A Theory of the Consumption Function, With and Without Liquidity Constraints

Christopher D. Carroll

May 18, 2001

Published as

Expanded version (for deeper treatment of theoretical issues) is available at [http://econ.jhu.edu/people/ccarroll/ATheryv3NBER.pdf](http://econ.jhu.edu/people/ccarroll/ATheryv3NBER.pdf)

**Keywords:** consumption, precautionary saving, uncertainty, Permanent Income Hypothesis, liquidity constraints

**JEL Classification Codes:** A23, B22, D91, E21

Christopher D. Carroll is Professor of Economics, Johns Hopkins University, Baltimore, Maryland. His e-mail address is <ccarroll@jhu.edu>. The programs that generated the results in this paper are available at [www.econ.jhu.edu/people/carroll/ccarroll.html](http://www.econ.jhu.edu/people/carroll/ccarroll.html). The author is grateful to Carl Christ and to the JEP editors for valuable comments on earlier versions of this paper.
Fifteen years ago, Milton Friedman’s 1957 treatise *A Theory of the Consumption Function* seemed badly dated. Dynamic optimization theory not been employed much in economics when Friedman wrote, and utility theory was still comparatively primitive, so his statement of the “permanent income hypothesis” never actually specified a formal mathematical model of behavior derived explicitly from utility maximization. Instead, Friedman relied at crucial points on intuition and verbal descriptions of behavior. Although these descriptions sounded plausible, when other economists subsequently found multiperiod maximizing models that could be solved explicitly, the implications of those models differed sharply from Friedman’s intuitive description of his ‘model.’ Furthermore, empirical tests in the 1970s and ’80s often rejected these rigorous versions of the permanent income hypothesis, in favor of an alternative hypothesis that many households simply spent all of their current income.

Today, with the benefit of a further round of mathematical (and computational) advances, Friedman’s (1957) original analysis looks more prescient than primitive. It turns out that when there is meaningful uncertainty in future labor income, the optimal behavior of moderately impatient consumers is much better described by Friedman’s original statement of the permanent income hypothesis than by the later explicit maximizing versions. Furthermore, in a remarkable irony, much of the empirical evidence that rejected the permanent income hypothesis as specified in tests of the 1970s and ’80s is actually consistent both with Friedman’s original description of the model and with the new version with serious uncertainty.
There are four key differences between the explicit maximizing models developed in the 1960s and '70s and Friedman’s model as stated in *A Theory of the Consumption Function* (and its important clarification in Friedman (1963)).

First, Friedman repeatedly acknowledged the importance of precautionary saving motives induced by uncertainty about the future level of labor income. In contrast, the crucial assumption that allowed subsequent theorists to solve their formal maximizing models was that labor income uncertainty had no effect on consumption, either because uncertainty was assumed not to exist (in the “perfect foresight” model) or because the utility function took a special form that ruled out precautionary motives (the “certainty equivalent” model).

Second, Friedman asserted that his conception of the permanent income hypothesis implied that the marginal propensity to consume out of transitory “windfall” shocks to income was about a third. However, the perfect foresight and certainty equivalent models typically implied an MPC of 5 percent or less.

Third, Friedman asserted that the “permanent income” that determined current spending was something like a mean of the expected level of income in the very near-term: “It would be tempting to interpret the permanent component [of income] as corresponding to the average lifetime value . . . It would, however, be a serious mistake to accept such an interpretation.” He goes on to say that households in practice adopt a much shorter ‘horizon’ than the remainder of their lifetimes, as captured in the assumption in Friedman (1963) that people discount future income at a “subjec-
tive discount rate” of 33-1/3 percent. In contrast, the perfect foresight and certainty equivalent models assumed that future income was discounted to the present at market interest rates (say, 4 percent).

Finally, as an interaction between all of the preceding points, Friedman indicated that the reason distant future labor income had little influence on current consumption was “capital market imperfections,” which encompassed both the fact that future labor income was uninsurably uncertain and the difficulty of borrowing against such income (for example, see Friedman (1963) p. 10).

It may seem remarkable that simply adding labor income uncertainty can transform the perfect foresight model into something closely resembling Friedman’s original framework; in fact, one additional element is required to make the new model generate Friedmanesque behavior: Consumers must be at least moderately impatient. The key insight is that the precautionary saving motive intensifies as wealth declines, because poorer consumers are less able to buffer their consumption against bad shocks. At some point, the intensifying precautionary motive becomes strong enough to check the decline in wealth that would otherwise be caused by impatience. The level of wealth where the tug-of-war between impatience and prudence reaches a stalemate defines a ‘target’ for the buffer stock of precautionary wealth, and many of the insights from the new model can best be understood by considering the implications and properties of this target.

A final insight from the new analysis is that precautionary saving behavior and liq-
Utility constraints are intimately connected. Indeed, for many purposes the behavior of constrained consumers is virtually indistinguishable from the behavior of unconstrained consumers with a precautionary motive; average behavior depends mainly on the consumer's degree of impatience, not on the presence or absence of constraints. As a result, most of the existing empirical studies that supposedly test for constraints should probably be reinterpreted as evidence on the average degree of impatience. Furthermore, future studies should probably focus more directly on attempting to measure the average degree of impatience rather than on attempting to detect constraints.

1 The Modern Model(s) of Consumption

Current graduate students rarely appreciate how difficult it was to forge today's canonical model of consumption based on multiperiod utility maximization. The difficulty of the enterprise is attested by the volume of literature devoted to the problem from the 1950s through the ’70s, beginning with the seminal contribution of Modigliani and Brumberg (1954). The model that eventually emerged has several key characteristics. Utility is time separable; that is, the utility that consumption yields today does not depend on the levels of consumption in other periods, past or future. Future utility is discounted geometrically, so that utility one period away is worth $\beta$ units of this period’s utility, utility two periods away is worth $\beta^2$, and so on, for some $\beta$ between 0 and 1. Furthermore, the utility function must satisfy various criteria of plausibility.
like decreasing marginal utility, decreasing absolute risk aversion, and so on. Finally, the model must incorporate a mathematically rigorous description of how noncapital income, capital income, and wealth evolve over time.

One of the unpleasant discoveries in the 1960s and '70s was that when there is uncertainty about the future level of labor income, it appears to be impossible (under plausible assumptions about the utility function to derive an explicit solution for consumption as a direct (analytical) function of the model’s parameters. This is not to say that nothing at all is known about the structure of optimal behavior under uncertainty; for example, it can be proven that consumption always rises in response to a pure increment to wealth. But an explicit solution for consumption is not available.

1.1 The Perfect Foresight/Certainty Equivalent Model

Economists’ main response to this problem was to focus on two special cases where the model can be solved analytically: The “perfect foresight” version in which uncertainty is simply assumed away, or the “certainty equivalent” version in which consumers are assumed to have quadratic utility functions (despite unattractive implications of quadratic utility like risk aversion that increases as wealth rises, and the existence of a ‘bliss point’ beyond which extra consumption reduces utility).

The perfect foresight and certainty equivalent solutions are very similar; for brevity, I will summarize only the perfect foresight solution, in which the optimal level of consumption is directly proportional to total “wealth,” which is the sum of market
wealth $W_t$ and ‘human wealth’ $H_t$,

$$C_t = k_t(W_t + H_t),$$

(1)

where market wealth $W_t$ is real and financial capital, while human wealth is mainly current and discounted future labor income (though in principle $H_t$ also includes the discounted value of transfers and any other income not contingent on saving decisions; henceforth I refer to these collectively as ‘noncapital’ income). The constant of proportionality, $k_t$, depends the time preference rate, the interest rate, and other factors.

A simple example occurs when consumers care exactly as much about future utility as about current utility ($\beta = 1$); the interest rate is zero; and there is no current or future noncapital income ($H_t = 0$). In this case, the optimal plan is to divide existing wealth evenly among the remaining periods of life. If we assume an average age of death of 85, this model implies that the marginal propensity to consume out of shocks to wealth for consumers younger than 65 should be less than $(1/20)$, or 5 percent – since the change in wealth will be spread evenly over at least 20 years. Furthermore, the theory implies that the MPC out of unexpected transitory shocks to noncapital income (“windfalls”; e.g. finding a $100 bill in the street) is the same as the MPC out of wealth, because once the windfall has been received, it is theoretically indistinguishable from the wealth the consumer already owned. When the model is made more realistic by allowing for positive interest rates, consumers younger than 65, etcetera, it still implies that the average MPC should be quite low, generally less than 0.05.
In contrast, Friedman (1963) asserted that his conception of the permanent income hypothesis implied an MPC out of transitory shocks of about 0.33 for the typical consumer.\textsuperscript{3} Friedman (1963) provided an extensive summary of the existing empirical evidence tending to support the proposition of an MPC of roughly a third.\textsuperscript{4} From today’s perspective, however, the most surprising aspect of Friedman’s (1957, 1963) arguments is that their main thrust is to prove an MPC much less than one (to discredit the ‘Keynesian’ model that said consumption was roughly equal to current income), rather than to prove an MPC significantly greater than 0.05.

The 15 years after the publication of \textit{A Theory of the Consumption Function} produced many studies of the MPC. Particularly interesting were some natural experiments. In 1950, unanticipated payments were made to a subset of U.S. veterans holding National Service Life Insurance policies; the marginal propensity to consume out of these dividends seems to have been between about 0.3 and 0.5. Another natural experiment was the reparations payments certain Israelis received from Germany in 1957-58.\textsuperscript{5} The marginal propensity to consume out of these payments appears to have been around 20 percent, with the lower figure perhaps accounted for by the fact that the reparations payments were very large (typically about a year’s worth of income). On the whole, these studies were viewed at the time as supporting Friedman’s model because the estimated MPCs were much less than one.

The change in the profession’s conception of the permanent income hypothesis in the 1970s from Friedman’s (1957, 1963) version to the perfect foresight/certainty
equivalent versions (with their predictions of an MPC of 0.05 or less) is nicely illustrated by a well-known paper by Hall and Mishkin (1982) that found evidence of an MPC of about 0.2 using data from the *Panel Study of Income Dynamics* (PSID). Rather than treating this as evidence in favor of a Friedmanesque interpretation of the permanent income hypothesis, the authors concluded that at least 15-20 percent of consumers *failed* to obey the PIH because their MPCs were much greater than 0.05.

### 1.2 The New Model

The principal development in consumption theory in the last 15 years or so, starting with Zeldes (1984), is that spectacular advances in computer speed have allowed economists to relax the perfect foresight/certainty equivalence assumption and determine optimal behavior under realistic assumptions about uncertainty.

A preliminary step was to determine the characteristics of the income uncertainty that typical households face. Using annual income data for working-age households participating in the PSID, Carroll (1992) found that the household noncapital income process is well approximated as follows. In period $t$ a household has a certain level of ‘permanent noncapital income’ $P_t$, which is defined as the level of noncapital income the household would have gotten in the absence of any transitory shocks to income. Actual income is equal to permanent income multiplied by a transitory shock, $Y_t = P_t \epsilon_t$ where permanent income $P_t$ grows by a factor $G$ over time, $P_t = GP_{t-1}$. Each year there is a small chance (probability 0.005) that actual household income will be essentially
zero \( (\epsilon_t = 0) \), typically corresponding in the empirical data to a spell of unemployment or temporary illness or disability. If the transitory shock does not reduce income all the way to zero, that shock is distributed lognormally with a mean value of one and a standard deviation of \( \sigma_\epsilon = 0.1 \). Carroll (1992) and subsequent papers also find strong evidence for permanent as well as transitory shocks to income, also with an annual standard deviation of perhaps 0.1. However, because permanent shocks complicate the exposition without yielding much conceptual payoff, I will suppress them for the purposes of this paper and compensate by boosting the variance of the transitory component to \( \sigma_\epsilon = 0.2 \); for the version with both transitory and permanent shocks, see Carroll (1992). The PSID also shows the annual household income growth factor to be about \( G = 1.03 \) or 3 percent growth per year for households whose head is in the prime earning years of 25-50.

The next step in solving the model computationally is to choose values for the parameters that characterize consumers’ tastes. For the simulation results presented in this paper, I will assume a rather modest precautionary saving motive by choosing a coefficient of relative risk aversion of \( \rho = 2 \), toward the low end of the range from 1 to 5 generally considered plausible.\(^6\) I follow a traditional calibration in the macro literature and choose a time preference factor of \( \beta = 0.96 \) implying that consumers discount future utility at a rate of about 4 percent annually, and I make a symmetric assumption that the interest rate is also 4 percent per year.

We are now in position to describe how the model can be solved computationally.
As is usual in this literature, it is necessary to solve backwards from the last period of life. For simplicity, we will assume that the income process described above, with constant income growth $G$, holds for every year of life up to the last. (For a version with a more realistic treatment of the lifetime income profile, including the drop in income at retirement, see Carroll (1997)).

In the last time period, the solution is easy: The benchmark model assumes there is no bequest motive, so the consumer spends everything. Following Deaton (1991), define cash-on-hand $X$ as the sum of noncapital income and beginning-of-period wealth (including any interest income earned on last period’s savings). In the second-to-last period of life, the consumer’s goal is to maximize the sum of utility from consumption in period $T-1$ and the mathematical expectation of utility from consumption in period $T$, taking into account the uncertainty that results from the possible shocks to future income $Y_T$. For any specific numerical levels of cash-on-hand and permanent income in period $T-1$ (say, $X_{T-1} = 5$ and $P_{T-1} = 1.4$), a computer can calculate the sum of current and expected future utility generated by any particular consumption choice. The optimal level of consumption for $\{X_{T-1}, P_{T-1}\} = \{5, 1.4\}$ can thus be found by a computational algorithm that essentially tries out different guesses for $C_{T-1}$ and homes in on the choice that yields the highest current and discounted expected future utility.

Note that for each different combination of $\{X_{T-1}, P_{T-1}\}$, the utility consequences of many possible choices of $C_{T-1}$ must be compared to find the optimum, and for each $C_{T-1}$ that is considered, the numerical expectation of next period’s utility must
be computed. The solution procedure is basically to calculate optimal $C_{T-1}$ for a grid of many possible $\{X_{T-1}, P_{T-1}\}$ choices, and then to construct an approximate consumption function by interpolation (‘connect-the-dots’).

Once the approximate consumption rule has been constructed for period $T - 1$, the same steps can be repeated to construct a consumption rule for $T - 2$ and so on.

This begins to give the flavor for why numerical solutions are so computation intensive. Indeed, the problem as just described would be something of a challenge even for current technology. Fortunately, there is a trick that makes the problem an order of magnitude easier: Everything can be divided by the level of permanent income. That is, defining the cash-on-hand ratio as $x_t = X_t/P_t$ and $c_t = C_t/P_t$, it is possible to find the optimal value of the consumption-to-permanent-income ratio as a function of the cash-on-hand ratio, so that rather than solving the problem for a two-dimensional grid of $\{X_{T-1}, P_{T-1}\}$ points one can solve for a one-dimensional vector of values of $\{x_{T-1}\}$. This and a few other tricks turn the problem into one that can be solved with the amounts of computer power that began to be widely available in the 1980s.\textsuperscript{7}

The solution to the optimal consumption problem is depicted in Figure 1. The cash-on-hand ratio $x$ is on the horizontal axis. The optimal consumption ratio for a given cash-on-hand ratio is on the vertical axis. The solid lines represent the consumption rules for different time periods, showing how optimal consumption changes as the ratio of cash-on-hand to labor income increases.

Consumption in the last period $c_T(x)$ coincides with the 45 degree line, indicating
consumption equal to cash-on-hand. For very low levels of $x$, consumption in the second-to-last period $c_{T-1}(x)$ is fairly close to the 45 degree line; the consumer spends almost, but not quite, everything. This reflects the precautionary motive: Because there is a chance the consumer will receive zero income in period $T$, she will never spend all of her period-$T-1$ resources because of the dire consequences of arriving at $T$ with nothing and then possibly receiving zero income. Note the contrast with behavior at high levels of wealth; for example, at an $x_{T-1}$ of around 10 the figure shows $c_{T-1}$ of a bit more than 5 – indicating that at this large level of wealth the consumer divides remaining lifetime resources roughly evenly between the last two periods of life.

An important feature of this problem is that, if certain conditions hold (in particular, if consumers are ‘impatient’ in a sense to be described shortly), the successive consumption rules $c_T(x), c_{T-1}(x), c_{T-2}(x), \ldots, c_{T-n}(x)$ will ‘converge’ as $n$ grows large. The meaning of convergence is most easily grasped visually: In Figure 1, the rules $c_T(x)$ and $c_{T-1}(x)$ are very far apart, while the rules $c_{T-10}(x)$ and the converged consumption rule $c(x)$ (which can be thought of as $c_{T-\infty}(x)$) are very close.

The importance of convergence can best be understood by contrasting it with the alternative. Modigliani (1966) points out that in the certainty equivalent model, optimal behavior is different at every different age, so that one cannot draw many general lessons about consumption behavior from the rule for any particular age. In the model solved here, however, behavior is essentially identical for all consumers more than 10 years from the end of life, so analysis of the converged consumption rule yields insights
about behavior of most agents in the economy.

What is required to generate convergence? Deaton (1991) and Carroll (1996) show that the necessary condition is that consumers be impatient, in the sense that if there were no uncertainty or liquidity constraints the consumer would choose to spend more than her current income. Technically, the required condition is

$$(R\beta)^{1/\rho} < G,$$

(2)

where $\rho$ is the coefficient of relative risk aversion and $G$ is the income growth factor.

Consider the version of this equation where $G = \rho = 1$, so that consumers are impatient if $R\beta < 1$. In this case, impatience depends directly on the whether the reward to waiting, as determined by the interest rate factor $R$, is large enough to overcome the utility cost to waiting, $\beta$. Positive income growth ($G > 1$) makes consumers more impatient (in the sense of wanting to spend more than current income) because forward-looking consumers with positive income growth will want to spend some of their higher future income today. Finally, the exponent $(1/\rho)$ on the $R\beta$ term captures the ‘intertemporal elasticity of substitution,’ which measures the extent to which the consumer responds to the net incentives for reallocating consumption between periods.

The remainder of the paper will focus almost exclusively on implications of the converged consumption function. It is natural to wonder, however, whether we should expect these results to be useful in understanding the behavior of consumers whose permanent income paths over the lifetime do not resemble the “constant growth at
rate $G$ until death” specification used here. For instance, income can be predicted to decline at retirement! However, Carroll (1997) shows that when a model like this is solved with an empirically realistic pattern of income growth over the lifetime, the consumption function resembles the ‘converged’ consumption function examined here until roughly age 50. After 50, with retirement looming, the consumer begins saving substantial amounts and behavior begins more and more to resemble that in the perfect foresight model. Thus, the results in the remainder of the paper based on the converged consumption function are most appropriately represented as characterizing the behavior of moderately impatient households up to about age 50.8

At present, three further observations about the converged consumption function depicted in Figure 1 are important. (The general shape of the consumption function, and the validity of the points made here, are robust to alternative assumptions about parameter values, so long as consumers remain moderately impatient.)

First, the converged consumption function is everywhere well below the perfect foresight solution (the dashed line). Since precautionary saving is defined as the amount by which consumption falls as a consequence of uncertainty, the difference between the converged $c(x)$ and the dashed perfect-foresight line measures the extent of precautionary saving. The precautionary effect is large here because under our baseline parameter values, human wealth is quite large and therefore induces a lot of consumption by the perfect-foresight consumers. In contrast, consumers with a precautionary motive are unwilling to spend much on the basis of uncertain future labor income, so the large
value of human wealth has little effect on their current consumption.

The second important observation is that as \( x \) gets large, the slope of \( c(x) \) (which is to say, the marginal propensity to consume) gets closer and closer to the slope of the dashed perfect foresight line. That is, as wealth approaches infinity the marginal propensity to consume approaches the perfect foresight MPC. This happens because as wealth approaches infinity the proportion of future consumption that will be financed out of uncertain labor income approaches zero, so the labor income uncertainty becomes irrelevant to the consumption decision.

The final observation is that for periods before the last one the consumption function lies everywhere below the 45-degree line; that is, consumers choose never to borrow (which they would need to do in order to have \( c > x \) and to be above the 45-degree line), even though no liquidity constraint was imposed in solving the problem.

This last result deserves explanation. As noted above, in the second-to-last period, consumers will always choose to spend less than their cash-on-hand because of the risk of zero income in the last period of life. If we know that in period \( T - 1 \) consumption will be less than \( x \), then that implies that in period \( T - 2 \) the consumer will always behave in such a way to make sure that he arrives in \( T - 1 \) with positive assets, again out of the fear of a zero-income event in \( T - 1 \). Similar logic goes through recursively to any earlier period.

This mechanism for preventing borrowing may seem rather implausible, relying as it does on the slight possibility of disastrous zero-income events. However, essentially
the same logic works as long as income has a well-defined lower bound. For example, suppose the worst possible outcome were that income might fall to, say, 30 percent of its permanent level. In this case the recursive logic outlined above would not prohibit borrowing. But it would prevent the consumer from borrowing more than the amount that could be repaid with certainty out of the lowest possible future income stream. In this case, consumers would define their precautionary target in terms of the size of their wealth holdings in excess of the lowest feasible level. The distinctive features of the model discussed below would all go through, with the solitary difference that the average level of wealth would be lower (perhaps even negative).

This logic provides the simplest intuition for a fundamental conclusion: The precautionary saving motive can generate behavior that is virtually indistinguishable from that generated by a liquidity constraint, because the precautionary saving motive essentially induces self-imposed reluctance to borrow (or borrow too much).

2 Implications

2.1 Concavity of the Consumption Function and Buffer Stock Saving

Perhaps the most striking feature of the converged consumption function $c(x)$ depicted in figure 1 is that the marginal propensity to consume (the slope of the consumption function) is much greater at low levels of cash-on-hand than at high levels. In other
words, the converged consumption function is strongly concave. Thus, the first intuitive result that comes out of the analysis is that, as Keynes (1935) argued long ago, rich people spend a smaller proportion of any transitory shock to their income than do poor people.

Carroll (1996) shows that concavity of the consumption function also implies that impatient consumers will engage in ‘buffer-stock’ saving behavior. That is, there will be some target level of the cash-on-hand ratio $x^*$ such that, if actual cash-on-hand is greater than the target, impatience will outweigh prudence and wealth will fall, while if cash-on-hand is below the target the precautionary saving motive will outweigh impatience and the consumer will try to build wealth up back toward the target. As usual, this result is something that Friedman grasped intuitively: He refers repeatedly to the role of wealth as an ‘emergency reserve’ against uncertainty or a ‘balancing resource’; indeed, Mayer (1972), p. 70 summarizes Friedman’s version of the PIH succinctly: ‘It is basic to [Friedman’s] permanent income theory that households attempt to maximize utility by using savings as a buffer against income fluctuations.’

Buffer-stock saving behavior is a qualitative implication of the model. In order to determine the model’s quantitative implications (for example, what it predicts about the average value of the MPC), it is necessary to simulate a population of consumers behaving according to the converged consumption rule. The first row of Panel A of Table 1 provides a variety of statistics about average behavior when consumers are distributed according to the steady-state distribution generated by the baseline para-
metric assumptions. Columns two and three indicate that the mean and median of the wealth ratio are both about 0.4, or equal to about five months’ worth of permanent noncapital income (remember that the time unit is a year). The average marginal propensity to consume is 0.33, in the ballpark of both empirical estimates and Friedman’s (1957) statement of his conception of the permanent income hypothesis, but a long way from the approximately 0.04 implied by the perfect foresight model under our baseline parameter values.

The second row of Panel A presents results under the assumption that household noncapital income growth is 2 percent a year, rather than the baseline of 3 percent. Lower income growth makes people more “patient,” in the sense that the contrast between tomorrow’s and today’s income – and thus the temptation to borrow against future income – is not as great. The table shows that greater patience leads to a higher mean wealth ratio a lower average MPC.

The final row of Panel A presents results when predictable income growth is zero. With these extremely patient consumers, who cannot rely on future income gains at all, average wealth is much higher, and the average MPC is only about 0.06, not much greater than in the perfect foresight model.

These results confirm that if consumers are moderately impatient, their behavior in the modern model with uncertainty resembles Friedman’s conception of the permanent income hypothesis. Neither liquidity constraints nor myopia is necessary to generate the high average marginal propensity to consume that has repeatedly been found in
empirical studies and that Friedman (1957) deemed consistent with his conception of the permanent income hypothesis. Impatience plus uncertainty will do the trick.

The reason precautionary saving increases the MPC is because the precautionary motive relaxes as the level of wealth rises. To put it another way, an extra unit of cash-on-hand today means that one has a better ability to buffer consumption against income shocks in the future, and so there is less need to depress consumption to build up one’s precautionary assets. Thus, the decline in the intensity of the precautionary motive as cash-on-hand rises allows consumption to rise faster than it would in the absence of a precautionary motive – which is to say, the MPC out of cash-on-hand (and therefore the MPC out of transitory shocks to income) is higher.

Recall that another difference between Friedman and the subsequent models was in the rate at which consumers were assumed to discount future income. In the subsequent models, the mean expectation of future labor income was discounted to the present at a market interest rate (say, 4 percent). Friedman (1963) insisted that future labor income was discounted at a rate of around 33 percent. (A substantial body of empirical evidence confirms that the actual reaction of consumption to information about future income is much smaller than the perfect foresight and certainty equivalent models imply; see Campbell and Deaton (1989); Viard (1993); Carroll (1994); and the large literature that finds that saving responds much less than one-for-one to expected future pension benefits (Samwick (1995)).)

We can examine this controversy in the new model by determining how average con-
Sumption changes when expectations about the future path of income change. Suppose we have a population of consumers who have received their period $t$ income and are distributed according to the steady-state distribution of $x_t$ that obtains under the baseline parameter values. Now consider informing these consumers that henceforth growth will be $G = 1.02$ rather than 1.03. It turns out that under the baseline parameter values, consumers react to the news of the change in income growth as though they are discounting future noncapital income at a 39 percent rate - even higher than Friedman’s estimate of 33 percent! The reason for the high discount rate is that prudent consumers know it would be unwise to spend today on the basis of future income that might not actually materialize.

2.2 The Consumption Euler Equation

Robert Hall (1978) provided the impetus for a large empirical literature over the past two decades by pointing out that in the certainty equivalent model, the predictable change in consumption in a given period should be unrelated to any information that the consumer possessed in earlier periods; consumption should follow a ‘random walk.’ The logic was simple: forward-looking consumers who want to smooth their consumption should react immediately and fully to any information they possess about the future, because if consumption doesn’t react fully and immediately, it will have to finish reacting at some future date, which implies a failure to smooth consumption between the present and that future date. Thus the only reason for a change in consumption is
the arrival of new, previously unknown information. So changes in consumption will not be related to predictable changes in income, or indeed to any other predictable factor. The only exception to the unpredictability of consumption growth comes from the interplay between tastes and opportunities represented by the discount factor and the time preference factor: If interest rates are predictably high, consumption growth will be predictably faster, and vice versa for the time preference rate. A similar analysis shows that there is no relationship between the optimal rate of consumption growth and the average rate of income growth in the perfect foresight model.

Because the formal statement of Hall’s result relied on the mathematical optimality condition known as the Euler equation, tests of the kind Hall proposed became known as Euler equation tests. Many papers in the subsequent literature (though not all) found that consumption growth was strongly related to the predictable component of income growth. This apparent violation of the Euler equation was typically considered evidence either of myopia or of binding liquidity constraints, since the Euler equation does not apply to consumers who are constrained; constrained consumers may set consumption equal to income in every period, even if changes in income are predictable.

The fourth column of Panel A in Table 1 shows that the growth rate of aggregate consumption for our simulated consumers is equal to the predictable underlying growth rate of permanent income for all three values for the rate of income growth. This obviously conflicts with the implication of the perfect foresight and certainty equivalent models that consumption growth should be unrelated to predictable income growth.
The resolution can be found in the analogy between precautionary saving and liquidity constraints. The constrained consumer cannot borrow against future income to finance current consumption; the consumer with a precautionary saving motive chooses not to borrow for fear of the consequences of borrowing and then experiencing negative shocks to income. The ultimate effect is the same: consumption growth can be strongly tied to predictable income growth. Note also that, in both cases, the connection between consumption growth and predictable income growth arises only if consumers are impatient enough; for patient consumers, consumption is less than income anyway, and neither external nor self-imposed constraints have much effect.

Another way of understanding the equality of consumption growth and income growth comes from thinking about the implications of target saving behavior. If consumption growth were forever below income growth then eventually consumption would constitute a vanishingly small proportion of income, and the wealth/income ratio would grow without bound – a contradiction of the target saving proposition. Conversely, if consumption growth were forever above income growth, then consumption would eventually exceed income by an arbitrarily large amount, sending the wealth/income ratio toward negative infinity – again a violation of target saving.

Note further that the stability of wealth around the target level implies that the level of consumption must generally be close to the level of total income, capital and labor. Once again, the prediction of the modern buffer-stock model is close to Friedman’s original idea that people set their consumption to approximately the average value of
expected income over the next few years, and not very close to the implication of the perfect foresight or certainty equivalent models that consumption equals some small MPC multiplied by total remaining lifetime wealth, human and nonhuman.

2.3 Precautionary Saving and Liquidity Constraints

The results of Panel A suggest that a model with precautionary saving and impatient consumers can produce behavior that resembles behavior in a liquidity constrained version of the perfect foresight or certainty equivalent models. This raises the question of whether there are any important differences in behavior between constrained and unconstrained consumers.

The simplest form of liquidity constraint is one in which all borrowing must be collateralized by some marketable asset (as Friedman (1963) points out, human wealth cannot serve as collateral because it cannot be seized and sold - unless anti-slavery laws are repealed!) Appending a constraint requiring consumption to stay below total market wealth, however, has no effect in the model specified above, since the possibility of the dreaded zero-income-events means that these consumers never want to borrow anyway. However, one could plausibly argue that in modern developed countries, the social safety net prevents consumption from falling all the way to zero, mitigating the impact of unemployment spells. To capture a social safety net, suppose now that the worst possible event is an unemployment spell in which income drops to 50 percent of its usual level, an event that occurs with probability $p = 0.05$ representing a 5 percent
unemployment rate. What does optimal behavior look like with such a social safety net if consumers are strictly prohibited from borrowing against future labor income?

Panel B of Table 1 presents some answers. The mean and median amount of buffer-stock wealth are both now around 0.25, or about two months’ worth less of income than in the unconstrained case. Precautionary wealth is lower because the risk of zero income events has now been replaced with a comparatively generous unemployment insurance system. Note, however, that the average MPC is roughly the same as in Panel A. Furthermore, the effect on the MPC of making consumers more patient (by reducing income growth to $G = 1.00$) is also virtually identical to that in Panel A: for patient consumers, the MPC drops to about 6 percent. A final result is that at any given time about 7 percent of households have exactly zero wealth; these are households who have recently experienced bad income shocks and have not had time to rebuild their buffer stocks.

Of course, a complete inability to borrow is unrealistic. But Ludvigson (1996) presents evidence that lenders do attempt to limit the ratio of the borrower’s debt to income. Panel C therefore summarizes behavior in an economy in which lenders restrict consumers’ debt to a maximum of 30 percent of their permanent noncapital income. The effect of a constraint of this type is essentially just to shift the no-borrowing consumption function and wealth distribution to the left by almost exactly 0.3. Note that the steady-state average MPC essentially the same as when consumers were prohibited from borrowing at all. This goes against the grain of intuition, since one might think
that consumers who can borrow should be better able to shield their consumption against income shocks. But remember that precautionary motives are the only reason these impatient consumers do any saving at all. The “buffering capacity” of a given level of wealth depends on how much lower wealth could potentially be driven in the case of a bad shock, so allowing borrowing just shifts the whole consumption function and wealth distribution left, without changing steady-state consumption behavior.

Collectively, the results in Panels A through C of the table demonstrate that liquidity constraints are neither necessary nor sufficient to generate a high MPC. What is both necessary and sufficient is impatience, whether there are constraints or not. (Another point these panels demonstrate is that if the entire population is impatient, the model generates far less aggregate wealth than is observed empirically; see the discussion of this problem in the ‘limitations’ section right before the conclusion).

The point that the average MPC depends on impatience rather than on the presence or absence of constraints means that many traditional tests of liquidity constraints are questionable at best. For example, Campbell and Mankiw (1991) argue that differences across countries in the sensitivity of consumption growth to predictable income growth may reflect differences in the degree of liquidity constraints, while Jappelli and Pagano (1989) suggest that constraints may be stronger in countries in which consumption growth exhibits excess sensitivity to lagged income growth. It is not clear that either of these interpretations is valid. Instead, the warranted conclusion is probably that countries in which consumption exhibits excess sensitivity to lagged or current
income may have more households who are more impatient, and consequently inhabit the portion of the consumption function where the MPC is high, whether they are formally constrained or not.

If empirical evidence on the MPC is not informative about the importance of liquidity constraints, what kind of evidence would be informative? One example is work by Gross and Souleles (2000), who have obtained a database containing comprehensive credit information on a representative sample of consumers. They show that exogenous increases in households’ credit limits result in a substantial increase in actual total debt burdens; in fact, the observed behavior appears to be qualitatively similar to the simulation results presented in Panels B and C of table 1, in the sense that the data show that the debt load after an exogenous credit expansion appears to stabilize at a point that provides roughly the same amount of unused credit capacity as before the expansion in the credit line.

Another way to distinguish precautionary behavior from liquidity constraints is to start with the point, noted above, that the wealth distribution under strict constraints contains a mass of households at exactly zero wealth. Figure 2 presents the cumulative distribution function for data from the 1995 US Survey of Consumer Finances on the ratio of nonhousing wealth to permanent income for US consumers between the ages of 25 and 50 – the age range for which the baseline buffer-stock model has been claimed as a plausible description of behavior.\textsuperscript{13}

Although it is hard to see in the figure, there is indeed a small concentration of
households (about 2.5 percent of the population, as indicated in Table 1) at exactly the zero-wealth point, and a total of about 10 percent have net worth in the range from zero to two weeks’ worth (one paycheck) of their permanent income (not in table). (‘Permanent income’ is defined, a la Friedman, as actual income for the subset of households who said that their income in the survey year was ‘about normal.’)

However, about 20 percent of households in the figure actually have negative financial net worth (uncollateralized loans greater than total financial assets); these people obviously have not been completely constrained from borrowing. Furthermore, a strict prohibition against borrowing flies in the face of daily experience in modern America, where even household pets receive unsolicited offers of credit cards (and sometimes accept them! see Bennett (1999)).

For comparison to the empirical data, therefore, figure 3 presents the theoretical cumulative distribution function (CDF) from a final version of the model in which households are allowed to borrow, but only up to a maximum debt/income ratio of 30 percent and only at a real interest rate of 15 percent (which roughly matches typical credit card interest rates). The qualitative shape of the theoretical CDF is a nice match with the empirical CDF; indeed, Panel D of Table 1 shows that under baseline parameters the proportion of households with negative net worth in the simulations (about 30 percent) is actually larger than in the empirical data (20 percent). This indicates, perhaps surprisingly, that for these moderately impatient consumers, the optimal consumption plan involves borrowing (even at a 15 percent annual rate) for a
substantial fraction of the time.

However, one big problem for the model is evident from a closer look at the upper part of the empirical CDF (Figure 2). Although the empirical median wealth/income ratio, at about 0.3, is in the vicinity of the small values predicted by all the models in Table 1 under baseline parameter values, the upper part of the empirical distribution contains vastly more wealth than is implied by the model; Panel E of Table 1 indicates that the mean value of the empirical wealth-to-permanent-income ratio $w$ is much greater than its median, indicating the strong skewness of the distribution. Thus, while the presence of substantial numbers of impatient consumers may be essential for reproducing the empirical finding of a high average marginal propensity to consume, the presence of some patient consumers is also necessary if the model is to match the overall amount of wealth in the US. Whether a life cycle version of the model with patient and impatient consumers can match the entire distribution of wealth is a matter of ongoing debate; my own view is that the model certainly cannot match the behavior of the richest few percent in the distribution (unless a bequest motive is added), but may be able to match much of the rest.$^{14}$

3 Limitations

I have argued here that the modern version of the dynamically optimizing consumption model is able to match many of the important features of the empirical data on
consumption and saving behavior. There are, however, several remaining reasons for discomfort with the model.

One problem is the spectacular contrast between the sophisticated mathematical apparatus required to solve the optimal consumption problem and the mathematical imbecility of most actual consumers. We can turn, again, to Milton Friedman for a potentially plausible justification for such mathematical modelling. Friedman (1953) argued that repeated experience in attempting to solve difficult problems could build good intuition about the right solution. His example was an experienced pool player who does not know Newtonian mechanics, but has an excellent intuitive grasp of where the balls will go when he hits them. This parable may sound convincing, but some recent work I have done with Todd Allen (2001) suggests that it may sound more convincing than it should. We examine how much experience it would take for a consumer who does not know how to solve dynamic optimization problems to learn nearly optimal consumption behavior by trial and error. Under our baseline setup, we find that it takes about a million ‘years’ of model time to find a reasonably good consumption rule by trial and error. This result may sound preposterous, but we are fairly confident that our qualitative conclusion will hold up, because if there were some trial-and-error method of finding optimal consumption behavior without a large number of trials (and errors), such a method would also constitute a fundamental breakthrough in numerical solution methods for dynamic programming problems. We suspect that the total absence of trial-and-error methods from the literature on optimal
solution methods for dynamic optimization problems indicates that such methods are very inefficient, even compared to the enormous computational demands of traditional dynamic programming solution methods. We conclude by speculating that there may be more hope of consumers finding reasonably good rules in a “social learning” context in which one can benefit from the experience of others. However, even the social learning model will probably take considerable time to converge on optimal behavior, so this model provides no reason to suppose that consumers will react optimally in the short- or medium-run to the introduction of new elements into their environment.

As an example of such a change in the consumption and savings environment, consider the introduction of credit cards. In a trial-and-error economy, many consumers would need to try out credit cards, discover that their heavy use can yield lower utility if they lead to high interest payments, and communicate this information to others before there would be any reason to expect the social use of credit cards to approximate their optimizing use. This social learning process could take some time, and even the passage of a recession or two.

There certainly seems to be strong evidence that many American households are now using credit cards in nonoptimal ways. The optimal use of credit cards (at least as implied by solving the final optimizing model discussed above) is as an emergency reserve to be drawn on only rarely, in response to a particularly bad shock or series of shocks. However, the median household with at least one credit card holds about $7,000 in debt on all cards combined; that $7,000 is the balance on which interest is
paid, not just the transactions use (Gross and Souleles (2000)). Laibson, Repetto, and Tobacman (1999) argue that this pattern results from time-inconsistent preferences in which consumers have a powerful preference for immediate consumption. Their approach is discussed further in the paper in this symposium.

Another set of empirical findings that are very difficult to reconcile with the modern model of consumption presented here comes in the relationship between saving and income growth, either across countries or across households. A substantial empirical literature has found that much and perhaps most of the strong positive correlation between saving and growth across countries reflects causality from growth to saving rather than the other way around (see Carroll, Overland, and Weil (2000) for a summary). This is problematic because the model implies that consumers expecting faster growth should save less, not more (cf. the model simulations in Table 1). Carroll, Overland, and Weil (2000) suggest that the puzzle can be explained by allowing for habit formation in consumption preferences, but as yet, there is no consensus answer to this puzzle.

A final problem for the standard model is its inability to explain household portfolio choices. The “equity premium puzzle” over which so much ink has been spilled (for a summary see Siegel and Thaler (1997)) remains a puzzle at the microeconomic level, where standard models like the ones presented here imply that consumers should hold almost 100 percent of their wealth in the stock market (for simulation results, see, e.g., Fratantoni (1998), Cocco, Gomes, and Maenhout (1998), Gakidis (1998),
4 Conclusion

We shall never cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

- T.S. Eliot, “Four Quartets”

Few consumption researchers today would defend the perfect foresight or certainty equivalent models as adequate representations either of the theoretical problem facing consumers or of the actual behavior consumers engage in. Most would probably agree that Milton Friedman’s original intuitive description of behavior was much closer to the mark, at least for the median consumer. It is tempting therefore to dismiss most of the work between Friedman (1957, 1963) and the new computational models of the 1980s and ‘90s as a useless diversion. But a more appropriate view would be that solving and testing those first formal models was an important step on the way to obtaining our current deeper understanding of consumption theory, just as (in a much grander way) the development of Newtonian physics was a necessary and important predecessor to Einstein’s general theory.

Understanding of the quantitative implications of the new computational model
of consumption behavior is by no means complete. As techniques for solving and simulating models of this kind disseminate, the coming decade promises to produce a flood of interesting work that should define clearly the conditions under which observed consumption, portfolio choice, and other behavior can or cannot be captured by the computational rational optimizing model. Indeed, one purpose of this paper is to encourage readers to join in this enterprise - a process that I hope will be made considerably easier by the availability on the author’s website (see the address on the first page) of a set of Mathematica programs capable of solving and simulating quite general versions of the computational optimal consumption/saving problem described in this paper.
References


Footnotes

1 The uncertainty considered here is explicitly labor income uncertainty. Samuelson (1969) and Merton (1969) found explicit solutions long ago in the case where there is rate-of-return uncertainty but no labor income uncertainty, and showed that rate-of-return uncertainty does not change behavior much compared to the perfect-foresight model.

2 For a rigorous analysis of the relationship between constraints and precautionary behavior, see Carroll and Kimball (2001).

3 My definitions of ‘transitory’ and ‘permanent’ shocks (spelled out explicitly in the next subsection) correspond to usage in much of the modern consumption literature, but differ from Friedman’s (1957) usage. In fact, Friedman (1957) actually states that the MPC out of ‘transitory income shocks’ is zero, but Friedman (1963) was very clear that in his conception of the PIH, first-year consumption out of windfalls was about 0.33. The reconciliation is that such windfalls were not ‘transitory’ shocks in Friedman’s terminology. Terminology aside, Friedman’s quantitative predictions for how consumption should change, for example in response to a windfall, are clear,
so I will simply translate the Friedman model’s predictions into modern terminology without further remark, e.g. by stating that Friedman’s model implies that the MPC out of (my definition of) transitory shocks is a third.

4The evidence included aggregate time series regressions of consumption on current and lagged income; a comparison of the income elasticity of consumption in microdata with the times-series correlation properties of household income; an examination of time series patterns in microdata on saving rates and income; and several other tests.

5For an excellent summary of these studies by Bodkin (1959), Kreinin (1961), Landsberger (1966), and others see Mayer (1972).

6This choice of \( \rho \) implies that a consumer would be indifferent between consuming $66,666 with certainty or consuming $50,000 with probability 1/2 and $100,000 with probability 1/2. For \( \rho = 0 \), the consumer is not risk averse at all and would be indifferent between $75,000 with certainty and $50,000 with probability .5 and $100,000 with probability .5. For \( \rho = \infty \), the consumer is infinitely risk averse, and would choose $50,000.01 with certainty over equal probabilities of $50,000 and $100,000.

7For more details of some of these tricks, see the lecture notes on solution methods
for dynamic optimization problems on the author’s website.

8Recent work by Gourinchas and Parker (1999) finds the switchpoint to be between 40 and 45 rather than 50, but Cagetti’s (1999) similar work suggests a later switching age.

9In fact, Carroll and Kimball (2001) show that as the probability of the zero-income events approaches zero, behavior in the model with zero-income events becomes mathematically identical to behavior in the liquidity-constrained model.

10Carroll and Kimball (1996) provide a proof that uncertainty induces a concave consumption function for a very broad class of utility functions, including the constant relative risk aversion form used here.

11In this case the consumer is on the edge of failing the impatience condition (but the condition does hold because $(R\beta)^{1/\rho} = 0.9992 < 1.00$ under the baseline values for $(R, \beta, \rho) = (1.04, 0.96, 2)).$

12The procedure for calculating an average ‘effective’ interest rate is as follows. First, determine what aggregate consumption would be in period $t$ if consumers continued to
expect $G = 1.03$; call the result $C_t^{0.03}$. Next, find the converged consumption rule under the expectation that $G = 1.02$, and use it to determine how much consumption would be done if consumers’ expectations were suddenly switched to $G = 1.02$ permanently; call that result $C_t^{0.02}$. Finally, find the value of the interest factor $R$ such that, in the perfect foresight model, if growth expectations changed from $G = 1.03$ to $G = 1.02$ then consumption would change by $C_t^{0.03} - C_t^{0.02}$. Unfortunately, the answer that one gets from this methodology for the “effective” interest rate depends very much on how the change in income is distributed over time, its stochastic properties, the level of current wealth, and all of the other parameters of the model.

13Housing and vehicle wealth have been excluded on the grounds that the model does not pretend to capture the complexities associated with durable goods investment. See Carroll and Dunn (1997) for simulation results showing that even when durable goods are added to the model, buffer stock saving behavior emerges with respect to liquid asset holdings.

14See Huggett (1996), Dynan Skinner and Zeldes (1996), Quadrini and Rios-Rull (1997), Engen, Gale, and Uccello (1999), and Carroll (2000b) for several perspectives on this
question. For general equilibrium macro models which attempt to match both mi-
cro and macro data using mixed populations of patient and impatient consumers, see
Figure 1: Convergence of Consumption Functions $c_{T-n}(x)$ as $n$ Rises
Figure 2: Empirical CDF of Ratio of Net Worth to Permanent Income, 1995 SCF
Figure 3: Steady-State Wealth Distribution with Credit Card Borrowing
The table below presents the steady-state statistics for alternative consumption models. The table is divided into five panels: A, B, C, D, and E. Each panel presents the mean (\(\text{Mean} \ w\)) and median (\(\text{Median} \ w\)) levels of consumption growth for different income growth rates (\(G\)). The table also includes the mean MPC (\(\text{MPC} \ w\)) and the fractions with certain conditions (\(\text{Frac With} \ w < 0\) and \(\text{Frac With} \ w = 0\)).

### Table 1: Steady-State Statistics For Alternative Consumption Models

<table>
<thead>
<tr>
<th>Income Growth Factor</th>
<th>Mean Consumption Growth</th>
<th>Median Consumption Growth</th>
<th>Mean MPC</th>
<th>Frac With (w &lt; 0)</th>
<th>Frac With (w = 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(G=1.03)</td>
<td>0.43</td>
<td>0.40</td>
<td>1.030</td>
<td>0.330</td>
<td>0.000</td>
</tr>
<tr>
<td>(G=1.02)</td>
<td>0.52</td>
<td>0.48</td>
<td>1.020</td>
<td>0.276</td>
<td>0.000</td>
</tr>
<tr>
<td>(G=1.00)</td>
<td>2.26</td>
<td>2.06</td>
<td>1.000</td>
<td>0.064</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Panel A. Baseline Model, No Constraints**

| \(G=1.03\)          | 0.28                    | 0.24                     | 1.030   | 0.361               | 0.000               | 0.070               |
| \(G=1.02\)          | 0.36                    | 0.32                     | 1.020   | 0.301               | 0.000               | 0.051               |
| \(G=1.00\)          | 2.28                    | 2.06                     | 1.000   | 0.065               | 0.000               | 0.000               |

**Panel B. Strict Liquidity Constraints**

| \(G=1.03\)          | -0.03                   | -0.06                    | 1.030   | 0.361               | 0.611               | 0.000               |
| \(G=1.02\)          | 0.06                    | 0.01                     | 1.020   | 0.299               | 0.478               | 0.000               |
| \(G=1.00\)          | 1.94                    | 1.71                     | 1.000   | 0.064               | 0.023               | 0.000               |

**Panel C. Borrowing Up To 0.3 Allowed**

| \(G=1.03\)          | 0.11                    | 0.07                     | 1.030   | 0.327               | 0.320               | 0.058               |
| \(G=1.02\)          | 0.21                    | 0.16                     | 1.020   | 0.274               | 0.210               | 0.046               |
| \(G=1.00\)          | 2.11                    | 1.89                     | 1.000   | 0.064               | 0.007               | 0.002               |

**Panel D. Borrowing Up to 0.3 at \(R = 1.15\) Allowed**

| \(G=1.03\)          | 0.12                    | 0.29                     | 0.007   | 0.025               |

**Panel E. Statistics from the 1995 SCF**

Notes: Results in Panels A through D reflect calculations by the author using simulation programs available at the author’s website, http://www.econ.jhu.edu/people/carroll/ccaroll.html. In Panel A, no constraint is imposed, but income can fall to zero, which prevents consumers from borrowing. In Panels B through D, the worst possible event is for income to fall to half of permanent income. For comparison, Panel E presents the mean and median values of the ratio of nonhousing wealth to permanent income from the 1995 Survey of Consumer Finances for non-self-employed households whose head was aged 25-50; the measure of permanent income is actual measured household income for households who reported that their income over the past year was ‘about normal’, and whose reported income was at least $5000; other households are dropped. The program that generates these statistics (and figure 2) is also available at the author’s website.