

The Method of Moderation For Solving Dynamic Stochastic Optimization Problems

Christopher Carroll¹ Kiichi Tokuoka² Weifeng Wu³

¹Johns Hopkins University and NBER
ccarroll@jhu.edu

²International Monetary Fund
ktokuoka@imf.org

³Fannie Mae
weifeng_wu@fanniemae.com

June 2012

<http://econ.jhu.edu/people/ccarroll/papers/ctwMoM/>

The Basic Problem at Date t

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

$$\mathbf{y}_t = \mathbf{p}_t \theta_t \quad (2)$$

$R_t = R$ - constant interest factor = $1 + r$

$\mathbf{p}_{t+1} = \Gamma_{t+1} \mathbf{p}_t$ - permanent labor income dynamics (3)

$\log \theta_{t+n} \sim \mathcal{N}(-\sigma_\theta^2/2, \sigma_\theta^2)$ - lognormal transitory shocks $\forall n > 0$.

- “Bewley” problem if liquidity constraint, Γ_{t+1} nonstochastic
- “Friedman/Buffer Stock” problem if also permanent shocks

The Basic Problem at Date t

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

$$\mathbf{y}_t = \mathbf{p}_t \theta_t \quad (2)$$

$R_t = R$ - constant interest factor = $1 + r$

$\mathbf{p}_{t+1} = \Gamma_{t+1} \mathbf{p}_t$ - permanent labor income dynamics (3)

$\log \theta_{t+n} \sim \mathcal{N}(-\sigma_\theta^2/2, \sigma_\theta^2)$ - lognormal transitory shocks $\forall n > 0$.

- “Bewley” problem if liquidity constraint, Γ_{t+1} nonstochastic
- “Friedman/Buffer Stock” problem if also permanent shocks

Bellman Equation

$$\mathbf{v}_t(\mathbf{m}_t, \mathbf{p}_t) = \max_{\mathbf{c}_t} u(\mathbf{c}_t) + \mathbb{E}_t[\beta \mathbf{v}_{t+1}(\mathbf{m}_{t+1}, \mathbf{p}_{t+1})] \quad (4)$$

\mathbf{m} – ‘market resources’ (net worth plus current income)

\mathbf{p} – permanent labor income

where

$$u(c) = \left(\frac{c^{1-\rho}}{1-\rho} \right) \quad (5)$$

Normalize the Problem

$$v_t(m_t) = \max_{c_t} u(c_t) + \mathbb{E}_t[\beta \Gamma_{t+1}^{1-\rho} v_{t+1}(m_{t+1})] \quad (6)$$

s.t.

$$a_t = m_t - c_t$$

$$m_{t+1} = \underbrace{(R/\Gamma_{t+1}) a_t + \theta_{t+1}}_{\equiv \mathcal{R}_{t+1}}$$

where nonbold variables are bold ones normalized by \mathbf{p} :

$$m_t = \mathbf{m}_t / \mathbf{p}_t \quad (7)$$

$\Rightarrow c_t(m)$ from which we can obtain

$$\mathbf{c}_t(\mathbf{m}_t, \mathbf{p}_t) = c_t(\mathbf{m}_t / \mathbf{p}_t) \mathbf{p}_t \quad (8)$$

My First Notational Atrocity

... in this paper.

$$v_t(a_t) = \mathbb{E}_t[\beta \Gamma_{t+1}^{1-\rho} v_{t+1}(\mathcal{R}_{t+1} a_t + \theta_{t+1})] \quad (9)$$

so FOC is

$$u'(c_t) = v'_t(m_t - c_t). \quad (10)$$

or

$$c_t^{-\rho} = v'_t(a_t) \quad (11)$$

My First Notational Atrocity

... in this paper.

$$v_t(a_t) = \mathbb{E}_t[\beta \Gamma_{t+1}^{1-\rho} v_{t+1}(\mathcal{R}_{t+1} a_t + \theta_{t+1})] \quad (9)$$

so FOC is

$$u'(c_t) = v'_t(m_t - c_t). \quad (10)$$

or

$$c_t^{-\rho} = v'_t(a_t) \quad (11)$$

My First Notational Atrocity

... in this paper.

$$v_t(a_t) = \mathbb{E}_t[\beta \Gamma_{t+1}^{1-\rho} v_{t+1}(\mathcal{R}_{t+1} a_t + \theta_{t+1})] \quad (9)$$

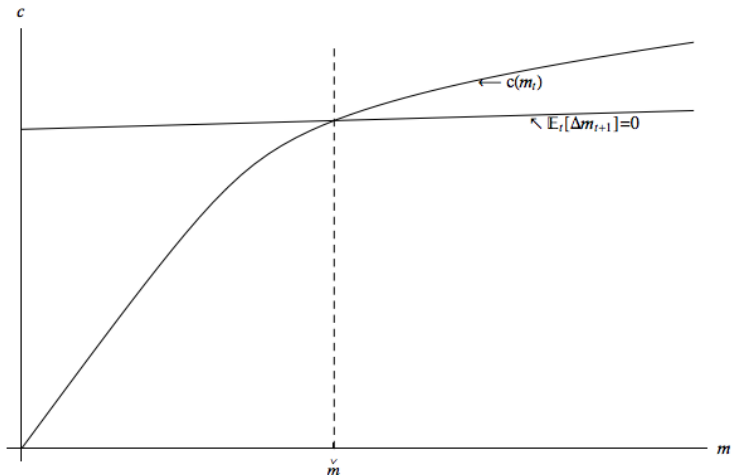
so FOC is

$$u'(c_t) = v'_t(m_t - c_t). \quad (10)$$

or

$$c_t^{-\rho} = v'_t(a_t) \quad (11)$$

Concave Consumption Function, Target Wealth Ratio



The Method of Endogenous Gridpoints

- Define vector of *end-of-period* asset values \vec{a}
- For each $a[j]$ compute $v'_t(a[j])$

Each of these $v'_t[j]$ corresponds to a unique $c[j]$ via FOC:

$$c[j]^{-\rho} = v'_t(a[j]) \quad (12)$$

$$c[j] = (v'_t(a[j]))^{-1/\rho} \quad (13)$$

But the DBC says

$$a_t = m_t - c_t \quad (14)$$

$$m[j] = a[j] + c[j] \quad (15)$$

So computing v'_t at a vector of \vec{a} values (easy) has produced for us the corresponding \vec{c} and \vec{m} values at virtually no cost!

From these we can interpolate to construct $\hat{c}_t(m)$

The Method of Endogenous Gridpoints

- Define vector of *end-of-period* asset values \vec{a}
 - For each $a[j]$ compute $v'_t(a[j])$

Each of these $v'_t[j]$ corresponds to a unique $c[j]$ via FOC:

$$c[j]^{-\rho} = v'_t(a[j]) \quad (12)$$

$$c[j] = (v'_t(a[j]))^{-1/\rho} \quad (13)$$

But the DBC says

$$a_t = m_t - c_t \quad (14)$$

$$m[j] = a[j] + c[j] \quad (15)$$

So computing v'_t at a vector of \vec{a} values (easy) has produced for us the corresponding \vec{c} and \vec{m} values at virtually no cost!

From these we can interpolate to construct $\hat{c}_t(m)$

The Method of Endogenous Gridpoints

- Define vector of *end-of-period* asset values \vec{a}
- For each $a[j]$ compute $v'_t(a[j])$

Each of these $v'_t[j]$ corresponds to a unique $c[j]$ via FOC:

$$c[j]^{-\rho} = v'_t(a[j]) \quad (12)$$

$$c[j] = (v'_t(a[j]))^{-1/\rho} \quad (13)$$

But the DBC says

$$a_t = m_t - c_t \quad (14)$$

$$m[j] = a[j] + c[j] \quad (15)$$

So computing v'_t at a vector of \vec{a} values (easy) has produced for us the corresponding \vec{c} and \vec{m} values at virtually no cost!

From these we can interpolate to construct $\hat{c}_t(m)$

The Method of Endogenous Gridpoints

- Define vector of *end-of-period* asset values \vec{a}
- For each $a[j]$ compute $v'_t(a[j])$

Each of these $v'_t[j]$ corresponds to a unique $c[j]$ via FOC:

$$c[j]^{-\rho} = v'_t(a[j]) \quad (12)$$

$$c[j] = (v'_t(a[j]))^{-1/\rho} \quad (13)$$

But the DBC says

$$a_t = m_t - c_t \quad (14)$$

$$m[j] = a[j] + c[j] \quad (15)$$

So computing v'_t at a vector of \vec{a} values (easy) has produced for us the corresponding \vec{c} and \vec{m} values at virtually no cost!

From these we can interpolate to construct $\hat{c}_t(m)$

The Method of Endogenous Gridpoints

- Define vector of *end-of-period* asset values \vec{a}
- For each $a[j]$ compute $v'_t(a[j])$

Each of these $v'_t[j]$ corresponds to a unique $c[j]$ via FOC:

$$c[j]^{-\rho} = v'_t(a[j]) \quad (12)$$

$$c[j] = (v'_t(a[j]))^{-1/\rho} \quad (13)$$

But the DBC says

$$a_t = m_t - c_t \quad (14)$$

$$m[j] = a[j] + c[j] \quad (15)$$

So computing v'_t at a vector of \vec{a} values (easy) has produced for us the corresponding \vec{c} and \vec{m} values at virtually no cost!

From these we can interpolate to construct $\hat{c}_t(m)$

The Method of Endogenous Gridpoints

- Define vector of *end-of-period* asset values \vec{a}
- For each $a[j]$ compute $v'_t(a[j])$

Each of these $v'_t[j]$ corresponds to a unique $c[j]$ via FOC:

$$c[j]^{-\rho} = v'_t(a[j]) \quad (12)$$

$$c[j] = (v'_t(a[j]))^{-1/\rho} \quad (13)$$

But the DBC says

$$a_t = m_t - c_t \quad (14)$$

$$m[j] = a[j] + c[j] \quad (15)$$

So computing v'_t at a vector of \vec{a} values (easy) has produced for us the corresponding \vec{c} and \vec{m} values at virtually no cost!

From these we can interpolate to construct $\hat{c}_t(m)$

The Method of Endogenous Gridpoints

- Define vector of *end-of-period* asset values \vec{a}
- For each $a[j]$ compute $v'_t(a[j])$

Each of these $v'_t[j]$ corresponds to a unique $c[j]$ via FOC:

$$c[j]^{-\rho} = v'_t(a[j]) \quad (12)$$

$$c[j] = (v'_t(a[j]))^{-1/\rho} \quad (13)$$

But the DBC says

$$a_t = m_t - c_t \quad (14)$$

$$m[j] = a[j] + c[j] \quad (15)$$

So computing v'_t at a vector of \vec{a} values (easy) has produced for us the corresponding \vec{c} and \vec{m} values at virtually no cost!

From these we can interpolate to construct $\hat{c}_t(m)$

The Method of Endogenous Gridpoints

- Define vector of *end-of-period* asset values \vec{a}
- For each $a[j]$ compute $v'_t(a[j])$

Each of these $v'_t[j]$ corresponds to a unique $c[j]$ via FOC:

$$c[j]^{-\rho} = v'_t(a[j]) \quad (12)$$

$$c[j] = (v'_t(a[j]))^{-1/\rho} \quad (13)$$

But the DBC says

$$a_t = m_t - c_t \quad (14)$$

$$m[j] = a[j] + c[j] \quad (15)$$

So computing v'_t at a vector of \vec{a} values (easy) has produced for us the corresponding \vec{c} and \vec{m} values at virtually no cost!

From these we can interpolate to construct $\hat{c}_t(m)$.

Problem: Extrapolating Outside the Grid (6 Gridpoints)

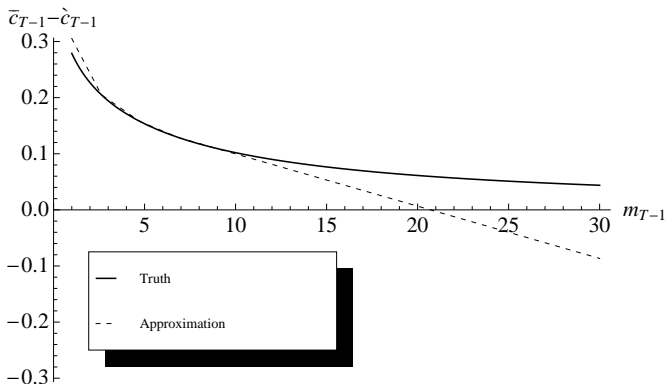


Figure: Oops

Call the 'true' problem that of a 'realist'

Contrasts with:

Optimist : Believes $\theta_{t+n} = \underbrace{1}_{=E_t[\theta_{t+n}]} \quad \forall n > 0$, consumes \bar{c}_t
Pessimist : Believes $\theta_{t+n} = \underline{\theta} \quad \forall n > 0$, consumes \underline{c}_t

Useful because:

- Both of these are perfect foresight problems
(→ how? Direct analytical solution (easy?))
- $\underline{c}_t < c_t < \bar{c}_t$

Call the 'true' problem that of a 'realist'

Contrasts with:

Optimist : Believes $\theta_{t+n} = \overbrace{1}^{=E_t[\theta_{t+n}]} \quad \forall n > 0$, consumes \bar{c}_t
Pessimist : Believes $\theta_{t+n} = \underline{\theta} \quad \forall n > 0$, consumes \underline{c}_t

Useful because:

- Both of these are perfect foresight problems
(→ how? Direct analytical solution (easy?))
- $\underline{c}_t < c_t < \bar{c}_t$

Call the 'true' problem that of a 'realist'

Contrasts with:

Optimist : Believes $\theta_{t+n} = \overbrace{1}^{=E_t[\theta_{t+n}]} \quad \forall n > 0$, consumes \bar{c}_t
Pessimist : Believes $\theta_{t+n} = \underline{\theta} \quad \forall n > 0$, consumes \underline{c}_t

Useful because:

- Both of these are perfect foresight problems
(→ how? Linear analytical solution (easy?))
- $\underline{c}_t < c_t < \bar{c}_t$

Call the 'true' problem that of a 'realist'

Contrasts with:

Optimist : Believes $\theta_{t+n} = \underbrace{1}_{=E_t[\theta_{t+n}]} \quad \forall n > 0$, consumes \bar{c}_t
Pessimist : Believes $\theta_{t+n} = \underline{\theta} \quad \forall n > 0$, consumes \underline{c}_t

Useful because:

- Both of these are perfect foresight problems
- (Does this really depend on the "true" problem?)
- $\underline{c}_t < c_t < \bar{c}_t$

Call the 'true' problem that of a 'realist'

Contrasts with:

Optimist : Believes $\theta_{t+n} = \underbrace{1}_{=E_t[\theta_{t+n}]} \quad \forall n > 0$, consumes \bar{c}_t
Pessimist : Believes $\theta_{t+n} = \underline{\theta} \quad \forall n > 0$, consumes \underline{c}_t

Useful because:

- 1 Both of these are perfect foresight problems
 - Linear analytical solution (easy!)
- 2 $\underline{c}_t < c_t < \bar{c}_t$

Call the 'true' problem that of a 'realist'

Contrasts with:

$$\begin{aligned} \text{Optimist} & : \text{Believes } \theta_{t+n} = \underbrace{1}_{=E_t[\theta_{t+n}]} \quad \forall n > 0, \text{ consumes } \bar{c}_t \\ \text{Pessimist} & : \text{Believes } \theta_{t+n} = \underline{\theta} \quad \forall n > 0, \text{ consumes } \underline{c}_t \end{aligned}$$

Useful because:

- 1 Both of these are perfect foresight problems
 - Linear analytical solution (easy!)
- 2 $\underline{c}_t < c_t < \bar{c}_t$

Call the 'true' problem that of a 'realist'

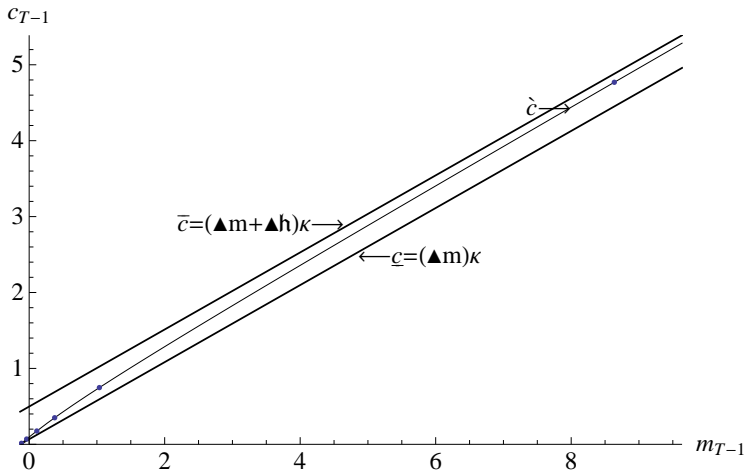
Contrasts with:

Optimist : Believes $\theta_{t+n} = \underbrace{1}_{=E_t[\theta_{t+n}]} \quad \forall n > 0$, consumes \bar{c}_t
 Pessimist : Believes $\theta_{t+n} = \underline{\theta} \quad \forall n > 0$, consumes \underline{c}_t

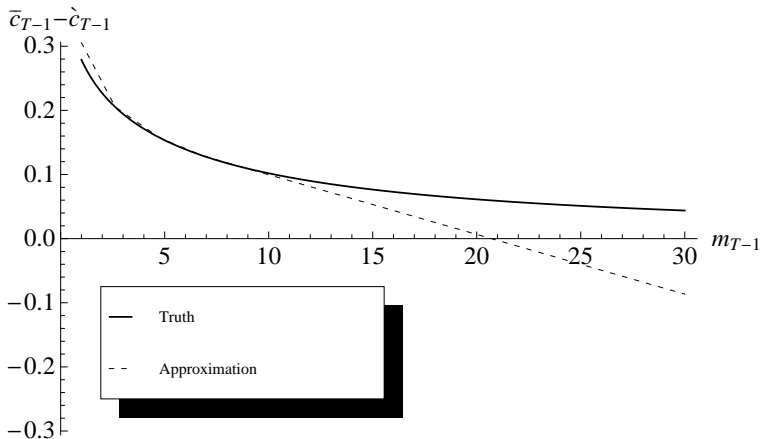
Useful because:

- ① Both of these are perfect foresight problems
 - Linear analytical solution (easy!)
- ② $\underline{c}_t < c_t < \bar{c}_t$

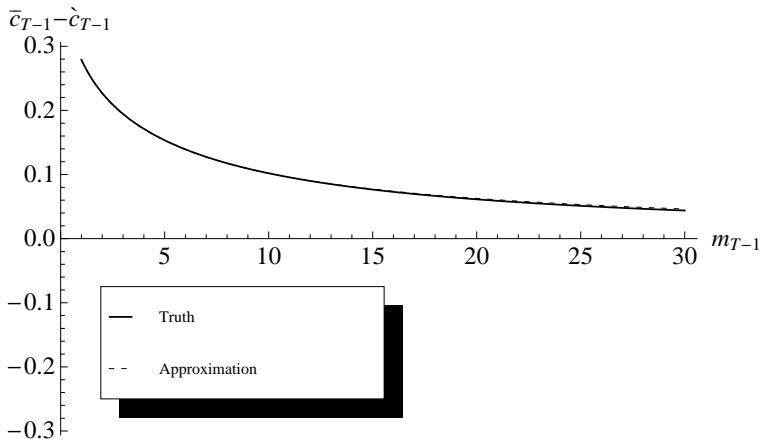
Illustration (6 Gridpoints)



Problem: Extrapolating Outside the Grid



Solution (Same 6 Gridpoints!)



Definitions and Solutions for Perfect Foresight Problem

$\underline{\kappa}_t$: MPC for perfect foresight problem in period t (16)

\underline{h}_t : end-of-period- t human wealth for 'optimist' (17)

$\underline{\bar{h}}_t$: end-of-period- t human wealth for 'pessimist' (18)

- Pessimist's solution:

$$\underline{c}_t(m) = \underbrace{(m + \underline{h}_t)}_{\equiv \blacktriangle m} \underline{\kappa}_t \quad (19)$$

- Optimist's solution:

$$\bar{c}_t(m) = (m + \bar{h}_t) \bar{\kappa}_t \quad (20)$$

$$= \underline{c}_t(m) + \underbrace{(\bar{h}_t - \underline{h}_t)}_{\equiv \blacktriangle \bar{h}_t} \bar{\kappa}_t \quad (21)$$

Definitions and Solutions for Perfect Foresight Problem

$\underline{\kappa}_t$: MPC for perfect foresight problem in period t (16)

\underline{h}_t : end-of-period- t human wealth for 'optimist' (17)

\underline{h}_t : end-of-period- t human wealth for 'pessimist' (18)

- Pessimist's solution:

$$\underline{c}_t(m) = \underbrace{(m + \underline{h}_t)}_{\equiv \blacktriangle m} \underline{\kappa}_t \quad (19)$$

- Optimist's solution:

$$\bar{c}_t(m) = (m + \underline{h}_t) \underline{\kappa}_t \quad (20)$$

$$= \underline{c}_t(m) + \underbrace{(\underline{h}_t - \underline{h}_t)}_{\equiv \blacktriangle \underline{h}_t} \underline{\kappa}_t \quad (21)$$

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

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

$$\hat{\varphi}_t(\mu_t) \equiv \left(\frac{\bar{c}_t(\underline{m}_t + e^{\mu_t}) - c_t(\underline{m}_t + e^{\mu_t})}{\blacktriangle h_t \underline{\kappa}_t} \right), \quad (22)$$

and we now define

$$\hat{\chi}_t(\mu_t) = \log \left(\frac{1 - \hat{\varphi}_t(\mu_t)}{\hat{\varphi}_t(\mu_t)} \right) \quad (23)$$

Approximating Precautionary Saving

$$\hat{c}_t = \bar{c}_t - \overbrace{\left(\frac{1}{1 + \exp(\hat{\chi}_t)} \right)}^{=\hat{q}_t} \blacktriangle \eta_t \underline{k}_t. \quad (25)$$

In the limit as μ approaches either ∞ or $-\infty$ this gives the 'right' answer.

Solving for the Value Function

Given analytical object

$$\mathbb{C}_t^T \equiv \text{PDV}_t^T(c)/c_t$$

The perfect foresight value function is:

$$v_t(m_t) = u(c_t)\mathbb{C}_t^T \quad (26)$$

which can be transformed as

$$\begin{aligned} \Lambda_t &\equiv ((1 - \rho)v_t)^{1/(1-\rho)} \\ &= c_t(\mathbb{C}_t^T)^{1/(1-\rho)} \end{aligned}$$

In the limit as μ approaches either ∞ or $-\infty$ this gives the 'right' answer.

Approximate Inverted Value Function

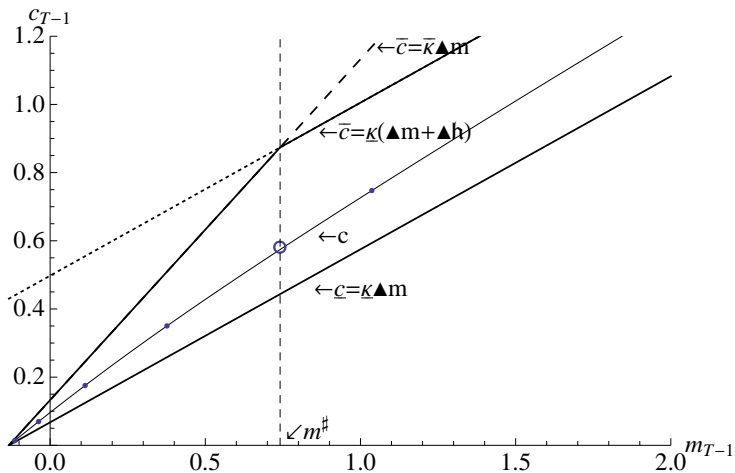
Same steps as for consumption function, because

$$\underline{v}_t < v_t < \bar{v}_t \quad (27)$$

Produce

$$\hat{v}_t(m) = u(\hat{\lambda}_t(m)) \quad (28)$$

A Tighter Limit: $c_t(m) < \min[\bar{\kappa}_t \blacktriangle m_t, \bar{c}_t(m_t)]$



The Tighter Upper Bound

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

$$\check{\gamma}_t(\mu_t) \equiv \frac{\bar{\kappa}_t - c_t(\underline{m}_t + e^{\mu_t})e^{-\mu_t}}{\bar{\kappa}_t - \underline{\kappa}_t}.$$

Combining Old and New Functions

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

we can define a well-behaved approximating consumption function

$$\check{c}_t = \mathbf{1}_{\text{Lo}}\check{c}_t + \mathbf{1}_{\text{Mid}}\check{c}_t + \mathbf{1}_{\text{Hi}}\hat{c}_t. \quad (29)$$

Extension To Case With Financial Risk Too

$$\log \mathbf{R}_{t+n} \sim \mathcal{N}(r + \phi - \sigma_r^2/2, \sigma_r^2) \quad \forall n > 0 \quad (30)$$

Merton (1969) and Samuelson (1969): Without labor income,

$$\kappa = 1 - \left(\beta \mathbb{E}_t[\mathbf{R}_{t+1}^{1-\rho}] \right)^{1/\rho} \quad (31)$$

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.

How Useful Is It?

When it works:

- Can approximate solution *globally* with very few gridpoints
- Can speed solution substantially

When Might It Not Work?

- Consumption With Nonconvexities
- Problems With No Perfect Foresight Solution
- How Many Such Problems Have A Solution With Shocks?
- *Should Work For Value Functions Fairly Generally*

How Useful Is It?

When it works:

- Can approximate solution *globally* with very few gridpoints
- Can speed solution substantially

When Might It Not Work?

- Consumption With Nonconvexities
- Problems With No Perfect Foresight Solution
- How Many Such Problems Have A Solution With Shocks?
- *Should* Work For Value Functions Fairly Generally

How Useful Is It?

When it works:

- Can approximate solution *globally* with very few gridpoints
- Can speed solution substantially

When Might It Not Work?

- Consumption With Nonconvexities
- Problems With No Perfect Foresight Solution
- How Many Such Problems Have A Solution With Shocks?
- *Should* Work For Value Functions Fairly Generally

How Useful Is It?

When it works:

- Can approximate solution *globally* with very few gridpoints
- Can speed solution substantially

When Might It Not Work?

- Consumption With Nonconvexities
- Problems With No Perfect Foresight Solution
 - How Many Such Problems Have A Solution With Shocks?
- *Should* Work For Value Functions Fairly Generally

How Useful Is It?

When it works:

- Can approximate solution *globally* with very few gridpoints
- Can speed solution substantially

When Might It Not Work?

- Consumption With Nonconvexities
- Problems With No Perfect Foresight Solution
 - How Many Such Problems Have A Solution With Shocks?
- *Should* Work For Value Functions Fairly Generally

How Useful Is It?

When it works:

- Can approximate solution *globally* with very few gridpoints
- Can speed solution substantially

When Might It Not Work?

- Consumption With Nonconvexities
- Problems With No Perfect Foresight Solution
 - How Many Such Problems Have A Solution With Shocks?
- *Should Work For Value Functions Fairly Generally*

How Useful Is It?

When it works:

- Can approximate solution *globally* with very few gridpoints
- Can speed solution substantially

When Might It Not Work?

- Consumption With Nonconvexities
- Problems With No Perfect Foresight Solution
 - How Many Such Problems Have A Solution With Shocks?
- *Should* Work For Value Functions Fairly Generally

References I

MERTON, ROBERT C. (1969): "Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case," Review of Economics and Statistics, 50, 247–257.

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