16 Theories of Consumption and Saving

Chapter 16 Contents

A. Topics and Tools

There is no topic in macroeconomics that has a longer, deeper, or more prominent literature than households' choice of how much of their income to consume and how much to save. As we saw earlier in the course, the theory of consumption is central to the model of Keynes's *General Theory*, which is often considered to be the origin of macroeconomics. Since then it has been the subject of countless theoretical and empirical studies.

Keynes treated consumption on a very "common sense" level. Like most other economists of his day, his methodology included neither abstract, mathematical theory nor detailed econometrics. Rather he relied almost entirely on intuition, as he demonstrates when he introduces the central principle of his consumption theory in Chapter 8:

> 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 their income increases, but not by as much as the increase in their income. Keynes (1936)

Keynes gives no basis for his theory in terms of utility maximization nor indeed gives any consideration of *why* a consumer would behave in the way he assumes. In place of rational-choice theory, Keynes relies on his "knowledge of human nature." Nor does he give any support using numerical data, but instead claims to glean support from "detailed facts of experience." How much economics has changed in 75 years!

While Keynes placed consumption theory at the center of the macroeconomic stage, he left it for future generations of economists to work out the microeconomic basis for his theory and competing theories. Keynes also inspired pioneers in the emerging field of econometrics to swarm over the newly invented national income and product statistics looking for verification or refutation of his model.

Keynes's basic model of consumption was that current consumption expenditures are determined mainly by current disposable income. The Keynesian consumption function is usually written in linear form: $C_t = a + bY_t$. The coefficient b, which Keynes called the "marginal propensity to consume" or MPC and which we would define concisely as $\partial C/\partial Y$, was to vie for the title of "most estimated coefficient" for several decades. Initial linear econometric consumption functions estimated by ordinary least squares produced results that conformed to Keynes's theory: consumption seemed to be closely related to current disposable income and the MPC seemed to be positive and less than one. However, Nobel-laureate Trygve Haavelmo used the consumption example prominently in pointing out the bias that is present in ordinary least-squares (OLS) estimation when shocks to the dependent variable (consumption, in this case) cause changes in the "independent" variable (income). Since Keynes's theory places aggregate demand at the center of output determination, aggregate consumption changes would be expected to affect aggregate income strongly. When the "Haavelmo problem" was accounted for, the corrected estimates of the MPC turned out to be considerably lower than OLS estimates.

At about the same time, Simon Kuznets (another Nobel winner) refined national-account measures of income and consumption and pointed out a paradox that could not be explained by the simple linear consumption function. The ***Kuznets paradox*** was that the percentage of disposable income that is consumed is remarkably constant in the long run, which suggests a proportional consumption function, i.e., that the intercept term a is equal to zero. However, estimates across individual households or using short-run aggregate time-series fluctuations in income and consumption consistently produce estimates implying that $a > 0$, which means that the share of income consumed declines as income rises. Explaining the Kuznets paradox became a major goal of consumption theorists in the 1950s.

One early approach was the "relative-income hypothesis," which asserted that a household's consumption depends not only on its current disposable income, but also on current income relative to past levels and relative to the income of other households. This hypothesis enjoyed considerable popularity in the 1950s, but then entered a long period of dormancy. Recent research in the growing field of behavioral economics has turned its attention on the effects of relative levels of consumption and has revived interest in models that revive some of the ideas underlying the relative-income hypothesis.

Two other theories pioneered by Nobel laureates, the life-cycle model associated with Franco Modigliani and the permanent-income hypothesis developed by Milton Friedman, were easier to reconcile with microfoundations of consumer choice. These two theoretical approaches have largely merged to become "modern consumption theory." In their original forms, they differed mainly in that the life-cycle theory emphasized natural variations in earnings over a finite lifetime whereas the permanent-income model stressed general variations in income over an indefinite horizon.

Although we call the model based on intertemporal utility maximization "modern," it is really a straightforward extension of standard microeconomic theory. Irving Fisher used this framework in his theory of interest developed in the 1920s. Modern macroeconomists have extended the basic theory in several ways. One of the truly modern extensions of the theory is the application of dynamic mathematical methods to the problem of utility maximization. A second major innovation is the modeling of uncertainty and expectations in a rigorous way. Finally, modern macroeconometricians have devised ingenious ways of testing the validity of intertemporal utility-maximization theory.

B. The Kuznets Paradox

Keynes called the relationship between aggregate consumption and current disposable income the "propensity to consume." He gave names to two measures of the sensitivity of consumption to income. The ***average propensity to consume*** (APC) is the ratio of consumption to income: C/Y; the ***marginal propensity to consume*** (MPC) is the amount by which consumption increases as current disposable income rises by a dollar, $\partial C / \partial Y$. Both the average and marginal propensities are generally believed to be between zero and one. The Kuznets paradox is an empirical anomaly that relates to the relative size of these two measures.

The linear Keynesian consumption function, which dominated early empirical work, is written as

$$C_t = a + bY_t. \tag{1}$$

The MPC in equation (1) is the constant b, since in a linear function the marginal effect (slope) is constant. The APC is $C_t / Y_t = (a + bY_t) / Y_t = b + a / Y_t$. How the APC varies as income changes depends on a. If $a > 0$, then the MPC < APC and people spend a decreasing share of their incomes as incomes rise. If $a = 0$, then the MPC = APC and spending is a constant proportion b of income.

Empirical estimation of equation (1) by ordinary least squares with aggregate time-series data generally yields a value of b in the neighborhood of 0.75 and a positive value of a. Thus, early empirical estimates led to the prevailing wisdom that the MPC was less than the APC. A common interpretation of this result is that saving was a "luxury" good, whose share of overall income rises as people received higher incomes.[1]

However, in an oft-cited but unpublished work based on his detailed reconstruction of historical data on economic aggregates, Simon Kuznets pointed out that the share of income consumed seemed to remain constant over almost a century of data spanning the latter half of the 19th century and the first half of the 20th.[2] If the APC > MPC as the OLS estimates of the linear consumption function suggest, then the share of income consumed should *decline* as income increases. Thus two kinds of empirical evidence seemed to lead to conflicting conclusions: short-run econometric studies found MPC < APC and long-run data showed that MPC = APC.

The conflict between short-run and long-run evidence is shown graphically in Figure 1. The long-run consumption function has a slope equal to the long-run APC (and MPC). The short-run consumption functions shown have a slope (MPC) that is smaller than the APC.

Economists working on consumption models also sought evidence from cross-section studies of the consumption expenditures of individual households. This evi-

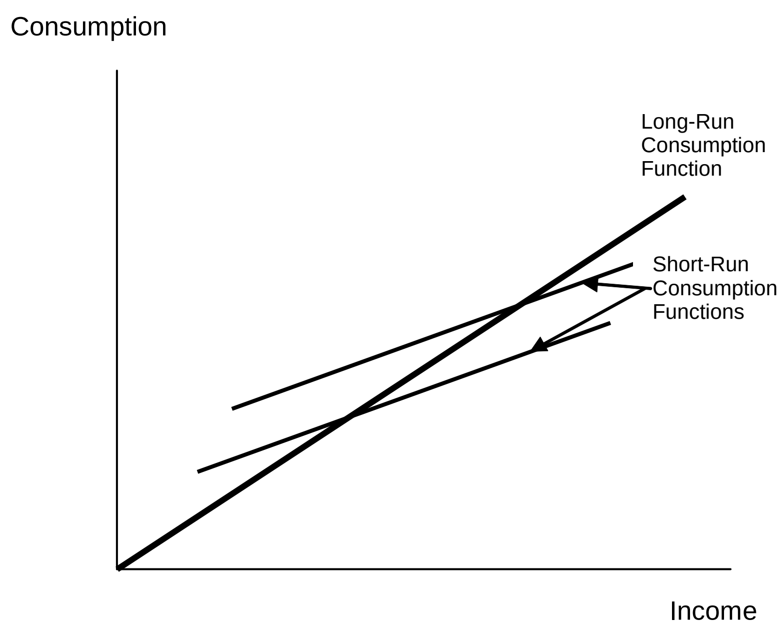
---

[1] Note the inconsistency between the idea of thinking of saving as a luxury good, or even a good at all, and the notion that saving is simply future consumption. Modern consumption theory, while not comfortable at all with the notion of saving as a luxury, achieves a similar result by introducing the possibility of liquidity-constrained consumers.

[2] The first publication of Kuznets's result was in a 1942 NBER Occasional Paper entitled *Uses of National Income in Peace and War*.

dence seemed to support the short-run econometric studies, showing that high-income households saved a larger fraction of their income than lower-income households did.

Figure 1. Short-run and long-run consumption functions.



The Kuznets paradox posed a challenge for theoretical modelers of consumption. Clearly the linear Keynesian consumption function was insufficient, since it could not explain why the MPC was less than the APC in the short run and across households, yet aggregate consumption was proportional to income over the long run. The early postwar theories that were devised with this paradox in mind eventually led economists to a model with very basic microeconomic roots. In the next two sections we examine briefly some mileposts on this path of theoretical evolution.

C. Relative-Income Hypothesis

One of the earliest attempts to reconcile these conflicting pieces of evidence about the consumption-income relationship was the ***relative-income hypothesis***, described by James Duesenberry (1949). Although this theory has vanished with hardly a trace from contemporary macroeconomics, it carried considerable influence in the 1950s and 1960s.[3] The reasons for its abandonment may have had less to do with its logic or conclusions than with its lack of conformity with assumptions that microeconomists commonly make about utility functions.

The relative-income model was formulated in two variants: a cross-section version and a time-series version. These variants correspond to the cross-section and time-series aspects of the Kuznets paradox. In both variants, consumption depends on current income *relative* to some income standard that the household sets based on its own past income or on the income of other households around it. In the cross-section version, Duesenberry appealed to the idea of "keeping up with the Joneses." He argued that a household's consumption would depend not just on its own current level of income, but on its income relative to those in the subgroup of the population with which it identifies itself. The household will attempt to align its consumption expenditures with those of other members of its group. Thus, households with lower income within the group will consume a larger share of their income to "keep up," while households with high incomes relative to the group will save more and consume less.

This hypothesis gained support from the observation that families with the same income seemed to consume systematically different amounts depending on the group to which they belonged. For example, survey evidence indicated that a black family with a given income would usually consume less than a white family with the same income. The relative-income hypothesis attributes this to the difference in their *relative* income within their respective groups. Because average incomes among whites were higher, the white family was presumed to consume more relative to its income in order to try to attain parity with other white families, while the black family feels less of this pressure among the group of black families. Thus, the two flatter lines labeled "short-run consumption function" in Figure 1 might represent the cross-section

[3] A quick perusal of the indexes from six leading macroeconomics texts failed to turn up a single reference to the relative-income hypothesis.

consumption functions of whites (the higher one) and blacks (the lower one). As incomes of both groups rise over time, both flatter lines would tend to slide up the steeper "long-run consumption function," with the average household in each group tending to spend a constant share of its income over time.

The time-series variant of the relative-income hypothesis is very similar to the cross-section version. The main difference is that instead of comparing their income to those of other households, each household is assumed to consider its current income relative to its own past income levels. A household that has in the past achieved income levels higher than its present levels would attempt to maintain the high consumption levels that it achieved earlier. Thus, when incomes fall, consumption would not fall in proportion. (Note that this is not totally inconsistent with our modern theory of consumption smoothing, though the basis for smoothing in the modern theory is the household's average lifetime income, not the highest level of past income.)

The result of this behavior for aggregate consumption is called a "ratchet effect." When incomes rise, consumption increases along the steeper long-run consumption function. However, when a recession hits and incomes decline, households reduce consumption less than proportionally and fall back along the flatter short-run consumption function. During the recovery, they move up along the flat line until they reach their highest attained level of consumption. After recovery, when incomes grow again, they proceed up the long-run line again until the next recession, when they fall back along a flatter line. Thus, consumption ratchets upward, staying relatively near its highest past value when income declines.

Although the relative-income hypothesis is quite successful in explaining the Kuznets paradox, it seems to have been relegated to the economic scrap heap. One important reason is that the cross-section variant involves ***interdependent utility functions*** in which one household's utility depends not only on its own consumption activities but also on those of other households. This greatly complicates the problem of modeling consumption behavior. Instead of being able to model each household's behavior in isolation, taking as given its income and market prices, one must model all households' consumption decisions together in a game-theoretic framework, taking into account how the behavior of other households affects each family's consumption behavior. Although modern developments in game theory have made such problems a little more approachable than they were in the 1950s, it is still a lot harder to build consumption models if utility is interdependent. Thus, the cross-section relative-income model may be described as "methodologically inconvenient."

Another reason that economists tend to avoid models with interdependent utility functions is that the socioeconomic, "as good as my neighbor" competition that it implies conflicts with economists' favorite depiction of "Homo economicus" as a self-contained, rational, maximizing machine. Once one opens up the possibility of cross-person utility dependence, extension of that idea makes the model so general that it is consistent with nearly any imaginable behavior. Thus, in place of the set of testable propositions about consumption behavior that come out of the individual utility-maximization model, interdependent utility functions may leave theorists with an untestable model that can explain any behavior imaginable through some combination of interdependence.

However, problems of intractability and conflict with economists' usual behavioral assumptions do not make a theory wrong. Indeed, recent theories of behavioral economics are breathing new life into this approach and we may see the relative-income hypothesis back on the stage as some point.

Ultimately, the abandonment of the relative-income hypothesis surely resulted partially, if not mostly, from the development of other, more attractive consumption models that were equally successful at explaining empirical phenomena such as the Kuznets paradox. We now turn to these models, which form the underpinning of modern consumption theory.

D. Life-Cycle Model and Permanent-Income Hypothesis

Two initially distinct theoretical paths that eventually merged into one are the ***life-cycle model*** developed by Franco Modigliani, Albert Ando, and Richard Brumberg in the mid-1950s and the ***permanent-income hypothesis*** introduced by Milton Friedman in 1957.[4] Both models emphasize consumption smoothing, though they vary a little in how they are set up. Later work showed that both could be viewed as special cases of the general intertemporal utility maximization model. Their relationship to one another is somewhat analogous to the Ramsey and Diamond growth models. The life-cycle model, like the Diamond overlapping-generations model, features a finite lifetime with a distinct period of retirement at the end. The permanent-income model, like the Ramsey model, has infinitely lived consumers.

[4] The seminal paper describing the life-cycle model is Modigliani and Brumberg (1954). The permanent-income model is laid out in Friedman (1957).

Life-cycle model

Modigliani's model emphasized how saving could be used to transfer purchasing power from one phase of life to another. In early life, labor income is usually low relative to later working years. Income typically peaks in the last part of the working life, then drops at retirement. Consumers who wish to smooth consumption would prefer to borrow during the early low-income years, repay those loans and build up wealth during the high-income years, then spend off the accrued savings during retirement.

Implicit in the life-cycle approach is the idea of a lifetime budget constraint that links consumption at various dates during the lifetime. The slope of the budget constraint, which determines the tradeoff between period t consumption and period $t + 1$ consumption, is $-(1 + r)$, where r is the real interest rate at which consumers lend and borrow. The position of the budget constraint depends on the present value of lifetime earnings, which is usually simply called ***wealth***. In terms of the modern utility-maximization model, wealth is

$$\Omega_0 = A_0 + \sum_{t=0}^{T} \frac{Y_t}{(1+r)^t}, \tag{2}$$

where Ω_0 is the stock of wealth (human and nonhuman) as of time zero, A_0 is the value of current nonhuman (financial or physical) assets, Y_t for $t = 0, 1, 2, \ldots, T$ is the expected stream of real labor income over the lifetime, and r is the real interest rate.

The early empirical tests of the life-cycle model were tests of whether wealth and the interest rate explained consumption better than current disposable income. Although some successful results were obtained, empirical work was bedeviled by the difficulty of measuring the stock of wealth accurately. In general, government statisticians are much more successful at measuring flows than stocks. Stocks are more difficult to measure for at least three reasons. First, because flows "move," it is easier to count them (and harder to hide them) than stocks of assets that may "hide" in someone's possession for many years. Second, income, sales, and expenditures are often taxed, which means that the government has good reasons for measuring these flows as accurately as possible.

The final reason that flows are easier to measure than stocks is that their value is usually easier to determine. Most economic variables are aggregated in terms of their dollar value. Each time a transaction occurs a dollar value is placed on the goods involved. Flows by definition involve current transactions and, thus, have a readily observable current value. Assets often change hands infrequently, so it can be difficult to assess their current market value. Prices are regularly quoted for assets that are traded on organized exchanges, such as equities (stocks), bonds, and gold, which makes it easy to establish their value. For other assets, such as real estate, tax collectors make regular estimates of market value. However, for a very large collection of assets, data collectors are forced either to use historical cost (the approach taken by accountants, which may drastically underestimate the value of structures and overestimate the value of such rapidly depreciating assets as computers) or to estimate market value based on whatever scanty information is available. The largest asset of most households in the economy is the earning power represented by the human capital of their members. Since historical cost is largely irrelevant here, this can only be estimated very crudely by trying to guess at their lifetime stream of future wage earnings and place a capital value on it by standard present-value techniques.

The difficulty of measuring wealth makes it very difficult to perform a reliable test of the life-cycle model. The most common approach is to include as wealth only a limited set of assets whose value is relatively easy to measure. In terms of equation (2), this amounts to using only the "visible" part of the A_0 term and neglecting both any unmeasurable components of A_0 *and* the potentially much larger unobservable summation of discounted future labor income. Since the appropriate concept of wealth is so much broader than the measures that are used in empirical applications, one would not necessarily expect a strong correlation between the measures used and consumption spending. Thus, the lack of robust statistical support for this version of the life-cycle model compared to the simple Keynesian function cannot be taken as a definitive refutation of the model.

Permanent-income hypothesis

Rather than focusing on the life cycle *per se*, Friedman discussed the general problem faced by households when their income fluctuates over time, whether due to life-cycle effects, business cycles, or other factors. He considered infinite-lived households and distinguished between a "normal" level of income that they expect over their lives, which he called ***permanent income***, and (positive or negative) deviations from that level, which he termed ***transitory income***.[5]

Similarly, Friedman distinguished ***permanent consumption***, which is the part of consumption that is planned and steady, from unexpected or irregular spending or ***transitory consumption***, such as unexpected medical bills or temporary college tuition expenses. Friedman argues that permanent consumption will be proportional to permanent income. Households will plan to spend in an average period a fraction (equal to one or slightly less) of their average lifetime income.

[5] Friedman's formal definition of permanent income was the amount a household could consume "without reducing its wealth." Since the household lives forever, this means intuitively that the household can in each period consume only the "interest" on its human and financial wealth and can never consume the principal. Thus, permanent income can be thought of as the annual return on households' stocks of human and nonhuman wealth.

He further assumed that both permanent and transitory consumption are independent of transitory income and that transitory consumption in any period is independent of permanent income. Thus, consumption consists of a planned part that depends on permanent income and an unplanned part that is totally independent of income. Transitory consumption can be identified with the random error term in a consumption-function regression. The focus of the permanent-income model, then, is the estimation of the relationship between consumption and a measure of permanent income.

In terms of the modern consumption model, permanent income can be thought of as the size of a constant annual flow of income that would have the same present value as the (possibly uneven) flow of income that is actually expected. If we know the future income path, we can calculate permanent income from the budget constraint as

$$\sum_{t=0}^{\infty} \frac{Y^p}{(1+r)^t} = A_0 + \sum_{t=0}^{\infty} \frac{Y_t}{(1+r)^t} \ , \tag{3}$$

where Y^p is permanent income. It can be shown that $Y^p = r\,\Omega$, where Ω is the wealth measure from equation (2).[6] This shows the close relationship between the life-cycle model, in which consumption is assumed to depend on wealth, and the permanent-income model, where consumption depends on permanent income.

Early empirical estimation of the permanent-income model relied on the rather shaky assumption that future income could be predicted as a stable linear function of current and past incomes. Under this ***adaptive-expectations model***, permanent income could be expressed as a linear function of current and past incomes. However, this model of expectations was often very inaccurate because it failed to distinguish between changes in income that people knew were permanent and those they knew were temporary. Although some supportive empirical results were reported, modern macroeconomists approach them with great skepticism.

Permanent vs. temporary changes in income

Both the life-cycle and permanent-income models make similar predictions about the consumption effects of permanent and temporary changes in a household's income. In the life-cycle model, an increase in income that is expected to be permanent

---

[6] To see this, notice that Y^p does not depend on t in the summation on the left-hand side. Thus, we can factor Y^p out of the summation. The remaining infinite summation of $1/(1+r)^t$ can be shown to converge to $1/r$. Multiplying both sides of the equation by r gives Y^p as r times the wealth expression on the right-hand side, which is Ω from the life-cycle model except with an infinite lifetime. This mathematical expression coincides with the idea of permanent income as an interest return on human and nonhuman wealth.

causes a large increase in lifetime wealth, since all future terms on the income side of the budget constraint rise along with the current term. Thus, consumption rises by about as much as income rises when the change is known to be permanent—the MPC out of a permanent change in income is near one.

A temporary increase in income affects only the current term in the lifetime-income summation, so it causes a relatively small change in wealth. As a result, households that smooth consumption will spread the temporary increase in income over the rest of their lives, increasing consumption in the current period (and in each future period) by about $1/T$ times the change in income, where T is the number of years left in the household's life. Thus, the MPC out of temporary changes in income is much smaller in the life-cycle model.

In the permanent-income model, permanent changes in income are changes in permanent income, and thus lead to large changes in consumption. Again, the MPC is near one. Temporary changes are transitory and, thus, have no direct effect on consumption. A one-time rise in income does raise lifetime wealth dollar for dollar, so it does have a small effect on permanent income: Since permanent income is the real interest rate times wealth, permanent income goes up by the interest rate times the amount of the temporary change in income. Thus, if (infinitely lived) households receive a one-time increase in income, they will consume the interest they can earn on that increase in the current year and in every future year. (Note that if an infinitely lived household consumed even a little of the principal each year in addition to the interest, it would eventually exhaust the principal and have to lower consumption.)

The life-cycle and permanent-income models explain the Kuznets paradox through the difference in the reaction of consumption to permanent and temporary changes in income. Recall that during long-run growth, Kuznets found that the MPC was quite high and equal to the APC. According to the life-cycle and permanent-income models, such changes are the result of long-run growth forces and are likely to be permanent. Thus, a high MPC and a proportional response to income are consistent with the predictions of the model.

Business cycles, on the other hand, are almost always quite short-lived. In response to temporary reductions (or booms) in their income, both models predict that households will try to spread the income reduction (increases) over their entire lives, lowering (raising) consumption only slightly. Thus, the flatter short-run consumption function with a smaller MPC results from the temporary nature of short-run fluctuations.

The dangers of ignoring macroeconomists: A policy faux pas

The empirical relevance of the life-cycle/permanent-income model was demonstrated by a remarkable failure of macroeconomic policy in the late 1960s. With spending on the Vietnam War and accommodating monetary policy sending the

U.S. macroeconomy into a raging boom and toward accelerating inflation, the Johnson administration sought temporarily to reduce aggregate demand. The method they chose was a temporary income-tax surcharge under which, in 1968, everyone would pay an extra 10% on top of their normal income-tax bill. The policy was explicitly announced as a temporary measure.

The government's economists used simple statistical models to estimate the "average" response of consumption to changes in income and thus predicted that consumption would decline significantly as a result of the surcharge, producing the desired decrease in demand. However, as the life-cycle and permanent-income models point out, consumers respond quite differently to a change in their incomes that they know is temporary than to an "average" one in which temporary and permanent elements are mixed together. Indeed, as predicted by the life-cycle/permanent-income model, consumers largely paid for the surcharge by dipping into their savings, so consumption spending declined only a little.

As a result of this miscalculation, the government greatly overestimated the deflationary effect of the surcharge. Demand remained unexpectedly strong and inflation continued to rise. As the life-cycle and permanent-income models make clear, the policy would have more nearly had the desired effect if people had believed the surcharge to be permanent. In retrospect, one can argue that the appropriate policy was to enact a "permanent" increase in taxes, then to repeal it a few years later when deflationary policy was no longer needed. Of course, if people correctly anticipated that the policy would be reversed, then they would treat the surcharge as temporary anyway, so the success of the "permanent" surcharge would depend on the government being able to fool the public. Experience suggests that governments (like everyone else) can, at best, fool some of the people some of the time!

E. Understanding Romer's Chapter 8

The mathematics in Chapter 8 are pretty tame compared to similar material on intertemporal consumption in the growth chapters you have already read. Romer treats consumption in discrete time and even ignores discounting through most of the chapter. If anything, you may be confused by the absence of the $1/(1 + r)^t$ expressions that you are familiar with from the Diamond overlapping-generations model.

Romer's equations (8.1) and (8.2) are simple representations of lifetime utility and the lifetime budget constraint over a finite lifetime when the interest rate and rate of time preference are both zero. The Lagrangian expression (8.3) is formed as the objective function we are maximizing minus the Lagrange multiplier (λ) times the constraint (written as an expression that is equal to zero when the constraint is satisfied).

The first-order conditions for a maximum are found by partially differentiating expression (8.3) with respect to each C_t and with respect to λ and setting these partial derivatives equal to zero. The partial derivatives with respect to each C_t are $u'(C_t) - \lambda$. Setting this expression equal to zero gives equation (8.4). This equation does not allow us to solve directly for C_t, but the fact that it holds for every period implies that the marginal utility of consumption must have the same value (λ) in every period. Since marginal utility is the same decreasing function of consumption in each period, marginal utility will only be equal across periods if consumption is also equal in every period, thus the household will smooth consumption.

Note that this result corresponds exactly to our results from the Ramsey and Diamond growth models. In those models, individuals would choose a rising, constant, or falling path of consumption depending on whether the interest rate was higher than, equal to, or less than the marginal rate of time preference. Since both rates are assumed to be zero here, they are equal and consumers prefer a flat consumption path with consumption equal in every period.

As in the Ramsey and Diamond models, the *level* of the consumption path is determined by the budget constraint. Given that consumers want a flat consumption path, what level of flat consumption can they afford? In this case, the algebra is simple. Since there is no discounting, they can consume $1/T$ of their lifetime earnings in each of the T remaining periods of their lives. This is shown in Romer's equation (8.5).

Empirical consumption functions

The discussion of the empirical estimates of the marginal propensity to consume that begins on page 368 is not as difficult as it looks. Equation (8.8) probably looks as though it dropped out of nowhere, but with a little analysis it is quite easy to understand. This analysis follows Friedman directly, so you might want to go back and review the permanent-income section of this chapter if you get bogged down in the notation or terminology.

One novelty here is the statistical concept of ***covariance***. The covariance between two variables is a generalization of the concept of variance of a single variable and a close relative of the familiar concept of correlation. Like correlation, covariance measures the degree to which the two variables tend to move together, independently, or in opposite directions.[7] We often prefer to use a correlation coefficient rather

[7] The formal relationship between covariance and correlation is that $\text{cov}(x, y) = \text{corr}(x, y) \times \text{s.d.}(x) \times \text{s.d.}(y)$, where s.d. stands for the standard deviation of a variable, the square root of its variance.

than covariance to express this relationship because the correlation coefficient does not depend on the units in which the variables are measured. However, we need the units to be retained in this application, so we use covariance rather than correlation.

The first expression in (8.8) is the formula for a least-squares regression coefficient in the case of one explanatory variable. This is the formula that we would use to estimate *b* in equation (8.7) by ordinary least squares. In the second equation of (8.8), he has substituted permanent plus transitory for total income and permanent income for consumption. The final line results from applying the assumption that permanent income and transitory income are uncorrelated, which means that their covariance is zero. In the numerator,

$$\operatorname{cov}\left(Y^T + Y^P, Y^P\right) = \operatorname{cov}\left(Y^T, Y^P\right) + \operatorname{cov}\left(Y^P, Y^P\right) = \operatorname{var}\left(Y^P\right)$$

because $\operatorname{cov}(Y^T, Y^P) = 0$ and the covariance of any variable with itself is just the variable's variance. In the denominator, the variance of the sum of two random variables equals the sum of their variances *only if the two variables are independent*. In this case, permanent and transitory income are independent, hence the third equation holds.

What does this equation mean? It simply means that if the world really operates according to the permanent-income model, the estimated MPC out of current income in an empirical consumption function will equal the share of the sample variation in income that is due to variations in permanent income. If all income fluctuations during the sample were permanent, then the MPC would be one (because Var $(Y^T) = 0$). If all variation in income were due to transitory movements, then the estimated MPC would be zero. If the fluctuations were due half to permanent and half to transitory changes in income, then the MPC would be estimated to be one-half. Note that the estimate of the MPC does not really tell us anything about the consumption response to any *particular* kind of income change. It merely tells us which kind of income change predominated during the sample period. Basing forecasts or policy recommendations on an empirical consumption function such as this can lead to the kind of policy blunder that we examined above—the 1968 income-tax surcharge.

Uncertainty, rational expectations, and consumption

We made an obviously unrealistic assumption in the first section of Romer's chapter (and in our growth models). Namely, we assumed that households know their lifetime income paths with certainty. A more reasonable though more analytically challenging approach would be to assume that households base their consumption choices on expectations about their future income paths. Although these forecasts are usually assumed to be correct on average ("unbiased," in statistical terms), they may be too high or too low in any given year, so they differ from "perfect foresight." Moreover, as new information about the economy becomes available over time, people should use that information to revise their forecasts.

The way that expectations are embodied in nearly all modern macroeconomic models is through the concept of ***rational expectations***. Expectations variables have always been problematic for macroeconomists, especially in empirical work. We have no reliable data series on what values people expect macroeconomic variables to take in the future, so in order to use expectations in an empirical model we must "solve out" the expectations variables, replacing them with some function of observable variables. We have encountered rational expectations in our consideration of modern theories of aggregate supply. Robert Lucas used the concept to demonstrate his famous and controversial proposition about the ineffectiveness of systematic monetary policy rules.

Beyond their importance in theoretical models, rational expectations are well suited to empirical work. Under rational expectations, agents' expectations of a future variable are assumed to correspond to a conditional mathematical expectation based on the predictions of the model itself. These mathematical expectations, in turn, can be approximated by using a least-squares regression to estimate an optimal forecasting equation for the variable, using as right-hand variables all the important variables that could be observed by agents at the time that the expectations were formed. Robert Barro produced the first tests of the ineffectiveness of anticipated monetary policy using this kind of empirical implementation of rational expectations.[8]

We denote the rational expectation of a variable y_t that is formed based on variables that are observable in period $t - 1$ as $E_{t-1}[y_t]$. As noted above, rational expectations have two important properties. First, rational expectations are unbiased, so

$$E_{t-1}[y_t] = E[y_t \mid X_{t-1}],$$

where the expression on the right-hand side refers to the mathematical expectation of y_t conditional on the set of variables X_{t-1}. The set of conditioning variables X_{t-1} is called the ***information set*** on which the expectation is based. In principle, it should include any variable that is observable at time $t - 1$, though empirical applications must always make do with a small set of variables that seem especially relevant.

The second important property of rational expectations is that the forecast error $y_t - E_{t-1}[y_t]$ is statistically independent of any variable that appears in the information set X_{t-1}. Any information that is available at the time a rational expectation is formed should be incorporated in the forecast itself. Thus, it cannot be correlated

[8] The two papers are Barro (1977) and Barro (1978).

with the forecast error. This important condition can be tested, since we can conduct statistical tests to see whether or not two variables are correlated.

In order to make models with rational expectations distinct from the assumption of ***perfect foresight***, agents' expectations must sometimes be wrong. If all variables in the model have exactly predictable relationships to one another—i.e., if the model is ***deterministic***—then agents' rational expectations will always be exactly correct. To create models in which agents' forecasts are sometimes wrong, we add random disturbance terms to some of the equations of the model, making the model ***stochastic***. Macroeconomic modeling changed in a fundamental way in the 1970s with the development of stochastic modeling techniques to incorporate rational expectations.

Quadratic utility

One further aspect of Romer's Chapter 8 that requires discussion is his choice of the quadratic utility function in equation (8.10) rather than the constant-relative-risk-aversion function he used in Chapter 2. The reason that Romer adopts the quadratic utility function here is discussed at the bottom of page 374 and the top of page 375. If utility is quadratic then uncertainty enters the model in a particularly simple way. Utility maximizers make decisions based on marginal utility—the partial derivatives of the utility function with respect to consumption at various dates. Because marginal utility is linear when utility is quadratic, and because the mathematical operation of taking an expectation is a "linear operator," people with quadratic utility functions operate according to the principle of ***certainty equivalence***. This means that they make the same decisions in maximizing expected utility that they would have made in maximizing utility in a world in which they *knew for certain* that the values of future variables would be at the levels that they *expect* them to be. Certainty equivalence makes the analysis of stochastic models much easier, so quadratic utility functions are common in such models.

A drawback of the quadratic utility function is that it implies the existence of a ***bliss point***. The graph of a quadratic function is a parabola, which has a peak. If the level of consumption exceeds this peak level, then marginal utility is negative, which is unrealistic. Thus, we must think of the quadratic utility function not as a function that is valid for all values of consumption, but as one that approximates the true utility function over a subset of the range of consumption values that lie to the left of the peak.

F. Empirical Tests of the Random-Walk Hypothesis

The neoclassical permanent-income/life-cycle model leads to a counterintuitive conclusion: the shape of a consumer's time path of consumption should be independent of the shape of his or her time path of income. This remarkable result follows directly from our analysis of consumption in the Ramsey growth model. Recall that the growth rate of consumption $\dot{C}/C$ depends only on the interest rate, the current level of *k*, and the parameters of the consumption function ρ and θ. That means that whether consumption is rising or falling does not depend on whether the consumer's income is rising or falling. The effect of income on the consumption path lies entirely in its effect on the level (not on the growth rate at each date). But the level of the consumption path depends only on the present value of *lifetime* income, not on when the income is received. Thus, the timing of income is totally irrelevant for consumption behavior and the theory implies that there is no association at all between rises and falls in income and rises and falls in consumption. (This also mirrors Friedman's assumption of zero correlation between transitory income and transitory consumption.)

We need to modify this result slightly when we incorporate uncertainty about income into the model. If a household's income in period *t* is higher than it was expected to be, this will usually change its perception of lifetime income, perhaps by a little or perhaps by a lot depending on whether the change is expected to be long-lasting or quite temporary. An *unexpected* change in income usually changes the household's expectation of the present value of its lifetime wealth, so it would cause the household's planned consumption path from that time forward to shift upward or downward to reflect that change. Consumption in period *t* is the first section of the higher or lower anticipated consumption path, but it is the only one that we actually observe; in period $t + 1$ there may be additional shocks to the income stream and the consumption path may have shifted again. However, notice that only *unexpected* changes in income would cause the consumption path to shift. Changes in income at date *t* (relative to $t - 1$) that are correctly anticipated at $t - 1$ will not cause changes in consumption at time *t* relative to the anticipated path formulated at $t - 1$.

This independence of consumption changes from expected changes in income is knows as the ***random-walk hypothesis*** of consumption. Notice that the random-walk hypothesis is not a separate theory but rather an implication of the neoclassical model. It was first explored in a seminal study by Robert Hall (1978).[9]

[9] See Romer's curious footnote number 6 on page 375. In personal correspondence, David Romer has claimed that the "prominent macroeconomist" in the footnote can be identified from information elsewhere in the text. Can you figure it out?

Romer discusses the random-walk hypothesis beginning on page 372. He starts with the simple discrete-time model with a zero rate of time preference and a zero interest rate. The zero interest rate implies that the consumer can trade current for future consumption at a one-for-one rate in the market. The zero rate of time preference assures that a consumer who is smoothing consumption is indifferent about making that trade.

Romer's equation (8.11) expresses the first-order condition for utility maximization: the marginal utility of consumption in period one must equal the consumer's expectation of the marginal utility of consumption in each future period. Assuming rational expectations, we can associate the consumer's expectations with the mathematical expectation of the variable. Carrying the expectation operator through the expression on the right-hand side of equation (8.11) yields the conclusion in equation (8.12) that the expected value of consumption in all future years equals the level of consumption in period one.

This simple result implies that the consumer chooses a perfectly flat consumption path from years 1 through T. It is so simple because of the assumptions that both the interest rate and the rate of time preference are zero. As we know from our consumption analysis in the Ramsey growth model, utility-maximizing consumers choose a rising, flat, or falling time path of consumption depending on whether the interest rate is greater than, equal to, or less than the consumer's rate of time preference. Thus, the model is somewhat more complicated when these assumptions are relaxed, but the main idea still holds: changes in consumption from one period to the next do not depend on correctly anticipated changes in income.

Romer's equation (8.19) shows what *does* cause changes in consumption from period 1 to period 2. The expression in parentheses on the right-hand side is the *change* in the household's expectation of its future income that occurs based on information that becomes available in period two. Thus, changes in expectations of income (even if they are changes that are expected to happen in future periods) will cause the household to revise its consumption path and make second-period consumption differ from first-period consumption.

Testing the random-walk hypothesis

Hall's test of the random-walk hypothesis is based on an equation similar to (8.12). He does not make the assumptions of a zero interest rate and a zero rate of time preference, so his equation is a little more complicated than (8.12). He considers several possible forms for the utility function, but the one that works out to be most convenient (because of certainty equivalence) is a quadratic function

$$u(c) = -\tfrac{1}{2}(\bar{c} - c_t)^2, \tag{4}$$

where $\bar{c}$ is the "bliss level of consumption" at which the quadratic utility function reaches its hypothetical peak. Because he allows for positive time preference and positive interest, Hall's lifetime utility is given by

$$E\left[U\right]=E\left[\sum_{t=1}^{T}\frac{1}{\left(1+\rho\right)^{t}}u(c_{t})\right] \tag{5}$$

and the budget constraint is

$$\sum_{t=1}^{T}\frac{E_{1}\left[C_{t}\right]}{\left(1+r\right)^{t}}=A_{0}+\sum_{t=1}^{T}\frac{E_{1}\left[Y_{t}\right]}{\left(1+r\right)^{t}}. \tag{6}$$

If the utility function is as shown in equations (4) and (5) and the budget constraint is as in (6), then consumption follows the following time path:

$$c_{t+1}=\bar{c}\ \frac{r-\rho}{1+r}+\frac{1+\rho}{1+r}\ c_{t}+\varepsilon_{t+1}\equiv\beta_{0}+\gamma c_{t}+\varepsilon_{t+1}. \tag{7}$$

The coefficient $\gamma = (1 + \rho)/(1 + r)$ in equation (7) should look somewhat familiar. The ratio of $1 + \rho$ to $1 + r$ appeared (in inverse form) in the consumption relationship for the Diamond overlapping-generations model. (See Romer's equation (2.47) on page 79.)

The disturbance term ε_{t+1} in equation (7), which is treated as a random variable, captures the effect on the consumption path of changes in income that the household finds out about in period $t + 1$. Thus, equation (7) expresses consumption in period $t + 1$ as a constant term β_0 plus a coefficient γ times period t consumption, plus a random term that reflects *new* information received at $t + 1$, and is therefore uncorrelated with anything that was known at time t.

Equation (7) is ideally suited for analysis using linear regression, because the most important assumption that we make in a regression is that the random error term is not correlated with the independent variables that appear on the right-hand side of the equation. According to the random-walk hypothesis, this assumption is valid as long as all the variables on the right-hand side (just c_t for now) are known at time t.

If consumption follows the relationship shown in equation (7) and if $r = \rho$, then consumption is called a ***random walk***, which is where the hypothesis gets its name. The logic behind this label is that in each period, the value of consumption can be thought of as "taking a step" relative to its past position. Notice that $r = \rho$ implies $\gamma = 1$ and $\beta_0 = 0$, so the right-hand side of (5) reduces to $c_t + \varepsilon_{t+1}$. In other words, consumption in period $t + 1$ equals consumption in t plus a random step. Following a path on which one takes a sequence of random steps leads one on a "random walk."

Hall sums up the key testable implication of the random-walk theory as follows:

> No information available in period t apart from the level of consumption c_t helps predict future consumption, c_{t+1}, in the sense of affecting the expected value of marginal utility. In particular, income or wealth in periods t or earlier are irrelevant, once c_t is known. Hall (1978)

This implies that if we were to consider adding more variables to equation (7) as possible explanatory variables for future consumption, "no variable observed in period t or earlier will have a nonzero coefficient if added to this regression" Hall (1978).

Hall tested this hypothesis by estimating equation (7) using quarterly U.S. data on per-capita consumption of nondurables and services from 1948 to 1977.[10] He tested whether including additional variables that were known at time t on the right-hand side of equation (7) yielded coefficients that differed significantly from zero. The most obvious variables to add are values of income at time t and earlier. These variables proved to have no marginal predictive power for c_{t+1}, i.e., their coefficients in Hall's regression were not statistically different from zero. This result supports the random-walk implication of the permanent income-life cycle model.

However, when Hall added the market value of corporate stock (as a measure of one kind of wealth), he found that values of this variable that were known at time t did have some marginal predictive value for future consumption. Since the hypothesis implies that *any* variable known at time t should have a zero coefficient if added to (7), this outcome was not consistent with the hypothesis.

Hall downplayed the importance of this rejection of the random-walk hypothesis by stressing that lagged consumption is more effective than lagged wealth as an explanatory variable and that the most obvious candidate as an additional explanatory variable—lagged income—fails to have predictive power, just as the theory implies. However, other macroeconomists interpreted this result as being more damaging to the theory. The evidence is, thus, somewhat mixed on whether changes in economic conditions that were anticipated in advance cause changes in the consumption path.

Later tests and excess sensitivity of consumption to income

Another testable empirical question is whether the amount of consumption change due to a given *unexpected* income change is consistent with the amount predicted by the life cycle-permanent income model. Marjorie Flavin (1993) studied this

[10] The reasons why including consumption of durable goods creates problems for the analysis are discussed below.

by examining the average degree to which unexpected changes in income affected lifetime income, then relating the change in consumption to the change in expected lifetime income. She found substantial evidence of ***excess sensitivity*** of consumption to current changes in income. Changes in income seemed to cause consumption to change by more than would be consistent with the average amount of change in lifetime income that should result.

Flavin's result triggered several hypotheses to attempt to explain excess sensitivity. One explanation that is tested implicitly in two studies described in Romer's chapter (Campbell and Mankiw (1989) and Shea (1995)) is that some agents are subject to ***liquidity constraints***. The usual version of the life cycle-permanent income theory assumes that households are able to borrow and lend freely at the same real interest rate r. However, we know that many low-income households find it difficult to borrow large amounts of money at any interest rate, even if they have good reason to expect that their lifetime income will be high enough to allow them to repay the loan. Such households will be unable to borrow enough to smooth consumption. The closest they can come to smooth consumption is often to consume all of their current income and wait for higher future incomes in order to achieve their higher desired level of lifetime consumption. Thus, households subject to binding liquidity constraints may consume all of their disposable incomes, but still be below their desired consumption levels. Temporary changes in income will thus cause much larger changes in consumption for these households than would be predicted by their effect on lifetime income.

As suggested by the two studies examined by Romer on pages 375–379, the evidence on liquidity constraints is mixed. A well-known paper by Stephen Zeldes (1989) finds substantial evidence of constraints among low-income households. He tests for differences in the excess sensitivity of consumption between households with very low levels of nonhuman wealth and those with higher levels. He finds that the low-wealth households exhibit much stronger excess sensitivity, which is consistent with the hypothesis that excess sensitivity is caused by liquidity constraints if low-wealth households are more likely to be constrained than those with higher nonhuman wealth to use as collateral against loans.

Another suggested explanation for the sensitivity of overall consumption expenditures to changes in income is expenditures on ***durable goods***. Recall that Hall included only nondurables and services in his consumption variable. The reason is that expenditures on durables include an element of saving, since a durable good purchased in the current period yields utility services throughout its service life.

In macroeconomics, consumption is modeled from two sides. On one side, it is a flow of goods and services that yields utility to households. One the other, it is an expenditure that drains income within the limitations of the budget constraint. For the first side, we denote the utility derived from period t consumption as $u(C_t)$; on the other side, we subtract C_t / $(1 + r)^t$ from the resources available for other uses in the budget constraint. The use of the very same C_t on both sides is not problematic as long as consumption is paid for in the same time period that it yields utility. However, once we allow the household to purchase durable goods, the tight temporal connection between expenditure and utility is broken.

Exactly what concept of consumption belongs in the utility function? Strictly speaking, one gains utility from the flow of *consumption services* provided by consumption goods, not the goods themselves. When the consumption good in question is nondurable (or, more obviously, a service), then the flow of utility-yielding services and the purchase of the good occur in the same period. When the good is durable, the utility-yielding services accrue to the user over many years, not just in the year of the purchase.

However, since a household that buys a consumer durable pays for it all at once, the expenditure (and thus the demand from the viewpoint of producers and of aggregate demand in the economy) occurs at the beginning, when it is produced and purchased.[11] Such consumption expenditures are the variable that belongs in the budget constraint. Thus, we have two distinct concepts of consumption: services and expenditures, whose timing may differ when the goods in question are durable. This means that a household that desires an increase in consumption services, perhaps because its budget constraint shifted out unexpectedly, may increase consumption spending by more (if it buys a new durable good) or less (if it uses a durable good it already owns more intensively) than it increases consumption services.

This problem is most extreme in the case of housing. It would obviously be inaccurate to model households' consumption of owner-occupied housing as occurring entirely in the year in which a house is bought. Many households would appear to consume housing only once in their lives, consuming over $100,000 worth of housing all at once and never consuming any more! Because of the absurdity of this example, national income accountants treat housing differently than other consumer goods. Purchases of (new) housing units are considered investment, even if households purchase them for their own use (consumption). Each month, the national income accountants pretend that the household rents the house from itself, with an estimated monthly rental price being charged as consumption of (housing) services on one side and as imputed income from owner-occupied housing on the other. Thus, in the housing case, national accounts correct for the asynchrony between spending and consumption services by inventing a rental transaction.

[11] Note that the household's spending on the durable good occurs at the time that it is bought even if the household buys on credit. The good belongs to the consumer, who may simultaneously incur a debt.

This problem also applies to a lesser degree to all other consumer durables. Automobiles, refrigerators and other appliances, furniture, and clothing are among the important categories of goods that are purchased all at once but yield their service flows over a long period. Even stocking-up trips to Costco could be thought of as saving if the inventories of goods are used over a long period. However, the national income accountants in most countries treat consumption of non-housing durables in the same way they treat nondurables and services—as though consumption spending and consumption services occurred at the same time.

Because the consumption statistics that are published in the national income and product accounts reflect expenditures on durables, not the flow of services from durables, economists may err if they assume that increases in this consumption variable directly reflect the utility that households obtain from durables in any given period. There are two approaches to solving this problem. One is to estimate service lives and depreciation patterns for major categories of durable goods, then estimate a stock of consumer durables that increases with investment and decreases with depreciation. Consumption services from durables are then assumed to be proportional to the stock.

This procedure raises many measurement problems. It is likely that the estimated services lives and depreciation patterns will be at best very crude approximations. Households may also vary the intensity with which they use their stock of consumer capital (as when they drive their cars more or less during one year than another). The potentially serious inaccuracy of the method of service-flow imputation has led most modern modelers of consumption behavior to follow Hall's example and to ignore durables completely and to focus on consumption expenditures on services and nondurables.

G. Works Cited in Text

Barro, Robert J. 1977. Unanticipated Money Growth and Unemployment in the United States. *American Economic Review* 67 (1):101-115.

———. 1978. Unanticipated Money, Output, and the Price Level in the United States. *Journal of Political Economy* 86 (4):549-580.

Campbell, John Y., and N. Gregory Mankiw. 1989. Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence. *NBER Macroeconomics Annual* 4:185-216.

Duesenberry, James S. 1949. *Income, Saving, and the Theory of Consumption Behavior*. Cambridge, Mass.: Harvard University Press.

Flavin, Marjorie. 1993. The Excess Smoothness of Consumption: Identification and Interpretation. *Review of Economic Studies* 60 (3):651-666.

Friedman, Milton. 1957. *A Theory of the Consumption Function*. Princeton, N.J.: Princeton University Press.

Hall, Robert E. 1978. Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy* 86 (6):971-987.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money*. New York: Harcourt, Brace.

Modigliani, Franco, and Richard Brumberg. 1954. Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data. In *Post-Keynesian Economics*, edited by K. K. Kurihara. New Brunswick, N.J.: Rutgers University Press.

Shea, John. 1995. Union Contracts and the Life-Cycle/Permanent-Income Hypothesis. *American Economic Review* 85 (1):186-200.

Zeldes, Stephen P. 1989. Consumption and Liquidity Constraints: An Empirical Investigation. *Journal of Political Economy* 97 (2):305-346.