The Missing Motivation in Macroeconomics†

By George A. Akerlof*

Macroeconomics changed between the early 1960s and the late 1970s. The macroeconomics of the early 1960s was avowedly Keynesian. This was manifested in the textbooks of the time, which showed a remarkable unity from the introductory through the graduate levels.1 John Maynard Keynes appeared, posthumously, on the cover of Time.2 Even Milton Friedman was famously—although perhaps misleadingly—quoted: “We are all Keynesians now.”3 A little more than a decade later Robert Lucas and Thomas Sargent (1979) had published “After Keynesian Macroeconomics.” The love-fest was over.

The decline of the old-style Keynesian economics was due in part to the simultaneous rise in inflation and unemployment in the late 1960s and early 1970s. That occurrence was impossible to reconcile with the simple nonaccelerationist Phillips curves of the time.

But Keynesian economics also declined because of a change in economic methodology. The Keynesians had emphasized the dependence of consumption on disposable income and, similarly, of investment on current profits and current cash flow.4 They posited a Phillips curve, where nominal—rather than real—wage inflation depended upon the unemployment rate, which was used as an indication of the looseness of the labor market. They based these functions on their own introspection regarding how the various actors in the economy would behave. They also brought some discipline into their judgments by estimating statistical relations.5

But a new school of thought, based on clas-

1 See, for example, Paul A. Samuelson (1964), Thomas F. Dernburg and Duncan M. McDougall (1967), and Gardner Ackley (1961). The econometric model of Lawrence R. Klein and Arthur S. Goldberger (1955) provides a useful synopsis of the variables that the early Keynesians thought most important for a macroeconomic model, and how they would be included.

† Presidential Address delivered at the one hundred eighteenth meeting of the American Economic Association, January 6, 2007, Chicago, IL.

* Department of Economics, University of California at Berkeley, 549 Evans Hall, Berkeley, CA 94720 (e-mail: akerlof@econ.berkeley.edu). This paper is based on a long-term research program with Rachel Kranton on the implications of identity for economic behavior. Our previous joint papers (Akerlof and Kranton 2000, 2002, 2005) have explored implications outside of macroeconomics of utility functions dependent on people’s notions of what ought to be. Some of this paper—especially Section III (“The Missing Motivation: Norms”) and Section IX (“Economic Methodology”)—has been directly taken from our joint manuscript: The Missing Motivation: Economics Made Human (Akerlof and Kranton 2006). I am especially grateful to Professor Kranton for extending to me the invitation to join this project, after she had the initial insight in the spring of 1996 that concerns regarding identity were missing from economic theory. I have also benefited from conversations with Robert Shiller, with whom I am coauthoring work on behavioral macroeconomics. In addition, I especially wish to thank Robert Akerlof and Janet Yellen for invaluable advice. I also want to thank Roland Bénabou, Alan Blinder, Louis Christofides, Stephen Cosslett, Ernst Fehr, David Hirshleifer, Houston McCulloch, John Morgan, George Perry, Antonio Rangel, Paola Sapienza, Robert Solow, Dennis Snower, and Luigi Zingales, and seminar participants at the IMF, the World Bank, Ohio State University, Vanderbilt University, the University of California at Berkeley, the Munich Behavioral Economics Summer Camp, the 2006 Macroeconomics and Individual Decision Making Conference of the NBER and the Federal Reserve Bank of Boston, and at the Social Interactions, Identity, and Well-Being, and Institutions, Organizations, and Growth groups of the CIAR. I am also grateful to Marina Halac for invaluable research assistance and to the Canadian Institute for Advanced Research and to the National Science Foundation under Research Grant SES 04-17871 for invaluable financial support.

2 Time, December 31, 1965. His appearance on the cover was especially remarkable because Time covers are rarely posthumous. Keynes had died in 1946.

3 But in a later disclaimer, Friedman said, almost surely correctly, that he had been quoted out of context. See http://www.libertyhaven.com/thinkers/miltonfriedman/miltonexkeynesian.html, which quotes Friedman (1968), Dollars and Sense, 15.

4 The treatment of consumption in The General Theory, as we shall see below, was typical of such thinking. Keynes first discusses the dependence of consumption on current income, which he clearly sees as the primary determinant of current consumption; but, in addition, he makes a long list of other factors that will alter the relation between consumption and current income.

5 A good example of this methodology can be seen in Alban W. Phillips’s (1958) mixture of light theory and statistical analysis in his estimation of the relation between wage inflation and unemployment.
classical economics, objected to the casual ways of these folks. New Classical critics of Keynesian economics insisted instead that these relations be derived from fundamentals. They said that macroeconomic relationships should be derived from profit-maximizing by firms and from utility-maximizing by consumers with economic arguments in their utility functions.

The new methodology had a profound effect on macroeconomics. Five separate neutrality results overturned aspects of macroeconomics that Keynesians had previously considered incontestable. These five neutralities are: the independence of consumption and current income (the life-cycle permanent income hypothesis); the irrelevance of current profits to investment spending (the Modigliani-Miller theorem); the long-run independence of inflation and unemployment (natural rate theory); the inability of monetary policy to stabilize output (the rational expectations hypothesis); and the irrelevance of taxes and budget deficits to consumption (Ricardian equivalence). These results fly in the face of Keynesian economics. They undermine its conclusions about the behavior of the economy and the impact of stabilization policy.

The discovery of these five neutrality propositions surprised macroeconomists. They had not suspected that radically anti-Keynesian conclusions were the logical outcome of such seemingly innocuous maximizing assumptions.

I. Neutralities and Preferences

How did macroeconomists react to the discovery of the five neutralities? On the one hand, the New Classical economists viewed their neutrality results as a telltale: that Keynesian economists of the previous generation had been thinking in the wrong way. In their view, scientific reasoning was producing a new, leaner, more precise economics.

On the other hand, Keynesian economists, for the most part, reacted differently. In due course they came to view the neutralities as logically impeccable. These New Keynesians accepted the methodological dictums of the New Classical economics: that constrained maximization of profit and utility functions is the appropriate microfoundation for macroeconomics. They also viewed the neutralities as having a certain sort of generality. The neutralities do commonly describe equilibria of competitive economies with complete information, irrespective of people’s preferences—as long as those preferences correspond to economists’ typical descriptions of them. The Keynesians then resurrected some—but not all—of the Keynesian conclusions by adding a variety of frictions to the New Classical model. Those frictions include credit constraints, market imperfections, information failures, tax distortions, staggered contracts, uncertainty, menu costs, and bounded rationality. This formulation preserves many (but not all) Keynesian conclusions regarding cyclical fluctuations and macroeconomic policy.

This lecture will suggest a new stance in regard to each of the five neutralities. Like New Classical and New Keynesian economics, it will derive behavior from utility and profit maximization. That captures the purposefulness of economic decisions. But this lecture will also question the generality of the preferences that lead to the five neutralities. There is a sense in which those preferences are very narrowly defined. They have important missing motivation—since they fail to incorporate the norms of the decision makers. Those norms reflect how the respective decision makers think they and others should or should not behave, even in the absence of frictions. Preferences reflecting such norms yield a macroeconomics with important remnants of the early Keynesian thinking. They also yield a macroeconomics that, in important details, cannot be obtained only with frictions.

We shall see that, with such preferences, even in the absence of frictions, each of the five neutralities will be systematically violated. Specifically:

- A realistic norm regarding consumption behavior will make consumption directly dependent on current income, in violation of the neutrality of consumption given wealth;
- A realistic norm will make investment directly dependent on cash flow, in violation of Modigliani-Miller;
- A realistic norm will make wages and prices dependent on nominal considerations and thus violate natural rate theory;

6 Of course, it took some time for the implications of these neutrality results to be fully appreciated. For example, life-cycle consumption and Modigliani-Miller were initially considered as nothing more than useful codicils to Keynesian thinking.
A realistic norm will make income and employment dependent on systematic monetary policy, and thus violate rational expectations theory; and

A realistic norm will make current consumption dependent on the current generation’s social security receipts, in violation of Ricardian equivalence.

Additionally, insofar as the behavior assumed by the early Keynesians differed from the behavior that produces the neutralities, there is likely to be a bias in favor of the Keynesians. The Keynesians based their models on their observation of motivations, rather than on abstract derivations. If there is a difference between real behavior and behavior derived from abstract preferences, New Classical economics has no way to pick up those differences. In contrast, models with norms based on observation will systematically incorporate such behavior—although, of course, as with any method, there is the possibility for error.

Inclusion of the “missing motivations in macroeconomics” then combines the observations of the Keynesians with the intentionality of economic decisions in New Classical economics. Such a synthesis yields the best of the two approaches.

Two Disclaimers.—Before beginning in earnest, let me offer two brief disclaimers. First, none of the behavior revealing of the norms that are introduced in this lecture will be new. On the contrary, I have purposefully chosen phenomena that have been emphasized since The General Theory by macroeconomists who have followed Keynes in voicing their continuing doubts about classical interpretations of macroeconomic behavior.

Second, this lecture will discuss different norms that respectively correspond to the five neutralities. I shall assume that these norms are exogenous. Such assumptions of exogeneity are standard in economic analysis. In a given problem in a given time frame, some terms are assumed constant, while others are allowed to vary. I ask you to withhold your doubts regarding whether such exogeneity is a correct assumption or not. The incorporation of such endogeneity is the next step—not the first step—in the study of the effect of norms on macroeconomics, especially since such endogeneity may sometimes dampen, but will rarely nullify, the conclusions of this lecture.

II. The Five Neutrality Results

For clarity, this section will now give an overview of each of the five neutrality results.

A. Dependence of Consumption on Wealth, Not Income

Standard theory tells us that, under only somewhat special conditions, consumption depends on wealth, which is the value of current assets plus the discounted value of future earnings. Thus there is no tendency for people to make their expenditures conform to the pattern of their income receipts (as long as their wealth is given).

Changes in the pattern of current income that leave overall wealth constant are neutral in their effects on current consumption.

B. The Modigliani-Miller Theorem

One version of the Modigliani-Miller Theorem says that a firm’s investment strategy is totally independent of its liquidity position. Thus, for example, a corporation with an unexpected windfall will not spend any additional investment dollars. Instead, it will pass the windfall on to shareholders or seek other financial investments, since it will make only those investments whose risk-adjusted rate of return exceeds the rate of return on capital.

Changes in the firm’s finances will thus be neutral in their effect on current investment.

C. Natural Rate Theory

According to Natural Rate Theory, there is some single rate of unemployment that is the only level that could be permanently maintained without ever-increasing inflation or ever-increasing deflation. A fiscal/monetary policy mix that sought to maintain employment that was any higher would result in permanently increasing inflation. A fiscal/monetary mix that

---

7 See Friedman (1957) and Franco Modigliani and Richard Brumberg (1954).
8 See Modigliani and Merton H. Miller (1958).
9 See Edmund S. Phelps (1968) and Friedman (1968).
sought to maintain employment that was any lower would result in permanently decreasing inflation. Fiscal/monetary mixes that yield different levels of long-term (steady) inflation will thus be neutral in their effects on long-term unemployment.

D. Rational Expectations

According to Rational Expectations Theory, a systematic response of monetary policy to the business cycle will have no effect on the stability of the macroeconomy. Wage and price setters will foresee the systematic component of the money supply; they will raise or lower prices and wages exactly proportionally, and thereby neutralize its effect on demand. The stability of the economy is thus neutral with respect to the systematic reaction of monetary policy to the business cycle.

E. Ricardian Equivalence

According to Ricardian Equivalence, under somewhat special conditions, a representative consumer who receives a lump-sum intergenerational transfer (for example, in the form of a social security payment) will not spend a single dime extra. Instead, she will pass on the whole extra income, dollar-for-dollar, to her heirs, who will have to pay the higher tax bills necessary to retire the increased debt incurred in funding the transfer to the previous generation. The transfer is neutral in its effect on current consumption.

III. The Missing Motivation: Norms

Each of the neutralities is based on the assumption that the respective decision makers are utility maximizers. But in each case the utility functions of the decision makers have been very narrowly described. They depend only on real outcomes. For example, in the consumption-neutrality models, utility depends on consumption and leisure; in Modigliani-Miller, it depends only on the discounted real return to shareholders.

But as early as the beginning of the twentieth century, Vilfredo Pareto pointed out that such characterizations of utility missed important aspects of motivation. According to Pareto, people typically have opinions as to how they should, or how they should not, behave. They also have views regarding how others should, or should not, behave. Such views are called norms, and they may be individual as well as social. The role of norms can be easily represented in people’s preferences by modifying the utility function to include losses in utility insofar as they, or others, fail to live up to their standards.

Sociology has a further concept that gives an easy and natural way to add those norms to the utility function. Sociologists say that people have an ideal for how they should or should not behave. Furthermore, that ideal is often conceptualized in terms of the behavior of someone they know, or some exemplar whom they do not know. The standard utility function is then modified by adding a loss in utility, dependent on the distance of behavior from that ideal. Religion and religious identity give us a good example of such norms. Consider the Gospels. What do they describe? The life of Christ. How should a Christian behave? “His life and conversation ought to be worthy of the Gospel of Christ [emphasis added].” How is a good Christian supposed to feel when she has not lived up to her conception of that ideal? Ashamed.

10 See Lucas (1972), Thomas J. Sargent (1973), and Lucas and Sargent (1979).
11 See Robert J. Barro (1974) for the modern reincarnation of these ideas, first discovered by Ricardo.
12 This section, including much of its exact wording, has been taken from a joint manuscript with Rachel Kranton (Akerlof and Kranton 2006). I should emphasize that these insights have been developed jointly. The initial instigation of our project is wholly due to Kranton. It is impossible for me to say which ideas or wordings are mine and which are hers.
13 See Pareto (1920), George C. Homans and Charles P. Curtis (1934) give an excellent summary of Pareto that is fully consistent with the emphasis here. Jon Elster (1989) also presents a similar conception of norms.
14 For example, the protagonist of the novel Rice Mother (Rani Manicka 2002) did not believe she should wear red with black.
15 See http://www.orthodoxytoday.org/articles/StBasil-Behavior.php.
16 Of course, there are many interpretations of the Gospel, and some of them are even contradictory. But that does not affect whether the person should be ashamed or not. She
A. Importance of Norms in Motivation: Some Examples

But religion is only one of the many realms where people have such an ideal. To appreciate the ubiquity of norms in motivation, it is useful to see some further examples. Those examples will demonstrate that people tend to be happy when they live up to how they think they should be; and they are, correspondingly, unhappy when they fail to live up to those norms.

For the audience for this lecture, most of whom are professors, teaching provides an especially familiar example. We have a view of what it means to be a good teacher. On our lucky days, when we live up to our standards and our classes go well, we tend to be happy; on our off days, when something goes awry in class, we may even feel quite miserable.

Such motivation in the workplace is the rule, rather than the exception. Most workers, like teachers, care about the conduct of their jobs. Randy Hodson (2001), who surveyed ethnographies of the US workplace, found that most employees care about their dignity at work. They want to conceive of what they do as useful. And they feel a lack of dignity if they are thwarted, either by their own actions or by the actions of others. Those who are unable to get such satisfaction are likely to show their displeasure by acting up in some way or other.

Such motivation in the workplace is the rule, rather than the exception. Most workers, like teachers, care about the conduct of their jobs. Randy Hodson (2001), who surveyed ethnographies of the US workplace, found that most employees care about their dignity at work. They want to conceive of what they do as useful. And they feel a lack of dignity if they are thwarted, either by their own actions or by the actions of others. Those who are unable to get such satisfaction are likely to show their displeasure by acting up in some way or other.

Studs Terkel’s *Working* (1972) captures in a single volume much of the ethnographic findings summarized by Hodson. Terkel interviews people from many different occupations about their feelings about their jobs and concludes that people “search for daily meaning as well as daily bread” (1972, xi). Some of the interviewees are successful in this search: like the stone mason, who cruises his Indiana county and basks in pride as he not infrequently passes his past work. At the opposite extreme is an Illinois steelworker, whose work denies him the dignity he seeks. He takes out his frustration at work by being disrespectful, and, after hours, by getting into tavern brawls. Most workers are somewhere between these extremes, but in all cases, following Terkel, they have a feeling for how they should behave at work. It is not just about the money; it is also about living up to an ideal about who they think they should be.

Such belief regarding how people should behave, and their behavior in accordance with such belief, goes beyond the workplace. It affects disparate areas, from playing golf to life in the family. Betty Friedan’s *Feminine Mystique* gives what may be as good a description of norms and their impact on people’s lives as can be found anywhere—in this case regarding the norms for middle-class women of the previous generation. Here is a brief sample of her description:

“Millions of women lived their lives in the image of those pretty pictures of the American suburban housewife, kissing their husbands goodbye in front of the picture window, depositing their station-wagonsful of children at school, and smiling as they ran the new electric waxer over the spotless kitchen floor.... Their only dream was to be perfect wives and mothers; their highest ambition was to have five children and a beautiful house, their only fight to get and keep their husbands.... They gloried in their role as women, and wrote proudly on the census blank: ‘Occupation, housewife’ (Friedan 1963, 18).

Most women lived up to these norms. Some of these were dissenters, like Friedan herself, who disagreed with them, but felt compelled, nevertheless, to follow a norm with which they disagreed. Friedan says they suffered from “the problem without a name.” In our terms, they were losing utility because they were failing to live up to what one part of them thought they should do.

We may appeal to religious texts, to work ethnographies, and, like Friedan, to women’s magazines to see the role of norms. But is there yet harder data, some form of natural experiment, that indicates the importance of norms? The sociologist Erving Goffman has found such an example. He observed the behavior of children of different ages when they were brought to the local merry-go-round. Because appropriate activity differs by age, the children should have predictably different reactions. For the toddlers, riding a wooden horse is an accomplishment. They show their joy at fulfilling what they should do with smiles and waves as they pass by. In contrast, for older children, there is
a gap between their conception of how they should behave and riding the merry-go-round. However much they may enjoy it, they also feel the need to distance themselves from an activity that is so age inappropriate. They manifest this distance by riding a frog, rather than a “serious” animal like a horse; alternatively they show off by standing up “dangerously” during the ride. In some way or other they play the clown.

Behavior at the merry-go-round is, of course, just the stuff of kids. But Goffman supplements it with a totally serious example. In surgical operations, because of their inexperience, medical students are given tasks that are ridiculously easy.\(^17\) They respond in the same way as the older children at the merry-go-round: they also act the clown.\(^18\)

In economics, as elsewhere, $500 bills do not just lie on the street. If living up to norms is such an important motivation, it must show up in many economic examples, even if it is not identified in exactly our language. Gary S. Becker’s *Economics of Discrimination* (1957) offers an example of now-standard economics that can also be interpreted in terms of such norms. Becker’s theoretical innovation was to modify plain-vanilla economic utility by the introduction of a *discrimination coefficient*. He defined that as the loss in utility incurred by exchange with someone from a different race—for example, the loss of a white from an exchange with a black. The natural interpretation is that the discrimination coefficient represents the loss in utility for the white from physically engaging in an exchange with a black. But this representation of the utility function can also be interpreted in terms of norms. There is a code as to how blacks and whites should behave toward each other. The white has a view that she *should not* deal with a black. She loses utility equal to the value of the discrimination coefficient—not from the physical association—but ipso facto from the violation of the code. There is reason to believe that such norm-based interpretation better reflects the nature of discrimination than a physical exchange–based theory. In the pre–Civil Rights period, when Becker was writing, there can be no doubt that discrimination, and the code that upheld it, was stronger in the South than in the North. Yet exchanges between blacks and whites were surely much more common in the South than in the North. At least one statistic reflects such a difference: there were significantly lower levels of residential segregation by race in the South than in the North.\(^19\)

### B. Summary

Our examples are illustrative of behavior that is pervasive. Sociology is dense in examples of people’s views as to how they and others should behave, their joy when they live up to those standards, and their discomfort and reactions when they fail to do so.

We now turn to examining the role of norms in each of the five macroeconomic neutrals.\(^20\) In each case we shall ask whether

---

\(^{17}\) Goffman (1961) observed the behavior of such students in medical operations.

\(^{18}\) Another example, the Milgram experiment (Stanley Milgram 1963, 1965) demonstrates the strength of such motivation—by showing the lengths that people will take to do what they think they *should* be doing. To see this interpretation of this experiment, which is only one of many ways of viewing it, it is useful to give a brief description. On arrival, subjects were told that they were involved in a learning experiment. They were put in the role of the “teacher,” who *should* administer shocks to a “learner” whenever he gave a wrong answer. The subjects are led to identify with their role as teacher in this experiment, and feel that they *should* obey the experimenter. Rather than being another subject, and, rather than being wired, as it appeared, actually the learner was an unwired, trained confederate of the experimenter. Subjects were then instructed to administer shocks of escalating voltage as the learner made errors. A surprising fraction of subjects escalated their shocks to the maximum 450 volts—even though such a dosage in real life would have been lethal. There are many different versions of the experiment, but the version where the confederate grunts and moans at 75 volts, asks to be let out of the experiment at 150 volts, and refuses to give any more answers at 300 volts, is typical. Here more than 60 percent of subjects went all the way. Nor is such motivation limited to the laboratory. The rampage of the Nazi Reserve Police Battalion #101 in Poland during World War II (Christopher R. Browning 1999) gives a real-world mirror of the behavior Milgram obtained in the laboratory. Like Milgram’s subjects, the members of this unit, were just *Ordinary Men* (Browning’s title). They were recruited from the most prosaic civilian occupations.

\(^{19}\) See Douglas S. Massey and Nancy A. Denton (1993, table 3.1, 64).

\(^{20}\) Some years ago, at a conference in Spoleto, Italy, Edmund Phelps gave a still-unpublished lecture wondering why the economics of the twentieth century had failed to discover what was central to most of the arts, which was the role of subjectivity. This paper is about the direct relevance of such subjectivity for macroeconomics. I have very much benefitted from enjoyable conversations with Professor
people’s views as to how they should behave will enter their utility function. In each case, we shall see that such views will nullify the respective neutrality result. Indeed, we shall also see that in each case there will be a natural norm broadly consistent with Keynesians’ views of economic behavior.\footnote{For each of the five neutralities we see that the inclusion of broader preferences, inclusive of norms, will bring Keynesian behaviors back to life. But, of course, that does not mean that the competitive forces and the maximizing behaviors responsible for the five neutralities are not important as well.}

\section*{IV. Ricardian Equivalence}

We shall begin our detailed discussion with Ricardian equivalence. It was chronologically the last of the neutralities to be appreciated by modern economists. But it is also the simplest. That makes it the best place to begin.\footnote{That appreciation is of course due to Barro (1974).} If there is missing motivation in the utility function, it should be easiest to see here.

A very simple model demonstrates the essence of Ricardian equivalence, as it was rediscovered by Robert Barro after a lapse of almost two centuries.\footnote{This model is quite close to Ricardo’s original discussion. It is a considerable simplification of Barro’s model. His model had a sequence of overlapping generations, each of which lived for two periods. Barro’s contribution was not only to show Ricardian equivalence in the two-generation model, but also its extension to a sequence of generations when parents’ utility depended only on their own utility and the utility of their own children. Ricardo’s discussion, which is close to the two-generation model here, was then subsequently rediscovered. There is no uncertainty, and all taxes are lump-sum. This proposition may be generalized, for example, following Barro, to a model with $m$ overlapping generations, each of which has different consumption when young and old. Each parent derives utility from his own consumption and the utility of his child.}

In the model, there are just two periods, periods 1 and 2. There are just two people, a parent and her child. The utility of the parent depends directly upon her own consumption, in period 1; it also depends upon the utility of her child. That utility depends upon his consumption, in period 2.

The parent’s utility function can be expressed simply as $U_1(c_1, U_2(c_2))$, where $c_1$ is the consumption of the parent, $c_2$ is the consumption of the child, $U_1$ is the utility of the parent, and $U_2$ is the utility of the child. The parent chooses her consumption in period 1 to maximize her utility. Whatever wealth remains, she bequeaths to her child.

Ricardian equivalence takes the following form in this model. Suppose that the government gives a transfer, which we shall call a social security payment, to the parent in period 1; but then in period 2 it taxes the child to retire the debt caused by this transfer.\footnote{The tax and the transfer are both lump-sum.} In this case, the consumption of a parent who maximizes the utility function $U_1$ and who leaves a bequest to her child will be unaffected by her receipt of social security.

The logic of this result is simple. With and without social security the discounted value of consumption of the parent and of the child is constrained by the discounted value of the family’s earnings (plus its initial wealth). Social security leaves that constraint unchanged. If the parent found $(c_1, c_2)$ to be the optimal division of consumption between herself and her child in the absence of a social security payment, this same division of consumption between herself and her child will optimize her utility with a social security payment.

A vast literature explains why such Ricardian equivalence is unlikely to be empirically descriptive.\footnote{The conventional wisdom is, of course, that social security will affect aggregate savings. Martin Feldstein (1974) and Feldstein and Anthony Pellechio (1979) act as if increases in social security of the current generation will result in increased consumption, so that the next generation will have a lower capital stock.} The long list of reasons includes (a) infinite, rather than finite, horizons; (b) strategic bequests to obtain the attention of one’s heirs while alive; (c) childless families; (d) uncertainty, including bequests made because of uncertainty about the age of death; (e) differential borrowing rates between the government and the public; (f) growth of the economy in excess of the interest rate, allowing steady debt issuance; (g) lack of foresight regarding the effect of social security on future taxes; (h) foreign ownership of debt; (i) tax distortions;\footnote{I take this list mainly from the review article by John J. Seater (1993).} (j)
constraints on the consumption of parents (so they do not leave bequests); (k) myopia of the parents regarding children’s future tax payments.29

The preceding list gives empirical reasons for failure of Ricardian equivalence; but, lengthy as it is, it still ignores its theoretical challenge. According to that challenge, under economists’ standard assumptions, with perfect certainty and with perfect foresight, Ricardian equivalence will occur. Such a result had previously been unsuspected by economists.30

Two possible conclusions can be drawn from this surprise. On the one hand, we might continue to assume that classical assumptions describe economic behavior. The five neutralities that are the subject of this paper concern the realignment to macroeconomics that occurred as economists gained understanding of the consequences of classical assumptions from the mid-1950s to the mid-1970s.

Economists may have been correct in drawing the conclusion that the early Keynesian economics was too simplistic and naive. But they could have drawn another conclusion from this surprise. In this view, Ricardian equivalence is a telltale: we do not believe, even in the presence of perfect foresight and perfect certainty, that the parent will make an equal and opposite offset of her social security transfer in terms of an increased bequest to her child. Something must be missing from the motivation in Barro’s model; otherwise, it would not have given rise to results that are so surprising.

B. Douglas Bernheim and Kyle Bagwell (1988) give further evidence suggesting that Ricardian equivalence is such a telltale. They show how the same logic would apply to a network of gift-givers. Remarkably, any member of such a network will be indifferent whether she receives an extra dollar or any other participant in the network is the recipient. Such conclusions, suspect as they are, suggest a problem with the model beyond the lack of realism involved in perfect foresight and perfect certainty. They also suggest missing motivation.

James Andreoni (1989) has put his finger on what that missing motivation might be.31 A bequest is a type of gift. The parent will receive utility from giving such a gift. Ricardian equivalence will fail if the parent has utility from gift-giving. With a social security transfer, more money is hers, and the same consumption allocation to herself entails a greater gift to her child. With declining marginal utility for bequest-giving, she will then divide an increased social security transfer between additional consumption for herself and an additional bequest to her child.32

Andreoni thus describes the utility missing from the standard utility function as that arising from the “warm glow” from giving. Such a characterization may be accurate. It also sounds as if it is very close to classical assumptions—that there is nothing fundamentally different about this additional motivation. But this segment of the utility function is, in fact, very different from economists’ usual characterization of motivation. We know that the “warm glow” does not come from the utility the parent

---

29 This was Ricardo’s own reason for dismissal of the argument. He said that the parent would alter her bequest because she would not take into account the added tax payments of the child (see Gerald P. O’Driscoll Jr. 1977). Uncertainty regarding the size of the future tax payments is different from such myopia, in which the payment is altogether ignored. But, with quadratic utility and expected utility maximization, uncertainty regarding the child’s future tax payments will have no effect on the size of the parent’s bequest.

30 For example, Feldstein (1974) and Feldstein and Pellechio (1979) engage in no theoretical soul-searching regarding the negative effects of social security on current savings. There is a voluminous literature (see Roberto Ricciuti 2003) examining the empirical validity of Ricardian equivalence. Largely because of the problem of endogeneity, it is difficult to come to firm conclusions regarding its empirical validity. There are studies with findings both for and against such crowding out.

31 See also John Laitner (2002), Laitner and Henry Ohlsson (2001), Alan S. Blinder (1975) and Michael D. Hurd (1989), who have also modeled the bequest motive as coming from the utility of the parent from giving the bequest.

32 Formally, she trades off the marginal utility of her own consumption against the marginal utility from gift-giving and the marginal utility she gets from her child’s consumption. In making this trade-off, she takes due account of the fact that one unit of consumption today is traded off against \((1 + r)\) units of consumption next period.
derives from her own consumption; nor, yet more tellingly, does it derive from the utility of her child (as the child’s utility depends on its own consumption). It enters the utility function as a separate term.

What, then, could account for a “warm glow”? Parent-to-child bequests are a form of gift. If there is any type of economic transaction that is governed by norms, it is the giving of gifts. Parent-to-child bequests also occur within families. Therefore, they should also be affected by the norms of family life. We have already seen one example of such norms (Friedan’s portrait of the proper place of women in the early 1960s).

The norms of family life are not constant. They vary by culture. They also change over time. As the nature of the ideal family has shifted, so has the ideal bequest. Actual bequests have changed in tandem. For example, the ideal sixteenth century Anglo-Saxon family was dynastic. The lineage passed from father to oldest son. Fathers then left the bulk of their estates to their oldest sons. In the twenty-first century, in the ideal family, siblings are equal. Most bequests are now evenly divided between them.

Summary.—Economic outcomes, such as the consumption of the parent and the utility of the child, are one determinant of bequests. But another possible determinant is parents’ views regarding how they should behave toward their children. Just as Friedan’s suburban housewives waxed their floors, because they thought that is what housewives should do, parents who leave bequests derive a warm glow from bequests because that is what they think they should do for their children. Ricardian equivalence then illustrates how odd neutralities can occur in models that fail to take such norms into account.

A comment by David Romer (2001, 539) tells us where we should venture next. He has remarked that “quantitatively important” violations of Ricardian equivalence and of the permanent income/life-cycle hypothesis occur for the same reasons. Ricardian equivalence is not important for us as an empirical aspect of macroeconomics. There are so many reasons other than the role of norms for its violation. But it does give us an initial window on the type of motivation missing in classical macroeconomics. Inclusion of such motivation will give us a new perspective on the consumption function. It allows us to return to a view in which consumption will depend on current income, just as its inclusion makes it natural to believe that social security transfers will affect savings and consumption, even in a world without frictions.

V. Consumption and Current Income

This takes us to the second neutrality. According to this result, other than its contribution to a consumer’s wealth, current income has no independent effect on the consumption of a utility-maximizing consumer.

Milton Friedman (1957) derived such consumption-income neutrality in the two-period model of Irving Fisher. In this model, the consumer chooses her consumption between two periods. She maximizes her intertemporal utility function, given by the function \( U(c_1, c_2) \): \( c_1 \) denotes her current consumption in the first period; \( c_2 \) denotes consumption in the second period. If she maximizes \( U(c_1, c_2) \), a dollar of income earned today will have the same effect on her current consumption as a discounted dollar earned in the next period. Thus, her consumption will depend only on the discounted value of her current and future income and the rate of interest. This proposition is easy to prove. It generalizes to many different commodities and to many different time periods,

\[ Y_1 \text{ in period 1, income } Y_2 \text{ in period 2, and she can borrow and lend at the rate of interest } r. \]
and, with quadratic utility, to uncertain incomes. In standard terminology, the value of her discounted income is called her wealth; the amount of that wealth that can be spent without its depletion is called permanent income. An alternative expression of Friedman’s hypothesis is that consumption depends on permanent rather than on current income.

The permanent income hypothesis may be in accordance with most standard economic models. Nevertheless, it contradicted prior thinking about the consumption function. Keynes, and his followers, believed that current income played an especially important role in the determination of current consumption.

“The fundamental psychological law [emphasis added], upon which we are entitled to depend with great confidence both a priori from our knowledge of human nature and from the detailed facts of experience, is that men are disposed, as a rule and on the average, to increase their consumption as income increases, but not by as much as the increase in income” (Keynes, The General Theory, 1936, 96).

It is true that The General Theory discussed a long list of other factors that could affect consumption. The list was sufficiently rich to include not only current income, but also all the other determinants of wealth, such as expected future income and the rate of interest. But that does not make Keynes’s theory identical to Friedman’s. In the Keynesian theory, consumers are more sensitive to current income than to other changes in income that have similar effect on the consumer’s wealth.

A. Empirical Results and Their Explanation

A large number of tests have demonstrated the excess sensitivity of consumption to current income, in concert with the Keynesian consumption function. For example, John Y. Campbell and N. Gregory Mankiw (1989) nested both Friedman’s view that consumption depends solely on wealth and the simplified Keynesian view that consumption depends solely on income. They suppose that a fraction of consumers \( \lambda \) are pure Keynesians, while a fraction \( (1 - \lambda) \) behave according to the permanent income hypothesis; they estimate \( \lambda \) from the extent to which consumption overreacts to changes in income that would be predictable from past changes in income and consumption. Usefully, then, \( \lambda \) gives a natural measure of the departure from the permanent income hypothesis. The estimates of \( \lambda \) are significant statistically and also of significant magnitude economically: between 40 and 50 percent (depending upon whether three or five periods are used to predict the change in current income).

Other studies corroborate such excess dependence on current income: John Shea (1985), for union members whose contracts specified their
future wages; David W. Wilcox (1989), for social security recipients who had been earlier notified of changes in cost-of-living adjustments; Jonathan Parker (1999), for payers of social security taxes with predictable inter-year changes; Nicholes S. Souleles (1999), for changes in disposable income net of tax refunds; and James Banks, Richard Blundell, and Sarah Tanner (1998), and Bernheim, Jonathan Skinner, and Steven Weinberg (2001), for retirees.

Textbooks explain such excess sensitivity by a variety of frictions, particularly borrowing constraints. For example, Rudiger Dornbusch and Stanley Fischer (1987) say: "Given that the permanent income hypothesis is correct [sic], there are two possible explanations." They are liquidity constraints for consumers and myopia in their projections of future income. Thus, we see the realignment that occurred because of the life-cycle permanent income hypothesis: excess sensitivity may occur, but only in the presence of credit constraints or myopia. Such a view cannot have been adopted because of its empirical support. Few studies have tested this proposition, but those that do have rejected it. For example, credit constraints cannot explain the reduction in consumption of retirees. And, neither myopia nor credit constraint can explain the reduction in union members’ consumption at the time of wage declines scheduled in their union contracts (Shea 1995, 996).

The adoption of the permanent income/life-cycle hypothesis then must rest on theoretical, not empirical, reasons. But the theory fails to take into account norms regarding what people think they should or should not consume. Such a norm-based theory will nest Keynes’s psychological law. Consumption-income neutrality will occur only in a singular special case.

B. Consumption and the Role of Norms

Why should consumption be overly sensitive to income? This section presents an argument in three steps. First, sociology gives motivations for consumption that are very different from the reasons for it in the life-cycle model. A major determinant of consumption is what people think they should consume. Second, what people think they should consume can often be viewed either as entitlements or as obligations. Finally, in turn, current income is one of the major determinants of these entitlements, and obligations.

Sociology of Consumption.—The motivation emphasized by sociologists for consumption is very different from that in the life-cycle model. Sociologists describe consumption as largely determined by the norms regarding what people should consume. These norms, in turn, are dependent upon the individual’s situation and also who she thinks she is.

Two examples illustrate such dependence on norms. Following Pierre Bourdieu (1984), people’s consumption of cultural goods—the literature they read, the music they hear, and the art they buy—reflects not just their individual tastes. The upper class should not make lower-class choices. Correspondingly, the lower class should avoid appearing above their station. The epithet "lace curtain Irish" illustrates. To the users of this phrase, those lace curtains were indicative of those violating their social place. Weber’s analysis of the relation between religion and savings further reflects the role of people’s views regarding who they should be. In The Protestant Ethic and the Spirit of Capitalism, Weber describes Calvinists as aspiring to be “worldly ascetics.” He concludes that “economic acquisition is no longer subordinated to man as the means for satisfaction of his material needs.” Here the purpose of saving is to live up to an ideal. The Calvinists are thrifty because they think they should not be consuming. That turns the motivation of the life-cycle

40 See Dornbusch and Fischer (1987, 284).
41 I am extremely grateful to Robert Akerlof for help in formulating the argument of this section.
42 Bourdieu views this as important because of the role of such differential consumption in the transmission of class structure from one generation to the next. The focus on consumption as a reflection of who people want to be can be seen throughout the sociology of consumption. On the lowbrow-highbrow scale, a study by Ian Woodward (2003) is at the opposite end of the spectrum from Bourdieu: Woodward asked Australian housewives about the reasons for their choice of furniture. Some went for comfort; others, for aesthetics. But they also indicated, with a surprising degree of moral fervor, that their choices reflected who they wanted to be.
44 Weber (1958, 53).
model on its head. There people save only because of their desire for consumption in retirement.

Luigi Guiso, Paola Sapienza, and Luigi Zingales (2003, 2006) have statistically affirmed Weber’s hypothesis that religion is correlated both with attitudes toward savings and with actual savings. In addition, they have more generally affirmed the quantitative significance of culture for savings and consumption; in their regressions, variables reflecting culture have as much power as variables derived from the life-cycle hypothesis in explaining cross-country savings ratios. 45

Consumption Entitlements and Obligations.— While sociology is useful in giving us the general insight that consumption depends on cultural norms, we need to be more specific. What is the nature of those norms? They can frequently be described in two ways: as entitlements and, also sometimes, as obligations to spend. Again some examples will illustrate.

First, oddly, people have obligations to spend. Social history is full of the obligation to keep up appearances. Most Wall Street bankers, for example, do not live like mothers on welfare. They do not want to. But, even if they did, it would occasion gossip. It is not what they should do. History is replete with stories of the debt of aristocrats struggling to maintain their social obligations.46 As just one example, the debts to British merchants by Southern planters, who were keeping up with the Joneses of the eighteenth century, are considered a significant factor underlying the Southern support of the American Revolution.47

In addition to obligations to spend, there are also entitlements. The lost-ticket paradox of Amos Tversky and Daniel Kahneman (1981, 457) gives an illustration. Eighty-eight percent of respondents to a questionnaire said they would buy a $10 theater ticket if they arrived at a theater to see a play and found that they had lost a $10 bill. In contrast, only 46 percent said they would buy a new $10 ticket in the same situation if they had lost a previously purchased ticket.

Tversky and Kahneman explain this difference by “mental accounts,” but an explanation in terms of entitlements is equally valid. Tversky and Kahneman say that those who have lost the $10 bill do not connect that loss to the play. In their mental account, its cost is just $10. But those who have lost the ticket see themselves as paying for it twice. In their mental account, its cost is $20. Those with the lost ticket then tend to opt out, because they see $20 as too much to pay to see the play. But the difference in behavior for those who lost the ticket and those who lost the $10 bill could also have been interpreted in terms of entitlements. Most people want to think of themselves as responsible human beings. When they lose the ticket, they do not feel entitled to just buy another one. That is not the type of person they aspire to be.

We should also observe that it is not coincidental that the lost ticket paradox could be explained both by mental accounting and by norms. Formally, any model of mental accounting can be translated into a model of norms: just replace the rules of mental accounting as the norms that people think they should follow.48

But even though norms and mental accounting may be equivalent, interpretations in terms of norms are important for this lecture. Mental accounting has the connotation, whether rightly or wrongly, of being a heuristic for quick decisions. Such a heuristic will, of course, sometimes result in cognitive error. Whether rightly or wrongly, most economists would dismiss cognitive error as unimportant. Why? because in their view people are smart about what they want, and their decisions are also very purposeful. But norms cannot be dismissed so easily. As I argued earlier, people feel strongly about adherence to them. Their

45 Guiso, Sapienza, and Zingales (2006, 39) report regressions of savings ratios on GDP growth, dependency ratios, and responses to the question: “Do you consider it especially important to encourage children to learn thrift and savings?” A one-standard-deviation difference to GDP growth and to attitude toward thrift both produce a 1.8-percentage-point difference in the savings ratio. (A one-standard-deviation difference in the dependency ratio, which could be the result both of cultural differences and of life-cycle considerations, produces a 3.2-percentage-point difference.)

46 See, for example, David Cannadine (1977).

47 See Woody Holton (1999).

48 But it turns out that there is quite possibly a substantive difference between the two interpretations. With the mental accounting interpretation the losers of the ticket could be induced to buy one, if only a wise friend would make them aware of the logical problems of their reasoning. In contrast with the norms interpretation the friend cannot be so helpful. Buying a new ticket is a departure from the person’s norm, and she loses utility by it.
absence from utility constitutes the missing motivation of macroeconomics.

The Link of Entitlements and Obligations to Current Income.—It remains to relate current spending to current income. Norms may be complex. But a web of evidence still reveals a strong association between current income and entitlements and obligations to spend. Such a link, in turn, produces the excess sensitivity of consumption on current income in Keynes’s Psychological Law.

A few examples follow.

- It is common practice in the United States for parents, even for rich ones with no budget constraint, to expect their children to assume financial independence after their graduation from college. They are indicating their belief in the norm that the child is entitled to spend what she earns. (Most parents, of course, give their children a helping hand as they seek their independence. But that does not mean that they do not also strongly believe that their children should live on their earnings, since that norm is only one of their motivations.)

- In a thought experiment, consider a woman living on $50,000 a year who learns that her uncle will die in one year leaving her $2,000,000. Even if she has considerable savings in the bank, it would be unseemly for her to run down her savings in anticipation of the bequest. She is not entitled to do so. She should stick to spending from her current income. This gives another example in which norms regarding entitlements to spend are related to current income, in violation of the life-cycle hypothesis.

- People’s expenditures are supposed to reflect their stations in life, and those stations usually reflect their earnings. Thus, for example, college students with little earnings are supposed to live that way—like college students. Their current spending is supposed to reflect their current earnings, not what they will be earning in the future. (At the other extreme, as an obligation, the college president is often expected to live in the presidential mansion.)

- Preliminary results from an experiment by John Morgan and myself illustrate another relation between entitlement and earnings. In this experiment, subjects were asked to donate to a charity before and after completing a task. Those who were asked for the donation afterward were more likely to keep the money than those who were asked beforehand. Those who had completed the task felt that they had earned the money and were thus entitled to keep it for themselves.49

- The mental accounting model by Hersh M. Shefrin and Richard H. Thaler (1988) is especially useful in our quest for a Keynesian consumption function. Norms take many forms, so their formal model is not unique.50 But it does illustrate a possible link between consumption and current income. In this model, people have three separate mental accounts: current income, current assets, and future income combined with pension wealth. As consumers exhaust one of these accounts and begin to use the next one for their current consumption, they incur a discontinuous “penalty.” Those penalties are psychological in nature—this is a model of mental accounting—and they take the form of a loss in utility.51 Corresponding to Shefrin and Thaler’s assumptions regarding the nature of these costs, as consumption rises, consumers will first finance it wholly from current income; then, from current assets; and, finally, from future income and retirement wealth.

As we discussed earlier, it should be no surprise that there is an exact translation of such a model into one with norms regarding entitlements to consume. The rules of mental accounting become the norms regarding how money should be spent. The basic norm is that consumption should come from current income.

49 These are the results for females. The men gave almost nothing so their differentials are irrelevant. The women gave on average about 10 percent of their earnings. Those who were asked to donate before the task gave twice as much as those who were asked afterward. The task lasted 40 minutes and was to highlight phrases in a manuscript to be used in making an index.

50 Shefrin and Thaler themselves are explicit about the possibility of other models.

51 We should also note that the Shefrin-Thaler model has elements not discussed in the text. In general, the discontinuous penalties from mental accounting are one reason why consumption might be at a corner solution in one of the three mental accounts. Shefrin and Thaler have another reason. They view saving as taking willpower, which entails a cost in terms of lost utility. The less people save the less of this costly willpower they need to expend. This gives another reason why consumption might be on one of the boundaries of the mental accounts. It is useful to remember that at one of the boundaries, consumption will conform to current income.
And the discontinuous penalties correspond to the losses of utility due to respective deviations from that norm. In particular, Shefrin and Thaler assumed that there is no such cost at all if consumption comes only from current income. That means that current income can be considered as consumers’ entitlement to spend, since any consumption that is less than current income entails no deviation at all from the norm regarding the account that should finance it.

- Shefrin and Thaler give an impressive array of econometric facts in support of their model. Insofar as these facts support their mental accounting model, they also equally well support its reinterpretation—with the norm that current income is an entitlement to spend. Those facts include: differential savings out of windfall and current income;\(^52\) a less than one-to-one displacement of discretionary saving by employee pension contributions;\(^53\) undersaving for retirement;\(^54\) and a marginal propensity to consume out of fully anticipated bonuses that is much greater than the marginal propensity to consume out of monthly income.\(^55\)

- Retired people are commonly believed to tailor their consumption to a concept of income rather than to the value of their assets. Shefrin and Statman (1984) have viewed this as another form of mental accounting. They also present considerable evidence regarding such behavior.

C. Summary

Considerable evidence suggests that people’s views regarding what they are entitled to spend play a major role in their consumption choices. It also suggests strongly that current income plays a special role in those entitlements. Shefrin and Thaler have explained such patterns by mental accounting. A reinterpretation of their model shows that they also could have explained this behavior in terms of norms. Once again we see that the current versions of the life-cycle hypothesis have left out missing motivation that easily justifies the excess sensitivity of consumption to income in Keynes’s psychological law.

VI. Investment and Cash Flow

The debate concerning investment has been surprisingly close to the debate about consumption. The early Keynesians emphasized two variables as determinants of investment: current cash flow (with profits as a major component), and the firm’s current holdings of liquid assets. Each of these variables is a measure of funds available to firms for investment without seeking outside finance.\(^56\) In contrast, the later literature denied any special role of liquidity in the investment function.

The first such questioning came from Modigliani and Miller, who assumed that managers maximize shareholder value and that markets are frictionless and competitive. In this case, a firm’s financial position plays no role in the value of the firm. The argument for this independence proceeds as follows. By construction, Modigliani and Miller show how a competitive equilibrium changes if a firm increases its debt and buys back shares. In the new equilibrium, investment will be unchanged, and shareholders will offset the increase in the firm’s debt by a compensating increase in the bonds in their respective private portfolios. The reason the equilibrium changes in this way is straightforward: if the markets for debt cleared in the old equilibrium, they will again clear in the new. If managers’ choice of investment maximized shareholder value in the old equilibrium, the same choice of investment maximizes it in the new. Investment is therefore independent of the firm’s current financial position, including its current liquidity position and its current cash flow.

The advent of \(q\)-theory similarly questioned a special place for current variables, such as cash flow and liquid asset holdings in the investment decision. In the original version of the theory, James Tobin (1969) suggested that a firm’s optimal investment strategy arbitrages between


\(^{54}\) Shefrin and Thaler (1988, 626–27). Especially, they say that there would be vast undersaving in the absence of social security and forced private pensions to prevent it. There is some ambiguity regarding whether there is undersaving in the presence of these institutions to counteract it.


\(^{56}\) See, especially, John R. Meyer and Edwin Kuh (1957).
the value at which it can sell a unit of its capital and its investment costs to produce a new unit of capital. In this case, firms should invest up to the point where the marginal cost of a new unit of capital is the valuation of such a unit of capital in the stock market. That valuation is the market value of the firm’s shares divided by its capital stock, called the \( q \)-ratio. If markets are efficient, \( q \) is also the expected discounted value of current and expected future profits per unit of capital.\(^{57}\) Since \( q \)-theory says that firms should invest in capital up to the point where the cost of an extra unit of capital stock is equal to the present discounted value of the stream of earnings from a unit of capital, again, as in Modigliani-Miller, investment is independent of the firm’s finance decision.\(^{58}\)

The empirical testing of \( q \)-theory also has a striking parallel to the empirical testing of the consumption function. Just as Campbell and Mankiw showed that there was excess sensitivity to current income in the consumption function, Steven M. Fazzari, R. Glenn Hubbard, and Bruce C. Petersen (1988) showed that investment depends not just upon \( q \), but also upon the current cash flows. Furthermore, as in the standard explanation of excess consumption sensitivity, Fazzari, Hubbard, and Petersen similarly suggest that credit constraints are responsible for the dependence of investment on cash flow. They continue with the Modigliani-Miller/\( q \)-theory assumption that managers maximize stockholder value. But they posit that the difference in information between managers and financiers results in a wedge between the cost of internal and external financing. This is clearest for firms that are credit constrained—so that credit-constrained firms will be especially sensitive to available liquidity.\(^{59}\) But, as with credit-constraint explanations of consumption, empirical evidence, such as there is, rejects this hypothesis. Steven N. Kaplan and Zingales (1997) analyzed the subsample of firms that Fazzari, Hubbard, and Petersen had considered most likely to be credit constrained. They find credit constraint to be rare. Furthermore, they also found that those firms with the least constraint had the greatest sensitivity to cash flows.\(^{60}\)

There is, thus, remarkable similarity between the consumption function and the investment function. In both cases, economic theory suggested rejection of earlier views regarding the role of current flow variables—current income in the case of consumption, cash flow in the case of investment. In both cases, empirical investigation showed the existence of excess sensitivity to the current flow variable. In both cases, these rejections support the previous Keynesian theory. In both cases, economists have sought to explain the divergence between practice and theory by the presence of credit constraints. In both cases, the empirical evidence, such as it is, does not support the case that credit-constraint explanations explain the theoretical anomaly.

### A. Theory of Excess Sensitivity of Investment to Cash Flow

Whatever the similarities, consumption and investment differ in one major respect. In the case of investment, economists are already aware of a fundamental reason why investment will depend on current cash flow. Modigliani-Miller and \( q \)-theory both assume that managers...

---

\(^{57}\) See Andrew B. Abel (1979), Lawrence H. Summers (1981), and Fumio Hayashi (1982).

\(^{58}\) This should not be a surprise, because the assumptions of this version of \( q \)-theory are in accord with Modigliani-Miller: competitive financial markets and investment that maximizes shareholder value. Thus, the firm’s current financial position should play no role in investment. In \( q \)-theory, current profits are just one component of the stream of current and future profits that determine the value of \( q \). In this sense, they play no special role in the determination of investment. This de-emphasis of current cash flow (and thus current profits) in investment is analogous to the denial of any special role of current income in the permanent income hypothesis.

\(^{59}\) See, for example, Fazzari, Hubbard, and Petersen (1988). Stewart C. Myers and Nicholas S. Majluf (1984) also argued that cash flow would affect investment when managers had information not available to investors.

\(^{60}\) An examination of the investment spending of firms with cash windfalls from winning or settling lawsuits supports this finding (Olivier J. Blanchard and Florencio Lopez-de-Silanes 1993). These firms had no problems regarding credit constraints; yet they invested in projects they would not have otherwise pursued. Another striking finding also shows excess sensitivity of investment to cash flow. In 1986, when the price of oil declined dramatically, non-oil subsidiaries of oil companies cut their investment relative to the median in their industry (Owen Lamont 1997). But because this study examines the investment implications of a fall, rather than of a rise, in the price of oil, it is not useful in resolving the role of credit constraint.
maximize shareholder value. In the now-standard theory of the firm, the interests of the shareholders and the interests of the managers are viewed as different. The managers are only the agents of the owners, and accordingly they maximize their own interests instead. Such incentives are said to turn the managers into “empire-builders,”61 who will use the resources they control to increase their own domains.

Empire-building can result from two types of motivations. On the one hand, managers may have only strict economic interests in mind: they care only about their take-home pay, and their effort on the job. Such managers, for example, will be biased in favor of investments whose operation or construction enhances their firm-specific human capital, and thereby increases their bargaining power.

On the other hand, empire-building may be pursued as a goal of its own, for its own sake. We saw earlier that most workers have views regarding how they should or should not perform their jobs. Accompanying such views, most managers and workers will have the further view that the firm should be investing in those jobs. For this reason, the agents making the investment decision are likely to engage in empire-building. We can represent such motivation by adding a term to the utility function of the agent–decision maker. Her utility function will not only depend on her own pecuniary returns and her expenditure of effort. It will also include an additional term reflective of her norms. She will lose utility insofar as the firm’s investment fails to live up to her ideal of what she thinks it should be. In this case, the typical norm is that she thinks that the firm should engage in investment that will enhance her job performance.

Following the logic of Michael Jensen (1986, 1993), empire-building, accompanied by the abdication of corporate oversight in favor of management interests, explains a correlation between investment and cash flow. Furthermore, this correlation will occur regardless of the motivation for the empire-building, whether for purely economic reasons as in the principal-agent model, or, instead, because of managers’ norms for how they think they should behave. Jensen has given many instances of lax corporate oversight in favor of management interests. For example, he has cited the excess exploration and drilling operations of oil companies when retained earnings were high, from 1975 to 1981,62 and the maintenance of low-return operations in many US industries, as in the investments of General Motors throughout the 1980s.63 In Jensen’s views, shareholders would have fared better if profits had been returned to them, giving them the option of investing at a higher rate of return, or perhaps if profits had been used for takeovers outside the industry. To cure what he calls the “failure of corporate internal control,” Jensen has also suggested that firms should issue large amounts of debt, perhaps even by going private. In that case, the added debt obligations act as a brake on excess investment. Regarding investment behavior, Jensen is then on the same page as Keynesian economists such as Klein and Goldberger. They refer to “the preference of many businessmen for internal as opposed to external financing” (1955, 12–13) and also consider it the major reason for the dependence of investment on cash flow.

B. Sociology of the Corporation

Once again, we have seen a neutrality result that depends on the goals of the respective decision makers. Accordingly, the norms of corporate decision makers are central to the sociology of the corporation. For example, Dirk M. Zorn (2004) has examined how the locus of control has changed in large US firms over the past 40 years. He has shown how this control has shifted away from those with a production or a sales orientation to those with a financial orientation.64 Empirically, this is seen in the rise of the chief financial officer. Prior to the 1960s, corporate finances were handled by corporate treasurers, whose duties were mainly restricted to keeping the accounts and producing the budgets. Now, most large corporations have replaced them by a CFO. With the change in title has come a change in function. CFOs are typically central to major decisions. Such a change affects investment decisions. If they are committed to their missions, managers with sales or

---

64 That distinction was emphasized earlier, for example, by Neil Fligstein (1990).
production orientations will be empire-builders. In contrast, the role of the conscientious CFO is to curb those enthusiasms. Fifty years have elapsed since the publication of Modigliani-Miller. According to Zorn, when it first appeared, it did not describe the investment decision of large corporations. Now, quite possibly, changes in corporate decision-making since that time make it more realistic.65

C. Summary

The investment decision demonstrates once again that the respective neutrality result depends on the objective function of the decision makers.

VII. Natural Rate Theory

We now turn to natural rate theory. Once again, the debate concerns the behavior of economic decision makers. The early Keynesians viewed wage setters, and possibly also price setters, as setting nominal wages and prices, respectively, without taking full account of inflationary expectations. In contrast, New Classical revisionists have assumed that wage and price setters care only about relative wages or prices, and therefore wage and price setting will fully incorporate inflationary expectations. Such behavior yields a long-run neutrality result with severe limits on the ability of monetary and fiscal policy to affect unemployment and output. When wage and price setters care only about relative wages and relative prices, accelerating inflation will occur if unemployment is below a critical level, called the natural rate; accelerating deflation will occur if unemployment is above it.

As we shall see, such spirals occur because, at high levels of demand, the representative firm will wish to set the price of its product relative to the price of other firms’ products—which we call its real price—in excess of unity. A standard natural rate model illustrates why this occurs. That model assumes that in each period the typical firm sets a desired real price for the following period; in each period it also makes a bargain with its labor regarding next period’s real wages. Next period’s nominal price and nominal wage are then respectively set by adjusting this desired real price and this bargained real wage according to inflationary expectations. When demand is higher, the desired real price of the representative firm is higher for two reasons: on the demand side, because the demand for its product is higher, and, on the cost side, because the bargained real wage is higher. That bargained real wage is higher both because the typical employee’s opportunity costs, which take into account her chances of being unemployed, are higher, and because the firm’s desire for her labor is higher. Since the firm’s owners, customers, and workers care only about real prices or real wages, a given level of real aggregate demand will be associated with a given real wage bargain between the firm and its workers, and a given desired real price for the firm’s product. If unemployment is sufficiently low—below the natural rate—that desired real price will be in excess of unity. If unemployment is above the natural rate, it will be less than unity.

It is now easy to explain the inflationary and deflationary spirals in natural rate theory. Consider what happens when the representative firm wishes to set its price above that of other firms. In this case, actual inflation will exceed expected inflation. With such a positive gap between actual and expected inflation, inflationary expectations will rise, as inflationary expectations are adjusted upward to conform to reality. But the firm’s desired real price, and therefore the difference between actual and expected inflation, will be unchanged as long as unemployment is constant. There will be no abatement in the rise in expected inflation. Inflationary expectations will be forever increasing, and inflation will rise with it, as nominal prices and wages adjust the real wage bargains and the desired real prices for these increasing inflationary expectations. By similar logic, if unemployment is above the natural rate, there will be a deflationary spiral. The natural rate is the only sustainable level of unemployment without accelerating or decelerating inflation. It corresponds to the exact level of demand where firms wish to set a real price of exactly one.

65 Curiously, the rise of the CFO may have substituted one overenthusiasm (from the point of view of shareholders) for another. There is considerable division regarding whether mergers and acquisitions have positive returns to the buyer. Robert Bruner’s meta-analysis (2002) of many different studies concludes that, on balance, the returns to bidders have been zero. This is a poor return for an activity that has involved so much corporate time and initiative. Furthermore, if some opportunities can be identified as having positive returns, then, to reach an average return of zero, the marginal merger and acquisition has negative payoff.
A. Acceptance of Natural Rate Theory

Most macroeconomists do not just view natural rate theory as a useful null hypothesis. They also see it as a description of reality. Such a view is revealed in textbook presentations. Economists accept natural theory for theoretical and empirical reasons.

Theoretically, they view the assumptions of natural rate theory as realistic. A standard criterion for an economic model is that participants in the economy care only about real outcomes. That is the fundamental assumption of natural rate theory. Also, unlike our other neutrality results, natural rate theory is insensitive to deviations due to "frictions," such as imperfect information, taxes, myopia, or transaction costs. As long as these "frictions" can be expressed solely in real terms, the neutrality result of natural rate theory will be robust.

Empirical considerations have also been influential in economists’ acceptance of natural rate theory. The original Phillips curve showed a close fit between the rate of change of nominal wages and the inverse of the unemployment rate for 97 years of British data, between 1861 and 1957. There was no inflation adjustment in this equation. In the United States in the late 1960s/early 1970s, the theory is also consistent with most econometric estimates. The textbooks thus typically present natural rate theory as a "just-so" story. It runs as follows. The previous Keynesian economists had posited a Phillips curve without a dependence on inflationary expectations. Friedman (1968) and Phelps (1968) perceived that such a theory could not result from models where the participants in the economy are concerned only with real variables. They modified the relationship so that wage and price equations would be affected one for one by inflationary expectations. Such judicious use of economic theory explained the otherwise-mysterious finding of the simultaneous increases in inflation and unemployment of the late 1960s/early 1970s. The theory is also consistent with most econometric estimates.

B. Nominal Considerations in Wage Behavior

We now turn to the same question regarding wages that we asked concerning consumption and investment. Is there “excess sensitivity” relative to the respective neutrality? Natural rate theory is based on the assumption that wages and prices are set only with real considerations in mind. “Excess sensitivity” here about the Phillips curve, including the natural rate of unemployment itself, is considered to be estimated with great imprecision. Akerlof, William T. Dickens, and George L. Perry (2000) show a range of estimates for both wage and price equations with many different specifications. These estimates, particularly when made for periods of low inflation, show considerable variation in the sum of the coefficients on lagged inflation, dependent on the specification. Another bit of evidence that suggests such estimates will be sensitive to specification comes from the high standard errors on the natural rate itself (Douglas Staiger, James H. Stock, and Mark W. Watson 1997); it would be surprising that the sum of lagged coefficients could be estimated precisely if another component of the Phillips curve, the natural rate, could be estimated only with very low precision. Gordon’s own estimates show very different values for this sum of coefficients. Of course, there is a theoretical reason why estimates of such a sum should not be robust. With rational expectations, rather than a simple mechanical theory of formation of inflationary expectations, Sargent (1971) shows that there is no theoretical reason that they should sum to one.

See, for example, Robert J. Gordon (1977, table 3, lines 6 and 7, 260).

Given the importance of such findings, it is remarkable that their robustness to specifications of time period, data, and exact specification of the Phillips curve has never been subjected to tough tests—even though everything else
takes the form that nominal considerations affect real wage or price setting in some way or other.

Evidence of one form of violation of the assumptions of natural rate theory is especially stark. That evidence concerns downward wage rigidity. Such wage behavior can easily be perceived statistically by examining distributions of wage-changes. These distributions are characterized by a bunching of wage changes at exactly zero; there are some wage changes just above zero in these distributions, but almost no wage changes just below.\(^68\) Careful studies have documented such wage stickiness in Australia, Canada, Germany, Japan, Mexico, New Zealand, Switzerland, the United States, and the United Kingdom.\(^69,70\) There seems to be no way to account for such nominal wage rigidity with the basic assumptions underlying natural rate theory; that participants in the economy care only about real prices and real wages.

Wage stickiness also explains a macroeconomic observation that is an anomaly for natural rate theory. Unemployment was so massive in the Great Depression that inflation should have been below inflationary expectations Phillips curve, such high unemployment would have caused a deflationary spiral. Data on inflation are available for 12 countries for this long period. With any natural-rate adaptive-expectations Phillips curve, such high unemployment would have caused a deflationary spiral. Data on inflation are available for 12 countries for the Great Depression. Not a single one of them shows such a spiral.\(^71\) For example, the United States experienced rapid deflation from 1929 to 1933, but inflation systematically neither rose nor fell for the next decade. The predictions of natural rate theory are thus grossly violated. But sticky wages offer a good explanation for such behavior. For example, a dynamic simulation of the US economy with money wage rigidity and with Depression-level unemployment fits the data all but exactly (Akerlof, Dickens, and Perry 1996).\(^72\)

Nominal wage rigidity may not only be statistically perceptible; it can also be macroeconomically important, even outside of Great Depressions. Nominal wage rigidity imparts a long-run trade-off between unemployment and long-run inflation. This trade-off is of sufficient size that it should deter central banks from targeting very low levels of inflation. For example, simulations of the US economy (Akerlof, Dickens, and Perry 1996) show that an increase of the inflation target from 0 to 2 percent will permanently reduce unemployment by 1.5 percentage points.\(^73\)

Norms as Explanation for Sticky Money Wages.—It seems to be impossible, or all but impossible, to explain the existence of sticky money wages, without relaxation of the basic assumption that the utility functions of employers or of employees contain real arguments. A simple and natural amendment to the standard model explains such sticky money wages: that employees have a norm for what wages should be. According to that norm, they will lose utility from a money wage decline. Sticky money wages then result, as the bargains between employers and employees reflect the presence of this ideal in the utility function.

Indeed, the study by Bewley (1999) gives

\(^68\) These distributions have accumulations at zero, and they are also asymmetric: there are more wage changes above zero than below zero. This suggests that the accumulations at zero do not occur just because there is a menu cost for changing wages.


\(^70\) See, for example, Anthony P. O’Brien (1989) and Christopher Hanes (2000).


\(^72\) There are other possible reasons for this failure of the standard predictions from natural rate theory. Inflationary expectations may not have been adaptive; the failure of deflation to accelerate could be due to expectations that the price level would return to some normal level. In the United States, the National Recovery Act, which encouraged firms to increase prices, and unionization, which gave a fillip to wages, could also have affected the trade-off between inflation and unemployment. But since unemployment was so very high for so very long, and since the absence of accelerating deflation was so universal across countries, this still seems to be a dog that did not bark. It seems to point to a problem with natural rate theory.

\(^73\) See Akerlof, Dickens, and Perry (1996, table 4).
direct evidence that such a norm exists and is responsible for wage stickiness. His extensive open-ended interviews sought to elicit why employers failed to cut money wages in the Connecticut recession of 1991–1992. Bewley concludes that, even though substitute labor was easily available, employers were reluctant to cut wages because of the negative effects of such cuts on morale. He says that managers were afraid that cuts in money wages would cause workers no longer to “identify” with their companies. There might be no immediate consequences during the recession. But employers thought that such cuts would cause workers to shirk after the recession had ended. They also feared that their best workers would be more likely to quit. These stories indicate that work—

eries are not thinking about their wages only in real terms, relative to the price level or the wages received by others. They also have a special aversion to cuts in wages below their current nominal levels.75

**Norms about Wage Increases.**—The motivation underlying resistance to money wage cuts is so obvious, and the facts are so unexceptionable, that most macroeconomists accept the possibility that money wages are sticky. Even so, they rarely appreciate the broader implications of such violation of the assumptions of natural rate theory. Their adjusted model is that price and wage decisions are made only with real considerations in mind, but desired wage changes will be truncated insofar as they entail money wage decreases. To my mind, such a view entails a theoretical error. As we have seen, the existence of money wage rigidity occurs because workers have a norm, which affects their utility function, that their employers should not make such cuts. The message of this finding is that norms in the utility function yield at least one clear violation of natural rate theory. That suggests the further empirical possibility that workers (and also employers and customers) may also have other norms regarding what nominal wages (and prices) should be. All such violations are exceptions to natural rate theory, and yield reasons for long-run trade-offs between inflation and unemployment.

Money wage rigidity is then potentially only the tip of an iceberg. If there is one way in which nominal wages enter utility functions, because of employees’ norms regarding what their employers should or should not do, there could also be many other ways.

There is another natural way whereby such norms could enter utility functions: employees may not only have a norm that they should not take wage cuts. They may also have norms regarding the nominal rate of increase of their wages or salaries. For example, employees may believe that their employer should give them a nominal raise.

There is little research on the existence of such norms. The two questionnaire studies that have investigated it obtain strong and mutually reinforcing results. Eldar Shafir, Peter Diamond, and Tversky (1997) asked respondents to comment on a vignette about two young women who take their first jobs with the same initial income. Specifically they asked respondents who will be better off: Barbara, who receives a 5-percent raise in the presence of 4-percent inflation; or Ann, who receives a 2-percent raise when inflation is zero; 79 percent of respondents correctly said that Barbara would be worse off than Ann economically. Nevertheless, 64 percent of respondents also said that Barbara would be happier.76 Such responses are contrary to the natural rate hypothesis that employees only care about real returns. But an easy explanation for this phenomenon occurs if

74 In more detail, Bewley (1999, 1–2) summarizes his findings: “Other theories fail in part because they are based on unrealistic psychological assumptions that people’s abilities do not depend on their state of mind and that they are rational in the simplistic sense that they maximize a utility that depends only on their consumption and working conditions, not on the welfare of others. Wage rigidity is the product of more complicated employee behavior, in the face of which manager reluctance to cut pay is rational. Worker behavior, however, is not always rational and completely understandable. A model that captures the essence of wage rigidity must take into account the capacity of employees to identify with their firm and to internalize its objectives. This internalization and workers’ mood have a strong impact on job performance and call for material, moral, and symbolic reciprocation from company leadership.”

75 Following the argument by Raj Chetty and Adam Szeidl (2006), some employers may have been concerned with the fact that their employees had fixed mortgages that they would find difficult to pay with cuts in nominal wages. This puts the violation of natural rate theory in another place: why were these financial contracts in nominal rather than in real terms?

Barbara and Ann both think that their employer should give them a nominal wage increase.

Another study, with a different form of questionnaire, independently found a similar response. Robert Shiller found that 49 percent of a sample of the general public either fully or weakly agreed with the following statement: “If my pay went up, I would feel more satisfaction in my job, more sense of fulfillment, even if prices went up as much.” An additional 11 percent of the general public were undecided, while only 27 percent completely disagreed. As in the case of Ann and Barbara, such opinions are consistent with the view that workers think their employers should give them a nominal wage increase: they will be disappointed when it does not occur. Shiller’s finding may be similar to the public’s view of Ann and Barbara. But, as he reports, it is also in stark disagreement with the view of professional economists that underlies natural rate theory. Ninety percent of economists weakly or strongly disagreed with the statement; 77 percent were in complete disagreement.77

Such norms—regarding the wage or salary increase that employees think they should receive—can be economically consequential. They cause the long-run inflation-unemployment trade-off to be downward sloping. With such a norm, at higher levels of inflation workers will not experience disappointment from receiving lower nominal wage increases than they think they should receive; therefore, at higher inflation, ceteris paribus, wage bargains will result in lower real wages, which will reduce the relative price that the firm wants to set, and therefore raise the rate of sustainable employment. There is a need for further research following Shafir, Diamond, and Tversky and Shiller regarding whether workers have norms regarding the nominal wage increases they think they should receive.

High Inflation.—The opinions expressed regarding Barbara and Ann, and also the opinions of Shiller’s respondents, suggest that the long-run trade-off between inflation and employment is upward sloping. These answers were elicited in the United States and thus are reflective of respondents’ views in an environment where inflation has been low. But if inflation is very high and therefore also very salient, the answers to such questionnaires could be very different. And they could impart a very different shape to the trade-off between macroeconomic demand and steady-state inflation.78 In such cases, people may gain satisfaction only from wage and salary increases that exceed inflation. Such norms regarding how employers should behave will then necessitate higher real wages (to maintain the same level of satisfaction) at higher levels of inflation. The long-run inflation-employment relation will then be downward sloping. Such behavior gives a much stronger rationale, even than current rational-expectations credibility models (Barro and Gordon 1983; Kenneth Rogoff 1987), why central banks should maintain price stability. Failure to appreciate this realistic possibility again may be another case in which the absence of norms from utility functions has unduly blinkered macroeconomic thinking.79


78 Bankruptcy and financial considerations become especially important when inflation is very high. It is also worth noting, at least parenthetically, that high levels of bankruptcy at times of high inflation are themselves a symptom of money illusion. Such bankruptcies reflect the non-indexation of financial contracts.

79 In addition to the two questionnaire studies I have mentioned, indexed contracts give another indicator for the existence of nominal notions concerning what wage increases should or should not be. Economists are often surprised at the small fraction of union contracts that are indexed at all. (Louis N. Christofides and Amy Chen Peng 2004, for example, analyzed a sample of almost 12,000 Canadian union contracts from 1976 to 2000. The mean length of these contracts was slightly more than two years (25 months). Only 19 percent of these contracts were indexed.) But even when such indexation occurs, their form violates the condition that they were struck with only real considerations in mind. For an imperfect index such as the CPI, which reflects both supply shocks and demand shocks, the optimal COLA adjustment will be less than one, but it will always be (almost) symmetric for positive and negative deviations of inflation from a threshold. (See Jo Anna Gray 1978, Ronald G. Ehrenberg, Leif Danziger, and Gee San 1983, and David Card 1986 for the derivation of optimal indexation.) But COLA adjustments are only positive. (Card 1986, S146, has expressed this in terms of a formula: w(t) = w(t) + max{0, α[p(t) − p]}, where w(t) is the nominal wage, w(t) is the nominal target, p(t) is the actual price level, and p is the threshold.) Thus, the form of the contract violates optimality. In practice, this violation is also bitter. For example, in roughly one-third of a large Canadian sample of indexed contracts, inflation was always below the threshold (see Christofides and Peng 2004, 11, fn. 19). Thus, the form of indexed contracts, when they exist, shows that union wage negotiators think that COLA adjustments should never be negative. The form of indexed contracts gives another robust indicator that, indeed, wage setters...
C. Prices

We have just seen that employees’ norms regarding nominal wages may affect bargained real wages, and therefore cause trade-offs between long-run inflation and long-run unemployment. Similarly, customers’ norms regarding price levels and price changes may also cause long-run trade-offs between output and inflation.

Indeed, models by Katsuhito Iwai (1981), Julio J. Rotemberg (1982), and Andrew S. Caplin and John Leahy (1991) all have long-run trade-offs between inflation and unemployment. Each of these models assumes that there are real costs to nominal price changes. If, instead, there were real costs to real price changes, the assumptions of natural rate theory would still be satisfied, and no such trade-off would occur. These models then pose the question why there should be such real costs from nominal price changes. Iwai, Rotemberg, and Caplin and Leahy all respectively assume that there is a “menu” cost in making these changes known. But the physical costs of making such changes, as in the printing of new menus, are trivially small. Norms regarding price changes, however, give an alternative reason why these costs might—indeed—be of sufficient size to induce a significant long-run trade-off between inflation and unemployment. Customers may think that firms should not raise prices. In that case, price increases (or increases of greater size) are likely to induce angry customers to search for alternative suppliers. At higher steady-state inflation, firms will be changing their nominal prices more, and therefore will face more elastic demands for their product. Producers’ natural microeconomic response to this increased elasticity—a lower price for their product—will produce a macroeconomic trade-off between inflation and aggregate demand.

Just as sticky money wages indicated that employees have norms regarding wage change, similarly, sticky prices indicate that customers have norms regarding price change. Thus, the extensive evidence on price stickiness reveals violation of the assumptions of natural rate theory, and also the existence of norms regarding price change. Like wage changes, price changes also agglomerate at zero. Dennis Carlton (1986) has shown that prices are often sticky for significant periods of time. Furthermore, prices seem to be especially sticky in customer markets. Alan Kackmeister (2002) has compared price changes at the end of the nineteenth century to such changes a bit more than a century later. Price changes of specific goods at retail stores were recorded from June 1889 to September 1891; Kackmeister revisited the same commodities and their price change for a comparable period, from June 1997 to September 1999. Price change in the late twentieth century was five times more frequent than a century earlier. Furthermore, in the nineteenth century, the average spell of constant price for an individual good was very long. It was approximately 80 months. Such constancy of prices can easily be explained by customer norms regarding price change. The customers have a notion of the price that they ought to pay at stores where they are continued and knowing customers. Kackmeister suggests that the decline in long-term customer relationships is one factor responsible for greater frequency of price change today.

Emi Nakamura and Jón Steinsson (2005) give an economic reason why customers would have such a norm that firms should not change prices. They view consumer purchases as habit-forming. Thus, by buying a particular brand, or patronizing a particular store, consumers are putting themselves in a position where they can be exploited. Their loyalty puts the firm in a

---

80 Marika Karanassou, Hector Sala, and Dennis J. Snower (2003) find considerable long-run trade-off between inflation and unemployment in a model with nominal price staggering and money growth.

81 See also Blinder and Don Choi (1990) and Blinder et al. (1998).

82 The meaning of customer markets was especially explored by Arthur Okun (1978).

83 I derive this result from Kackmeister’s data in the following way. He finds that in the nineteenth century, only 5 percent of items changed their prices per month. This means that the average spell of constant prices would have been 20 months (the inverse). But that is a biased statistic for the average length of time between price changes for an item on the shelf. The difference between the average spell of employment or unemployment and the average spell being experienced by an individual suggests a rule of thumb ratio for four to one. Using this ratio as a rule of thumb suggests that the spell between price changes averaged over the individual items on the shelf would be 80 months.
position where it can take advantage of the consumer by raising prices. Firms then make an implicit contract with their customers: they will not change their prices unjustifiably. Since such an implicit contract is easier to make (and enforce) regarding nominal prices than real prices, the implicit guarantee is in nominal terms. Nakamura and Steinsson have also discovered a phenomenon that suggests strikingly that firms do behave this way. Goods in store 126 (chosen for its completeness of data) of Dominicks Finer Foods chain frequently go on sale; when the sale ends, their nominal price returns to the exact same level. Such behavior is consistent with the view that consumers think that prices should not change (for whatever reason); and that they are also likely to retaliate (change brands) when prices do change.

I should also remark that in countries where inflation is very high, customers will expect price changes to occur frequently, and possibly be of large magnitude. The inhibitions against price changes when inflation is low are eroded at high inflation. Thus, while norms concerning prices give a negative long-run trade-off between inflation and unemployment at low inflation, at high inflation that trade-off could very well be reversed.

D. Summary

To summarize, there is considerable evidence of violation of the assumptions and predictions of natural rate theory. Wages and prices are nominally rigid; there were no deflationary spirals in the Great Depression; and questionnaire respondents act as if they have a positive like for nominal wage increases. This evidence suggests that wage earners and customers have views on what wages and prices should be. The reflection of such views in utility functions produces trade-offs between inflation and unemployment. Those trade-offs have significant implications for economic policy. On the one hand, central banks should avoid very low targets for inflation. On the other hand, they should avoid high inflation, where the trade-offs between inflation and unemployment may be reversed.

VIII. Rational Expectations Theory

Our discussion of rational expectations piggybacks on our previous discussion of the natural rate.

According to rational expectations theory, insofar as the central bank changes the money supply systematically in response to employment conditions, the public will foresee that response and change prices and wages exactly to compensate. The public’s anticipation will then exactly offset the response. Monetary policy is neutral.85

There are two key assumptions underlying this neutrality. The obvious one is rational expectations. To some, rational expectations regarding the effects of the money supply on prices and wages would seem to be beyond the sophistication of most wage and price takers, and also of most wage and price setters.

Even in the case where all those involved in buying and selling goods and labor services have rational expectations, however, the neutrality results of rational expectations theory require also that nominal considerations do not enter into the setting of either wages or prices. The previous descriptions of the ways in which nominal wages and prices enter into preference functions, via employees’ views of the wages that ought to be received and consumers’ views of the prices that ought to be paid, give further reason why the neutrality results of rational expectations will be violated. If prices and wages are affected by people’s notions of what their nominal values should be, monetary policy can be effective in stabilizing output—and possibly in raising its long-run level—even in the presence of rational expectations.

IX. Economic Methodology

We have seen that the absence of norms plays a key role in each of the five neutralities. Why have economists made such systematic omissions? The omission of norms from macroeconomics, as well as from economics more

84 Also, COLA clauses are asymmetrically positive. See footnote 79 above.

85 Empirically there is a theoretical puzzle of excess sensitivity to monetary shocks (Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans 1998). Christina and David Romer (1989) have shown that such a response occurs with lags that would be surprisingly long if expected monetary shocks were always neutralized.
generally, can be explained by economists’ adherence to positive economics.\textsuperscript{86} Friedman’s (1953) essay on positive economics describes the methodological implications of such belief. In particular, he says that economic theorists should strive for parsimonious modeling. According to Friedman, they should even forsake realistic assumptions in pursuit of such parsimony. Maximization models with only objective arguments of utility have been defined as more parsimonious than models where people, additionally, lose utility insofar as they, or others, fail to live up to their standards. As a result, whatever the empirical validity or relevance of such norms, positive economics has a methodological bias against their consideration. It privileges models without norms.

The prescriptions of positive economics regarding the conduct of empirical investigation compound the bias against norms. Friedman says that economists should not pay heed to the stated intentions of decision makers, which would include their norms as to how they and others should behave. Instead, empirical work should test only hypotheses that economists consider to be based on parsimonious models.

If economic tests had great power, then it would be easy, of course, to follow Friedman’s dictum of making more and more refined tests of hypotheses with decreasing parsimony. If norms really do affect behavior, this method would reject models without norms and in due course would arrive at models where people’s views regarding how they should behave affect decision making. But economic tests lack power. All economic models are very imprecise in their specification of the independent variable, the nature of the dependent variables, the nature of leads and lags, and the nature of residuals. Yet worse, most economic problems involve simultaneity (as in supply and demand), making establishment of causality difficult. In almost any instance, such a large number of models can be fitted statistically that it is extremely hard—and perhaps impossible—to statistically reject all the variants of models without norms. As a result, the program of positive economics—with its initial nulls of models based only on utility with objective variables verified only by statistical hypothesis testing—has severe bias against explanations of economic phenomena where norms play a role.

Summers (1986) illustrates the severity of this bias. The conventional test of the efficient markets hypothesis—that stock prices are the expected value of future returns—looks for autocorrelations of the excess returns on stocks relative to bonds. Following Summers, it would take approximately 5,000 years of data with such a test to obtain as much as 50 percent rejection of an alternative model where stock prices are more than 30 percent away from their fundamentals 35 percent of the time. With such lack of power, nulls are important. When they are not rejected, alternative theories, such as those with norms, are not even considered. This lecture has illustrated such reversion to normless nulls. Consumption behavior, investment behavior, and wage and price behavior—the three most important components of most macro models—all display excess sensitivity relative to respective neutralities. All of these violations could be easily explained by norms. Yet in each case economists have sought to explain such violations of classical theory by norm-less models.

In contrast to reliance on statistical testing, disciplines other than economics typically put much greater weight on a naturalistic approach. This approach involves detailed case studies. Such observation of the small often has been the key to the understanding of the large. To me, the most dramatic example of such a relation between the small and the large occurs in the structure of life itself. Francis Crick and James D. Watson\textsuperscript{87} conjectured correctly that if they could describe the crystalline structure of a single DNA molecule, they would have unlocked the secret of life. The duality between the structure of the DNA molecule and the way in which organisms are generated and reproduced is one of the most beautiful findings of human knowledge. It indicates the sense in which Crick and Watson were, indeed, profoundly correct.

What are the implications for social science? Positive economics, with its emphasis on statistical analysis of populations, would suggest that...
the intensive study of a single molecule would be an all-but-worthless anecdote. In the case of DNA, we know that the exact opposite is true: because DNA is a template that determines all of the cells of the organism, and also its reproduction, one molecule may not tell all, but it does tell a great deal. Form follows function.

Is there some reason to believe that economic behavior and economic units are any different? Economic decisions may not be as duplicable as biological processes, but the basic reason why science intensively studies the microscopic applies to economics as well. The individual economic unit, be it a firm, a consumer, or an employee, behaves the way it does for a reason. And if these actors behave as they do for a reason, we can expect to find those reasons from the structures that we see in close observation; and because of those structures their behavior will also tend to be duplicated. This duality between duplicability and structure explains why much of science concerns very close observation, as it also explains why the study of even a single part of a single DNA molecule will be revealing.

Standard economic methodology says that it is impossible to infer motivation of individual actors from intensive case studies. Anthropologists and sociologists listen carefully to individuals in such studies. When people follow the norms, they use them to explain their actions; when, on the other hand, they violate the norms, they become the subject of local gossip. Those case studies are revealing because—like a language, which dictates how one should speak—the norms are common knowledge. In this lecture, we have seen one prominent example of the use of such knowledge: Bewley’s interviews uncovered the common understanding of the norms regarding wage cuts among Connecticut employers in the early 1990s.

Summary.—Positive economics systemically denies that norms can be understood from intensive case study. Precedence given to models without norms because they are by definition more parsimonious and statistical tests of low power then jointly create a firewall against consideration that norms play a role in determining behavior. For these reasons, current economic methodology inherently has created a biased economics. In contrast, a more naturalistic approach would prescribe a different methodology. In this case, economists would observe decision makers as closely as possible, with the express intent of characterizing their motivation, and would use such characterization as the basis for modeling of economic structure. Indeed, sociological and anthropological ethnographers do precisely that: they depict their subjects’ motivation from close observation.

X. Endogeneity of Norms

It is now time to discuss the endogeneity of the norms. There is a special reason for its consideration. Robert Lucas discovered that, with endogenous rational expectations regarding inflation, monetary policy that was intended to stabilize the macroeconomy would, instead, be exactly neutral. Similarly, is it not possible that endogeneity of the norms, like Lucas’s endogeneity of inflationary expectations, will cause the neutralities again to hold? We shall discuss this question regarding all five neutralities. For the most part, we find that the type of government interventions being considered are usually of such frequency, or of such order of magnitude, that they should provoke relatively little change in the norms. Endogeneity of the norms should have little effect, then, on our previous conclusions.

A. Ricardian Equivalence

Let’s begin by returning to Ricardian equivalence, which is still the simplest case. We found that if people have a norm regarding the amount of their bequest, then lump-sum transfers to an older generation will not be neutral. There remains the possibility that the source of the warm glow to the older generation is not the total bequest, but instead the bequest to the younger generation net of the transfer. In this case, if the transfers change, then the norm changes. Ricardian equivalence will again be valid. While such changes in norms with the size of transfers are a theoretical possibility, they also seem highly unlikely. The size of the transfers involved—especially for those rich enough to make large nonaccidental bequests—would seem to be too small to warrant such a
sophisticated calculation. Our earlier discussion did discuss at least one change in the norms regarding bequests, but that resulted from a very large change in people's orientation. It resulted from changes in their conception of the family—of their own place within it and of the place of their heirs. That also occurred over a very long run—over the course of centuries.

B. Life-Cycle Hypothesis

Regarding the life-cycle hypothesis, we argued that consumption depends upon current income because norms regarding how much people think they should spend are linked to it. But such a norm would be highly unlikely to change as a result of the use of fiscal and monetary policy for stabilization. In the first place, such stabilization will make the adherence to the norm less costly, not more costly, in purely economic terms. Furthermore, macroeconomic sources are responsible for only a small fraction of the variation in individual incomes. As a result, there is further reason why the role of current income in norms is unlikely to change as a result of macroeconomic stabilization.

C. Cash Flow and Investment

The rise of the CFO suggests that norms regarding investment have changed in large US firms. Quite possibly, this change occurred because firms realized the need for financial controls that compared the returns on inside and outside options. Such an endogenous response would make Modigliani-Miller correct. But, following Zorn (2004), this change took 40 years. In the meantime, in the short run, following our earlier logic, investment would have depended on cash flow. And, of course, even in the long run the CFO, who is only one voice among many in corporate decisions, may not be fully effective.

D. Natural Rate Hypothesis and the Role of Rational Expectations

Regarding the natural rate hypothesis and also the rational expectations hypothesis, we saw that they will no longer hold if norms of price and wage setting have nominal components. Regarding prices and wages, the most powerful evidence in favor of norms comes from employees' resistance to money wage cuts and customers' resistance to nominal price increases. As long as inflation is low, it is doubtful that small changes in inflation will affect such norms. People seem to find it easier to think in nominal, rather than in real, terms. Indeed the facilitation of such thinking is one of the benefits of money according to the textbook mantra on its three uses: for transactions, as a store of value, and as a unit of account. Money is useful as a unit of account especially if people think in nominal, rather than in real, terms. As a result, as long as inflation is low, people are unlikely to forsake making calculations in nominal terms, especially regarding the norms of what wages or prices should be. Of course, if inflation increases to high levels, the norms for wages and prices and the method of calculating those norms will change. Exactly how they change—with the possibility that they under-adjust to increases in inflation when it is low and over-adjust when it is high—should be empirically investigated.

E. Where Do the Norms Come From?

We do not know the general answer to the question where norms come from. This lecture has tried to make the case that norms, such as they are, could potentially play an important role in macroeconomics. Hopefully, then, it has added to the motivation for research on their microfoundations.88

XI. Conclusion

This lecture has shown that the early Keynesians got a great deal of the working of the economic system right, in ways that are denied by the five neutralities. As quoted from Keynes earlier, they based their models on “our knowledge of human nature and from the detailed facts of experience.” They used their intuitions regarding the norms of how consumers, investors, and wage and price setters thought they

88 This lecture has been very much influenced by the insights of the Ph.D. thesis of Robert Akerlof (2006) on preferences for beliefs. His thinking on this subject has influenced many of the sections of this paper, especially on consumption and the endogeneity of norms.
should behave. There is systematic reason why such knowledge and experience are likely to be accurate: by their nature, norms are generated and known by a whole community. They are known to those who abide by them, and those who observe them as well.

We have shown ways in which macroeconomic variables will be affected by norms. The neutralities say that consumption should have no special dependence on current income; investment should be independent of current cash flow; wages and prices should not depend on nominal considerations. The very construction of those neutralities denies the possibility that peoples’ decisions might be influenced by their views regarding how they, and how others, should behave. In practice, however, the neutralities are systematically violated. Insofar as economists have felt it necessary to explain these violations, they have appealed to a variety of different frictions, such as myopia and credit constraint. In so doing, they have failed to consider that those violations would occur even in the absence of those frictions: they will occur because of decision makers’ norms.

The incorporation of norms based on careful observation imparts an appropriate balance to macroeconomics. The New Classical research program was correct in viewing models of the early Keynesians as too primitive. They had not been sufficiently attentive to the role of human intent in choices regarding consumption, investment, wages, and prices. But that research program itself has failed to appreciate the extent to which the Keynesians’ views of macroeconomics were also reflective of reality, since they were based on experience and observation. A macroeconomics with norms in decision makers’ objective functions combines the best features of the two approaches. It allows for observations regarding how people think they should behave. It also takes due account of the purposefulness of human decisions.

REFERENCES


Bauer, Thomas, Holger Bonin, and Uwe Sunde. 2003. “Real and Nominal Wage Rigidities and the Rate of Inflation: Evidence from Ger-


ation, and Non-Contingent Wage Adjustment.” Unpublished.


Jensen, Michael C. 1993. “The Modern Industrial Revolution, Exit, and the Failure of In-


Myers, Stewart C., and Nicholas S. Majluf. 1984. “Corporate Financing and Investment Decisions When Firms Have Information That


