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 various mathematical and computational techniques that, together, speed the solution by many orders of magnitude compared to "brute force" methods. The problem is solved with and without liquidity constraints, and the infinite horizon solution is obtained as the limit of the finite horizon solution. After the basic consumption/saving problem with a deterministic interest rate is described and solved, an extension with portfolio choice between a riskless and a risky asset is also solved. Finally, a simple example shows how to use these methods (via the statistical 'method of simulated moments' (MSM for short) to estimate structural parameters like the coefficient of relative risk aversion (*a la* Gourinchas and Parker (2002) and Cagetti (2003))).

2 The Problem

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

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

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

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

$$\mathbf{y}_t = \mathbf{p}_t \boldsymbol{\theta}_t \tag{2}$$

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 $\mathbf{p}_t$).

---

[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.

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

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

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

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

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

Income evolves according to:

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

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

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

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:

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

3 Normalization

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

---

[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).

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

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

Now define nonbold variables as the bold variable divided by the level of permanent income in the same period, so that, for example, $m_T = \mathbf{m}_T/\mathbf{p}_T$; and define $\mathrm{v}_T(m_T) = \mathrm{u}(m_T)$.[3] For our CRRA utility function, $\mathrm{u}(xy) = x^{1-\rho}\mathrm{u}(y)$, so (8) can be rewritten as

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

Now define a new (dissected) optimization problem:

$$\mathrm{v}_t(m_t) = \max_{c_t} \quad \mathrm{u}(c_t) + \beta\,\mathbb{E}_t[\mathcal{G}_{t+1}^{1-\rho}\mathrm{v}_{t+1}(m_{t+1})]$$

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

$$a_t = m_t - c_t$$
$$k_{t+1} = a_t \tag{10} \quad \text{\{eq:vNormed\}}$$
$$b_{t+1} = \underbrace{(\mathsf{R}/\mathcal{G}_{t+1})}_{\equiv \mathcal{R}_{t+1}} k_{t+1}$$
$$m_{t+1} = b_{t+1} + \boldsymbol{\theta}_{t+1},$$

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

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

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

and so on back to all earlier periods. Hence, if we solve the problem (10) which has only a single state variable m_{T-1}, we can obtain the levels of the value function, consumption, and all other variables from the corresponding permanent-income-normalized solution objects by multiplying each by $\mathbf{p}_T$, e.g. $\mathbf{c}_T(\mathbf{m}_T, \mathbf{p}_T) = \mathbf{p}_T\mathrm{c}_T(\mathbf{m}_T/\mathbf{p}_T)$ (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 future reference it will be useful to write the problem (10) in the traditional way, by substituting b_{t+1}, k_{t+1}, and a_t into m_{t+1}:

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

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

---

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

4 The Usual Theory, and a Bit More Notation

4.1 Steps

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

1. **Arrival**: Incoming state variables (e.g., k_t) are known, but any shocks associated with the period have not been realized and decision(s) have not yet been made

2. **Decision**: All exogenous variables (like income shocks, rate of return shocks, and predictable income growth $\mathcal{G}$) have been realized (so that, e.g., m_t's value is known) and the agent solves the optimization problem

3. **Continuation**: After all decisions have been made, their consequences are revealed by evaluation of the continuing-value function at the values of the 'outgoing' state variables.

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

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

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

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

4.2 The Usual Theory, Notated

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

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

and because the `Envelope` theorem tells us that

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

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

$$\mathrm{u}^c(c_t) = \mathrm{v}_t^m(m_t) \tag{14}$$

and rolling forward one period,

$$\mathrm{u}^c(c_{t+1}) = \mathrm{v}_{t+1}^m(a_t \mathcal{R}_{t+1} + \boldsymbol{\theta}_{t+1}) \tag{15}$$

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

$$\mathrm{u}^c(c_t) = \mathbb{E}_{t\to}[\beta \mathrm{R}\mathcal{G}_{t+1}^{-\rho}\mathrm{u}^c(c_{t+1})]. \tag{16}$$

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

$$\mathrm{v}_{\leftarrow(t+1)}(k_{t+1}) = \mathbb{E}_{\leftarrow(t+1)}[\ \mathcal{G}_{t+1}^{1-\rho}\mathrm{v}_{t+1}(\overbrace{\mathcal{R}_{t+1}k_{t+1} + \boldsymbol{\theta}_{t+1}}^{m_{t+1}})] \tag{17}$$
$$\mathrm{v}_{\leftarrow(t+1)}^k(k_{t+1}) = \mathbb{E}_{\leftarrow(t+1)}[\mathrm{R}\mathcal{G}_{t+1}^{-\rho}\mathrm{v}_{t+1}^m(m_{t+1})]$$

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

$$\mathrm{u}^c(c_t) = \mathrm{v}_{t\to}^a(m_t - c_t). \tag{18}$$

4.3 Summing Up

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

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

$$\mathrm{v}_{\leftarrow t}(k_t) = \mathbb{E}_{\leftarrow t}[\mathrm{v}_t(\overbrace{k_t \mathcal{R}_t + \boldsymbol{\theta}_t}^{m_t})] \tag{19}$$

$$\mathrm{v}_t(m_t) = \max_{\{c\}}\ \mathrm{u}(c) + \mathbb{E}_t[\mathrm{v}_{t\to}(\overbrace{m_t - c}^{a_t})] \tag{20}$$

and

$$\mathrm{v}_{t\to}(a) = \beta \mathrm{v}_{\leftarrow(t+1)}(\overbrace{k_{t+1}}^{a}) \tag{21}$$

5 Multiple Control Variables

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

5.1 Theory

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

$$\mathfrak{R}(\varsigma) = \mathsf{R} + (\mathbf{R} - \mathsf{R})\varsigma. \tag{22}$$

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

$$\mathrm{v}_t(m_t) = \max_{\{c_t, \varsigma\}} \ \mathrm{u}(c_t) + \mathbb{E}_t[\beta \mathrm{v}_{t+1}((m_t - c_t)(\mathsf{R} + (\mathbf{R} - \mathsf{R})\varsigma) + \boldsymbol{\theta}_{t+1})] \tag{23}$$

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

But as a notational choice, there is good reason to designate the realization as happening in $t + 1$. A standard way of motivating stochastic returns and wages is to attribute them to "productivity shocks" and to assume that the productivity shock associated with a date is the one that affects the production function for that date.

5.2 Stages Within a Period

Solving simultaneously for the two variables ς and c can be computationally challenging. Fortunately, there is a simple solution: Break the problem into two 'stages.'[4]

As demonstrated in (23), the mathematical solution for the optimal portfolio share is the same whether we conceive the shocks as occuring at the end of t or the beginning of $t + 1$. But our tripartite dissection of the problem above into $\{t_{\leftarrow}, t, t_{\rightarrow}\}$ steps was motivated partly by a desire to invent a notation that can allow for construction of standalone 'stages' whose only connection to the subsequent problem is through the continuation-vaue function.

To illustrate this point, consider the standalone problem of an 'investor' whose continuation value function $\mathrm{v}_{[\varsigma]_{\rightarrow}}$ depends only on the amount of liquid assets $\ell_{\rightarrow}$ with which they end up after the realization of the stochastic $\mathbf{R}$ return. The expected value that the

---

[4]cite mnw and ael papers.

investor will obtain from any combination of initial $\ell_\leftarrow$ and ς is the expectation of the continuation value function over the liquid assets that result from the portfolio choice:

$$v_{[\varsigma]}(\ell_\leftarrow) = \max_\varsigma \; \mathbb{E}\left[v_{[\varsigma]_\rightarrow}\left(\mathfrak{R}(\varsigma)\ell_\leftarrow\right)\right] \qquad (24) \quad \text{\{eq:vMidStgShr\}}$$

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

$$v_\leftarrow(\ell_\leftarrow) = \mathbb{E}[v_{[\varsigma]_\rightarrow}(\overbrace{\hat{\mathfrak{R}}(\ell_\leftarrow)\ell_\leftarrow}^{\ell_\rightarrow})]. \qquad (25)$$

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

$$v_{t_\rightarrow}(b_t) = \beta v_{\leftarrow(t+1)}(\kappa_{t+1}) \qquad (26)$$

while the dynamic budget constraint for m changes to

$$m_t = \kappa_t + \boldsymbol{\theta}_t \qquad (27)$$

and the problem in the decision step is now

$$v_t(m_t) = \max_c \; u(c) + \mathbb{E}_t[v_{t_\rightarrow}(m_t - c)] \qquad (28)$$

while value in the arrival step is now

$$v_{\leftarrow t}(\kappa_t) = \mathbb{E}_{\leftarrow t}[v_t(m_t)] \qquad (29)$$

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

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

the version where the portfolio share is chosen at the beginning of the period, and $v_{\leftarrow t[c\varsigma]}$ is initial value for the the problem where the share choice is at the end.

While the investor's problem cannot be solved using the endogenous gridpoints method,[5] we can solve it numerically for the optimal ς at a vector of $\boldsymbol{a}$ (aVec in the code) and then construct an approximated optimal portfolio share function $\grave{\hat{\varsigma}}(a)$ as the interpolating function among the members of the $\{\boldsymbol{a}, \boldsymbol{\varsigma}\}$ mapping. Having done this, we can now calculate a vector of values that correspond to aVec

$$\boldsymbol{v}_{\rightarrow} = v_{t_{\rightarrow}}(\boldsymbol{a}, \grave{\hat{\varsigma}}(\boldsymbol{a})), \tag{30}$$

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

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

$$\boldsymbol{c} \equiv \left(v^a_{t_{\rightarrow}}(\boldsymbol{a})\right)^{-1/\rho}, \tag{31}$$

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

5.3 Application

In specifying the stochastic process for $\mathbf{R}_{t+1}$, we follow the common practice of assuming that returns are lognormally distributed, $\log \mathbf{R} \sim \mathcal{N}(\phi + r - \sigma_{\mathbf{r}}^2/2, \sigma_{\mathbf{r}}^2)$ where ϕ is the equity premium over the thin returns r available on the riskless asset.[6]

As with labor income uncertainty, it is necessary to discretize the rate-of-return risk in order to have a problem that is soluble in a reasonable amount of time. We follow the same procedure as for labor income uncertainty, generating a set of $n_{\mathbf{r}}$ equiprobable shocks to the rate of return; in a slight abuse of notation, we will designate the portfolio-weighted return (contingent on the chosen portfolio share in equity, and potentially contingent on any other aspect of the consumer's problem) simply as $\mathfrak{R}_{i,j}$ (where dependence on i is allowed to permit the possibility of nonzero correlation between the return on the risky asset and the $\boldsymbol{\theta}$ shock to labor income (for example, in recessions the stock market falls and labor income also declines).

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

$$
\begin{aligned}
v^a_{t_{\rightarrow}}(a_t, \varsigma_t) &= \beta \left(\frac{1}{n_{\mathbf{r}} n_{\boldsymbol{\theta}}}\right) \sum_{i=1}^{n_{\boldsymbol{\theta}}} \sum_{j=1}^{n_{\mathbf{r}}} \mathfrak{R}_{i,j} \left(c_{t+1}(\mathfrak{R}_{i,j} a_t + \boldsymbol{\theta}_i)\right)^{-\rho} \\
v^\varsigma_{t_{\rightarrow}}(a_t, \varsigma_t) &= \beta \left(\frac{1}{n_{\mathbf{r}} n_{\boldsymbol{\theta}}}\right) \sum_{i=1}^{n_{\boldsymbol{\theta}}} \sum_{j=1}^{n_{\mathbf{r}}} (\mathbf{R}_{i,j} - \mathbf{R}) \left(c_{t+1}(\mathfrak{R}_{i,j} a_t + \boldsymbol{\theta}_i)\right)^{-\rho}.
\end{aligned}
\tag{32}
$$

---

[5]Because $v^\varsigma_{t[\varsigma]\rightarrow}$ is not invertible with respect to ς, see [references to MNW and AEL's work].

[6]This guarantees that $\mathbb{E}[\mathbf{R}] = \boldsymbol{\varphi}$ is invariant to the choice of σ_ϕ^2; see LogELogNorm.

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

Details of the construction of the interpolating approximation are given below; assume for the moment that we have the approximations $\hat{\mathrm{v}}_{t\rightarrow}^{a}$ and $\hat{\mathrm{v}}_{t\rightarrow}^{\varsigma}$ in hand and we want to proceed. As noted above in the discussion of (23), nonlinear equation solvers can find the solution to a set of simultaneous equations. Thus we could ask one to solve

$$
\begin{aligned}
c_t^{-\rho} &= \hat{\mathrm{v}}_{t\rightarrow}^{a}(m_t - c_t, \varsigma_t) \\
0 &= \hat{\mathrm{v}}_{t\rightarrow}^{\varsigma}(m_t - c_t, \varsigma_t)
\end{aligned}
\tag{33}
$$

{eq:FOCwrtw}

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

There is a better way. Define the problem

$$
\tilde{\mathrm{v}}_{t\rightarrow}(a_t) = \max_{\varsigma_t} \quad \mathrm{v}_{t\rightarrow}(a_t, \varsigma_t)
$$

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

$$
0 \leq \varsigma_t \leq 1
$$

where the tilde over $\tilde{\mathrm{v}}(a)$ indicates that this is the v that has been optimized with respect to all of the arguments other than the one still present (a_t). We solve this problem for the set of gridpoints in `aVec` and use the results to construct the interpolating function $\grave{\mathrm{v}}_t^{a}(a_t)$.[7] With this function in hand, we can use the first order condition from the single-control problem

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

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

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

---

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

5.4 Implementation

Following the discussion from section 5.1, to provide a numerical solution to the problem with multiple control variables, we must define expressions that capture the expected marginal value of end-of-period assets with respect to the level of assets and the share invested in risky assets. This is addressed in "Multiple Control Variables."

5.5 Results With Multiple Controls

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

But Figure 2 which plots the optimal portfolio share as a function of the level of assets, exhibits several interesting features. First, even with a coefficient of relative risk aversion of 6, an equity premium of only 4 percent, and an annual standard deviation in equity returns of 15 percent, the optimal choice is for the agent to invest a proportion 1 (100 percent) of the portfolio in stocks (instead of the safe bank account with riskless return R) is at values of a_t less than about 2. Second, the proportion of the portfolio kept in stocks is *declining* in the level of wealth - i.e., the poor should hold all of their meager assets in stocks, while the rich should be cautious, holding more of their wealth in safe bank deposits and less in stocks. This seemingly bizarre (and highly counterfactual – see Carroll (2002)) prediction reflects the nature of the risks the consumer faces. Those consumers who are poor in measured financial wealth will likely derive a high proportion of future consumption from their labor income. Since by assumption labor income risk is uncorrelated with rate-of-return risk, the covariance between their future consumption and future stock returns is relatively low. By contrast, persons with relatively large wealth will be paying for a large proportion of future consumption out of that wealth, and hence if they invest too much of it in stocks their consumption will have a high covariance with stock returns. Consequently, they reduce that correlation by holding some of their wealth in the riskless form.

6 The Infinite Horizon

All of the solution methods presented so far have involved period-by-period iteration from an assumed last period of life, as is appropriate for life cycle problems. However, if the parameter values for the problem satisfy certain conditions (detailed in Carroll (2023b)), the consumption rules (and the rest of the problem) will converge to a fixed rule as the horizon (remaining lifetime) gets large, as illustrated in Figure ??. Furthermore, Deaton (1991), Carroll (1992; 1997) and others have argued that the 'buffer-stock' saving behavior that emerges under some further restrictions on parameter values is a good approximation of the behavior of typical consumers over much of the lifetime. Methods for finding the converged functions are therefore of interest, and are dealt with in this section.
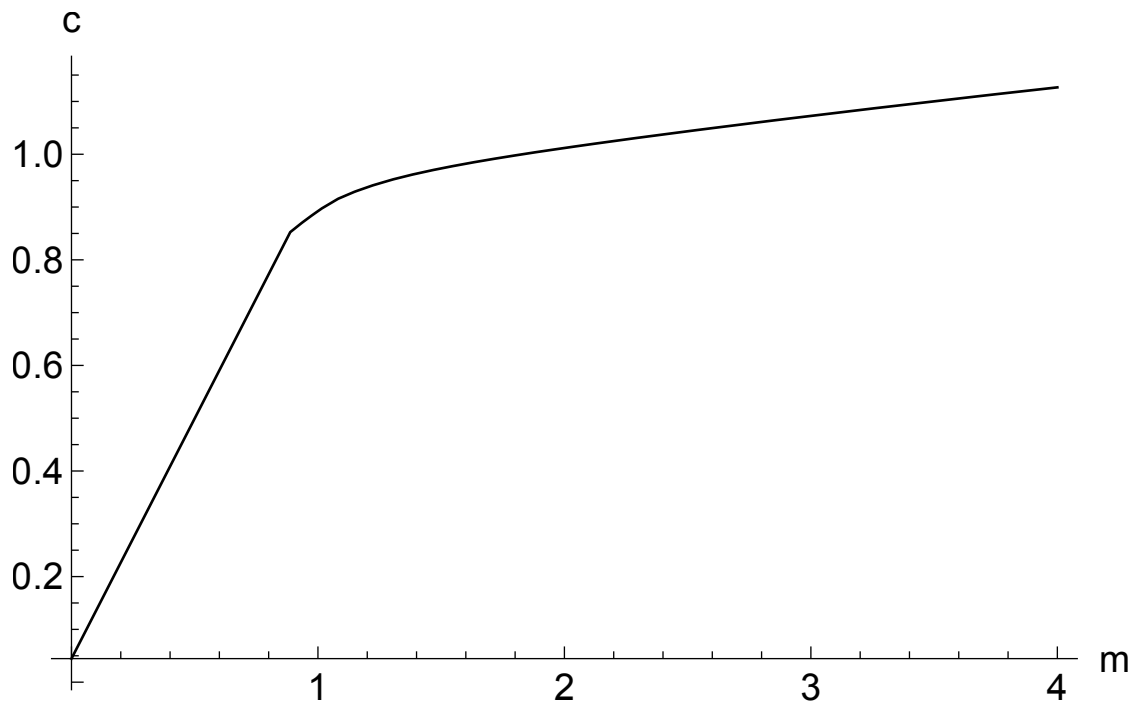
Figure 1 $c(m_1)$ With Portfolio Choice

Limit as a approaches ∞
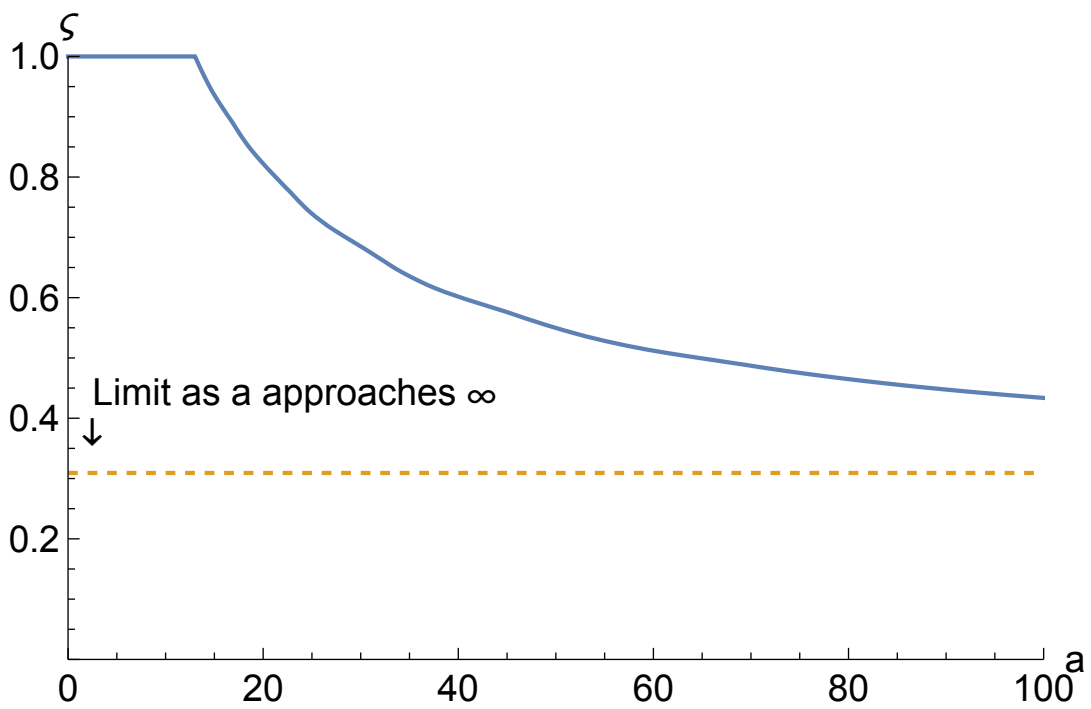↓

Figure 2 Portfolio Share in Risky Assets in First Period $\varsigma(a)$

11

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

6.1 Convergence

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

A natural metric is defined by the unique 'target' level of wealth that Carroll (2023b) proves will exist in problems of this kind under certain conditions: The $\hat{m}$ such that

$$\mathbb{E}_t[m_{t+1}/m_t] = 1 \text{ if } m_t = \hat{m} \tag{34}$$

where the accent is meant to signify that this is the value that other m's 'point to.'

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

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

7 Structural Estimation

This section describes how to use the methods developed above to structurally estimate a life-cycle consumption model, following closely the work of Cagetti (2003).[8] 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.

7.1 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

---

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

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

$$v_t(\mathbf{p}_t, \mathbf{m}_t) = \max_{\{c\}_t^T} \quad u(\mathbf{c}_t) + \mathbb{E}_{t\to} \left[\sum_{n=1}^{T-t} \beth^n \mathcal{L}_t^{t+n} \hat{\beta}_t^{t+n} u(\mathbf{c}_{t+n}) \right] \tag{35}$$

subject to the constraints

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

where

$\mathcal{L}_t^{t+n}$: probability to $\mathcal{L}$ive until age $t+n$ given alive at age t

$\hat{\beta}_t^{t+n}$: age-varying discount factor between ages t and $t+n$

Ψ_t : mean-one shock to permanent income

$\beth$: 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:

$$\Xi_s = \begin{cases} 0 & \text{with probability } \wp > 0 \\ \boldsymbol{\theta}_s/\wp & \text{with probability } (1-\wp), \text{ where } \log \boldsymbol{\theta}_s \sim \mathcal{N}(-\sigma_{\boldsymbol{\theta}}^2/2, \sigma_{\boldsymbol{\theta}}^2) \end{cases} \tag{36}$$
$$\log \Psi_s \sim \mathcal{N}(-\sigma_{\boldsymbol{\Psi}}^2/2, \sigma_{\boldsymbol{\Psi}}^2)$$

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), $\wp = 0.5/100$, $\sigma_{\boldsymbol{\theta}} = 0.1$, and $\sigma_{\boldsymbol{\Psi}} = 0.1$.[9] 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 3).[10] The interest rate is assumed to equal 1.03. The model parameters are included in Table 1.

The structural estimation of the parameters $\beth$ and ρ is carried out using the

---

[9]Note that $\sigma_{\boldsymbol{\theta}} = 0.1$ is smaller than the estimate for college graduates estimated in Carroll and Samwick (1997) ($= 0.197 = \sqrt{0.039}$) which is used by Cagetti (2003). The reason for this choice is that Carroll and Samwick (1997) themselves argue that their estimate of $\sigma_{\boldsymbol{\theta}}$ is almost certainly increased by measurement error.

[10]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).
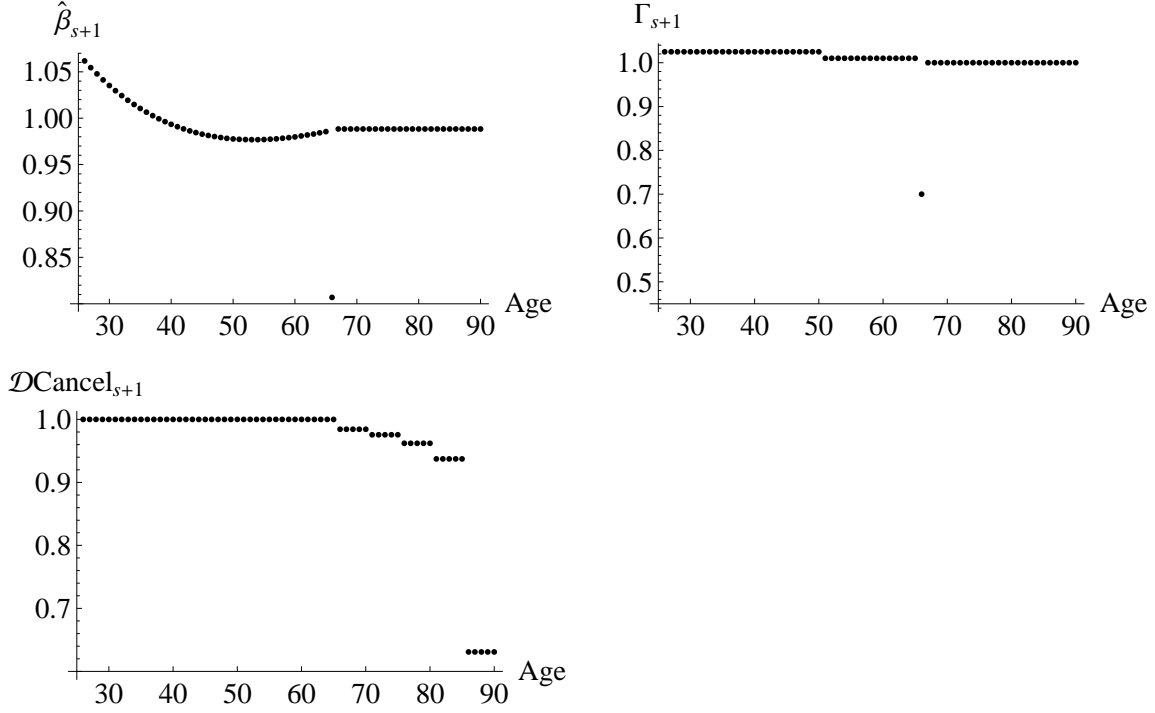
Figure 3 Time Varying Parameters

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

7.2 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 $56-60$.[11] 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

---

[11]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.

Table 1 Parameter Values

$\sigma_{\boldsymbol{\theta}}$	0.1	Carroll (1992)
$\sigma_{\boldsymbol{\Psi}}$	0.1	Carroll (1992)
$\wp$	0.005	Carroll (1992)
$\mathcal{G}_s$	figure 3	Carroll (1997)
$\hat{\beta}_s, \mathcal{L}_s$	figure 3	Cagetti (2003)
R	1.03	Cagetti (2003)

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 model which can be written in Bellman form as:

$$\mathrm{v}_t(m_t) = \max_{c_t} \quad \mathrm{u}(c_t) + \beth\mathcal{L}_{t+1}\hat{\beta}_{t+1} \, \mathbb{E}_t[(\boldsymbol{\Psi}_{t+1}\mathcal{G}_{t+1})^{1-\rho}\mathrm{v}_{t+1}(m_{t+1})]$$

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

$$a_t = m_t - c_t$$

$$m_{t+1} = a_t \underbrace{\left(\frac{\mathsf{R}}{\boldsymbol{\Psi}_{t+1}\mathcal{G}_{t+1}} \right)}_{\equiv \mathcal{R}_{t+1}} + \boldsymbol{\theta}_{t+1}$$

with the first order condition:

$$\mathrm{u}^c(c_t) = \beth\mathcal{L}_{t+1}\hat{\beta}_{t+1}\mathsf{R} \, \mathbb{E}_t \left[\mathrm{u}^c \left(\boldsymbol{\Psi}_{t+1}\mathcal{G}_{t+1}c_{t+1} \left(a_t\mathcal{R}_{t+1} + \boldsymbol{\theta}_{t+1} \right) \right) \right]. \tag{37}$$

The first substantive tock 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 (37) following the procedure in sections **??** and 'Recursion.' The latter routine is slightly complicated by the fact that we are considering a life-cycle model and therefore the growth rate of permanent income, the probability of death, the time-varying discount factor and the distribution of shocks will be different across the years. We thus must ensure that at each backward iteration the right parameter values are used.

Correspondingly, the first part of the `StructEstimation.py` file begins by defining the agent type by inheriting from the baseline agent type `IndShockConsumerType`, with the modification to include time-varying discount factors. Next, an instance of this "life-cycle" consumer is created for the estimation procedure. The number of periods for the life cycle of a given agent is set and, following Cagetti, (2003), we initialize the wealth to income ratio of agents at age 25 by randomly assigning the equal probability values to 0.17, 0.50 and 0.83. In particular, we consider a population of agents at age 25 and follow their consumption and wealth accumulation dynamics as they reach the age of 60, using the appropriate age-specific consumption functions and the age-varying

parameters. The simulated medians are obtained by taking the medians of the wealth to income ratio of the 7 age groups.

To complete the creation of the consumer type needed for the simulation, a history of shocks is drawn for each agent across all periods by invoking the `make_shock_history` function. This involves discretizing the shock distribution for as many points as the number of agents we want to simulate and then randomly permuting this shock vector as many times as we need to simulate the model for. In this way, we obtain a time varying shock for each agent. This is much more time efficient than drawing at each time from the shock distribution a shock for each agent, and also ensures a stable distribution of shocks across the simulation periods even for a small number of agents. (Similarly, in order to speed up the process, at each backward iteration we compute the consumption function and other variables as a vector at once.)

With the age-varying consumption functions derived from the life-cycle agent, we can proceed to generate simulated data and compute the corresponding medians. Estimating the model involves comparing these simulated medians with empirical medians, measuring the model's success by calculating the difference between the two. However, before performing the necessary steps of solving and simulating the model to generate simulated moments, it's important to note a difficulty in producing the target moments using the available data.

Specifically, defining 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

$$\min_{\xi} \sum_{\tau=1}^{7} |\varsigma^{\tau} - \mathbf{s}^{\tau}(\xi)| \tag{38}$$ {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 (38) 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:

$$\min_{\xi} \sum_{i}^{N} \omega_i |\varsigma_i^{\tau} - \mathbf{s}^{\tau}(\xi)| \tag{39}$$ {eq:StructEstim}

where ω_i is the weight of household i in the entire population,[12] 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

---

[12]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.
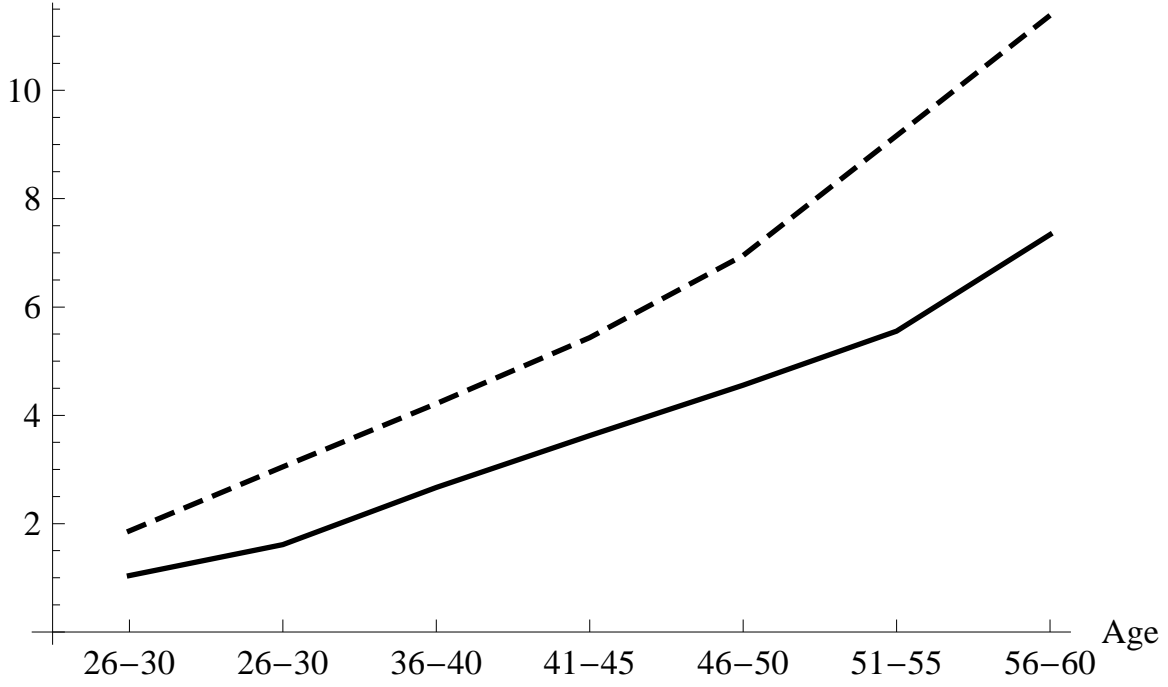
Figure 4 Wealth to Permanent Income Ratios from SCF (means (dashed) and medians (solid))

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 4. 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, `smmObjectiveFxn`, calls the `simulate_moments` function to create the objective function described in (39), 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

17

3. returns the value of equation (39)

We delegate the task of finding the coefficients that minimize the `smmObjectiveFxn` function to the `minimize_nelder_mead` function, which is defined elsewhere and called in the second part of this file. This task can be quite slow and rather problematic if the `smmObjectiveFxn` function has very flat regions or sharp features. It is thus wise to verify the accuracy of the solution, for example by experimenting with a variety of alternative starting values for the parameter search.

The final object defined in this first part of the `StructEstimation.py` file is `calculateStandardErrorsByBootstrap`. As the name suggsts, the purpose of this function is to compute the standard errors by bootstrap.[13] This involves:

1. drawing new shocks for the simulation

2. drawing a random sample (with replacement) of actual data from the SCF

3. obtaining new estimates for ρ and $\beth$

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

7.2.1 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 et al. (2017) notes, a researcher employing such structural empirical methods may be interested in how alternative assumptions (such as misspecification or measurement bias in the data) would "change the moments of the data that the estimator uses as inputs, and how changes in these moments affect the estimates." The authors provide a measure of sensitivity for given estimator that makes it easy to map the effects of different assumptions on the moments into predictable bias in the estimates for non-linear models.

In the language of Andrews et al. (2017), section 7 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 (39). 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.

7.3 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

---

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

the positional arguments `estimate_model` and `compute_standard_errors` to true.[14]
Results are reported in Table 2.[15]

Table 2 Estimation Results

ρ	$\beth$
3.69	0.88
(0.047)	(0.002)

The literature on consumption and saving behavior over the lifecycle in the presenece of labor income uncertainty[16] 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 5 provides a plot of the plug-in estimate of the sensitivity measure described in 7.2.1. 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

---

[14]The procedure is: First we calculate the ρ and $\beth$ estimates as the minimizer of equation (39) using the actual SCF data. Then, we apply the `Bootstrap` function several times to obtain the standard error of our estimates.

[15]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.

[16]For example, see Gourinchas and Parker (2002) for an exposition of this.
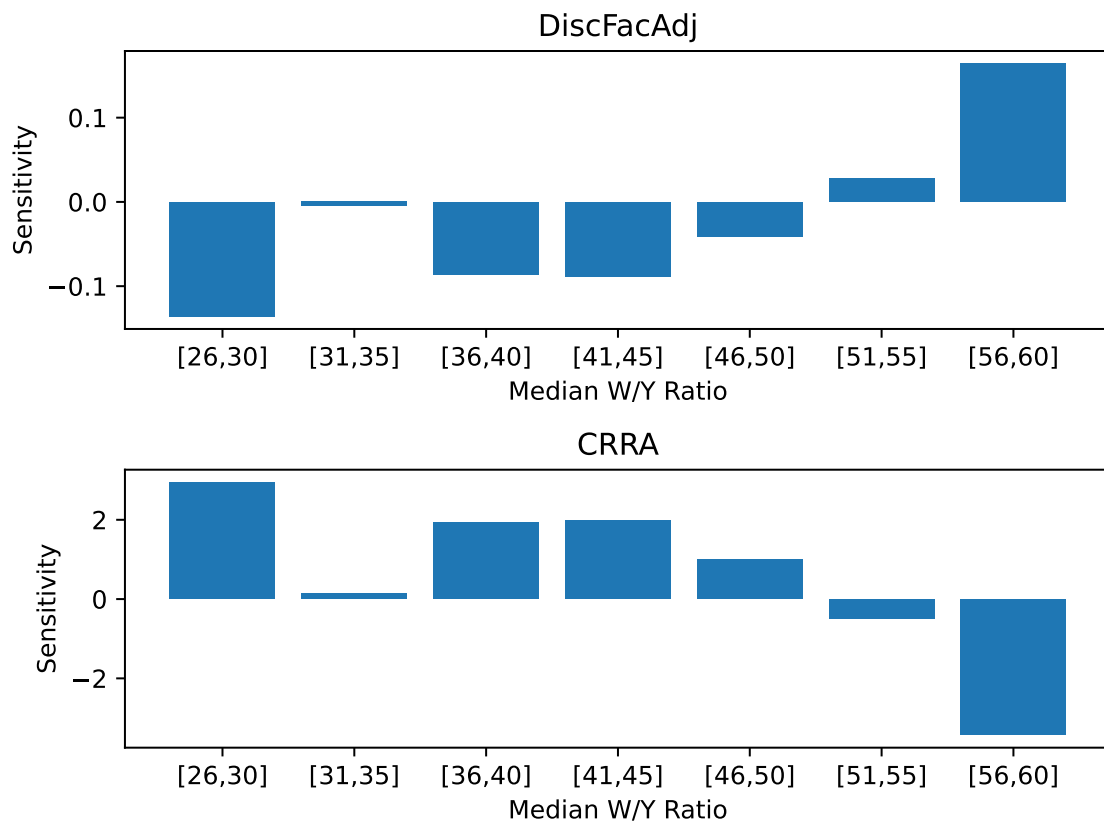
Figure 5 Sensitivty of Estimates $\{\rho, \beth\}$ regarding Alternative Mismeasurement Assumptions.

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 6 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 (39)). Interestingly, there is a large rather flat region; or, more formally speaking, there exists a broad set of parameter pairs which leads to similar simulated wealth to income ratios. Intuitively, the flatter and larger is this region, the harder it is for the structural estimation procedure to precisely identify the parameters.
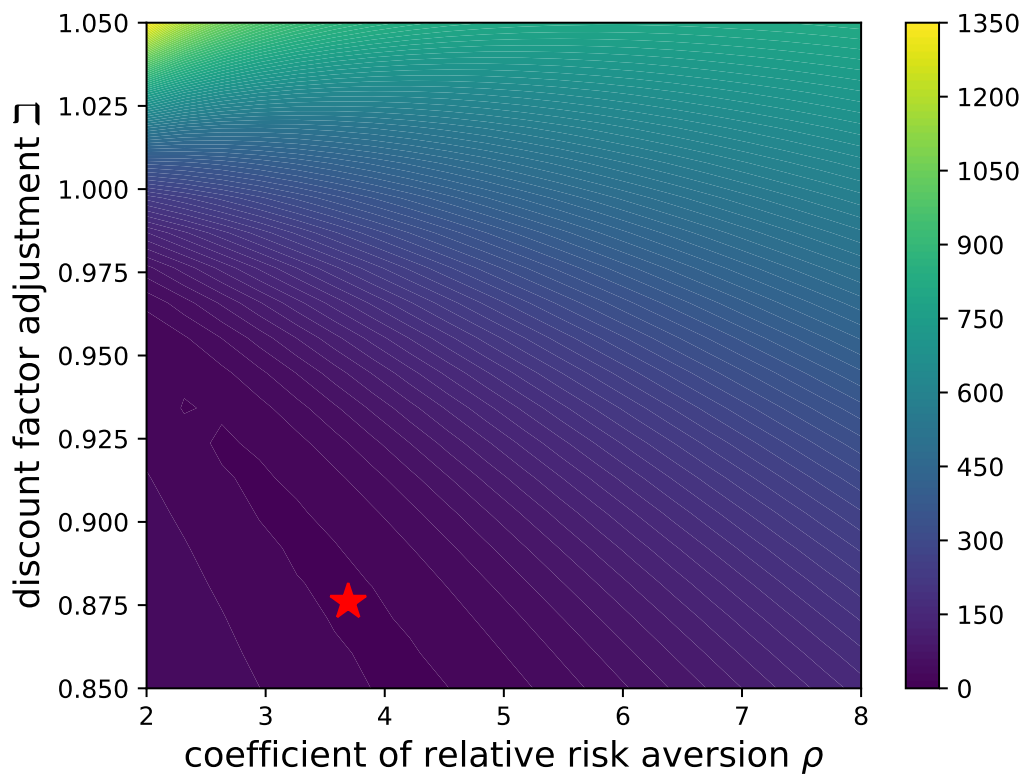
Figure 6 Contour Plot (larger values are shown lighter) with $\{\rho, \beth\}$ Estimates (red dot).

{fig:PlotContourMe

Appendices

A SCF Data

Data used in the estimation is constructed using the SCF 1992, 1995, 1998, 2001 and 2004 waves. The definition of wealth is net worth including housing wealth, but excluding pensions and social securities. The data set contains only households whose heads are aged 26-60 and excludes singles, following Cagetti (2003).[17] Furthermore, the data set contains only households whose heads are college graduates. The total sample size is 4,774.

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

Normal income levels in the SCF are before-tax figures. These before-tax permanent income figures must be rescaled so that the median of the rescaled permanent income of each age group matches the median of each age group's income which is assumed in the simulation. This rescaled permanent income is interpreted as after-tax permanent income. Rescaling is crucial since in the estimation empirical profiles are matched with simulated ones which are generated using after-tax permanent income (remember the income process assumed in the main text). Wealth / permanent income ratio is computed by dividing the level of wealth by the level of (after-tax) permanent income, and this ratio is used for the estimation.[18]

---

[17]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.

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

References

ANDREWS, I., M. GENTZKOW, AND J. M. SHAPIRO (2017): "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 132, 1553–1592.

ATTANASIO, O., J. BANKS, C. MEGHIR, AND G. WEBER (1999): "Humps and Bumps in Lifetime Consumption," *Journal of Business and Economic Statistics*, 17, 22–35.

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

CARROLL, C. D. (1992): "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 1992, 61–156, https://www.econ2.jhu.edu/people/ccarroll/BufferStockBPEA.pdf, https://www.econ2.jhu.edu/people/ccarroll/BufferStockBPEA.pdf.

——— (1997): "Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, CXII, 1–56.

——— (2002): "Portfolios of the Rich," in *Household Portfolios: Theory and Evidence*, Cambridge, MA: MIT Press, https://www.econ2.jhu.edu/people/ccarroll/richportfolios.pdf, https://www.econ2.jhu.edu/people/ccarroll/richportfolios.pdf.

——— (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, 91, 312–320, 10.1016/j.econlet.2005.09.013, https://www.econ2.jhu.edu/people/ccarroll/EndogenousGridpoints.pdf.

——— (2023a): "Solving Microeconomic Dynamic Stochastic Optimization Problems," *Econ-ARK REMARK*, https://llorracc.github.io/SolvingMicroDSOPs.

——— (2023b): "Theoretical Foundations of Buffer Stock Saving," *Revise and Resubmit, Quantitative Economics*, https://zenodo.org/badge/latestdoi/304124725.

CARROLL, C. D., AND M. S. KIMBALL (1996): "On the Concavity of the Consumption Function," *Econometrica*, 64, 981–992, https://www.econ2.jhu.edu/people/ccarroll/concavity.pdf, https://www.econ2.jhu.edu/people/ccarroll/concavity.pdf.

CARROLL, C. D., AND A. A. SAMWICK (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40, 41–71, https://www.econ2.jhu.edu/people/ccarroll/papers/nature.pdf.

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

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

GOURINCHAS, P.-O., AND J. PARKER (2002): "Consumption Over the Life Cycle," *Econometrica*, 70, 47–89.

DEN HAAN, W. J., AND A. MARCET (1990): "Solving the Stochastic Growth Model by Parameterizing Expectations," *Journal of Business and Economic Statistics*, 8, 31–34, Available at http://ideas.repec.org/a/bes/jnlbes/v8y1990i1p31-34.html.

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

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

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

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

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

TANAKA, K., AND A. A. TODA (2013): "Discrete approximations of continuous distributions by maximum entropy," *Economics letters*, 118, 445–450, 10.1016/j.econlet.2012.12.020.

ZELDES, S. P. (1989): "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *Quarterly Journal of Economics*, 104, 275–298.