1 Introduction

Now let’s turn to consumption. We will begin with Friedman on the consumption function. I am going to go over Friedman for two reasons.

First, his way of thinking about consumption is interesting in its own right. But also he sets the agenda for modern macroeconomics here by making a very surprising theoretical finding.

That finding is very simple. But it suggested that the standard way of thinking about macroeconomics at that time was wrong.

I think that economists accepted Friedman too readily without realizing the radical nature of what he was saying.

Let’s begin our review of the consumption function.

The consumption function began with Keynes. Keynes proposed a consumption function in which consumption was a function of current income with a positive marginal propensity to consume.

Following Keynes, the natural thing to do was to estimate the consumption function to see whether it existed in reality, as Keynes had said it would. These estimates provide the background to Friedman’s consumption function.

So economists began looking at consumption and income in the late 1930s and early 1940s.

Plotting points of \((C_t, Y^d_t)\) – that is (consumption, disposable income) pairs. Smithies found a line as follows.

Now if you think about this consumption function for longer than 3 minutes you realize there is something theoretically wrong about it.

It could not be a consumption function for any very poor country. For example, it could not be the consumption function for Bangla Desh because the yearly income in Bangla Desh is close to zero.

Their consumption function, according to this consumption function _indicating Y close to zero_ would be in this region with the net result that they would be borrowing continually.

In any case to continue the history of the consumption function, before such “logical” considerations intruded themselves, an empirical problem was noticed with this consumption function.
Simon Kuznets noticed that over long periods of time the consumption/income ratio was constant.
He found:

<table>
<thead>
<tr>
<th>Year Interval</th>
<th>C/Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1869-99</td>
<td>.867</td>
</tr>
<tr>
<td>1884-1913</td>
<td>.867</td>
</tr>
<tr>
<td>1904-13</td>
<td>.879</td>
</tr>
</tbody>
</table>

So the consumption-income ratio looks very constant.
In contrast, estimates of consumption functions from budget studies, which use cross-section plots of the consumption of households with their respective incomes, also yielded significant positive intercepts for consumption.
The question was why this might be the case.
Three explanations for this were given:
1. The Modigliani-Duesenberry hypothesis that consumption depended on previous peak income.
2. The Modigliani life-cycle hypothesis.
3. Friedman’s permanent income hypotheses.
Friedman’s solution is the most elegant.

2 Consumption under Certainty

Suppose a consumer who will receive income of $Y_1$ in period 1 and $Y_2$ in period 2, and for whom the rate of interest is $r$. Let’s call this person Mrs. Tian.
This person will then maximize her lifetime utility function, which is

$$U(C_1, C_2)$$

subject to the constraint that the present discounted value of her consumption will not exceed the present discounted value of her income.
She will then maximize $U$ subject to her intertemporal budget constraint

$$C_1 + \frac{C_2}{(1 + r)} = Y_1 + \frac{Y_2}{(1 + r)}.$$
This budget constraint just says that the PDV of her consumption is the PDV of her income.

Let’s put $C_1$ on this axis and $C_2$ on that axis.

She is maximizing utility subject to her budget constraint. So she chooses $(C_1, C_2)$ to be on the highest feasible indifference curve, as pictured. There is a remarkable result here.

According to Keynes: $C = C(Y, r)$, where $Y$ is current income.

But according to Friedman

$$C = C(W, r),$$

where $W$ is Friedman’s definition of wealth.

These are very different conclusions, and Friedman has contradicted Keynes.

Who is right? For this model Friedman is right.

How do we know?

Friedman defines wealth as the PDV of Current and Future income.

$$W = Y_1 + [Y_2/(1 + r)].$$

Wealth, $W$, is the intercept of the budget line and $-(1 + r)$ is the slope of the budget line.

The slope and the intercept of the budget line, and the utility (indifference) function together determine current consumption $C_1$. 

3
This yields two insights. Here is the first insight. If short-run income varies, consumption should still be close to constant.

Why? because $C$ depends upon $W$ – not on $Y$ – and $W$ is the discounted sum of many different incomes.

The second insight is that in the long-run $W$ and $Y$ vary together. Therefore $C$ appears as a function of long-run income. For example if income in every period doubles, with $r$ constant, $W$ will also double.

And then Kuznets' finding is also explained. If utility is homothetic, so that there are the same slopes along rays from the origin, then a doubling of wealth at the same rate of interest will cause a doubling of consumption. Consumption will then be proportional to wealth, or the consumption/permanent income ratio will be roughly constant.

Now the Friedman-Fisher model that we have just posed is clearly unrealistic.

It is unrealistic because Mrs. Tian does not know her future income. One would suspect, and Keynes did suspect, that the introduction of uncertainty would change Mrs. Tian's fundamental method of decision making.

We will return to the role of uncertainty later in the lecture.

It is remarkable that a logical argument, such as comes from Friedman's invocation of Irving Fisher, could prove Keynes wrong.

The assertion here is not only that Keynes was close to making a logical error.

Not only did Keynes make that logical error.

So did all of his followers. Those followers include many economists who seemed at the time to have a very good understanding of economics.

We went back and looked at what Keynes actually did say, and this is the most important line about consumption from The General Theory.

The fundamental psychological law.... is that men are disposed, as a rule and on the average, to increase their consumption as income increases, but not by as much as the increase in income.

This means that Keynes based his case not on economics, but on psychology.

We should have been more precise. Keynes' full sentence also has an additional clause. With that clause it reads:

"The fundamental psychological law, upon which we are entitled to depend with great confidence both a priori from our knowledge of human nature and from the detailed facts of experience, is that men are disposed, as a rule and on the average, to increase their consumption as income increases, but not by as much as the increase in income."

So Keynes views the dependence of consumption on current income as derived from something that we know from human nature, and that it also corresponds to our experience, which shows that this is how people usually behave.

Here is a possible characterization of the two schools of thought.

Friedman thinks that people's consumption can be derived only from something called the utility that people get from that consumption.

But that is only one of the factors that affects utility.

In addition to weighing current against future utility, consumers also have a notion as to how appropriate it is for them to consume a certain amount.
They have a view as to how much they should or should not spend. If they receive a given amount of money that they classify as income, they think that they have earned it.

And therefore they think that they have a right to spend that much money.

A compromise between the Keynesian and the Friedman view of consumption would have consumers weigh these two motives for consumption.

Their consumption would be partially determined by optimization of expected utility.

But it would also be partially determined by consumers’ considerations of how much they think that it is appropriate for them to consume.

This corresponds to the theme that I have mentioned earlier. Decisions are determined by the utility returns from them. They are also determined by what i think that i should or should not do.

I think that most of us have two notions about our consumption. We ask ourselves two questions when we go shopping. One question is how much utility we get from what we buy. We weigh that against our current income. But the second question is how much do we think it is appropriate to spend.

A rationalization then of the Keynesian view would be that people think that current income has a very large influence on how much people think is appropriate to spend.

Ricardian Equivalence

A test of this notion is whether it answers Barro’s paradox of Ricardian equivalence.

We shall give you a brief review of the proposition of Ricardian equivalence.

If the government gives a lump-sum transfer to a parent in the current generation and then later increases taxes on the child to pay for it the parent will not change her consumption at all.

To be precise, the parent will not change her consumption as long as her bequest to the child exceeds the amount of the transfer. The reason for this neutrality result is that the opportunity set between utility for the parent and utility for the child will be totally unchanged. Therefore the parent should choose exactly the same utility for herself and for her child with and without the government transfer.

There are very few economists who believe that people behave this way. It suggests that social security should have only minor effects on consumption.

But why not?

There are at least 11 (or so) practical reasons, such as lack of perfect foresight about the future taxes that would falsify Ricardian equivalence.

But each of those reasons for violation of Ricardian equivalence is fairly obvious. Barro could not have published his original paper if there were not also some additional fundamental reason for violation of Ricardian equivalence.

Nor, presumably, would a generation of economists have found Ricardian equivalence interesting.
It would have been all too obvious. The surprise result of Barro's paper is that standard economic methodology gives us a result which almost no one believes would be true.

They do not think it would be true even in the absence of all of the 11 (or so) obvious distortions.

That methodology says that all demand and supply curves (including the consumption function) should be derived from utility functions with only economic considerations.

But here is a reason that, if true, would account both for why Ricardian equivalence might not be true, and also why it would be a surprise.

A bequest is a form of gift. If there is any form of activity that is determined by norms regarding what people think they should or should not do, it is the giving of gifts.

If social security gives more money to the parent, Ricardian equivalence says that the bequest will be larger by exactly the same amount.

The increase in the bequest, or the gift, from the parent to the child will be one-to-one.

Let's suppose that the parent thinks she should make a bequest. Then her utility is higher the larger her bequest.

In this case, if the parent derives utility from making such a bequest, but if the marginal value to the parent declines as it increases in size, the social security transfer will not be neutral in its effect.

The more social security she receives, the more she will leave to her offspring. But the increased bequest will not exactly match the increase in social security. It will be less.

In terms of sociology how the parent frames her view of what she should do, which includes how much she thinks she should spend of what she considers to be her money will determine the size of her bequest.

When she receives a higher social security check she thinks that she has more money, and that it is appropriate for her to divide that extra money between her own expenditure and the giving of a bequest.

The size of her bequest depends then upon how much money she thinks is hers, and how she thinks she should or should not spend that money.

That is an observation that is straight from Pareto.

In my view economists have repeatedly found the consequences of classical economics surprising because they never appreciated the limitations of the classical model. They have never appreciated the role of norms in people's utility functions. If this is true it does not just apply to the consumption function, but potentially to all of the behavioral relations in macroeconomics.

I see the absence of norms as explaining a great deal of what I consider to be puzzling about modern macroeconomics.

I offer you this because whether or not you accept it, it is potentially very useful to have a perspective that is nonstandard. That allows you better to appreciate what is right or possibly wrong about the standard model.

I cannot guarantee that these views will be widely accepted.
Having discussed the main theme in consumption under certainty, now we give a more formal derivation so that we can have see both forests and trees.

Individual who lives for $T$ periods has utility

$$ U = \sum_{t=1}^{T} u(c_t), \ u'>0, \ u''<0. $$

His budget is

$$ \sum_{t=1}^{T} c_t \leq A_0 + \sum_{t=1}^{T} Y_t, $$

where $A_0$ is his endowment and $Y_t$ is his income in the period $t$.

The Lagrangian for his maximization problem is

$$ L = \sum_{t=1}^{T} u(c_t) + \lambda (A_0 + \sum_{t=1}^{T} Y_t - \sum_{t=1}^{T} c_t). $$

The first-order condition for $c_t$ is

$$ u'(c_t) = \lambda. $$

As a result,

$$ c_1 = c_2 = \ldots = c_T $$

or

$$ c_t = \frac{1}{T}(A_0 + \sum_{t=1}^{T} Y_t) \text{ for any } t. $$

This implies the individual’s consumption in a given period is NOT determined by his income in that period, BUT by his income over his entire lifetime. This is Friedman’s Permanent Income Hypothesis (PIH).

### 3 Consumption under Uncertainty

Now we move on to uncertainty.

As we said before, Friedman’s dependence of consumption on wealth did not deal with uncertainty of future income.

Will such dependence make a difference? There is a guess that it will, but that conjecture is probably wrong. With the standard presentation of uncertainty it makes probably little difference.

Here is how uncertainty would be entered into Friedman’s model according to the standard model.

The consumer would choose $c_1$ to maximize

$$ E[U(c_1, \tilde{c}_2)] $$

where $\tilde{c}_2$ is a random variable which must conform to the constraint that

$$ PDV[\text{consumption}] = PDV[\text{income}]. $$

Future income is stochastic, however, and this makes future consumption stochastic. So
\[ PDV(c_1, \tilde{c}_2) = PDV(y_1, \tilde{y}_2), \]

or

\[ c_1 + [\tilde{c}_2/(1 + r)] = y_1 + [\tilde{y}_2/(1 + r)]. \]

There is now a result due to Theil.
Suppose utility is quadratic.
Suppose, for example, utility is of the form

\[ U = ac_1 - bc_1^2 + ac_2 - bc_2^2. \]

Then the choice of \( c_1 \) that maximizes

\[ E[U(c_1, \tilde{c}_2)] \]

is only dependent on \( y_1, E(y_2) \), and \( r \). In other words the choice of \( c_1 \) is independent of the distribution \( y_2 \) about its mean – provided utility is quadratic.

This is indeed such a surprising result that we will talk about it at some length.

First, let me give you an alternative statement of the result.
Let there be two persons with the same quadratic utility function.
We will call them Mr. and Mrs. Tian.
There is the same rate of interest for these two persons.
They also have the same income in period one.

Mrs. Tian knows her income with certainty in period 2. It is \( \tilde{y}_2 \). In contrast, Mr. Tian knows \( E(\tilde{y}_2) = \tilde{y} \). But he is uncertain about the actual value of \( y_2 \), which is a random variable.

Then the result above claims that Mr. and Mrs. Tian will choose the same consumption in Period 1. Most persons do not know this result and consider it counterintuitive, or perhaps we should say wrong. So I am going to spend some time going over this so that if this is the case you see why your intuition is messing you up.

Let me illustrate with the simplest possible example.
A child on her birthday is allowed all the candies she wants. Her utility is \( U = ac - bc^2 \), where \( c \) is her candy consumption. And so she chooses

\[ \partial U/\partial c = 0 \Rightarrow a - 2bc = 0 \]

or \( c = a/(2b) \).

Now let’s contrast this with an uncertainty story. This girl’s brother on his birthday gets to choose surprise packages of candy of different sizes. (He has to eat all the candy contained in the package because his parents will not let him waste “food.”) With probability 1/2 a package of size \( c \) contains \( c + \varepsilon \). With probability 1/2 a package of size \( c \) contains \( c - \varepsilon \). This boy chooses packages to maximize his utility. If he does, he chooses \( c \) so that

\[ E(MU(\tilde{c})) = 0 \]
where $MU$ is his marginal utility, and $\tilde{c}$ is the stochastic amount of candy in his box. This amounts to choosing so that

$$1/2[a - 2b(c + \varepsilon)] + 1/2[a - 2b(c - \varepsilon)] = 0.$$ 

The first term here is his marginal utility if his box has a positive $\varepsilon$ candy-shock; the second is if his box gives him a $-\varepsilon$ candy-shock. This marginal condition amounts to:

$$a - 2bc = 0.$$ 

So that the brother chooses $c = a/(2b)$. His expected candy consumption then is exactly the same as his sister’s. The presence of uncertainty for the brother makes absolutely no difference in the amount of his candy consumption.

However, this DOES NOT mean that people like uncertainty. Indeed in our example the brother does not like his uncertainty. We can calculate his total utility and compare it to his sister’s. The brother’s expected utility will be

$$E(U) = a^2/4b - bc^2.$$ 

In contrast his sister’s utility is: $U = a^2/4b$.

So the brother with the uncertainty would be willing to pay his parents $bc^2$ worth of utility not to give him surprise packages.

So you see here an example in which Friedman is basically correct. The introduction of uncertainty does not fundamentally alter decision making.

The key reason that the choice is basically unchanged is clear in this candy example.

Since marginal utility is linear:

$$E[MU(\tilde{c})] = MU[E(\tilde{c})]$$

As a result the condition that $E[MU(\tilde{c})] = 0$ does not depend at all on the distribution of consumption, only on its expected value.

A slightly more complicated proof goes through in the case of Mrs. Tian and Mr. Tian, who is each deciding on her/his respective level of consumption.

Mrs. Tian’s optimal consumption choice $(c_1^*, c_2^*)$ consumption must satisfy two conditions:

1. $c_1^* + [c_2^*/(1 + r)] = y_1 + [y_2/(1 + r)]$.
2. $MU(c_1) = (1 + r)MU(c_2)$, or $a - 2bc_1^* = (1 + r)[a - 2bc_2^*]$.

The budget line is the first equation. The tangency condition is the second. These are two linear equations in two unknowns, $(c_1^*, c_2^*)$. We can solve for them.

Now let’s look at Mr. Tian’s problem.

His consumption choice must first satisfy the budget constraint:

$$c_1 + [\tilde{c}_2/(1 + r)] = y_1 + [\tilde{y}_2/(1 + r)]$$
so
\[ c_1 + [E(\tilde{c}_2)/(1 + r)] = y_1 + [E(\tilde{y}_2)/(1 + r)]. \]

And for the optimum,
\[ MU(c_1) = (1 + r)E[MU(\tilde{c}_2)] \]

and because \( MU \) is linear \( E[MU(\tilde{c}_2)] = MU[E(\tilde{c}_2)] \) and \( MU(c_1) = (1 + r)MU[E(\tilde{c}_2)] \). So we have
\[ a - 2bc_1 = (1 + r)[a - 2bE(\tilde{c}_2)] \]

We now have two linear equations for Mr. Tian, for \( c_1 \) and for \( E(\tilde{c}_2) \).

These are the exact same equations that we had for Mrs. Tian, in terms of \( c_1^* \) and \( c_2^* \).

Therefore Mr. and Mrs. Tian choose exactly the same value of \( c_1 \).

Thus with quadratic utility the presence of uncertainty in income has no effect on the choice of first-period consumption.

As you saw in the proof the key ingredient to this result is that \( MU \) is linear, so uncertainty has no effect on marginal conditions.

That is why quadratic utility yields certainty equivalence.

With quadratic utility, since Marginal Utility is linear, the marginal condition is unaffected by the presence of uncertainty.

Let me now make three comments on Certainty Equivalence.

Comment 1. This first comment answers a question that you must have asked yourself.

Question: Why should we care about a result dependent on a special utility function?

Answer: There is a sense in which all second differentiable functions are approximately quadratic. In this sense certainty equivalence is also approximately true. With a second-differentiable utility function the effect of uncertainty on the optimal choice of \( c_1 \) becomes negligible relative to the spread of the uncertainty, as uncertainty becomes small.

This theorem is a result of a Taylor series expansion and has been proved by Edmond Malinvaud in Econometrica.

In macroeconomics we are very frequently using some approximation based on calculus. We have already seen quite a few of these approximations by Sargent and Taylor and in near-rationality.

I find that students usually have a fear of approximation. For that reason I am going to take a few minutes on the meaning of calculus, at least for economics.

I want to give you a dramatic example that shows you that approximations really do work.

Suppose that we have a nice second differentiable function \( f(x) \).
\[ f(x_0 + \varepsilon) = f(x_0) + \varepsilon f'(x_0) + \varepsilon^2/2 f''(y) \]

where \( x_0 \leq y \leq x_0 + \varepsilon \).
This theorem is often taken to say that in the limit as $\varepsilon \to 0$,
\[ f(x_0 + \varepsilon) = f(x_0) + \varepsilon f'(x_0) \]
because the remainder term, $\varepsilon^2 / 2f''(y)$ becomes arbitrarily small relative to $\varepsilon$, as $\varepsilon \to 0$.

Comment 2. Quadratic utility is not homothetic\(^1\) and homotheticity was needed to derive Friedman’s other result. Friedman derived the result that
\[ C = k(r, u)W. \]

That consumption is proportional to wealth, with that proportion dependent on the interest rate and $r$ and the utility function $u$. He had assumed that utility function was homothetic.

As long as the utility function is second-differentiable then it is approximately quadratic and certainty equivalence is likely to be a somewhat good approximation.

Comment 3: Certainty equivalence is not the same thing as risk neutrality. Certainty equivalence says that Period 1 actions are independent of uncertainty.

In contrast, risk neutrality says that people do not care whether or not there is uncertainty.

You may recall that the little boy with the surprise birthday candy, and the little girl with the certain amount of birthday candy chose the same size packages. So there was certainty equivalence. But the little boy had less expected utility, so he was not risk neutral.

Having completed our discussion of Friedman’s Chapter II. We shall now proceed to Friedman’s Chapter III. Then we will also discuss Bob Hall’s Rational Expectations Permanent Income Hypothesis.

In that I will use the notion of certainty equivalence that we developed last time.

in Chapter III Friedman sets up his macro model of the consumption function. I want to go over that because Friedman has interesting ways of looking at things that are both modern and old at the same time.

In this Chapter Friedman sets up his macro model of consumption, based at least to some extent on his micro analysis of Chapter II.

He says:

the long-run relationship between consumption and income is:
\[ c_p = k(r, u)y_p, \]
where $c_p =$ permanent consumption
\[ y_p = \text{permanent income}. \]
He defines $y_p = rW$.

$r$ is of course the interest rate and $u$ is the utility function.

We derived this result last time for homothetic utility functions.

\(^1\)We say a preference relation is homothetic if $x \sim y \iff ax \sim ay$. 
Consumption with a given utility function and a given interest rate is proportional to wealth. So permanent consumption is proportional to permanent income.

Friedman then forgets about the things that are in the parentheses, so

\[ c_p = ky_p. \]

However, in the short run there are random deviations of various sorts from this relationship.

People have random deviations of their consumption from their permanent levels so current consumption is

\[ c_c = c_p + c_t \]

Current consumption is:

\[ c_c = (\text{permanent consumption})c_p + (\text{transitory consumption})c_t \]

Similarly, in any period current or actual income may deviate from its long-run value by something random.

\[ y_c = y_p + y_t \]

So current income is also the sum of a permanent component and a transitory component. And in modern language the transitory components are —pure noise” so that

\[ \rho(y_p, y_t) = \rho(c_p, c_t) = \rho(y_t, c_t) = 0 \]

These assumptions constitute the Friedman model.

Given these assumptions an estimated short-run consumption function will have a positive intercept and a lower slope than the long-run consumption function.

For Friedman, this lower slope and positive intercept correspond to the empirical findings from budget-studies and time-series estimates of the consumption function. So this "model" of consumption seems to fit the facts.

We can explain the model in four different ways, all of which have the same meaning. We shall list these four ways.

Explanation I is a story about horse races.

The key to this explanation is Friedman’s sentence on page 35:

"The winners in any particular set of races may well be better off than the losers, but they are also likely to have more than their share of good luck."

I am going to review for you how that sentence is at the heart of Friedman’s story.

If you think about this sentence for a moment you will see what it means.

Consider two runners who run a lot of races. Call them Mrs. Tian and Mr. Tian.
Let’s suppose that both Mrs. Tian and Mr. Tian are real stars. Together they win all the races.

Mrs. Tian wins half the time; Mr. Tian wins half the time. Let’s suppose that Mrs. Tian runs the mile in 4 minutes when she is fast, and in 4.5 minutes when she is slow. Similarly, Mr. Tian runs the mile in 4 minutes when he is fast, and in 4.5 minutes when he is slow.

When Mrs. Tian wins, she is running the 4 minute mile, not the 4.5 minute mile. Similarly, when Mr. Tian wins, he is running the 4 minute mile, not the 4.5 minute mile. Now let’s return to Friedman’s statement.

1. The winners are on average faster than the losers. Yes, both Mrs. Tian and Mr. Tian are indeed very fast runners.

2. But the winners are also having more than their share of luck. Yes. Why? Because when Mrs. Tian wins she is (usually) faster than her average. Similarly, when Mr. Tian wins he is (usually) faster than his average.

Let me be specific about the analogy.

\[ s_c = s_p + s_t \]
\[ s_p \] = average speed
\[ s_c \] = current speed in a particular race
\[ s_t \] = transitory speed

When Mr. and Mrs. Tian win, their transitory speed will be negative. The average speed of each is 4:15. So \( s_p = 4:15 \). When Mrs. Tian, or Mr. Tian is fast, their speed is 4:00 minutes. And the transitory speed is -0.15 minutes.

We shall use this analogy to look at current income.

Let’s make a graph with \( c_c \) and \( c_p \) on the vertical axis and \( y_c \) and \( y_p \) on the horizontal axis.

On average, persons with high current incomes have high current incomes for two reasons: (1) because their permanent income is high; (2) because they are lucky.

For persons with high current income, say \( y''_c \), average transitory income is positive.

Because their average transitory income is positive, their average permanent income \( y''_c \) will be less than their current income \( y''_c \).

Similarly, for persons with low current income, their income is low for two reasons: (1) because their permanent income is low; (2) because their transitory income is low. Persons with low current income \( y'_c \) have low permanent income. But, on average, they have also had bad luck, so that the permanent income of those with the low current income \( y'_c \) is \( y'_p \). And \( y'_p \) is greater than \( y'_c \).

Now let’s add to the diagram Friedman’s key behavioral relation: \( c_p = ky'_p \)

Persons with low current income \( y'_c \) have on average permanent income \( y'_p \) and therefore their current consumption will be \( ky'_p \).

The econometrician records that those people have current income \( y'_c \) and average consumption \( ky'_p \).

Similarly, people with high current income \( y''_c \) on average have high permanent income \( y''_p \).

Their average consumption is \( ky''_p \).
The econometrician who sees the average level of consumption for those people with income $y_0$ will see a point like this. And if the econometrician takes a regression of current consumption on current income, she gets a line like this. This line has a positive intercept and a lower slope than the true long-run consumption function.

This explains why econometricians think that they see consumption functions with these positive intercepts even though the "true" relation is: $c_p = ky_p$. Let me now give you Explanation II of Friedman’s consumption function.

We will call this the Friday binge interpretation. I will give you Friedman’s view of the matter. Suppose that you get paid $100$ every Friday, at which time you spend $40$ of your income. On all the other six days of the week you spend $10$ of your income.

Your "true" consumption function is: $c_p = y_p$.

You hire a stupid Keynesian econometrician, like those who estimated the first consumption functions, to estimate your consumption function. The stupid econometrician sees that 6 days of the week you receive 0 and spend $10$. He happily marks 6 x’s for 0 income and $10$ expenditure. He sees one day with $100$ income and $40$ expenditure. Then the stupid econometrician connects the x’s.

And he calls it a consumption function. How does that relate to Friedman?

Friedman is an old-style National Bureau of Economic Research economist. The old-style people at the NBER, like isley Clair Mitchell, Geoffrey Moore, and Arthur Burns spent their lives worrying about trends and cycles.

In this view

$y_p = $ trend income
$y_d = $ deviation from trend income
$y_c = $ current income
\[ c_p = \text{trend consumption} \]
\[ c_t = \text{deviation from trend consumption} \]
\[ c_c = \text{current consumption}. \]

According to Friedman, over the course of the business cycle, the econometrician has the Friday-binge problem. By that I mean that looking at plots of \((y_c, c_c)\) all you can find is a meaningless regression telling how people time their purchases over the course of the business cycle.

Such regressions are not revealing as to whether an increase in income in the short run or in the long run will result in corresponding increases in consumption. Thus, in no way, are they a consumption function.

The third interpretation of Friedman is what I am sure almost all of you have seen in your econometrics courses. This is a classic example of the errors in variables problem.

I will give that to you very briefly because I am sure that it is so familiar.

The true model is:

\[
\begin{align*}
c_p &= ky_p \\
c_c &= c_p + c_t \\
y_c &= y_p + y_t \\
\rho(y_p, y_t) &= \rho(c_p, c_t) = \rho(y_t, c_t) = 0
\end{align*}
\]

But the econometrician estimates the coefficient \(k\) by hypothesizing that

\[ c_c = a + by_c + \varepsilon. \]

The coefficient \(k\) is now mis-estimated because, as you can verify, there is a correlation between the error term, which is \(-ky_t + c_t\) and the independent variable, which is \(y_c = y_p + y_t\).\(^2\)

\(^3\)

We have one more interpretation of Friedman.

Explanation IV.

There is a further explanation of the error. In addition to an errors-in-variables problem the econometrician has made a specification error. The econometrician has mis-specified the lag structure.

The true model is: \(c_c = ky_p + c_t\), where \(y_p\) is a weighted average of current and past incomes. In Friedman’s actual empirical work:

\[ y_p = (1 - \theta) \sum_{i=0}^{\infty} \theta^i y_{t-i} \]

\(^2\)FOOTNOTE:

The true relation is:

\[ c_c = c_p + c_t = ky_p + c_t. \] So \(c_c = k(y_c - y_t) + c_t = ky_c - ky_t + c_t. \) So the error term is \(-ky_t + c_t\). The error term is correlated with the independent variable which is \(y_c = y_p + y_t\). The two transitory components are correlated, which gives a statistical bias in estimating \(k\).

\(^3\)How do we know that \(y_t\) is negative for low values of \(y_c\) and positive for high values of \(y_c\)? There is a correlation between \(y_t\) and \(y_c\). We know that because: \(\text{cov}(y_t, y_c) = \text{cov}(y_t, y_p + y_t) = \text{var}(y_t) > 0\) Therefore for high values of \(y_c\), \(y_t > 0\), for low values of \(y_c\), \(y_t < 0\).
The omission of the lagged variables which should have been included in the consumption-income regression is a specification error which leads to a mis-estimate of the parameter $k$.

### 3.1 Hall on Random Walk

Let’s now turn to Hall’s article on consumption.

Hall’s article pictures consumption as a random walk with trend.

Hall, in fact, makes two related propositions.

Somewhat sloppily, his article is usually summarized by saying that he shows consumption follows a random walk.

To remind you, a variable $x_t$ is said to follow a random walk if the value of $x$ at $t$ is the value of $x$ at $t - 1$ plus a random term, which is uncorrelated with past random terms.

In equations:

$$x_t = x_{t-1} + \varepsilon_t$$

where $\varepsilon_t$ is an i.i.d. random variable with mean zero. If $E(\varepsilon_t) \neq 0$, then $x_t$ is said to follow a random walk with drift. Let me be more accurate as to what Hall actually says. He says

$$c_t = \alpha + \beta c_{t-1} + \varepsilon_t$$

where $\alpha$ is close to zero and $\beta$ is close to one, and no variable, besides last period’s consumption $c_{t-1}$ would be useful in predicting future consumption. In particular, this means that the residual $\varepsilon_t$ is uncorrelated with any variable whose value is known at $t - 1$, or earlier.

This view of consumption entails a deep insight. Statement I. Old view of consumption due to Keynes.

$$C = C(Y)$$

The new view of consumption:

$$c_t = \alpha + \beta c_{t-1} + \varepsilon_t$$

where $\alpha \leq 0$, and $\beta \geq 1$, and $\varepsilon_t$ is uncorrelated with series known in previous periods.

Statement II: Old view: if $Y$ changes, consumption changes. New View: if $Y$ changes, consumption changes only insofar as the changes in income were unanticipated. (Note how similar it is to the effect of money on real output in rational expectation school’s view)

Furthermore, the changes in consumption must be a white noise error term uncorrelated with past values. Otherwise, the changes would have been anticipated.

With this introduction, let’s begin to go over Hall’s article in some detail. We will use Hall’s notation rather than David Romer’s because it will be more useful for a few of the things that we have to say.
Here is the model. A consumer of age $t$, who expects to live to age $T$, maximizes her intertemporal utility function:

$$U = E\left[\sum_{j=0}^{T-t} \frac{1}{(1 + \delta)^j} u(c_{t+j})\right]$$

subject to the constraint

$$PDV(c) = PDV(\text{wages}) + A_t.$$ 

The interest rate is known with certainty, is constant in all periods, and is denoted $r$. $A_t$ is her assets at time $t$. The fundamental marginal equilibrium condition (called the Euler equation) from this maximization is:

$$u'(c_t) = [(1 + r)/(1 + \delta)]E_t[u'(c_{t+1})].$$

Let me give an explanation for this equality. Suppose $u'(c_t) < [(1 + r)/(1 + \delta)]E_t[u'(c_{t+1})]$ Then consider the following action: take away one unit of consumption at $t$ and add $(1 + r)$ units of consumption at $t+1$. The expected increase in capital $u$ maximizes her intertemporal utility function:

$$\text{maximize } u(c_t).$$

Therefore $u'(c_t)$ cannot be less than $[(1 + r)/(1 + \delta)]E_t[u'(c_{t+1})]$. if the consumer is maximizing.

Similarly, if $u'(c_t) > [(1 + r)/(1 + \delta)]E_t[u'(c_{t+1})]$, an increase of $c_t$ by one unit and a decrease of $c_{t+1}$ by $(1 + r)$ units will result in an increase in expected intertemporal utility $U$ by $u'(c_t) - [(1 + r)/(1 + \delta)]E_t[u'(c_{t+1})] > 0$.

In consequence if expected intertemporal utility is being maximized, then we must have

$$u'(c_t) = [(1 + r)/(1 + \delta)]E_t[u'(c_{t+1})].$$

There are 3 immediate and important corollaries.

We first get corollary 2.

**Corollary 1** $u'(c_{t+1}) = \lambda u'(c_t) + \varepsilon_{t+1}$, where $\lambda = (1 + r)/(1 + \delta), E_t[\varepsilon_{t+1} | \theta_t] = 0$. Here $\theta_t$ denotes the information set available at time $t$.

**Proof.** Hall’s key marginal condition is

$$E_t[u'(c_{t+1}) | \theta_t] = [(1 + r)/(1 + \delta)]u'(c_t).$$

Another way to write this, which is more precise in our notation is:

$$E_t[u'(c_{t+1}) | \theta_t] = [(1 + r)/(1 + \delta)]u'(c_t).$$

Now nothing stops me from adding and subtracting a term in curlicue brackets to the LHS and the RHS of this equation.

$$E_t[u'(c_{t+1}) | \theta_t] + \{u'(c_{t+1}) - E_t[u'(c_{t+1}) | \theta_t]\}
\quad = [(1 + r)/(1 + \delta)]u'(c_t) + \{u'(c_{t+1}) - E_t[u'(c_{t+1}) | \theta_t]\}
\Rightarrow u'(c_{t+1}) = [(1 + r)/(1 + \delta)]u'(c_t) + \varepsilon_{t+1}$$
where $\varepsilon_{t+1} = u'(c_{t+1}) - E_t[u'(c_{t+1}) | \theta_t]$.

But I know from our discussion of Sargent that what Rational expectations is all about is that expressions such as

$$E_t[\varepsilon_{t+1} | \theta_t] = 0.$$  

Why?

Look at $\varepsilon_{t+1}$. It is the deviation of the actual value of the Marginal Utility of consumption at $t+1$ from the expectation of the Marginal utility of consumption at $t+1$, with the expectations made at time $t$. By definition

$$E_t[\varepsilon_{t+1} | \theta_t] = u'(c_{t+1}) - E_t[u'(c_{t+1}) | \theta_t]$$

and this is zero given rational expectations.

How do we know that this is zero? Remember that earlier, exactly analogously, we showed in Sargent’s model

$$E_t\{[p_{t+1} - E(p_{t+1} | \theta_t)] | \theta_t \} = 0.$$  

Just formally substitute $u'(c_{t+1})$ for $p_{t+1}$ in this expression and the proposition is exactly the same. There was nothing special about $p_{t+1}$ that made the result specially true for that variable.

With rational expectations the expected deviation of a variable $x$ from its expected value based on prior information is always zero. For any variable $x$

$$E_t\{[x_{t+1} - E_t(x_{t+1} | \theta_t)] | \theta_t \} = 0.$$  

Why?

If this statement is not true then one could have made a better prediction of $x$ based on prior information.

This now takes us to Hall’s Corollary 3.

Using Romer’s notation now, if utility is:

$$u = c - ac^2$$

we find

$$c_{t+1} = \alpha + \beta c_t + \varepsilon_{t+1}$$

where $\alpha = [1/(2a)][(r-\delta)/(1+r)]$ and $\beta = (1+\delta)/(1+r)$. Note that if $\delta = r$, $\alpha = 0$ and $\beta = 1$, so we get exactly

$$c_{t+1} = c_t + \varepsilon_{t+1}.$$  

In this case consumption is an exact random walk.

Corollary 5 says that Corollary 3 is approximately correct even without quadratic utility.

Why? because all second differentiable functions are approximately quadratic, which is something we know from Taylor series expansion.
Hall’s approach to consumption is potentially quite useful since it gives us a way of checking whether people behave in a Keynesian way or whether they behave according to the Rational Expectations/Permanent income Hypothesis.

Yet more important his equation relating the marginal utility of consumption today to the expected marginal utility of consumption tomorrow is a vast simplification that enables us to solve for consumption streams even in the presence of stochastic incomes.

This allows us to evaluate such things as the deadweight losses from capital income taxation even when income is stochastic.

It turns out that these deadweight losses are an order of magnitude different when one takes into account the stochastic nature of income.

In a nutshell, Hall says it is exactly true if utility is quadratic and people have rational expectations and maximize their intertemporal utility function that a test for the RE/PIH comes from the regression:

\[ c_t = \alpha + \beta c_{t-1} + \gamma Z_{t-1} + \varepsilon_t. \]

A test of the RE/PIH is that

\[ \alpha = 0 \]
\[ \beta = 1 \]
\[ \gamma = 0 \]

for any series of \( Z_{t-1} \), which would be known to consumers at \( t - 1 \).

Hall does such a test.

His consumption is per capita consumption on non-durables and services in constant dollars.

His \( Z \)-variable is past income. To be precise it is per capita disposable income deflated by the implicit deflator for non-durables and services.

Hall gets good results that seem to support well the RE/PIH.

For example, he does one test with just disposable income lagged once as the \( Z \) variable.

He obtains:

\[ c_t = -16 + 1.024 c_{t-1} - 0.010 y_{t-1} \]

\[ R^2 = 0.9988, D.W. = 1.71. \]

Some Comments on Hall.

A great deal of work has been done, usually showing that the RE/PIH is rejected rather than accepted.

But it should not be surprising that it would be fairly easy to reject with Hall’s test.

After all, \( Z \) may be any variable at all.

There is a very large literature that has followed Hall, that purports either to attempt to accept or reject his hypothesis.
Point 1.

This may be the most important question about Hall's test: what is the economic meaning of a rejection of the RE/PIH using such a regression?

Suppose I reject the hypothesis on the basis of a statistical test. Does that tell us about the accuracy of the predictions of the RE/PIH? Or does it mainly reflect instead the power of the test?

If there is one single person somewhere in the world who does not follow the RE/PIH then you can reject it. The hypothesis is not true. Think about it: That person might even be your very own grandmother, or grandfather. It could even be your brother or sister.

But such a rejection is not necessarily interesting. Statistical failure to reject a hypothesis may tell you more about the amount of data you have and the power of your econometric test than whether the hypothesized behavior is (or is not) a good approximation to reality.

But that rejection is not interesting. This is a classic of one of those cases that you learned about in econometrics class, where what is important is the size of the coefficients, not just the statistical significance.

Many of the best empirical papers are not about statistical testing of some hypothesis: they are establishing the size of some variable. I will give you three examples:

1) Angrist and Krueger have, presumably, established that the returns to an extra year of education are about 6 percent. (Perhaps this is the reason why you are sitting here!)

2) Feldstein established that the value of social security wealth is very large--sufficiently large relative to the value of other assets that anyone would think that it would have a very important impact on savings.

3) Economists ignored for a shockingly long time the very large changes in income distribution that were occurring in the US and other Anglo-Saxon countries in the 1970s and 1980s. These changes were finally recognized by Frank Levy. The most interesting finding in the paper by Katz and Murphy that is down the reading list is not the test of some hypothesis, but instead the absolute size of the changes in relative incomes of young high school and young college graduates.

The words BIG and SMALL sometimes have meaning. Some of the best theoretical and also empirical economics is about establishing that some number is BIG, or some number is SMALL.

In other words: Hall's article is about hypothesis testing, but not all data work is hypothesis testing.

Much of what I really believe from analysis of data is more likely establishing whether some variable is big or small.

Point 2.
The second point that I want to make concerns the power of Hall’s test. I want to make this point for two reasons.  
First, I think that power of tests is an interesting and important topic.  
We are going to return to it at some length in going over a very fine paper by Larry Summers on the power of tests in financial markets.  
Second, I think that you probably see too little of a very nice concept that was developed at the University of Chicago, which is the stock adjustment model.  
And we can use this point to illustrate the stock adjustment model.  
For some reason or other, it turns out to be useful.  
The lack of power of the test may be a surprise.  
Just because I found that the coefficient of $c_{t-1}$ to be close to one, and the coefficients on other variables to be, arguably, close to zero does not necessarily mean that alternative consumption functions that are very different from the RE/PIH do not also conform to the data.  
We will use a very simple example to illustrate this point.  
This example as advertised earlier, is based on a stock adjustment consumption function.  
Long-run desired consumption is denoted $c^*_t$. And $c^*_t$ is proportional to last period’s income:

$$c^*_t = \theta y_{t-1}.$$  

And then, according to the hypothesis, consumption adjusts slowly to the long-run desired level

$$c_t - c_{t-1} = \alpha (c^*_t - c_{t-1}) = \alpha (\theta y_{t-1} - c_{t-1})$$  

$$\Rightarrow c_t = (1 - \alpha)c_{t-1} + \alpha \theta y_{t-1}$$

If $\alpha$ is small, the coefficient on $c_{t-1}$ is close to one and the coefficient on $y_{t-1}$ is close to zero. This is then an alternative model that will produce results that are very close to those of Hall.  
A formal way of putting it is that in this case Hall’s test has low power relative to this natural alternative.  
A study by Akerlof etc shows that Hall’s test has power only 234 out of 1000 against the standard Keynesian stock adjustment consumption function that it was meant to replace.  
To be just a bit imprecise, 0.234 is the fraction of the time that the RE/PIH would be rejected by Hall’s test if the true model for consumption were the alternative:

$$c_t - c_{t-1} = 0.1(0.8y_{t-1} - c_t) + \varepsilon_t.$$  

It could be that Hall’s consumption function and the Keynesian consumption function passed the same test because they are in fact very similar.
So before passing on, I want to demonstrate that the two consumption functions are, in fact, very different.

Consider what happens in Hall's example when there is a change in permanent income.

Suppose that permanent income is constant up to $t_0$ and then there is a discontinuous shock at $t_0$ and it is constant thereafter.

With Hall's consumption function we would find (immediate adjustment.)

With the stock adjustment function we find slow adjustment.

So the impulse response functions of the two hypotheses are very different.

Point 3.

In fact power is one of the central themes of this course.

We have seen and will continue to see many so-called results that are "derived" from rational maximizing behavior.

Hall's rational expectations permanent income hypothesis is one of them.

Just because these hypotheses are not rejected by tests does not make them true.

Indeed there may be deviation from the rational model that makes a very significant difference in the way that the economy works, but there may be very low power in testing such hypotheses against rational expectations.

In such cases standard economic methodology has been to accept the rational hypothesis and to reject the non-rational hypothesis.

That is a methodological error, and it can lead to very bad economic policy.

To give just one example i may not reject the hypothesis that the coefficient on inflationary expectations in a Phillips Curve is one.

We may also be able to reject that this coefficient is zero.

But it is extremely hard to reject that this coefficient would be, say, 0.7.

There is apt to be very little power to test that the coefficient would be one rather than 0.7.

But a coefficient of 0.7 would give economically a very different long-run trade-off between inflation and unemployment than a coefficient of one.

Point 4

This is regarding Hall's Rational Expectations/Permanent income Hypothesis.

I want you to compare the idea that consumption is a random walk to the idea that stock prices are a random walk. Superficially they sound the same. Both contain the word random walk. Presumably both depend on rationality. But in fact these theories are quite different.

I shall explain.

According to the random walk theory of stock prices, if the stock market was predictable then speculators could sit at home, chart its behavior on their PC's and make a profit.

But if such speculation did occur, profits on the stock exchange would be driven to zero, or near-zero.
So speculation will drive stock prices so that (almost) zero profits can be predicted from public information.

I want you to note the important difference between the stock-market random walk model and the consumption random walk model.

Suppose that a fraction $\lambda$ of the population decides not to maximize in the consumption model.

The rest of the population will not cause consumption to follow a random walk.

However, in the stock market suppose that a fraction $\lambda$ of the population does not maximize, while the remaining fraction $(1 - \lambda)$ is risk neutral.

In that case the returns on stocks net of bonds will be unpredictable from past events.

This is an important general distinction in economic theory.

Suppose that I go to a restaurant and I order tomato soup when I really want chicken soup. In that case the real equilibrium is changed.

I will eat my tomato soup when it arrives.

But if I do not take up the profitable opportunity to buy the KFC’s Franchise in wenzhou, Zhejiang, it does not matter because someone else, like Mr. Tian will buy it instead.

Point 5.

I think that it turns out that there are quite a few tests along the lines of Hall that do appear to have some power and that actually do reject the rational expectations/permanent income hypothesis.

Campbell and Mankiw is one of them.

They conducted a test that nests both Friedman’s view that consumption depends solely on wealth and the simplified Keynesian view, that consumption depends solely on income.

They suppose that a fraction of consumers $\lambda$ are pure Keynesians, while a fraction $(1 - \lambda)$ behave according to the permanent income hypothesis.

They estimate $\lambda$ from the extent to which consumption overreacts to predicted changes in income.

So here is the test:

they predict $\Delta y_t$ from past income.

They then take the regression:

$$\Delta c_t = \mu + \lambda \Delta y_t^p$$

where $\Delta y_t^p$ is the change in income predicted from past levels of income.

In some trials $\Delta y_t^p$ is the predicted change in income from past incomes and from past consumptions.

Usefully $\lambda$ gives a natural measure of the departure from the permanent income hypothesis. The estimates of $\lambda$ are both significant statistically and also of significant magnitude economically: between 40 and 50 percent (depending upon whether three or five periods are used to predict the change in current income). (NBER Working Paper 2924, p. 36, Table 1.)
Such dependence of consumption on current receipts has been found in at least six other studies.

John Shea, looked at expenditures of union members whose contracts specified their future wages.

David Wilcox looked at expenditures of social security recipients who had been earlier notified of changes in cost-of-living adjustments.

Jonathan Parker looked at the expenditures of payers of social security taxes with predictable inter-year changes.

Souleles looked at the dependence of expenditures on tax refunds.

Banks, Blundell and Tanner (1998), and Bernheim, Skinner and Weinberg (2001) for retirees.

These studies ALL find that changes in consumption depend upon the predictable component of changes in current receipts.

Point 6

I do not think that the importance of the dependence of consumption on expected income has been realized for graduate macroeconomics.

In graduate level macroeconomics behavior is supposed to be derived from maximization of some utility function.

But no (standard) utility function will yield dependence of consumption on current receipts or current income.\textsuperscript{4}

In the standard maximization problem consumption depends on wealth, not income.

I saw that with Friedman’s consumption function.

Standard methodology in macroeconomics does not even nest the possibility that people might want their consumption to have some dependence on current income (independent of wealth.)

But these findings suggest that a significant portion of consumption behavior is inconsistent, not just with the life-cycle model. These findings suggest much more generally that consumption behavior is inconsistent with a simple maximizing model.

I have added a footnote that consumption might be observed to depend upon current income if people wanted to consume more than their income but were faced with a borrowing constraint. But the footnote also explains that this is an unlikely explanation for much of the dependence of consumption on current income.

To obtain this dependence of consumption on current income there has to be something missing from maximization models.

Point 7

\textsuperscript{4}An exception occurs if consumers want to spend more than their incomes and cannot borrow. But in cases involving declines in consumption, rather than increases, the consumer should be saving for increased future consumption because income is decreasing. Borrowing is not an issue.
When we discussed certainty equivalence we made the point that if changes were very large then certainty equivalence would not be a good approximation.

The standard utility function is

\[ c^{1-\rho}/(1 - \rho) \]

The standard assumption is that people maximize such a utility function with a rate of discount \( \delta \) of 3 percent. \( \rho \) is the constant rate of risk aversion.\(^5\)

By the Euler equation:

\[
\frac{c_{t+1}}{c_t} = \left( \frac{1 + r}{1 + \delta} \right)^{1/\rho}.
\]

\[
\ln \left( \frac{c_{t+1}}{c_t} \right) = \frac{1}{\rho} \ln \left( \frac{1 + r}{1 + \delta} \right)
\]

\[
g_c = 1/\rho(r - \delta)
\]

Standard methodology is to assume that people are maximizing such a utility function.

I will now turn to work by Chris Carroll.

Carroll has found that a significant number of households, including very young households have very large variations in income. He reports that in the PSID (the Panel Study on Income Dynamics) in his sample of 1,238 households with no change in head of household over a 10-year period, 2/3 % would have labor income that was less than 10% of its permanent labor income in any given year. (Heads of household were between 24 and 63 for the entire period.) This may seem insignificant, but over a 15 year period it means that approximately 10% of the population will have had such a disaster.

Now what is this, and why would it make a significant difference?

These very large negative income observations are probably due to unemployment. Carroll thinks that these extreme low level observations of income are sufficiently common that they seriously impact consumption.

A third-order, rather than just a second-order approximation to the consumption function with CRRA yields:

\[
\Delta \ln c_{t+1} \approx a + b \mathbb{E}_t [\text{var} \ln c_{t+1}] + \varepsilon_{t+1}.
\]

He uses unemployment as a proxy for the conditional expectation of \( \text{var} \ln c_{t+1} \)–since he believes that the events that are likely to cause these very large changes in income, and therefore in consumption, will be unemployment.

He calls this approach a "buffer stock" approach–since the optimal policy is the following:

Above some target level of wealth, wealth is on average drawn down. If wealth is below that level, it will, on average, increase.

---

\(^5\)\( \rho \) is the inverse of the elasticity of substitution between consumption at \( t \) and consumption at \( t+1 \).
Where are consumption functions, and utility functions particularly important for economic policy?

Probably the most important use of utility functions comes in tax policy. One wants to know the deadweight losses from different tax policies.

The natural way to make such a calculation is to estimate the parameters of a utility function from observed consumption, and then to see what the impact on intertemporal utility will be.

Martin Feldstein has made the following proposition: capital-income taxation causes very large deadweight loss.

This is in fact one intellectual back-bone of tax-cut proposals for low capital gains taxes, and possibly for doing away with the estate tax.

The logic behind this is that if returns are fairly high then capital income taxation is taking away people's opportunity to save a little bit in their youth and be quite rich in their old age.

I will give you an example.

Feldstein has estimated the return to capital at 9 percent. (This of course may be disputed). Suppose that people could save and get this rate of return.

Consider someone of means, but these means could be fairly moderate too, who manages to salt away $100,000 while in their twenties. You could imagine doing that. You do not go to graduate school. You squirrel away $10,000 per year for ten years.

40 years later they retire, and they have earned 9 percent real rate of interest. My calculator says that this nest-egg will then be valued at $3.6 million. If old Mrs. Zhong has done this and is then ready to retire she will then have a sizeable nest-egg. If she continues to earn the 9% on her assets she will have about $325,000 per year for her retirement and still leave her kiddies or her University a $3.6 million bequest. Remember that if Mrs. Zhong meets Mr. Zhong at age 30 and Mr. Zhong has also done the same, they will have $7.2 million for their retirement.

Suppose, however, the government decides to tax the returns on this as income. Suppose that the income tax rate is 45 percent-35 percent Federal and 10 percent state. This reduces the rate of return from 9 percent to just shy of 5 percent. In that case the nest-egger will have done OK, perhaps, but she will have only $725,000 rather than $3.6 million at the end of her 40 years.

These utility functions can then be used to evaluate the deadweight losses due to capital income taxation.

As you can imagine the deadweight losses from capital income taxation are very large under the implied assumptions.

Capital income taxation has denied everyone the right to become rich to put away a little bit of money in their youth.

The economy could actually do that because the rate of return on capital is so high.

So the deadweight loss is huge.

The presence of uncertainty changes all of this.
This principle has been discovered by Steve Zeldes and Gregory Mankiw in the mid-1980’s.

Why would uncertainty make such a difference? It makes a specially large difference when workers cannot borrow against future income.

It makes a very large difference because the long-run rate of return is not so relevant any more to the savings of very young people. With unemployment uncertainty, young people have to lay away a nest egg whatever the rate of return may be, because they need to protect themselves while young against loss of consumption when they are unemployed.

Since they are going to save anyway, whatever the rate of capital income taxation, the deadweight loss in the presence of such uncertainty is much lower.

Uncertainty makes a difference because it means that young people will save against the possibility of such negative events.

Remember that in the standard model the marginal utility of consumption is \( \infty \) at \( c = 0 \). So the presence of large negative changes in income encourage quite a bit of saving at younger ages, and this will occur only because the MU of consumption has some chance of being very high.

The net result is that young people, if they optimize their saving, even if capital income is taxed will have a fair enough amount of saving, and the effects of capital income taxation will be comparatively mild.

Simulating the standard system, gives an estimate that the deadweight loss taking account of the variance in people’s income will be only one-tenth of the deadweight loss that occurs if I consider only the means.

4 The Interest Rate and Saving

An important issue concerning consumption involves its response to rates of return. Understanding the impact of rates of return on consumption is very important.

First we consider the interest rate and its impact on consumption growth.

Suppose interest rate is \( r \), then the budget constraint becomes

\[
\sum_{t=1}^{T} \frac{1}{(1 + r)^t} c_t \leq A_0 + \sum_{t=1}^{T} \frac{1}{(1 + r)^t} Y_t.
\]

To make things simple, we assume that utility function takes the CRRA form, i.e., \( u(c_t) = c_t^{1-\theta}/(1 - \theta) \), where \( \theta \) is the coefficient of relative risk aversion. In this case, the utility function becomes

\[
U = \sum_{t=1}^{T} \frac{1}{(1 + \rho)^t} \frac{c_t^{1-\theta}}{1 - \theta}.
\]

First-order condition implies

\[
\frac{1}{(1 + \rho)^t} c_t^{-\theta} = (1 + r) \frac{1}{(1 + \rho)^{t+1}} c_{t+1}^{-\theta}.
\]
Rearranging it we have
\[
\frac{c_{t+1}}{c_t} = \left( \frac{1 + r}{1 + \rho} \right)^{1/\theta}.
\]
This implies once the real interest rate and the discount rate are not equal, consumption need not be a random walk. Consumption is risking over time if \( r \) exceeds \( \rho \) and falling if \( r \) is less than \( \rho \).

5 Consumption and Risky Assets

Extending our analysis to account for multiple assets and risk raises new issues concerning both household behaviour and asset markets.

To start, let’s consider an individual’s optimization problem. Suppose he reduce his consumption in period \( t \) by an infinitesimal amount and using the resulting saving to raise consumption in period \( t+1 \). As the individual is optimizing, this change should have no impact on his expected utility. Thus we must have
\[
u'(c_t) = \frac{1}{1 + \rho} E_t \left[ (1 + r^{i}_{t+1}) u'(c_{t+1}) \right]
\]
for any \( i \), where \( r^i \) is the return on asset \( i \). Note that
\[
E[xy] = E[(x - E(x))(y - E(y))] = E[(x - E(x))(y - E(y))] + E(x)E(y) = \text{Cov}(x, y) + E(x)E(y),
\]
we have
\[
u'(c_t) = \frac{1}{1 + \rho} \left\{ E_t[1 + r^{i}_{t+1}] E_t[u'(c_{t+1})] + \text{Cov}_t(1 + r^{i}_{t+1}, u'(c_{t+1})) \right\} \text{ for all } i.
\]
If utility is quadratic, \( u(c) = c - ac^2/2 \), then marginal utility is \( 1 - ac \). Substituting it into the equation above, we have
\[
u'(c_t) = \frac{1}{1 + \rho} \left\{ E_t[1 + r^{i}_{t+1}] E_t[u'(c_{t+1})] - a \text{Cov}_t(1 + r^{i}_{t+1}, u'(c_{t+1})) \right\}
\]
This implies that in deciding whether to hold more of an asset, the individual is not concerned with how risky the asset is. This is because the variance of the asset’s return does not appear in the equation.

Intuitively speaking, a marginal increase in holding of an asset that is risky, but whose risk is not correlated with the overall risk the individual faces, does not increase the variance of the individual’s consumption. Thus in evaluating that marginal decision, the individual considers only the asset’s expected return.

The equation tells that the aspect of riskiness that matters to the decision of whether to hold more of an asset is the relation between the asset’s payoff and consumption.
Above discussion assume asset’s expected return is fixed. In reality, individual’s demands for assets determine these expected returns. If an asset’s payoff is highly correlated with consumption, its price must be driven down to the point where its expected return is high for individual to hold it.

To see its implication, suppose that all individuals are the same, and return to the first-order condition in equation above. Solving this for the expected return on the asset yields

\[ E_t[1 + r^i_{t+1}] = \frac{1}{E_t[u'(c_{t+1})]}[(1 + \rho)u'(c_i) + \alpha \text{Cov}_t(1 + r^i_{t+1}, u'(c_{t+1}))]. \]

This equation states that the higher the covariance of an asset’s payoff with consumption, the higher its expected return must be.

Note that for risk free asset, its covariance with consumption is zero, so the risk-free rate \( r \) must satisfy

\[ 1 + r_{t+1} = \frac{(1 + \rho)u'(c_i)}{E_t[u'(c_{t+1})]}. \]

Subtracting two equations yields

\[ E_t[r^i_{t+1} - r_{t+1}] = \frac{\alpha \text{Cov}_t(1 + r^i_{t+1}, c_{t+1})}{E_t[u'(c_{t+1})]}. \]

This equation states that the expected-return premium that an asset must offer relative to the risk-free rate is proportional to the covariance of its return with consumption.

This model is known as the Consumption Capital-Asset Pricing Model, or CCAPM. The coefficient from a regression of an asset’s return on consumption growth is known as its consumption beta. The central prediction of the consumption CAPM is that the premium that assets offer are proportional to their consumption betas.

### 5.1 Equity-Premium Puzzle

One of the most important implications of this analysis of assets’ expected returns concerns the case where risky asset is a broad portfolio of stocks. To see this, return to the Euler equation

\[ u'(c_i) = \frac{1}{1 + \rho} E_t[(1 + r^i_{t+1})u'(c_{t+1})] \text{ for any } i. \]

Suppose individuals have CRRA utility function, then the Euler equation becomes

\[ c_i^{\rho} = \frac{1}{1 + \rho} E_t[(1 + r^i_{t+1})c_{t+1}^{\rho}], \]

where \( \theta \) is the coefficient of relative risk aversion. Dividing both sides by \( c_i^{\rho} \) and multiplying both sides by \( 1 + \rho \), we have

\[ 1 + \rho = E_t[(1 + r^i_{t+1}) \frac{c_{t+1}^{\rho}}{c_i^{\rho}}]. \]
Let \( g^c_{t+1} \) denote the growth rate of consumption from \( t \) to \( t+1 \), or \((c_{t+1}/c_t) - 1\), ignore the time subscripts, we have

\[
E[(1 + r)(1 + g^c)^{-\theta}] = 1 + \rho.
\]

For the LHS, applying Taylor expansion around \( r = g = 0 \), yields,

\[
(1 + r)(1 + g)^{-\theta} \approx 1 + r - \theta g - \theta gr + \frac{1}{2}\theta(\theta + 1)g^2.
\]

Thus, replacing \( r \) with \( r^i \), \( g \) with \( g^c \), we have

\[
E[r^i] - \theta E[g^c] - \theta \{E[g^c]E[r^i] + Cov(r^i, g^c)\} + \frac{1}{2}\theta(\theta + 1)\{(E[g^c])^2 + Var(g^c)\}
\approx \rho.
\]

If the time period is short, then \( E[r^i]E[g^c] \) and \((E[g^c])^2\) are small relative to others. Omitting these terms and solving for \( E[r^i] \) yields

\[
E[r^i] \approx \rho + \theta E[g^c] + \theta Cov(r^i, g^c) - \frac{1}{2}\theta(\theta + 1)Var(g^c).
\]

This implies for any two assets, \( i \) and \( j \), the difference between the expected returns satisfies

\[
E[r^i] - E[r^j] = \theta Cov(r^i, g^c) - \theta Cov(r^j, g^c)
= \theta Cov(r^i - r^j, g^c).
\]

However, in a famous paper, Mehra and Prescott (1985) shows that it is difficult to reconcile observed returns on stocks and bonds with above equation. As Mankiw et al (1991) shows, the difference between the average return on the stock market and the return on short-term government debt—the equity premium—is about 6 percentage points. Over the same period, the std deviation of the growth of consumption(measured by real purchases of nondurables and services) is 3.6 percentage points, and the std. deviation of the excess return on the market is 16.7 percentage points; the correlation between these two quantities is 0.4. These figures imply that the covariance of consumption growth and the excess return on the market is 0.40(0.036)(0.167), or 0.0024.

Equation above therefore implies that the coefficient of relative risk aversion needed to account for the equity premium is the solution to

\[
0.0024 = \theta(0.0024)
\]

or \( \theta = 25 \). But this is an extraordinary level of risk aversion.

As Mehra and Prescott describe, other evidence suggests that risk aversion is much lower than this.

Another problem is that the equity-premium puzzle has become more severe. For example, from 1979 to 2003, the average equity premium is 7 percentage
points, which is higher than in Mehra and Prescott’s sample period. Furthermore, consumption growth has become more stable and less correlated with returns. This makes it even more difficult to explain.

This equity-premium puzzle has stimulated a large amount of research. And it is also a good thesis topic.

6 Hyperbolic Discounting and Its Implication for Consumption: Laibson

This now takes us to something different. I will review the article by Laibson, which is in your reader.

Let me start with a few comments on the general philosophy behind his article.

For standard economics, saving too little or too much, like involuntary unemployment, is an impossibility.

It is a straightforward contradiction of the assumptions of maximizing behavior.

Since saving is the result of individual utility maximization, it must, absent externalities, be just right.

A key theoretical innovation in Laibson’s analysis of savings is the recognition that individuals may maximize a utility function that is divorced from that representing their "true welfare."

Once this distinction is accepted, "saving too little" becomes a meaningful concept.

This is where the lemmings that I mentioned in the very first lecture make their return.

The idea is illustrated by the ancient myth of the lemmings.

Every few years the lemmings are said to converge in a death march, which ends with their final plunge into the sea.

The alleged behavior of those lemmings reveals a distinction that is common among psychologists.

But this distinction is rare for economists.

Unless the lemmings experience an unusual epiphany in that final plunge, their utility or welfare is given by one function; yet they maximize another.

Think about it.

The popular view of saving, that people undersave, is similarly described.

Determining whether people save too much or too little involves asking whether people, like the lemmings, have one (inter-temporal) utility function which describes their welfare, but maximize another.

Such evidence as there is suggests potentially large difference between the two concepts.

High positive rates of time discount are necessary to explain actual wealth-earnings ratios.
Yet, questionnaire responses on the consumption-saving tradeoffs that people think they ought to make reveals an intertemporal discount rate that is on average slightly negative.

The hyperbolic discount function, which Laibson uses to study intertemporal savings choices, can be used to formalize the distinction between the utility function that describes actual saving behavior and the utility function that measures the welfare resulting from that behavior.

This hyperbolic discount function captures the difficulty people have in exercising self-control.

It assumes that the discount rates used to evaluate tradeoffs between adjacent periods decline as the time horizon lengthens: individuals use high discount rates to evaluate options that require an immediate sacrifice for a future reward and lower discount rates when the same sacrifice is deferred into the future.

Thus, people are patient in making choices requiring gratification delays when those sacrifices are deferred; but impatient in delaying gratification in the short run.

Because present consumption is more salient than future consumption, individuals procrastinate about saving.

The hyperbolic function accords closely with experimental findings: Human and animal subjects are far less willing to delay gratification immediately than to commit to such delays in the future.

Intertemporal utility with such hyperbolic discounting seems to explain the phenomenon of under-saving.

Consider the division of the U.S. aged population ranked by income.

It is not until one gets to the fourth quintile, the 60th to 80th percentile that people have any significant non-social-security retirement income at all. (Burtless)

In addition, it seems that on retirement people make considerable cuts in their consumption.

I already mentioned the papers by Bernheim, Skinner and Sinberg in the US and by Richard Blundell and his co-authors in the U.K.

The standard utility function, as in Hall, predicts that there will be no break in consumption at retirement.

Laibson has a utility function that explains why people save so little. The utility function has the following property.

Looking forward from today:

The discount factor for current consumption is 1. There is no discount.

The current discount factor on consumption one period from now is \( \beta \).

The current discount factor on consumption two periods from now is \( \beta^2 \).

The current discount factor on consumption \( n \) periods from now is: \( \beta^n \).

\( \beta \) is less than one.

Now let’s go forward one period.

Next period, when I make my decision on consumption at that time, I will have similar discount factors looking forward.

Again, from the viewpoint of that period, the discount factor for current consumption is 1.
The discount factor on consumption one period from now is $\beta \delta$.
The discount on consumption two periods from now is $\beta \delta^2$.
The discount factor on consumption $n$ periods from now is: $\beta \delta^n$.
The meaning of such discounting is that the current time is something special.

That occurs naturally.
Today is the day that I go to the mall.
And that current purchase or that current need seems to be especially pressing and salient.

Today is the day my son says he needs Air Jordans.
All the other kids in his class have Air Jordans.
I have to buy him Air Jordans.
Such behavior will have very adverse effects on savings.
It turns out that it will have adverse effects on saving under two polar conditions.
The first is the easy one to see.
In the first case the consumer is naive about her behavior.
She says to herself.
There is no reason to save today because today’s consumption is especially valuable.

As I look forward I see that beyond today that I have discount factors of $\delta$, which are much less than my current discount factor between today and tomorrow.

Today’s consumption has a special salience relative to tomorrow’s consumption.
So I will wait until tomorrow to begin my saving plan, but tomorrow consumption will not be so necessary, so I will begin to save then.

I will give you the same thoughts in another context.
That muffin looks really good. I know that I am too fat, but i can begin dieting tomorrow. Furthermore i do not need to start on a diet today because i know that i will go on a diet tomorrow. But then tomorrow comes and you say the same thing to yourself, so you have never gone on that diet.

That is one version of the argument. Now let’s try another one. This is the one that Laibson actually uses in his study of savings.

That chocolate muffin with the chocolate chips is especially important to me today.
I am really aching for it. It will really not make much difference to my long-run weight and I would like it now. I might as well have one now.

After all, if I do not eat it now, I know that again tomorrow I will have a chocolate chip muffin.
I am all blubber today, and the small difference that it makes to my blubber tomorrow is not worth giving it up today.

Now let’s translate that into savings.
I could start saving today. But my consumption is especially important to me today.
It has special salience.
Furthermore, it makes no difference because tomorrow consumption will be especially salient, and if I have left money in the bank I will just spend it tomorrow. So I might as well spend it today when it has that extra value to me because this is today.

I am such a spendthrift that I might as well spend my money today when I really want it, because if I do not spend it today I will just spend it tomorrow anyway.

So how do you solve such a problem about what spending will be?

How would one formalize these thoughts in a utility function, and then what problems do these changes cause in our thinking and in our solution for consumption? So how do you solve such a problem about what spending will be?

First of all you set up today's utility function.

Today's utility function is:

$$u(c_t) + \sum_{i=1, \ldots, n-t} \beta^i u(c_{t+i}).$$

The individual will choose \( c_t \) to maximize the expected value of this utility function as a function of wealth at \( t \), \( W_t \), and income at \( y_t \):

$$u(c_t) + \sum_{i=1, \ldots, n-t} \beta^i u(c_{t+i}).$$

subject to the constraint that in the next period and in each subsequent period consumption will be chosen in the same way.

There is one further constraint on spending. I cannot exhaust my wealth. wealth behaves in the following way:

$$W_{t+1} = (1 + r)W_t - c_t + y_{t+1},$$

and \( W_t \geq 0 \) for all \( t \).

You want to solve for consumption today given your level of wealth, your stochastic income tomorrow, and also how you are going to behave tomorrow.

It turns out that this is not such a simple problem. The consumer does not simply maximize her utility of consumption today, given her current utility function.

Instead she has to have a trickier strategy.

What she must do is to maximize her current utility function today, given what her behavior will be tomorrow when she has a different utility function.

Then of course tomorrow she is doing the same thing.

She is maximizing her utility function of that time, but given how she thinks that she will again be behaving the day after tomorrow.

Now you might think that this was an impossible problem to solve, especially when income is random.

It turns out that it is not, because it can be solved in the following way.

The total solution to the problem is a set of rules.

Those rules give the amount of consumption at \( t \) as a function of the amount of wealth held at \( t \), or

$$c_t = c_t(W_t).$$
But I can give a criterion for whether given the rules that we have established for the future whether consumption at \( t \) as a function of wealth is too large or too small.

And that criterion enables you to solve a problem on the computer.

More accurately, it allows you to solve such a problem on a supercomputer.

Let’s work our way backwards.

Suppose that we have established the optimal amount of consumption at \( t + 1 \) as a function of wealth at \( t + 1 \).

If I know this function I can compute how much consumption I should have at \( t \) as a function of wealth at time \( t \).

How? I have an Euler condition.

That Euler condition is that

\[
u'(c_t) = E_t(\frac{\partial c_{t+1}(W_{t+1})}{\partial W_{t+1}}) \beta \delta + (1 - \frac{\partial c_{t+1}(W_{t+1})}{\partial W_{t+1}}) \delta [(1 + r)u'(c_{t+1})] \]

Let me begin by trying to make sense of this formula, and then I will try to indicate how one would use this in a simulation.

The amazing thing about this formula is that we can actually make sense of it.

Stare for a moment at the term in square brackets.

The whole problem in Laibson’s article, what makes it different, is the appearance of \( \beta \):

\( \beta \) is the special salience of today over tomorrow. If \( \beta \) is one then the term in square brackets will be exactly \( \delta \), and this formula is just the previous Euler equation that I derived earlier. Remember that here \( \delta \) is the discount factor, not the rate of discount.

Let’s go over the case where \( \beta \) is one and review what the equation says, then we will consider how the square bracket term modifies it.

What does the standard Euler equation say? Suppose I save a dollar today. I then lose MU of \( u'(c_t) \). With my \((1 + r)\) dollars next period I will be able to get additional utility \((1 + r)u'(c_{t+1})\), which is discounted by the discount factor \( \delta \).

In equilibrium, \( u'(c_t) \) should equal \((1 + r)\delta u'(c_{t+1})\). Remember here that \( \delta \) is the discount factor, not the discount rate. Now consider how this equality gets modified when there is the additional discount factor \( \beta \) gets added.

Suppose that I save a dollar.

I will then lose \( u'(c_t) \) of current utility.

I will gain \((1 + r)\) dollars tomorrow.

But of the extra dollars I have, I will spend only a fraction.

Tomorrow I will have \(1+r\) dollars.

Of the fraction of the extra dollar that I save today and that I spend tomorrow, I should discount by the special discount factor \( \beta \delta \), which is the special discount rate between today and tomorrow.
But of the fraction of that dollar that is deferred between tomorrow and the next day, I should discount only by the amount $\delta$, since that dollar is being deferred in expenditure between tomorrow and yet later times.

And the discount between tomorrow and those yet later times is only $\delta$.

So I discount the fraction that is actually spent tomorrow, $\partial c_{t+1}(W_{t+1})/\partial W_{t+1}$ by the factor $\beta\delta$ and the rest that is yet deferred I should discount only by the factor $\delta$.

This is the sense of Laibson’s Euler equation.

The use of this Euler equation allows for computation of the consumption rule even when income has stochastic terms and when there is a large amount of complication in the description of the income process and the nature of tax codes that exempt types of saving from taxes.

Let me indicate how you would then solve for consumption.

You do it by backward induction.

In the last period you spend all the wealth that is left.

Now suppose that you have wealth $W_{n-1}$ at $n - 1$.

Suppose that you spend $C_{n-1}$ at this level of wealth.

That will leave you with $(1 + r)(W_{n-1} - C_{n-1}) + \tilde{y}_n$ next period.

You can compute the expected value of the marginal utility of consumption at that level of income.

If the expected marginal utility is too high relative to $u(C_{n-1})$ you should consume less at $n-1$ and more at $n$, and if the expected marginal utility is too low relative to $u(C_{n-1})$ you should consume more.

In this way by trial and error you can fairly quickly on the computer establish the value of $c_{n-1}(W_{n-1})$.

Now work your way back, using the Euler conditions to establish each sequential function, $c_t(W_t)$.

The Euler equation establishes for you whether $c_t(W_t)$ is too high or too low.

So now that I know how one can establish the levels of consumption, what does Laibson do here?

The basic problem is that because of $\beta$ people save too little.

Laibson evaluates the effectiveness of defined contribution plans.

He finds that defined contribution plans greatly increase welfare.

They increase welfare because people then have an easier time beginning their saving.

Once they have begun their saving they have locked it away.

So they do not say to themselves: I will not save today because tomorrow I am going to waste it at the mall. I will not spend it next period because it is locked away.

Comment.

I like this article by Laibson not because of this extremely complicated game that the self of time $t$ is playing with all the other selves to determine its
consumption, but rather because it begins to capture some of the human-ness of people.

The reason that I need social security is precisely because in the absence of social security people will not save for themselves.

That is the probably the major reason that we have pension plans.

If we go back to Stiglitz' Economics 1 textbook the principle of modern economics, in contrast to the classics is that the moderns feel that they have to specifically model the problem at hand.

In this spirit Laibson is inductively figuring out what people do by paying attention to what people do, both in consumption and in psychological experiments and then trying to model it.

He may even be right about it.

As a final comment let me remark that Laibson's article is also relevant to Point 8 about Hall.

What is just as interesting in Hall as his test of the RE/PIH is the Euler condition.

The Euler condition allows us to do a lot of "work" in evaluating any problem involving saving and consumption over the life cycle.

It is very theoretical, but at the same time surprisingly useful. Because of Laibson's work the regulations regarding tax advantaged 401(k) savings plans have been changed.

If firms so choose, workers may now be automatically enrolled with an automatic default contribution.

Adoption of such plans significantly increases plan participation and many workers maintain their contributions at the level of the default.

Thaler and Benartzi have devised a savings plan to overcome workers' tendency to procrastinate and have tested it on an experimental basis at a mid-size manufacturing firm.

Employees were invited to join a savings plan allowing them to elect, in advance, the fraction of wage or salary increases to be set aside for savings.

Consistent with hyperbolic discounting, but not with the standard exponential model, workers chose relatively modest saving out of current income but committed to save large fractions of future wage and salary increases.

Within a short period of time, the average savings rate had doubled.

But now I want you to go back and consider something about Laibson.

He may have short-sighted consumption, but it is not the short-sighted consumption that most economists think that we do have.

Most economists think that the short-sightedness of consumption is that current consumption depends at least in part on current income. It does not just depend upon wealth.

In Laibson's model consumption only depends on current income because of spending constraints.

People cannot be myopic in the way that they feel that they should spend only what they currently earn or some fraction of it.

This indicates a way in which economic derivations from a priori reasoning seem to get everything wrong.
Laibson may have modeled present bias.
But it is not the present bias that most Keynesians thought that people had.
In that version people look at their current income, and that has a strong independent effect on their consumption expenditures.
Curiously, such an effect seems to appear robustly in empirical studies on consumption.