)

{fig:PlotOTm1RawV

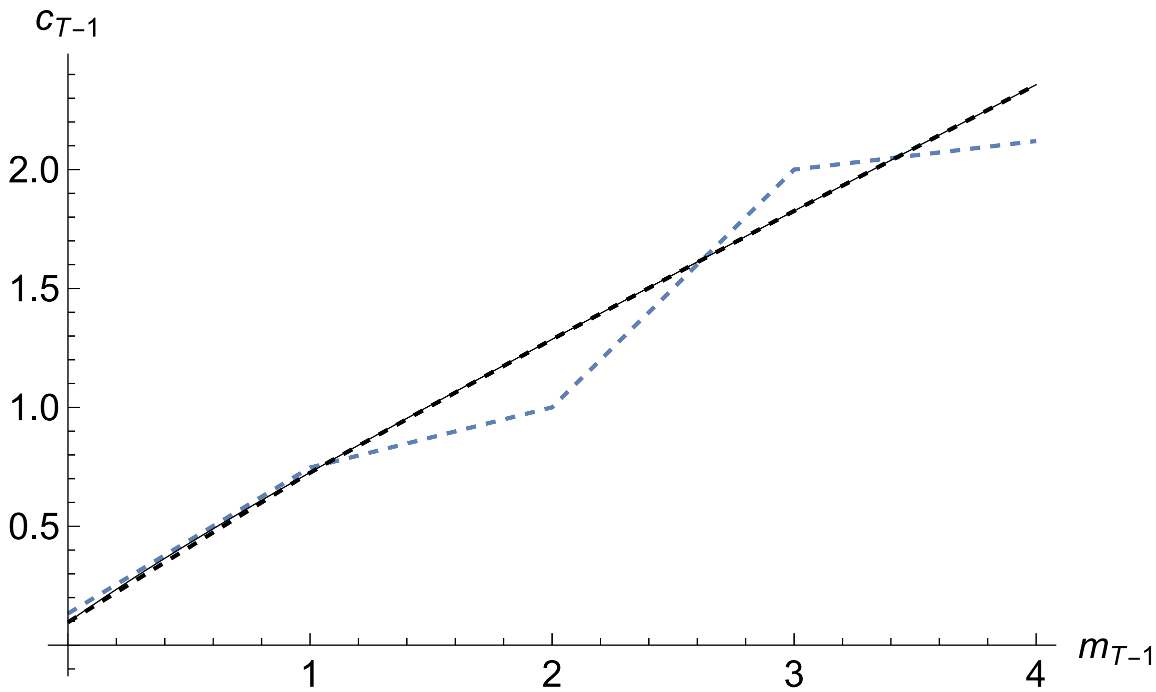

Figure 5 $\mathrm{c}_{T-1}(m_{T-1})$ (solid) versus $\grave{\mathrm{c}}_{T-1}(m_{T-1})$ (dashed)

{fig:PlotComparecT

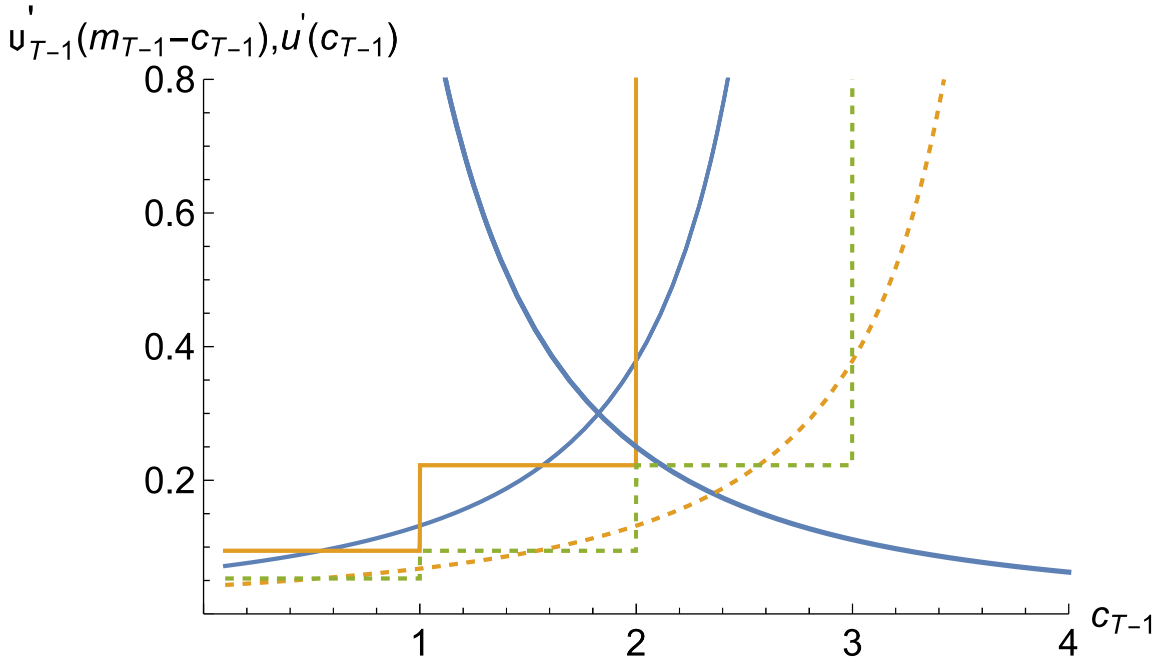

Figure 6 $\mathrm{u}^c(c)$ versus $\mathrm{v}^a_{T-1_{\rightarrow}}(3-c), \mathrm{v}^a_{T-1_{\rightarrow}}(4-c), \grave{\mathrm{v}}^a_{T-1_{\rightarrow}}(3-c), \grave{\mathrm{v}}^a_{T-1_{\rightarrow}}(4-c)$ {fig:PlotuPrimeVSC

unaffected by risk:

$$\mathbb{E}[-2(c-\text{¢})] = -2(\mathbb{E}[c]-\text{¢}). \tag{25}$$

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:

$$\begin{aligned} \mathrm{u}^c(c_{T-1}) &= \beta\,\mathbb{E}_{\rightarrow(T-1)}[\mathsf{R}\mathrm{u}^c(c_T)] \\ c_{T-1}^{-\rho} &= \mathsf{R}\beta\left(\frac{1}{n_\theta}\right)\sum_{i=1}^{n_\theta}\left(\mathsf{R}(m_{T-1}-c_{T-1})+\theta_i\right)^{-\rho}. \end{aligned} \tag{26}$$ {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 (26) 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 (26) 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 (26). The value of c_{T-1} for which the RHS and LHS of (26) 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 (26) under the assumption that $m_{T-1} = 4$, and the intersection of this curve with $\mathrm{u}^c(c_{T-1})$ 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 $\grave{\mathrm{v}}_{T-1_{\rightarrow}}(a_{T-1})$. Because we have constructed $\grave{\mathrm{v}}_{T-1_{\rightarrow}}$ as a linear interpolation, the slope of $\grave{\mathrm{v}}_{T-1_{\rightarrow}}(a_{T-1})$ between any two adjacent points $\{\boldsymbol{a}[i], , \boldsymbol{a}[i+1]\}$ 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 $\grave{\mathrm{v}}_{T-1_{\rightarrow}}(a_{T-1})$ follows a step function.

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

$$\mathrm{v}^a_{T-1_{\rightarrow}}(a_{T-1}) = \beta \mathsf{R} \left(\frac{1}{n_\theta}\right) \sum_{i=1}^{n_\theta} \left(\mathcal{R}_T a_{T-1} + \theta_i\right)^{-\rho} \tag{27}$$ {eq:vEndPrimeTm1

at the points in `aVec` yielding $\boldsymbol{v}^{\mathtt{a}}_{\mathtt{T-1}_{\rightarrow}}$ (the vector of expected end-of-period-$(T-1)$ marginal values of assets corresponding to `aVec`), and construct $\grave{\mathrm{v}}^a_{T-1_{\rightarrow}}(a_{T-1})$ 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).

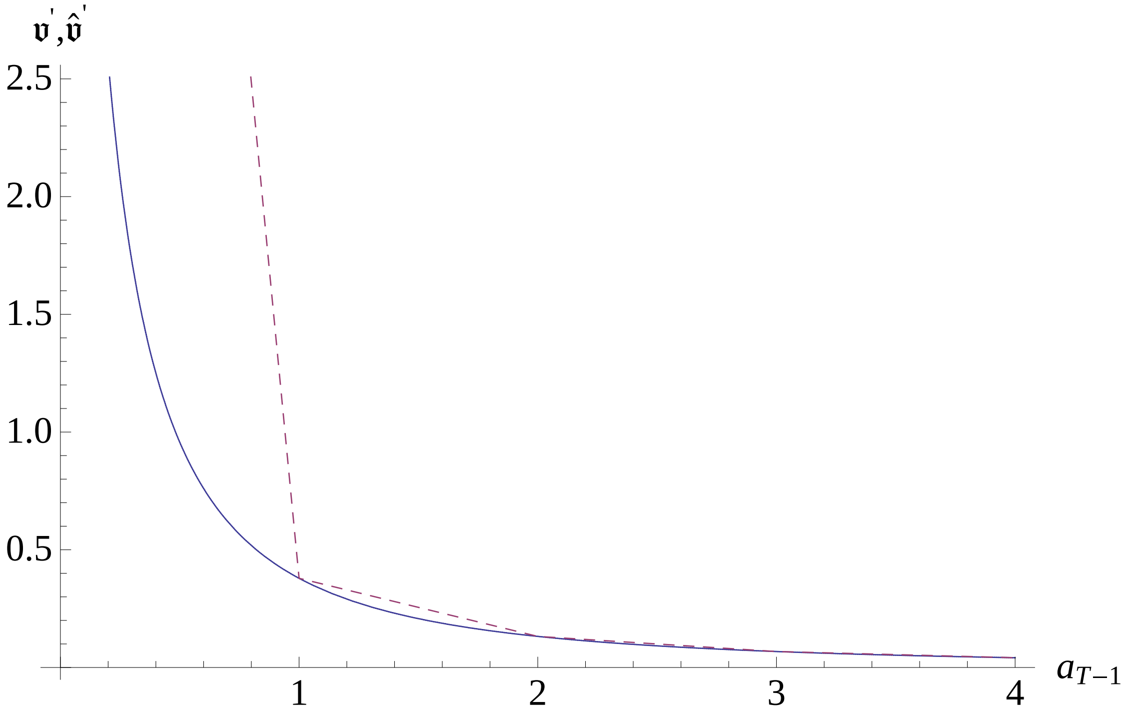

Figure 7 $\mathrm{v}^{a}_{T-1_{\rightarrow}}(a_{T-1})$ versus $\grave{\mathrm{v}}^{a}_{T-1_{\rightarrow}}(a_{T-1})$

{fig:PlotOPRawVSI

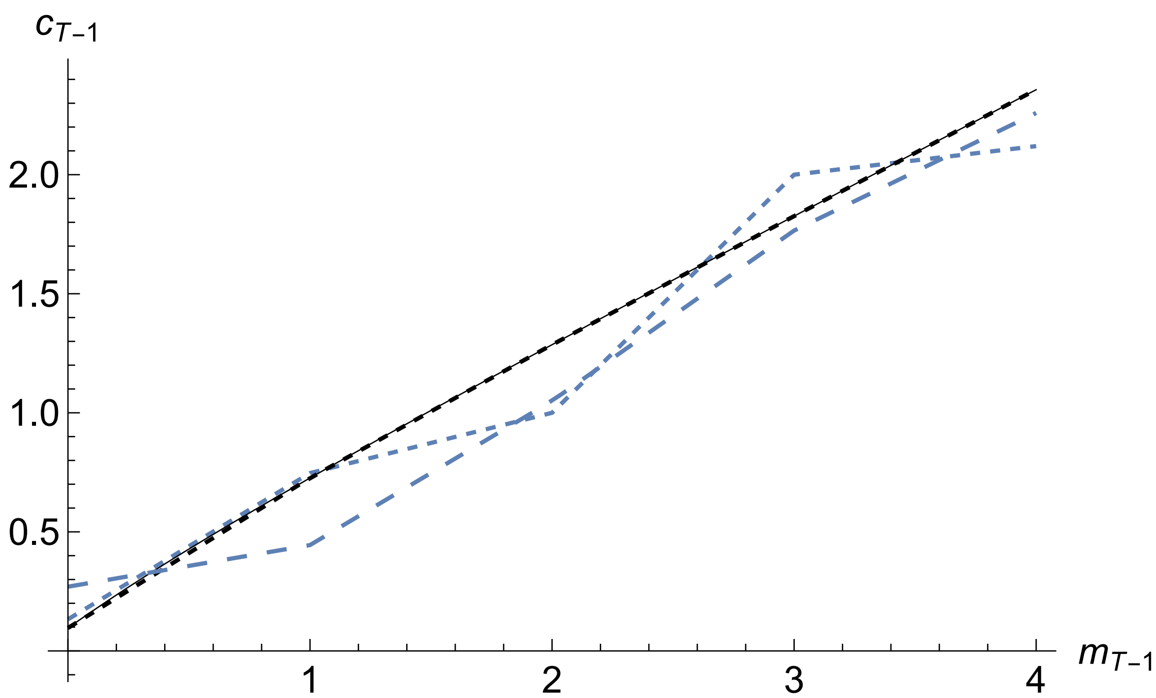

Figure 8 $\mathrm{c}_{T-1}(m_{T-1})$ (solid) Versus Two Methods for Constructing $\grave{\mathrm{c}}_{T-1}(m_{T-1})$

{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 $\mathrm{v}^{a}_{T-1_{\rightarrow}}(a_{T-1})$ 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:

$$c_{T-1}^{-\rho} = c_T^{-\rho}. \tag{28}$$

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 are 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 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 $\mathrm{v}^{m}_{T}(m_{T-1}) = \mathrm{u}^{c}(c_{T-1})$ this implies that

$$\mathrm{v}^{m}_{T-1}(m_{T-1}) = \left(\frac{m_{T-1}+1}{2}\right)^{-\rho}. \tag{29}$$ {eq:vPLin}

Of course, this is a highly nonlinear function. However, if we raise both sides of (29) to the power $(-1/\rho)$ the result is a linear function:

$$[\mathrm{v}^{m}_{T-1}(m_{T-1})]^{-1/\rho} = \frac{m_{T-1}+1}{2}. \tag{30}$$

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 problem 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

$\mathrm{v}^{a}_{T-1}(a_{T-1})$, we can define

$$\mathrm{c}_{T-1_{\rightarrow}}(a_{T-1}) \equiv \left(\mathrm{v}^{a}_{T-1_{\rightarrow}}(a_{T-1})\right)^{-1/\rho} \tag{31}$$ {eq:cGoth}

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

Note that this is *not* a consumption function. It is a 'consumed' function - it reveals the amount that must have been consumed for the consumer to have arrived at the end of the period with a given amount of assets.

5.7 The Natural Borrowing Constraint and the a_{T-1} Lower Bound

{subsec:LiqConstrS

This is the appropriate moment to ask an awkward question we have so far neglected: How should a function like $\grave{\mathrm{c}}_{T-1_{\rightarrow}}$ be evaluated outside the range of points spanned by $\{\mathtt{a}_{\mathtt{T-1}}[1], ..., \mathtt{a}_{\mathtt{T-1}}[\mathtt{n}]\}$ for which we have calculated the corresponding $\mathrm{c}_{T-1_{\rightarrow}}$ values used to produce our linearly interpolating approximation $\grave{\mathrm{c}}_{T-1_{\rightarrow}}$?

For most piecewise-linear interpolation implementations, when the interplating function is evaluated at a point outside the provided range of values used to construct the function, the algorithm silently performs extrapolation under the assumption that the slope of the function remains the same beyond the measured boundaries as within the nearest piecewise segment to the point.

The easiest answer would be linear extrapolation; for example, if the bottommost gridpoint is $a_1 = \boldsymbol{a}_{\mathtt{T-1}}[1]$ and the corresponding level of consumption is $c_1 = \mathrm{c}_{T-1_{\rightarrow}}(a_1)$ we could calculate the 'marginal propensity to have consumed' $\varkappa_1 = \grave{\mathrm{c}}^{a}_{T-1_{\rightarrow}}(a_1)$ and construct the approximation as the linear extrapolation below a_1:

$$\grave{\mathrm{c}}_{T-1_{\rightarrow}}(a_{T-1}) \equiv c_1 + (a_{T-1} - a_1)\varkappa_1 \tag{32}$$ {eq:ExtrapLin}

for values of $a_{T-1} < a_1$ To see that this approach will lead us into difficulties, consider what happens to the true (not approximated) $\mathrm{v}^{a}_{T-1_{\rightarrow}}(a_{T-1})$ as a_{T-1} approaches the value $\underline{a}_{T-1} = -\underline{\theta}\mathcal{R}_T^{-1}$. From (27) we have

$$\lim_{a_{T-1}\downarrow\underline{a}_{T-1}} \mathrm{v}^{a}_{T-1_{\rightarrow}}(a_{T-1}) = \lim_{a_{T-1}\downarrow\underline{a}_{T-1}} \beta\mathsf{R}\left(\frac{1}{n_\theta}\right)\sum_{i=1}^{n_\theta}\left(a_{T-1}\mathcal{R}_T + \theta_i\right)^{-\rho}. \tag{33}$$

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

These reflections reveal that the consumer faces a 'self-imposed' (or '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 $\underline{a}_{T-1}$).[9] The constraint is 'self-imposed' in the sense that if the utility function were different (say, Constant Absolute Risk Aversion), the consumer would be willing to borrow more than $\underline{\theta}\mathcal{R}_T^{-1}$ because a choice of zero or negative consumption in period T would yield some finite amount of utility.[10]

This self-imposed constraint cannot be captured well when the $\mathrm{v}^a_{T-1_\rightarrow}$ function is approximated by a piecewise linear function like $\grave{\mathrm{v}}^m_{T-1_\rightarrow}$, because there is no chance that the linear extrapolation below $\underline{a}$ will correctly predict $\mathrm{v}^a_{T-1_\rightarrow}(\underline{a}_{T-1}) = \infty$. To see what will happen instead, note first that if we are approximating $\mathrm{v}^a_{T-1_\rightarrow}$ the smallest value in `aVec` must be greater than $\underline{a}_{T-1}$ (because the expectation for any $a_{T-1} \leq \underline{a}_{T-1}$ is undefined).

Then when the approximating $\mathrm{v}^a_{T-1_\rightarrow}$ function is evaluated at some value less than the first element in `aVec`, the approximating function will linearly extrapolate the slope that characterized the lowest segment of the piecewise linear approximation (between `aVec[1]` and `aVec[2]`), a procedure that will return a positive finite number, even if the requested a_{T-1} point is below $\underline{a}_{T-1}$. This means that the precautionary saving motive is understated, and by an arbitrarily large amount as the level of assets approaches its true theoretical minimum $\underline{a}_{T-1}$.

The foregoing logic demonstrates that the marginal value of saving approaches infinity as $a_{T-1} \downarrow \underline{a}_{T-1} = -\underline{\theta}\mathcal{R}_T^{-1}$. But this implies that $\lim_{a_{T-1} \downarrow \underline{a}_{T-1}} \mathrm{c}_{T-1_\rightarrow}(a_{T-1}) = (\mathrm{v}^a_{T-1_\rightarrow}(a_{T-1}))^{-1/\rho} = 0$; that is, as a approaches its minimum possible value, the corresponding amount of c must approach *its* minimum possible value: zero.

The upshot of this discussion is a realization that all we need to do is to augment each of the $\boldsymbol{a}_{T-1}$ and $\boldsymbol{c}_{T-1}$ vectors with an extra point so that the first element in the list used to produce 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.

From there, we plot the lists that have been prepended with the natural borrowing contraint and the associated minimal level of consumption. Figure 9 shows the result. The solid line calculates the exact numerical value of the consumed function $\mathrm{c}_{T-1_\rightarrow}(a_{T-1})$ while the dashed line is the linear interpolating approximation $\grave{\mathrm{c}}_{T-1_\rightarrow}(a_{T-1})$. This figure illustrates the value of the transformation: The true function is close to linear, and so the linear approximation is almost indistinguishable from the true function except at the very lowest values of a_{T-1}.

Figure 10 similarly shows that when we generate $\grave{\grave{\mathrm{v}}}^a_{T-1_\rightarrow}(a)$ using our augmented $[\grave{\mathrm{c}}_{T-1_\rightarrow}(a)]^{-\rho}$ (dashed line) we obtain a *much* closer approximation to the true function $\mathrm{v}^a_{T-1_\rightarrow}(a)$ (solid line) than we did in the previous program which did not do the transformation (Figure 7). (The vertical axis label uses $\mathfrak{v}'$ as an alternative notation for what in these notes we designate as $\mathrm{v}^a_{T_\rightarrow}$).

[9] Another term for a constraint of this kind is the 'natural borrowing constraint.'

[10] 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.

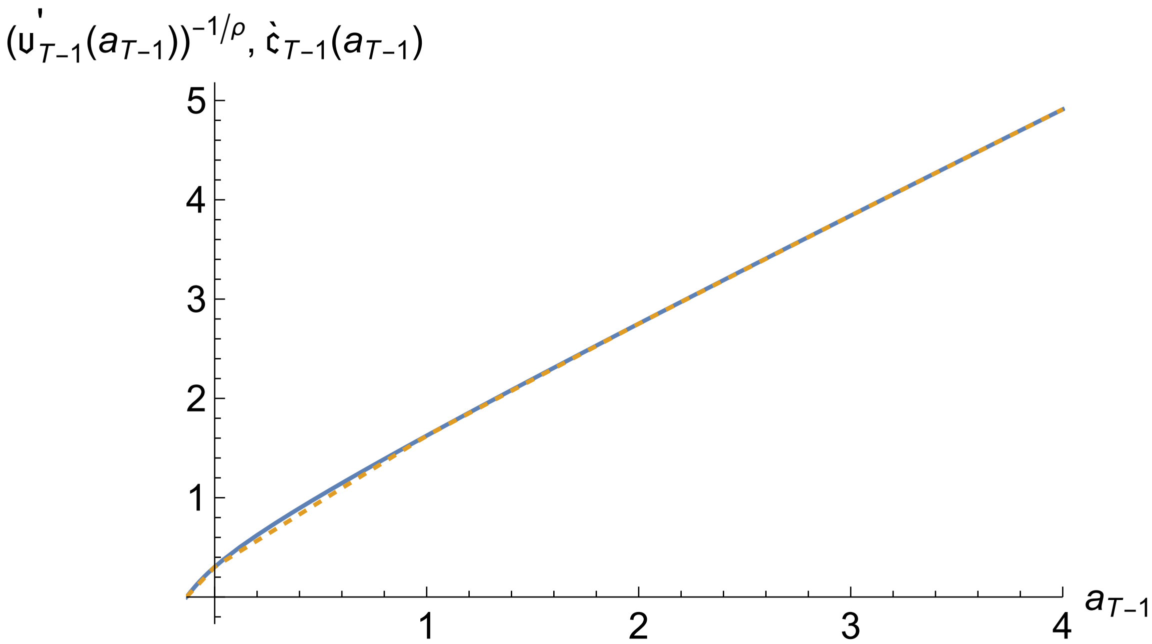

Figure 9 True $\mathrm{c}_{T-1_{\rightarrow}}(a_{T-1})$ vs its approximation $\grave{\mathrm{c}}_{T-1_{\rightarrow}}(a_{T-1})$

{fig:GothVInvVSGo

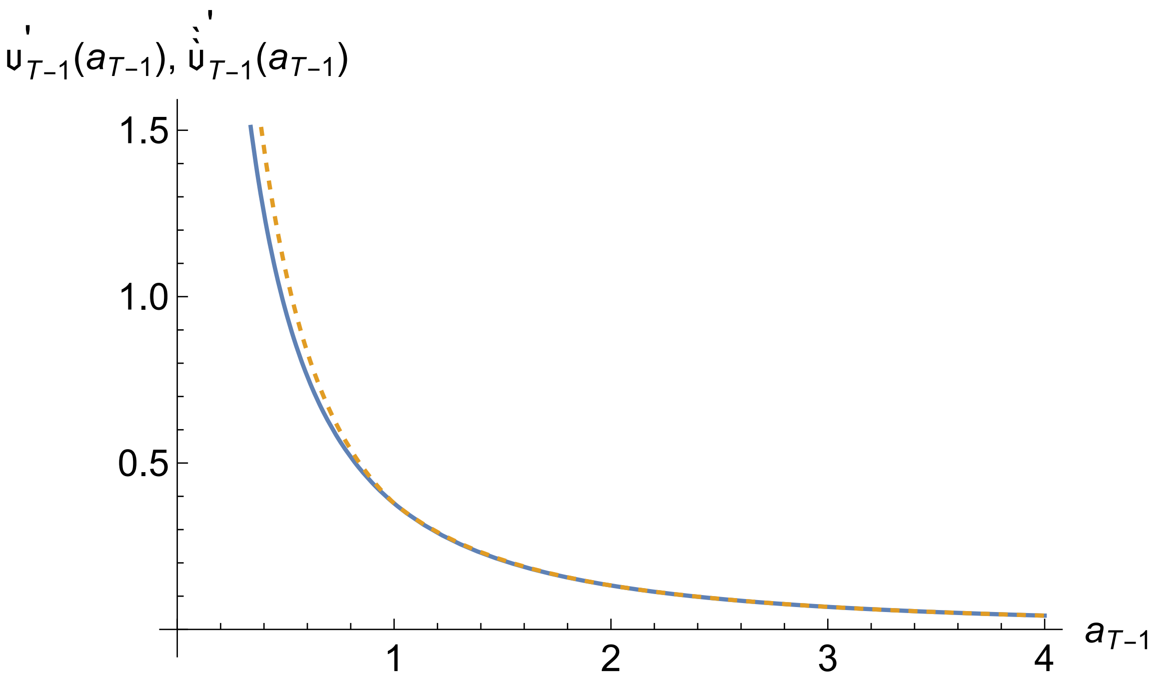

Figure 10 True $\mathrm{v}^{a}_{T-1_{\rightarrow}}(a_{T-1})$ vs. $\grave{\grave{\mathrm{v}}}^{a}_{T-1_{\rightarrow}}(a_{T-1})$ Constructed Using $\grave{\mathrm{c}}_{T-1_{\rightarrow}}(a_{T-1})$

{fig:GothVVSGoth

5.8 The Method of Endogenous Gridpoints

Our solution procedure for c_{T-1} still requires us, for each point in $\boldsymbol{m}_{\mathtt{T-1}}$ (`mVec` in the code), to use a numerical rootfinding algorithm to search for the value of c_{T-1} that solves $\mathrm{u}^c(c_{T-1}) = \mathrm{v}^a_{T-1_\rightarrow}(m_{T-1} - c_{T-1})$. Unfortunately, rootfinding is a notoriously computation-intensive (that is, slow!) operation.

It turns out that there is a way to completely skip the rootfinding step. The method can be understood by noting that any arbitrary value of $\boldsymbol{a}_{\mathtt{T-1}}$ (greater than its lower bound value a_1) will be associated with *some* marginal valuation as of the continuation ($\rightarrow$) stage of $T-1$ (that is, at the end of the period), and the further observation that it is trivial to find the value of c that yields the same marginal valuation, using the first order condition,

$$\mathrm{u}^c(\boldsymbol{c}_{\mathtt{T-1}}) = \mathrm{v}^a_{T-1_\rightarrow}(\boldsymbol{a}_{T-1}) \tag{34}$$ {eq:eulerTm1}

by using the inverse of the marginal utility function,

$$\begin{aligned} c^{-\rho} &= \mu \\ c &= \mu^{-1/\rho} \end{aligned} \tag{35}$$

which yields the level of consumption that corresponds to marginal utility of μ. Using this to invert both sides of (34), we get

$$\boldsymbol{c}_{\mathtt{T-1}} = \left(\mathrm{v}^a_{T-1_\rightarrow}(\boldsymbol{a}_{T-1})\right)^{-1/\rho} \tag{36}$$

where the $\rightarrow$ emphasizes that these are points on the 'consumed' function (that is, the function that reveals how much an optimizing consumer must have consumed in order to end the period with $a_{T-1,i}$).

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

$$\boldsymbol{m}_{\mathtt{T-1}} = \boldsymbol{c}_{\mathtt{T-1}} + \boldsymbol{a}_{\mathtt{T-1}}. \tag{37}$$

Rename gothic class: EndStep

These m_{T-1} gridpoints are "endogenous" in contrast to the usual solution method of specifying some *ex-ante* grid of values of m_{T-1} and then using a rootfinding routine to locate the corresponding optimal c_{T-1}. 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-specfied 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 $\boldsymbol{c}_{T-1_\rightarrow}$ values in hand, the notebook can generate a set of $\boldsymbol{m}_{\mathtt{T-1}}$ and $\boldsymbol{c}_{\mathtt{T-1}}$ pairs that can be interpolated between in order to yield $\grave{\mathrm{c}}_{T-1}(m_{T-1})$ at virtually zero computational cost![11]

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 nonlin-

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

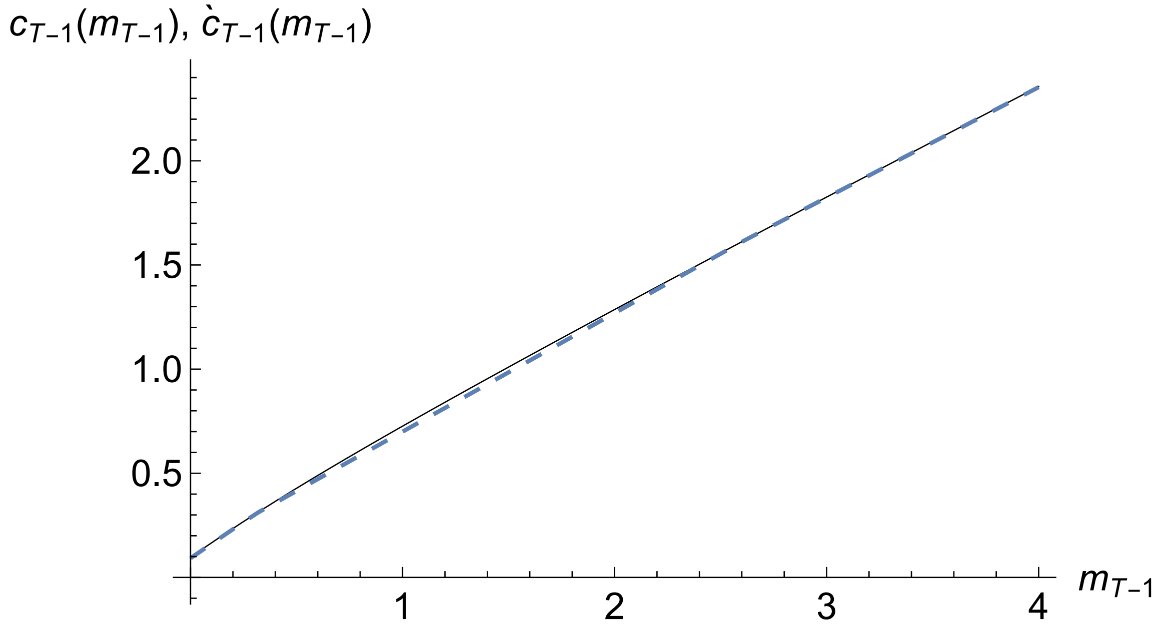

Figure 11 $\mathrm{c}_{T-1}(m_{T-1})$ (solid) versus $\grave{\mathrm{c}}_{T-1}(m_{T-1})$ (dashed)

{fig:ComparecTm1A

earity). Figure 11 plots the actual consumption function c_{T-1} and the approximated consumption function $\grave{\mathrm{c}}_{T-1}$ derived by the method of endogenous grid points. Compared to the approximate consumption functions illustrated in Figure 8, $\grave{\mathrm{c}}_{T-1}$ is quite close to the actual consumption function.

5.9 Improving the a Grid

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 $\grave{\mathrm{c}}_{T-1_\rightarrow}$ function tends to be farthest from its true value $\mathrm{c}_{T-1_\rightarrow}$ 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 than the uniform grid at low values of a. A pragmatic choice that works well is to find the values such that (1) the last value *exceeds the lower bound* by the same amount $\bar{a}_{T-1}$ as our original maximum gridpoint (in our case, 4.); (2) we have the same number of gridpoints as before; and (3) the *multi-exponential growth rate* (that is, $e^{e^{e^{\cdots}}}$ for some number of exponentiations n_θ) 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 (captured by the arguments `minval`, `maxval`, and `size`) and returns the new grid incorporating the multi-exponential approach outlined

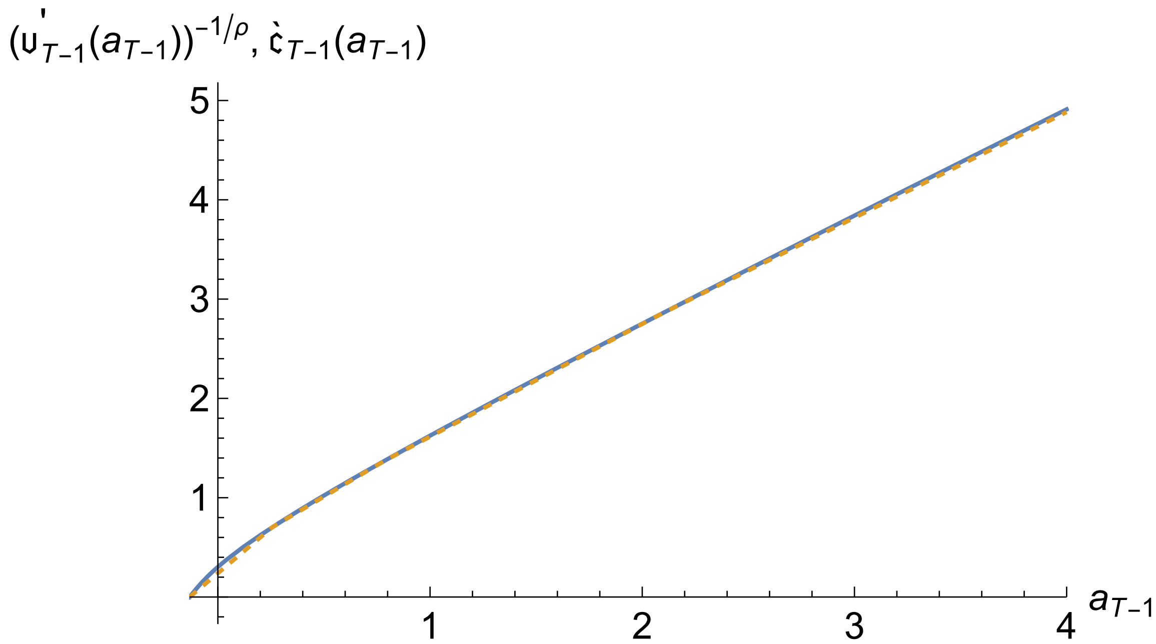

Figure 12 $\mathrm{c}_{T-1_{\rightarrow}}(a_{T-1})$ versus $\grave{\mathrm{c}}_{T-1_{\rightarrow}}(a_{T-1})$, Multi-Exponential `aVec`

{fig:GothVInvVSGo

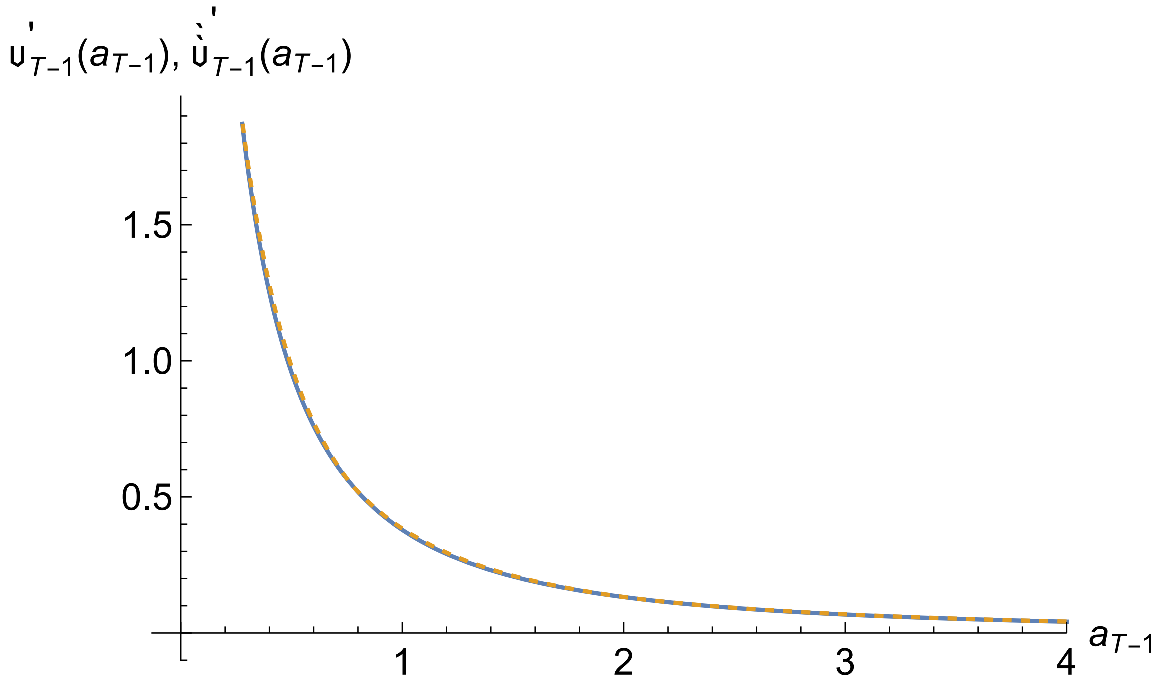

Figure 13 $\mathrm{v}^{a}_{T-1_{\rightarrow}}(a_{T-1})$ vs. $\grave{\grave{\mathrm{v}}}^{a}_{T-1_{\rightarrow}}(a_{T-1})$, Multi-Exponential `aVec`

{fig:GothVVSGothC

above. Then, a call is made to this function and the improved grid of assets is stored in the object `aVec_eee`. Lastly, the endogenous gridpoint method described in the previous section is performed using this new grid of assets. 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.

6 Structural Estimation

{sec:StructEst}

This section describes how to use the methods developed above to structurally estimate a life-cycle consumption model, following closely the work of Cagetti (2003).[12] The key idea of structural estimation is to look for the parameter values (for the time preference rate, relative risk aversion, or other parameters) which lead to the best possible match between simulated and empirical moments.

6.1 Life Cycle Model

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 β) for the 'pure' time preference factor, the value function for the revised problem is

$$\mathrm{v}_t(\mathbf{p}_t, \mathbf{m}_t) = \max_{\{\mathrm{c}\}_t^T} \quad \mathrm{u}(\mathbf{c}_t) + \mathbb{E}_t \left[\sum_{n=1}^{T-t} \beth^n \mathcal{L}_t^{t+n} \hat{\beta}_t^{t+n} \mathrm{u}(\mathbf{c}_{t+n}) \right] \tag{38}$$

{eq:lifecyclemax}

subject to the constraints

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

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

where

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

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:

$$\Xi_s = \begin{cases} 0 & \text{with probability } \wp > 0 \\ \theta_s/\wp & \text{with probability } (1-\wp), \text{ where } \log \theta_s \sim \mathcal{N}(-\sigma_\theta^2/2, \sigma_\theta^2) \end{cases} \tag{39}$$
$$\log \mathbf{\Psi}_s \sim \mathcal{N}(-\sigma_{\mathbf{\Psi}}^2/2, \sigma_{\mathbf{\Psi}}^2)$$

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 $\sigma_{\mathbf{\Psi}} = 0.1$.[13] 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 21).[14] The interest rate is assumed to equal 1.03. The model parameters are included in Table 1.

Table 1 Parameter Values

{table:StrEstParam

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

The structural estimation of the parameters $\beth$ and ρ 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

[13] 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 σ_θ is almost certainly increased by measurement error.

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

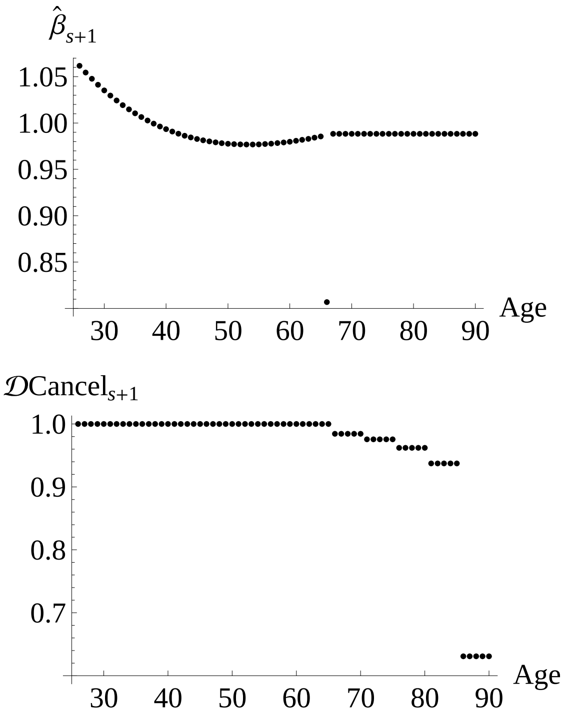

Figure 14 Time Varying Parameters

{fig:TimeVaryingPa

their corresponding commands. The next section elaborates on the procedure and its accompanying code implementation in greater detail.

6.2 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 ρ at which some measure of the model's outcome (like "median wealth by age") is as close as possible to an empirical measure of the same thing. Here, we choose to match the median of the wealth to permanent income ratio across 7 age groups, from age $26-30$ up to $56-60$.[15] 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

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

model which can be written in Bellman form as:

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

with the first order condition:

$$
\mathrm{u}^c(c_t) = \beth\mathcal{L}_{t+1}\hat{\beta}_{t+1}\mathsf{R}\,\mathbb{E}_t\left[\mathrm{u}^c\left(\boldsymbol{\Psi}_{t+1}\mathcal{G}_{t+1}\mathrm{c}_{t+1}\left(a_t\mathcal{R}_{t+1} + \theta_{t+1}\right)\right)\right]. \tag{40}
$$

{eq:FOCLifeCycle}

The first substantive step 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 (71) following the procedure in sections 5 and 6. 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 ξ as the set of parameters to be estimated (in the current case $\xi = \{\rho, \beth\}$), we could search for the parameter values which solve

$$\min_{\xi} \sum_{\tau=1}^{7} |\varsigma^{\tau} - \mathbf{s}^{\tau}(\xi)| \tag{41}$$ {eq:naivePowell}

where ς^{τ} and $\mathbf{s}^{\tau}$ are respectively the empirical and simulated medians of the wealth to permanent income ratio for age group τ. A drawback of proceeding in this way is that it treats the empirically estimated medians 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 (72) 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:

$$\min_{\xi} \sum_{i}^{N} \omega_i \, |\varsigma_i^{\tau} - \mathbf{s}^{\tau}(\xi)| \tag{42}$$ {eq:StructEstim}

where ω_i is the weight of household i in the entire population,[16] and ς_i^{τ} is the empirical wealth to permanent income ratio of household i whose head belongs to age group τ. ω_i is needed because unequal weight is assigned to each observation in the Survey of Consumer Finances (SCF). The absolute value is used since the formula is based on the fact that the median is the value 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 Survery 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 22. 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,

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

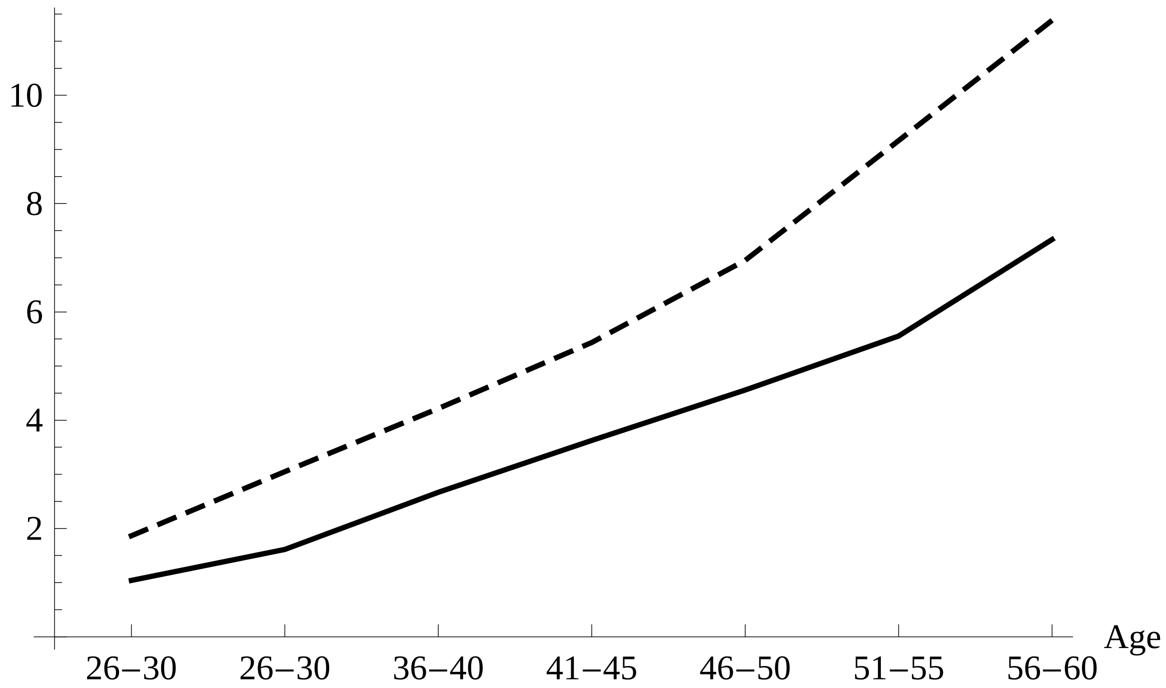

Figure 15 Wealth to Permanent Income Ratios from SCF (means (dashed) and medians (solid))

{fig:MeanMedianSC

smmObjectiveFxn, calls the simulate_moments function to create the objective function described in (73), 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 (73)

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 time demanding 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 suggsts, the purpose of this function is to compute the standard errors by bootstrap.[17] This involves:

1. drawing new shocks for the simulation

[17] 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 ρ and $\beth$

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

6.2.1 An Aside to Computing Sensitivity Measures

{subsubsec:sensmea

A common drawback in commonly used structural estimation procedures is a lack of transparency in its estimates. As Andrews, Gentzkow, and Shapiro (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, Gentzkow, and Shapiro (2017), section 9 is aimed at providing an estimator $\xi = \{\rho, \beth\}$ that has some true value ξ_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 (73). 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.

6.3 Results

The second part of the file `StructEstimation.py` defines a function `main` which produces our ρ and $\beth$ estimates with standard errors using 10,000 simulated agents by setting the positional arguments `estimate_model` and `compute_standard_errors` to true.[18] Results are reported in Table 2.[19]

Table 2 Estimation Results

{tab:EstResults}

ρ	$\beth$
3.69	0.88
(0.047)	(0.002)

[18]The procedure is: First we calculate the ρ and $\beth$ estimates as the minimizer of equation (73) using the actual SCF data. Then, we apply the `Bootstrap` function several times to obtain the standard error of our estimates.

[19]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.

The literature on consumption and savings behavior over the lifecycle in the presenece of labor income uncertainty [20] 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 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 23 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 ρ and $\beth$ 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 ρ and $\beth$ 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 $\beth$ 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 ρ is required to match the affected empirical moments.

To recap, the sensitivity measure not only matches our intuition about the inverse relationship between ρ and $\beth$ 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 24 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 ρ and $\beth$ which lead to the same deviations between simulated and empirical medians (equivalent values of equation (73)). 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.

7 Conclusion

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

[20] For example, see Gourinchas and Parker (2002) for an exposition of this.

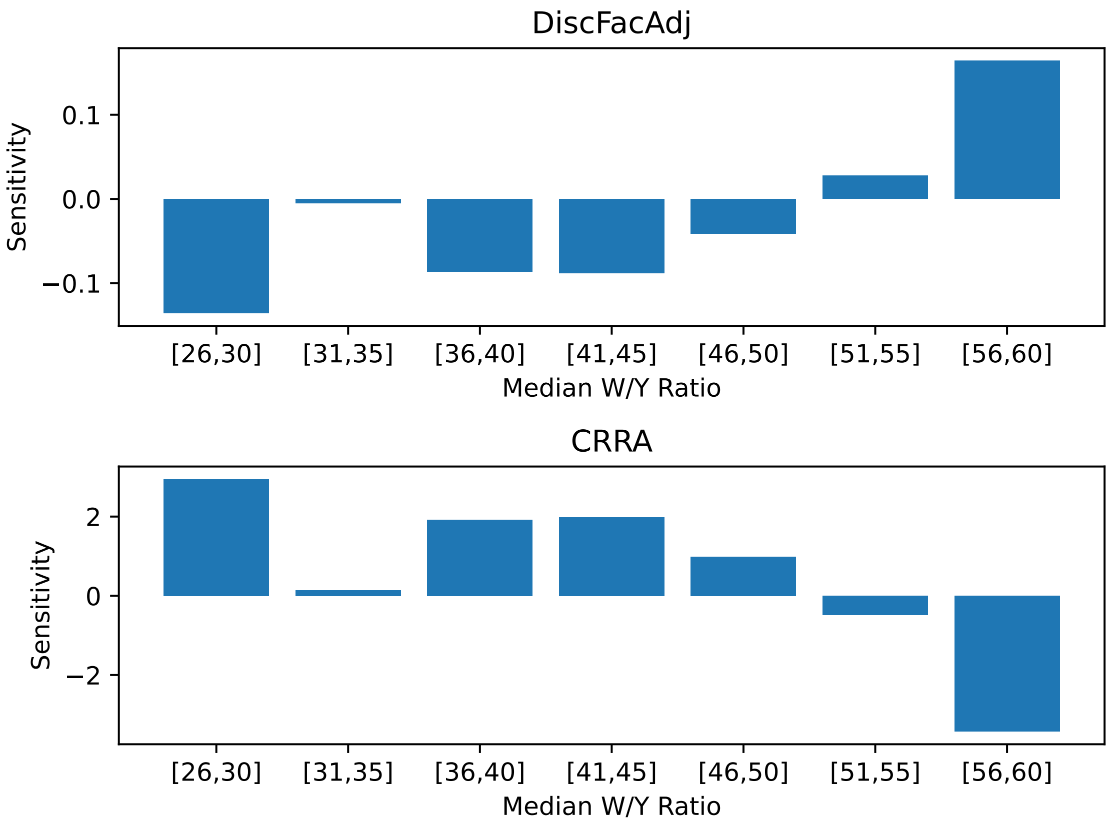

Figure 16 Sensitivty of Estimates $\{\rho, \beth\}$ regarding Alternative Mismeasurement Assumptions.

{fig:PlotSensitivityI

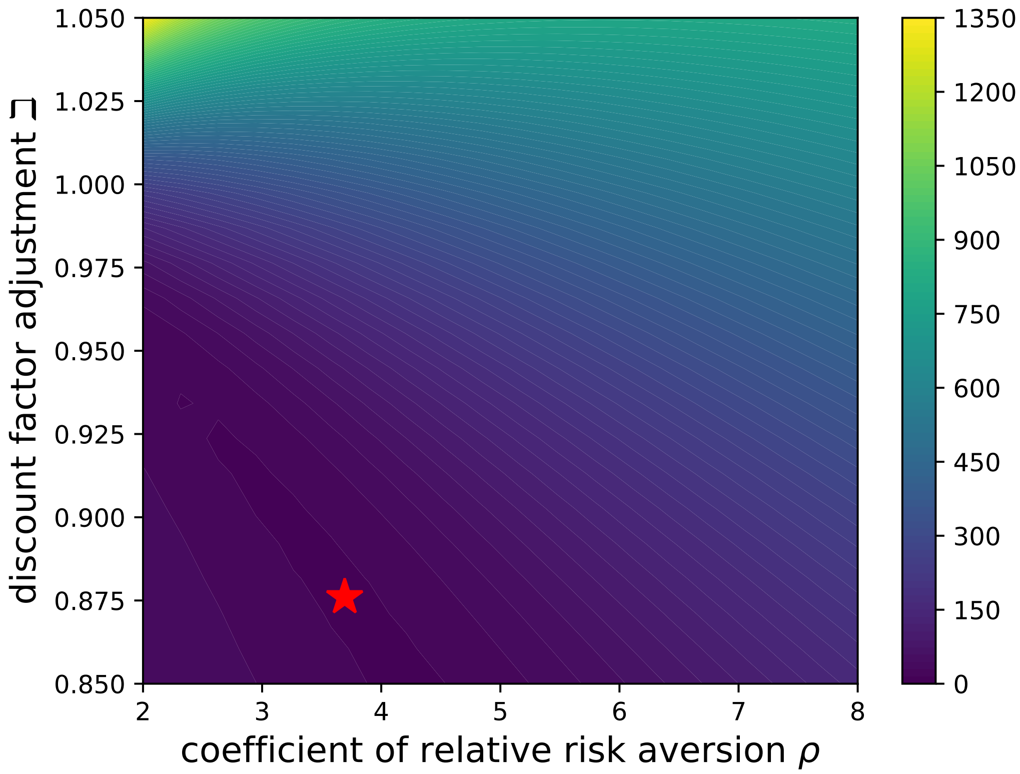

Figure 17 Contour Plot (larger values are shown lighter) with $\{\rho, \beth\}$ Estimates (red dot).

{fig:PlotContourMe

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

Appendices

A Further Details on SCF Data

{app:SCFdata}

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).[21] Furthermore, the data set contains only households whose heads are college graduates. The total sample size is 4,774.

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

Normal income levels in the SCF are before-tax figures. These before-tax permanent income figures must be rescaled so that the median of the rescaled permanent income of each age group matches the median of each age group's income which is assumed in the simulation. This rescaled permanent income is interpreted as after-tax permanent income. This rescaling is crucial since in the estimation empirical profiles are matched with simulated ones which are generated using after-tax permanent income (remember the income process assumed in the main text). Wealth / permanent income ratio is computed by dividing the level of wealth by the level of (after-tax) permanent income, and this ratio is used for the estimation.[22]

[21] 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.

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

References

Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro (2017): "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 132(4), 1553–1592.

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

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

Carroll, Christopher D. (1992): "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 1992(2), 61–156, 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), 1–56.

——— (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, 91(3), 312–320, https://www.econ2.jhu.edu/people/ccarroll/EndogenousGridpoints.pdf.

——— (2023a): "Solving Microeconomic Dynamic Stochastic Optimization Problems," *Econ-ARK REMARK*.

——— (2023b): "Theoretical Foundations of Buffer Stock Saving," *Revise and Resubmit, Quantitative Economics*.

Carroll, Christopher D., and Miles S. Kimball (1996): "On the Concavity of the Consumption Function," *Econometrica*, 64(4), 981–992, https://www.econ2.jhu.edu/people/ccarroll/concavity.pdf.

Carroll, Christopher D., and Andrew A. Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40(1), 41–71.

Deaton, Angus S. (1991): "Saving and Liquidity Constraints," *Econometrica*, 59, 1221–1248, https://www.jstor.org/stable/2938366.

——— (1992): *Understanding Consumption*. Oxford University Press, New York.

den Haan, Wouter J, and Albert Marcet (1990): "Solving the Stochastic Growth Model by Parameterizing Expectations," *Journal of Business and Economic Statistics*, 8(1), 31–34, Available at http://ideas.repec.org/a/bes/jnlbes/v8y1990i1p31-34.html.

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

HOROWITZ, JOEL L. (2001): "The Bootstrap," in *Handbook of Econometrics*, ed. by James J. Heckman, and Edward Leamer, vol. 5. Elsevier/North Holland.

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

KOPECKY, KAREN A., AND RICHARD M.H. SUEN (2010): "Finite State Markov-Chain Approximations To Highly Persistent Processes," *Review of Economic Dynamics*, 13(3), 701–714, http://www.karenkopecky.net/RouwenhorstPaper.pdf.

MERTON, ROBERT C. (1969): "Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case," *Review of Economics and Statistics*, 51, 247–257.

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

SAMUELSON, PAUL A. (1969): "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 51, 239–46.

TANAKA, KEN'ICHIRO, AND ALEXIS AKIRA TODA (2013): "Discrete approximations of continuous distributions by maximum entropy," *Economics letters*, 118(3), 445–450.

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