



Milton Friedman's Scientific Legacy

**Thomas S. Coleman, PhD
Executive Director,
Becker Friedman Institute for Research in Economics**

8 November 2012

By any measure, Milton Friedman has had a tremendous impact on our world. Friedman is one of the most recognizable economists of the 20th century. He is known to millions for *Free to Choose*, the book and television series he produced with his wife, Rose Director Friedman; for his 19 years of *Newsweek* columns; and for his 1976 Nobel Memorial Prize in economics. He was an untiring advocate of economic freedom, free markets, economic liberalism, and “the small man”.

But the source of Friedman’s continuing influence lies in the power of his ideas—precepts built on the twin pillars of sound economic theory and careful empirical analysis. History tells us that ideas matter, and have the power to change our world. Friedman’s ideas have changed the way economists, governments, and citizens understand our economy and society. As Friedman himself said, the function of economists is to provide a “stockpile of ideas” that are available as solutions when a crisis arises. And Friedman has provided an abundance of ideas.

Any discussion of Friedman’s legacy must include his lasting influence on both the economics profession and on the wider world. Both result from his profound and deep contributions to economic theory, methodology, empirical analysis, and economic history. By all accounts, Friedman was a powerful intellect and formidable opponent in debates. But he was principally a scientist who took seriously the goal of understanding the world around him. He rigorously challenged his own and others’ ideas with empirical evidence. Friedman himself seemed most proud of his intellectual contributions: “The thing I will really be proud of is if some of the work I have done is still being cited in the textbooks long after I am gone.”

Consumption Function

Ideas that Friedman developed and championed have become part of the economist’s lexicon. And these ideas continue to have importance and relevance to all of us today. Chief among them is *A Theory of the Consumption Function*, Friedman’s landmark 1957 book and a work that was cited when Friedman was awarded the 1976 Nobel Memorial Prize in Economic Sciences.

The consumption function is an important concept in economic theory, Keynesian macroeconomics in particular. Introduced by J.M. Keynes in 1936 in *The General Theory of Employment, Interest, and Money*, it represents the relation between aggregate income and consumer spending. Keynes presumed the relationship was stable, or at least stable enough to act as the fundamental building block for the multiplier, the mechanism by which an increase in aggregate expenditure produces a larger increase in demand.

In its simplest form, the consumption function would be a linear (technically affine) function:

$$C = c_0 + c_1Y$$

The coefficient c_1 is the marginal propensity to consume and measures how much of each additional dollar of income consumers spend (versus save). Note that the coefficient c_1 need not be a constant; it can depend on such factors as the level of interest rates, the average level of wealth, and distribution of income and wealth across the population, but it is presumed to be fairly stable over time.

Empirical studies, however, soon showed some important and challenging inconsistencies, particularly between observations from cross-sectional or household-level studies on the one hand and time series or aggregate data on the other. When examining behavior across individuals or households at a point in time, economists found that consumers with higher incomes saved more and their consumption showed much less than a one-for-one response—from an extra dollar of income, only a small portion was consumed and a large fraction was saved. And yet analysis revealed a different trend when they examined aggregate data over long time spans or across countries with widely differing incomes. Economists found a roughly constant share of national income being saved.

Taking an example from the present-day United States, household savings was 4.3% in 1929 and 4.2% in 2011, although income grew by roughly 14 times over that period. This means that although the absolute level of consumption increased in line with income, the share of consumption stayed roughly the same. In other words, the cross-sectional observations appeared to be inconsistent with both the aggregate observations and with the assumptions behind the Keynesian consumption function

Friedman showed that the Keynesian foundations of the consumption function were fundamentally flawed. His explanation was as simple as it was brilliant; his own words provide the simplest explanation:

*The central theme of this monograph can be illustrated by a simple hypothetical example. Consider a large number of men all earning \$100 a week and spending \$100 a week on current consumption. Let them receive their pay once a week, the paydays being staggered, so that one-seventh are paid on Sunday, one-seventh on Monday, and so on. Suppose we collect budget data for a sample of these men for one day chosen at random, defined income as cash receipts on that day, and defined consumption as cash expenditures. ... It may well be that the men would spend more on payday than on other days but they would also make expenditures on other days, so we would record the one-seventh with an income of \$100 as having positive savings, the other six-sevenths as having negative savings. Consumption might appear to rise with income, but, if so, not as much as income, so that the fraction of income saved would rise with income. [Chapter 9 of *A Theory of the Consumption Function*, Princeton University Press, 1957]*

So we have cross-sectional data showing that those with higher incomes save more and that the ratio of savings to income rises with income (the ratio of consumption to income

declines). But in reality there is no savings (considering all the men together) and presumably if the incomes for all these men rose from \$100 to \$120 they would all consume \$120 and there would still be no savings overall. In other words, at the aggregate level there would be no tendency for savings to rise with income and the propensity to consume out of income would be one.

Friedman's hypothesis was that many of the puzzling and inconsistent results regarding savings and consumption were simply a result of "inappropriate concepts of income and consumption." In the example of staggered paydays, individual savings rise and the ratio of consumption to income declines with rising income only because we are measuring daily income. If we used an appropriate measure of income, weekly income in the simple example, then there would be no rise in savings with income, and consumption (the ratio of consumption to income) would not change with income.

Building on the idea behind the simple example, Friedman introduced the concept of permanent income as distinct from the measured income reported by individual consumers or measured at the aggregate level. Although difficult to define, we might think of permanent income as corresponding roughly to the long-term expected income or lifetime income or wealth. As Friedman states, "the concept of permanent income is easy to state in these general terms, hard to define precisely." In the example above, the analog of permanent income would be individuals' \$100 weekly income. Measured income would be either \$100 or \$0, depending on which day we measure an individual's income.

The difference between measured income and permanent income is transitory income:

$$\text{measured income} = \text{permanent income} + \text{transitory income} .$$

Friedman's central hypothesis can now be simply stated, in two parts. First, consumers respond to changes in permanent income but largely ignore changes in transitory income. Second, measured income is a combination of both permanent and transitory income.

The result is that any attempt to relate consumption to measured income will not measure a behavioral relationship but a statistical artifact. Say that transitory income is a large component of measured income, as will likely be the case when examining individual household income for a large number of differing households. Some will have low income year after year, but some will simply have a poor year (unemployment, a poor bonus, or some other bad luck). Similarly some will have high income because of unusually good luck or other purely transitory reasons.

Among households with high measured incomes, there will be a large portion with high transitory income and these households will not increase their consumption in response to the

transitory income. So the measured consumption will not rise with measured income, or more accurately will do so much less than it would versus permanent income.

Consider, in contrast, a situation where transitory income is a small component of measured income. In this case, consumption would be expected to rise much more with income because a rise in measured income would represent primarily a rise in permanent income. This would often be the case for aggregate data because across the whole economy, there would be households with both high and low transitory income, and the transitory component would tend to average out across households. Changes over the years in measured aggregate income would thus tend to reflect primarily changes in permanent income.

This simple hypothesis can, at least conceptually, account for the seeming inconsistency of the cross-sectional and the time series observations. Across households at a point in time, those households with higher measured income tend to have a large component of higher transitory income, and consumption will generally not go up with transitory income. This means the cross-sectional response of consumption to higher measured income will be relatively small (the marginal propensity to consume is low; the ratio of consumption to income goes down as income goes up; the savings rate goes up as income goes up).

For aggregate income (averaging or aggregating across households at a point in time) the transitory income will tend to average out, so measured aggregate income and changes in aggregate income over time will tend to measure permanent income. This implies that for long time spans using aggregate income, the response of consumption to income will be higher than for cross-sectional observations.

This idea was not completely new—it can be traced back to the writings of the one of the Bernoulli clan in the 18th century. Friedman’s genius, however, was twofold—first, in taking this simple idea and fully exploring and exploiting its implications, and second, in truly testing the theory by marshaling empirical evidence in support of the hypothesis but also searching for evidence that would disprove it.

Friedman believed that ideas and hypotheses had to be put up against data. The hypothesis alone—the fact that in theory it could explain the difference between the cross-sectional and the time series data (the microeconomic and the macroeconomic data)—was not sufficient. A useful economic theory must be able to account for the quantitative observations. Much of *A Theory of the Consumption Function* was dedicated to measuring and testing the data to determine whether the permanent income hypothesis could in fact account for the differences between the microeconomic and macroeconomic data.

Methodology

Friedman took a simple hypothesis, explored its implications in order to develop novel predictions, and then confronted those predictions with empirical evidence. In doing so, he was simply being true to another of his seminal contributions to economic science, his 1953 essay “The Methodology of Positive Economics” (from *Essays in Positive Economics*). This work remains as fresh and relevant to modern economics as it was 60 years ago and, rightly, it remains on the reading list for graduate students today (at least at the University of Chicago).

In this essay, Friedman argued that the truth or value of economic theories is not tested by examining their assumptions, but by developing their predictions and testing those predictions with empirical observation. He asked, “Can a hypothesis be tested by the realism of its assumptions,” and his definitive answer was “No”.

Friedman argued for judging hypotheses and economic theories by their results, by the correspondence of their predictions with observations, rather than by their assumptions. This was the application to economics of the philosophy of science associated with the work of Karl Popper. Friedman’s agenda followed what Imre Lakatos would later call a “progressive research programme.” The approach emphasizes the value of a simple theory that can explain much with little input, judging by the results rather than the assumptions:

A hypothesis is important if it “explains” much by little, that is, if it abstracts the common and crucial elements from the mass of complex and detailed circumstance surrounding the phenomena to be explained and permits valid predictions on the basis of them alone. To be important, therefore, a hypothesis must be descriptively false in its assumptions; it takes account of, and accounts for, none of the many other attendant circumstance, since its very success shows them to be irrelevant for the phenomena to be explained.

“The Methodology of Positive Economics” was and remains controversial. Most economists would argue that complete disregard of assumptions is going too far. Nonetheless, the methodological program Friedman advocated has proved extremely fruitful, with *A Theory of the Consumption Function* being a leading example of its application.

Turning back to the consumption function, we should note that Friedman’s permanent income hypothesis is not the only theory of the consumption function. The life-cycle hypothesis, (credited to Modigliani and Brumberg in 1954) proposed consumption and savings as a function of wealth over people’s lifetimes, with younger workers saving more of their income and older workers spending or dissaving.

Permanent Income Today

Nonetheless, Friedman’s concept of permanent income remains powerful and relevant today. How much consumers spend out of increased income was a vital question when governments undertook the substantial fiscal stimulus in response to the financial crisis of 2007–2008, and

the question remains even more relevant today as governments consider fiscal austerity. The justification for fiscal stimulus is that during a recession every \$1 of government spending or tax cut would create \$1 or more of economic growth—the so-called fiscal multiplier. The worry during a period of fiscal austerity is that the reverse will occur, so every \$1 of government spending cuts or increased taxes will cut growth.

The justification behind the fiscal multiplier is that government spending or tax cuts will increase consumer income, consumers will spend out of that increased income, and this will lead to further increases in income, further spending, and so on. But the critical question is how much consumers will actually spend out of the increased income—a large portion (a high propensity to consume and low propensity to save) will lead to a larger multiplier. And so we return to the concept of permanent versus transitory income, asking whether consumers treat the increased income from government stimulus as permanent or transitory, because if it is transitory the propensity to consume will presumably be small.

The Natural Rate of Unemployment (NAIRU) and the Phillips Curve

Friedman's contributions to methodology and the theory of the consumption function are substantive and have stood the test of time, but they are only part of Friedman's scientific legacy. His presidential address to the American Economic Association ("The Role of Monetary Policy," *American Economic Review* 1968) is masterful, containing fundamental insights that still resonate today. A central argument concerned the Phillips curve, the view that macroeconomic policy could trade off higher inflation on the one hand for higher growth and lower unemployment on the other. (These ideas are due to Friedman and also to Edmund Phelps; see "Phillips Curve, Expectations of Inflation and Optimal Unemployment Over Time," *Economica* 34, August 1967.)

The naïve Phillips curve posits "a stable negative relation between the level of unemployment and the rate of change of wages—high levels of unemployment being accompanied by falling wages, low levels of unemployment by rising wages." (quoted from Friedman's 1976 Nobel Prize address, "Inflation and Unemployment"). The move from rising wages to rising prices in general is then easy to imagine. This posited relationship is superficially appealing because it seems like, with low unemployment, employers would have to raise wages. But it confuses nominal wages with real wages.

Friedman's argument around the Phillips curve and the natural rate of unemployment was simple, and he elaborated upon it in his Nobel address. At its core, unemployment depends on real wages but inflation may obscure the pattern of real wages and changes in wages. The first part of the argument asserts that there is a "natural rate" or equilibrium rate of unemployment implied by real quantities—real wages, preferences, demand, production functions, and so on. The word "natural" in this context does not have a normative meaning,

nor does it imply that it is unchanging. The term simply means the rate that will result from equilibrium in the markets. The natural rate has also come to be called the NAIRU or non-accelerating inflation rate of unemployment.

The second, and truly insightful part of the argument, showed that unexpected inflation may push employers and workers from the natural rate, with positive inflation causing a lower unemployment rate—the classic Phillips curve. But most importantly, this effect is only temporary and depends on the inflation being unexpected. This was a powerful, in fact fatal, argument against the Phillips curve as a policy tool that could be used to “fine tune” the economy and exploit a trade-off between higher inflation versus lower unemployment.

To see the second part of the argument, let us start by ignoring inflation. Consider an employer who experiences a rise in the price of his output relative to other goods—this effectively lowers the real wage and will induce an increase in the firm’s labor demand. Then consider an employee who experiences a rise in the wage relative to all other good. This effectively raises the real wage and will induce an increase in labor supply.

So far this is straightforward price theory—labor supply and demand both depend on real wages, with demand going up when real wages go down and supply going up when real wages go up. But then Friedman applied a key insight: Inflation can obscure changes in prices in just such a way that makes it seem that real wages go down for employers and up for employees. This factor induces employers to demand more labor, workers to supply more, and unemployment to go down.

Friedman’s argument that inflation can make real wages appear to go both down and up is simple, although on further consideration it is a deep insight into the purpose and power of the price mechanism and the costs of inflation in obscuring and degrading the signals provided by prices.

When an employer experiences a rise in the price of his output he will take this, at least partly, as a rise in the real price of his output. Again, Friedman says it best (from his 1976 Nobel address):

In an environment in which changes are always occurring in the relative demand for different goods, he [the employer] will not know whether this change is special to him or pervasive. It will be rational for him to interpret it as at least partly special and react to it, by seeking to produce more to sell at what he now perceives to be a higher than expected market price for future output. He will be willing to pay higher nominal wages than he had been willing to pay before. ... A higher nominal wage can therefore mean a lower real wage as perceived by him.

...

To workers the situation is different: what matters to them is the purchasing power of wages ... over all goods in general. ... A rise in nominal wages may be perceived by workers as a rise in

real wages and hence call forth an increased supply, at the same time that it is perceived by employers as a fall in real wages and hence calls for an increased offer of jobs.

When prices change because of unexpected inflation, employers and workers cannot easily and quickly determine whether they are seeing a change in relative prices or absolute prices. Indeed, by its very definition, unexpected inflation is not anticipated. As such, employers and workers will tend to interpret (incorrectly) the change in the price level as a change in relative prices (real wages, in this case). Employers will interpret it as a fall in real wages and workers as a rise in real wages. This will induce employers to increase labor demand and workers to increase labor supply. In other words, it will push employment up and unemployment down, below the equilibrium rate (the natural rate).

But employers and workers will not be fooled forever, probably not even for long. They will learn that the change in product prices and wages is not a change in real prices and wages but is only due to inflation. We should expect the unemployment level to temporarily fall below the natural rate in response to unexpected inflation but then rise back. There will be a Phillips curve (unemployment goes down when inflation goes up), but it will be only a short-run effect.

If the inflation were fully anticipated, we should not expect any response of unemployment to inflation. This is why Friedman argued that any policy maker attempting to exploit the short-run Phillips curve would have to generate not a steady level of inflation, but ever-accelerating and unexpected inflation.

Friedman's analysis is as relevant today as it was in the 1960s. There is still discussion of the relation between price changes (inflation) and growth. There is a widely-held presumption that high growth (and low unemployment) are related to or cause inflation, while slow growth and high unemployment inhibit inflation because of "deficient demand" or "slack in the economy." As for the original Phillips curve the argument is superficially appealing since it relates higher demand to increases in prices. Like the original Phillips curve, however, there is a tendency to conflate nominal prices and real price changes.

A Monetary History of the United States 1867–1960

I will not discuss this book, or Friedman's monetarism, at any length. Friedman is renowned as a monetarist and this book, coauthored with Anna Schwartz, still stands as one of the foundations of monetary economics. As the title says, this is a monetary history of the United States and in reading this book one recognizes the importance of history in understanding our current world. As George Santayana so famously said, "Those who cannot remember the past are condemned to repeat it."

The financial crisis of 2007–2008 is an event unique in our lives. But it is by no means unique in the history of the United States— we see it as unusual only because our experience is too limited. Prior to the Great Depression of the 1930s financial and banking panics occurred with some regularity — in the late 1870s, 1893-94, 1907-08, 1920-21. The data, narrative, and analysis of the long periods covered by *A Monetary History* illuminate current events in a way that a shorter, more limited history cannot.

Friedman and Schwartz’s analysis of the monetary aspects of the Great Depression (or Great Contraction, as they termed it) transformed economists’ view of those events. Contemporary economists, John Maynard Keynes among others, believed that monetary conditions in the United States were loose following the 1929 stock market crash. They took from this a lesson that monetary policy was ineffective in the face of a crisis such as befell the United States (and the rest of the world). Friedman and Schwartz showed, definitively, that this was not the case, that monetary policy was tight, and that Federal Reserve actions, or inactions, led to a substantial fall in money supply (money stock fell by roughly one-third from August 1929 to March 1933).

The warnings of Friedman and Schwartz did not fall on deaf ears. In a striking exchange at a celebration of Friedman’s 90th birthday in 2002, Ben Bernanke (then a governor rather than the chairman of the board of governors of the Federal Reserve) thanked Friedman and Schwartz for showing how the Federal Reserve’s actions and inactions in the 1930s set the conditions for the Great Depression and promised that the Federal Reserve would not make the same mistakes again. Then in 2008, when conditions required action, the Federal Reserve did as Friedman and Schwartz recommended, flooding the banking system with liquidity rather than letting monetary conditions deteriorate.

Conclusion

In the end, although Milton Friedman may be best remembered for his policy advocacy, it is the legacy of his ideas that we, as economists, revere. And once again, Friedman’s words provide a fitting summary:

“In order to recommend a course of action to achieve an objective, we must first know whether that course of action will in fact promote the objective. Positive scientific knowledge that enables us to predict the consequences of a possible course of action is clearly a prerequisite for the normative judgment whether that course of action is desirable.”

—Milton Friedman’s Nobel Prize Lecture, 1976

#