

# Cycles in Lending Standards?

---

---

John A. Weinberg

The lending activity of commercial banks has long received considerable attention as an important contributor to the performance of the economy. This attention has, perhaps, become sharper in the wake of the difficulties experienced by the banking industry in the 1980s. In recent years, the public perception of bank lending seems to have swung through a cycle. In the early 1990s, the prevailing view was that the bank loan market was experiencing a credit crunch in which banks set unreasonably high credit standards, denying credit to qualified borrowers.<sup>1</sup> By late 1994, with growth in bank loans picking up, some voiced concerns that banks were possibly becoming too loose in their standards for acceptable credit risks. These concerns appeared in the pages of the *American Banker* and other professional journals and in speeches by the Chairman of the Federal Reserve Board and the Comptroller of the Currency.<sup>2</sup>

Do swings from tightness to laxity in credit standards constitute an inherent part of bank lending activity? Some observers have suggested that such cycles can be caused by an imperfection in bank credit markets that results in a systematic tendency for banks to overextend themselves during general expansions of lending. In some expressions of this view, the imperfection is the result of government intervention in banking markets, while in others it results from the very nature of credit markets. In any case, the implied consequence is a cycle in lending behavior that is distinct from and may exert an independent influence on the general business cycle.

The existence of a systematic cycle in lending standards could have important public policy implications. If lending displays a bias toward too much

---

■ The author thanks Tom Humphrey, Mary Finn, Tony Kuprianov, and John Walter for their comments on an earlier draft. The views expressed in this article are the author's and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

<sup>1</sup> For general discussions of credit crunches and the experience of the early 1990s, see Bernanke and Lown (1991) and Owens and Schreft (1995).

<sup>2</sup> See, for instance, Dunaief (1995), Connor (1995), and Stevenson (1994).

risk during expansions, resulting in increased risks of bank failures and losses to bank insurance funds, then bank regulation might justifiably seek to smooth out lending cycles. Placing greater scrutiny and restrictions on bank lending during periods of loan growth would slow the expansions but limit ensuing losses.

Justifying a regulatory policy aimed at eliminating or reducing cyclical swings in credit standards and lending activity requires that expansions of credit be inherently excessive. Excessive, however, is a relative term. A relevant question in this regard concerns the behavior of an “ideal” credit market, in which there is no source of market failure or government-induced distortion. Would such a market produce cycles in lending standards? This is the question that this article explores. First, Section 1 describes the notion of cycles in standards in a bit more detail. The following sections examine a stylized model of credit market behavior. This model suggests that lending standards might naturally be expected to change with market conditions. In fact, the model serves to make the point that in a well-functioning credit market, an expansion of lending almost necessarily implies an easing of standards and the extension of credit to “riskier borrowers.” This is true when borrower riskiness is defined in terms of borrower-specific characteristics drawn, for instance, from a commercial borrower’s recent income statement or balance sheet. These characteristics tell only part of the story of the true expected profitability of a loan to a particular borrower. Also important are aggregate conditions that affect the demand for the borrower’s product or the supply of its inputs. These factors are typically not well captured by borrower-specific indicators of credit quality.

## 1. CYCLES IN CREDIT STANDARDS

In discussions of bank lending activity, the notion of cycles in lending standards typically begins with expansions; standards fall with heightening competition in expansions and rise in contractions as banks respond to their own capital shortfalls or the constraints of regulators. On the downside, this view is related to the notion of a credit crunch, which has received much attention in recent years. On the upside, the view often seems to embody the belief that expansions of credit are inherently excessive. In March of 1995, for instance, the *American Banker* reported on growing concerns about easing credit standards and on warnings from the loan officers’ professional association advising lenders to resist competitive pressures and maintain credit quality.<sup>3</sup> Such warnings come very close to the caution voiced by Federal Reserve Chairman Alan Greenspan in speeches given in late 1994.

If, indeed, there is a natural tendency for expansions of credit to push down lending standards, then it would naturally follow that expansions would lead,

---

<sup>3</sup> See Mathews (1995).

at least sometimes, to significant increases in losses on loans. Further, the contraction phase of the credit cycle could be worsened as banks find themselves with bad assets on their books. Hence, under this view the primary driving force of cycles in the credit markets is the propensity of lenders to succumb to an unrealistic optimism in good times, creating lending booms that sow the seeds of their own demise. Just such a description of the cycle appears in some discussions by credit professionals.<sup>4</sup> In this sort of description the expansion of lending could, itself, be the impulse that drives the cycle, as a spontaneous wave of optimism hits the lending community. Alternatively, the expansion could be an overreaction to other shocks to the economy that “legitimately” shift the supply of or demand for credit.

The view of cycles outlined above is one of fluctuations in a number of variables. In this view, the cycle in lending standards coincides with the cycle in the amount of lending, while movements in the amount of lending are followed by movements in loan losses. In particular, an increase in lending activity is followed by an increase in losses. Does this pattern appear in the data? Figure 1 displays the behavior of the growth rate of total loans and loan charge-offs as a fraction of total loans at insured banks and thrifts in the United States from 1950 to 1992.

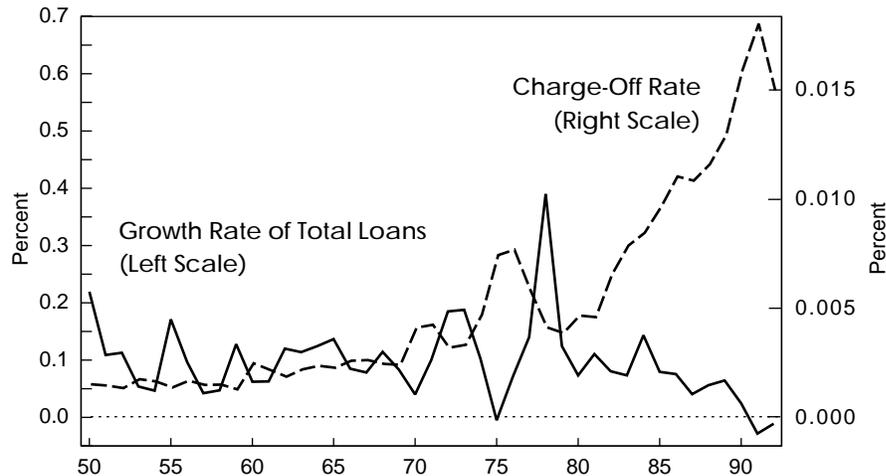
While the relationship is not striking, there do appear to be periods in which unusually strong loan growth preceded rising charge-offs. For instance, in 1972 and 1973, annual loan growth was about 18 percent compared to around 10 percent in 1971 and 1974. From 1973 through 1976, charge-offs, as a percent of total loans, grew from 0.33 percent in 1973 to 0.77 percent in 1976. It is important to note that accounting standards allow banks some discretion as to when to write off a nonperforming loan.<sup>5</sup> Consequently, charge-offs resulting from an episode of poor credit quality can be spread out over time, delaying and smoothing the apparent response of losses to an expansion of lending. The last several years of data in Figure 1 seem to reflect the downside of the cycle: an extended period of rising losses followed by a period of declining loan growth, culminating in two years of declines in the level of lending activity.

While not conclusive, the behavior of loan growth and charge-offs is not inconsistent with the notion of a cycle in lending standards. Alternatively, one might seek more direct evidence on lending standards. The Federal Reserve Board’s Senior Loan Officer Opinion Survey on Bank Lending Practices contains explicit questions on this topic. Schreft and Owens (1991) provide a detailed description of this survey evidence. Broadly stated, they find that loan officers’ self-professed tendency to tighten lending standards follows a cyclical pattern that tends to peak (attain the greatest tightening of lending standards)

---

<sup>4</sup> See, for instance, Mueller and Olson (1981) and Stevenson (1994).

<sup>5</sup> This issue is discussed in Darin and Walter (1994).

**Figure 1 Loan Growth and Charge-Offs**

just prior to or during general economic downturns. They find similar behavior in responses to questions about loan officers' general willingness and unwillingness to make loans. Further, peaks in the loan growth series in Figure 1 tend to occur at or around the troughs in the lending standards series studied by Schreft and Owens. In other words, the survey data suggest that low self-reported standards coincide with high growth in actual lending, while bankers report tightest credit around periods when lending growth is slow.

An additional source of information on lending standards could come from examining the characteristics of banks' borrowers. If borrowers' average financial conditions become weaker, then one might be able to conclude that standards have eased, especially if one can control for the general state of the economy. Cunningham and Rose (1994) perform such a study. They examine evidence on the financial conditions of small and mid-sized commercial borrowers. This evidence appears to be consistent with a general easing of standards from 1978 to 1988 and a tightening from 1988 to 1991. Except for 1978 itself, 1978 to 1988 was a period of fairly steady loan growth, with a few relatively strong years. By contrast, 1988 to 1991 was a period of generally declining loan growth. Hence, this study serves to confirm the general coincidence of changes in standards with movements in total lending activity.

The evidence discussed above is broadly consistent with the notion of cycles in lending standards. More difficult is the question of whether expansions represent a systematic tendency of banks to become too easy in the extension

of credit. There are at least two views as to what might drive lenders to display an excessively tolerant attitude toward credit risk. One view is that there is a fundamental imperfection in financial markets and, in particular, markets for bank credit. Under this view, the source of the imperfection seems to be that the credit quality of borrowers is difficult and costly to observe. Banks spend resources gathering information on borrower characteristics, but it is difficult for outside observers to verify the information obtained by the bank. Such limits to the flow of information form the basis of much of the recent work in the theory of banking. Limited information, however, does not necessarily imply a bias toward accepting greater risks. Indeed, if providers of funds feel that they are at an informational disadvantage, the cost of funds could be higher than in the case of perfect information. This could have the effect of making banks less willing to accept risks than they otherwise might be.

In some discussions of cycles in lending the supposed market imperfection simply seems to be a general failure of lenders to make good credit decisions. Sometimes this failure takes the form of basing decisions on the decisions of other lenders, rather than on an independent evaluation of market conditions. A model of banking markets that encompasses this view is presented by Rajan (1994). In that model, bankers are driven by a concern for their reputations, which could suffer if they fail to expand credit while others are doing so. Related to this view is the belief that competition drives lenders to ease their standards.<sup>6</sup> Hence, one might refer to this type of imperfection as the “herd mentality” problem.

The second, and perhaps more widely advanced view on the source of excessive risk tolerance by banks, is deposit insurance. It is well understood by now that federal deposit insurance has the potential of distorting banks’ attitudes toward credit risk, and indeed there is an extensive literature on this subject.<sup>7</sup>

There is nothing inherently cyclical in the distortion caused by deposit insurance. In fact, by placing a limit on the losses a bank can incur, deposit insurance is likely to have its greatest effect on incentives when the overall financial condition of banks is weak. This seems to run counter to the view of a lending cycle in which banks overextend in good times, making themselves vulnerable to adverse shocks. It is possible that the interaction of deposit insurance and the behavior of bank regulators could produce the type of cyclical behavior described. Such behavior could arise if the scrutiny of and restrictions on bank lending applied by examiners and regulators varied, after the fact, with observed performance of banks. While times are good, regulators might impose little interference, intervening only after significant losses have been incurred.

---

<sup>6</sup> This view is expressed in the discussion described by Randall (1994).

<sup>7</sup> A notable statement of the problem is given by Kareken (1983).

Some have argued that the “prompt corrective action” requirements for regulators in the Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA) embody just such a procyclical effect. On the other hand, recent, pre-FDICIA experience might be better characterized by a different pattern of regulatory response to changes in banks’ conditions. In particular, it is possible that regulatory forbearance during the 1980s resulted in too few restrictions on the activities of troubled banks.

Much of the research and writing on cycles in bank lending activity often has been driven by the observation that expansions in credit-risk exposure tend to be concentrated in a particular type of lending; lending for commercial real estate development in New England in the 1980s is a typical example.<sup>8</sup> From the point of view of an individual bank, loan concentration implies a loan portfolio that is, perhaps, not as well diversified as it might be. In an environment characterized by deposit insurance and bank regulation, concentrations could present a problem to regulatory agencies. The description of cycles driven by concentrated expansions of lending, however, is very similar to the more general description mentioned above; lenders take on excessive risks as the market is swept up in a lending euphoria that skews individuals’ evaluations of credit quality.

Each of the various views on market or regulatory failures driving cycles in lending behavior embodies some theory of the behavior of banks and financial markets. While each may have the ability to explain some aspect of observed behavior, the question remains: Compared to what? The following sections explore the implications of a benchmark model, in the absence of market imperfections or government intervention.

## 2. A MODEL OF A CREDIT MARKET

An explicit model of equilibrium in a market for loans is useful for interpreting observed lending patterns. The model presented below is one without many of the transaction costs and other frictions that are often thought to be important to banking and credit markets. While such frictions are probably important for an explanation of the institutional structure of these markets, they are not necessarily essential to every aspect of observed market behavior. Exploration of an “ideal” or frictionless model will help to uncover what aspects of observed behavior result from market frictions and what aspects arise simply from the competitive allocation of credit among heterogeneous borrowers.

First, it is useful to think of the activity in the model economy as taking place over two time periods. These two time periods might be thought of as a component of a more explicitly dynamic model, in which aggregate

---

<sup>8</sup> Randall (1994) summarizes the proceedings of a conference on this subject.

conditions evolve over time. A dynamic model would, ultimately, be better suited to closely matching the notion of cycles in lending standards. In the simple environment considered here, the forces that might drive cycles show up clearly in the comparative statics of the two-period version of the model. Since, in the model, lending standards arise as a characteristic of equilibrium, one can examine directly how changes in exogenous (to the decisions of market participants) market conditions change the equilibrium standards.

The demand side in the model economy's credit market is composed of a large number,  $N$ , of potential borrowers. Suppose that these are all business borrowers. Each borrower's business is particularly simple. A business requires a fixed amount of resources in the first period and, if successful, produces a fixed amount of output in the second period. If unsuccessful, the business produces nothing. Both input and output should be thought of as measured in monetary units, and, for simplicity, it is useful to set the value of the fixed amount of inputs required equal to one. The output of a successful business enterprise is denoted by  $Y > 1$ . This output, or revenue, should be thought of as being net of the opportunity cost of the business owner's time.

Borrowers' businesses vary in their likelihood of success. In general, probability of success might be thought of as depending on an array of borrower characteristics such as education, past business experience, and the nature of the product or service being produced. A detailed list of characteristics is beyond the scope of the model. Suppose that all of the relevant information about a borrower can be reduced to a single summary statistic, or "score," that can be expressed as a probability. Hence, a borrower's type is  $\phi$ ,  $0 \leq \phi \leq 1$ . An important market characteristic, then, is the distribution of individual borrower types, represented by a cumulative distribution function  $F(\phi)$ . That is,  $F(\phi)$  denotes the fraction of the population with type no greater than  $\phi$ . This fraction increases with  $\phi$ . Since  $\phi$  is a probability,  $F(0) = 0$  and  $F(1) = 1$ . This distribution function has an associated density, denoted by  $f(\phi)$  where  $f(\phi) \equiv F'(\phi)$ .

In addition to their individual types, businesses' prospects may depend on aggregate conditions. One way of introducing aggregate conditions into the model is to assume that the economy is subject to an aggregate technology shock that affects the output of successful firms. In this regard,  $Y$  should be thought of as a random variable, the realization of which is not known until firms observe their productive outcomes. Accordingly, a firm's output is the product of two random variables, its own success or failure and the aggregate shock to technology. The expected output for a firm of type  $\phi$  is  $\phi EY$ , where  $EY$  is the expected value of  $Y$ .

An important ingredient of any model of resource allocation among heterogeneous users is its information structure. The ability of market mechanisms to assign resources to their most productive users can depend on whether

individual productive capabilities are public or private information.<sup>9</sup> Similarly, if it is difficult for outsiders to observe a business's productive outcome, then the business may not be able to commit to payments contingent on that outcome. This limit on payment schemes can, in turn, limit the opportunities for gains from trade. In what follows, these complications are assumed away. This simplifying assumption is not based on a belief that informational frictions are not an important characteristic of financial market transactions. Indeed, such frictions are probably essential for understanding why financial institutions and contracts look the way they do. The assumption of perfect information allows the model to focus more directly on the implications of diversity among borrowers for the market allocation of credit.

Suppose that business owners have no funds of their own. Since there is an interval of time between the employment of inputs and the realization of output, providing a business with the funds to acquire inputs necessarily involves an extension of credit. If there are limited funds available for acquiring resources, or if there are alternative uses of funds that provide sufficient returns, then some businesses will operate and others will not. The credit market in this model economy allocates savers' funds to ultimate business borrowers and thereby determines which firms operate.

Business borrowers are assumed to be risk-neutral, caring only about the expected value of their profits. A business is willing and able to borrow funds if, after paying for the loan, it expects to cover at least the opportunity cost of its owner-manager's time. Since the resolution of uncertainty occurs between the extension and the repayment of the loan, the measure of the cost of credit relevant to the borrower's decision is the expected payment. This expected payment, or expected return from the lender's point of view, plays the role of the price in this market. That is, in order to attract funds, a borrower must be able to offer payments that have an expected value equal to the market return, denoted  $r$ .

Given a market return,  $r$ , any borrower whose expected output,  $\phi EY$ , is at least equal to  $r$  will be willing and able to profitably take a loan. Any such borrower can fashion a feasible repayment schedule that yields an expected return of  $r$  to lenders. Repayment schedules, here, are particularly simple. Since an unsuccessful business has no proceeds, no payment is made in that event. A successful business can make a repayment up to its realized output  $Y$ . Since  $Y$  is a random variable (common to all borrowers) that is realized before repayments are made, the repayment by a successful borrower can be contingent on  $Y$  as well as on the borrower's type,  $\phi$ . Hence, the repayment made by a successful borrower will be denoted by  $\rho(\phi, Y, r)$ , where  $r$  is added as an argument to indicate dependence on the required expected return.

---

<sup>9</sup> See Lacker and Weinberg (1993) for a discussion of private information in a model very similar to that in the present discussion.

The expected repayment from a type  $\phi$  borrower is  $\phi E_Y[\rho(\phi, Y, r)]$ , where the notation  $E_Y$  indicates expectation with respect to the aggregate random variable  $Y$ . An “acceptable” payment schedule is one that, given the borrower’s type, meets the market expected-return requirement. That is, to be acceptable, a schedule must satisfy

$$\phi E_Y[\rho(\phi, Y, r)] \geq r. \quad (1)$$

Recall that each loan is assumed to be one unit of funds. Therefore, given the required expected return,  $r$ , the demand for funds is simply given by the portion of the population of borrowers that can structure a payment schedule that yields an expected payment of at least  $r$ . The most a successful borrower can pay is the entire realized output,  $Y$ . A borrower who agrees to pay this maximum for all realizations of  $Y$  will be just indifferent between borrowing and remaining idle. The lowest-type borrower for whom obtaining a loan that the market finds acceptable will be worthwhile is the borrower for whom  $\phi EY = r$ , or  $\phi = r/EY$ . Hence, the demand for funds is the number of borrowers with  $\phi \geq r/EY$ . Accordingly, demand can be expressed as a function of the required return and the expected value of the aggregate technology shock:

$$D(r, EY) = [1 - F(r/EY)]N. \quad (2)$$

The right-hand side of this equation is simply the size of the population,  $N$ , times the fraction of borrowers with probabilities of success above the cut-off value,  $r/EY$ . Notice that this function has the usual property of a demand function: demand is decreasing in the cost of funds,  $r$ . In addition, all businesses are willing to borrow when  $r = 0$ , and none will be willing when  $r = EY$ .

Given a cost of funds, the cut-off value of the individual probability of success resembles a credit standard. This probability of success is assumed to be a function of the observable characteristics of the borrower. Hence, setting a minimum level for  $\phi$  amounts to establishing a standard based on borrowers’ characteristics.

In addition to borrowers, the economy is populated by other individuals who are endowed with funds in the first period, but may seek to save some of those funds for consumption purposes in the second period. Depending on the existence of alternatives, some or all of these funds may be saved in the form of loans to the business borrowers described above.

The behavior of savers will not be as carefully described as that of borrowers. For purposes at hand, it is sufficient to specify a function,  $S(r)$ , that determines the aggregate savings available to business borrowers. This function could take a variety of forms. One possibility is that  $S(r)$  is a constant value (perfectly inelastic). This would occur if loans to businesses constituted the only outlet for savings and if savers’ preferences were such that their desired

savings were independent of the rate of return.<sup>10</sup> Alternatively, if the business borrowers represent only a small portion of the economy's demanders of funds, then the supply of funds to this small sector can be treated as perfectly elastic; as much funding as is demanded will be forthcoming so long as borrowers are able to pay an expected return at least as great as that available elsewhere in the economy.

A last possibility for the behavior of savers arises if, for instance, they use first-period funds for both first- and second-period consumption, and if the only means of saving is through loans to businesses. In this case, savings increase with  $r$ . The key maintained assumptions about  $S(r)$  are that  $S(0) < N$  and  $S(EY) > 0$ . The first of these two assumptions assures that if credit is free, demand exceeds supply. The second implies that there are at least some rates of return for which supply exceeds demand.

In addition to the expected return, savings may depend on other aggregate conditions. For instance, one might imagine variation in savers' first-period endowment. This endowment might be, in part, the result of an aggregate shock in the first period that may, in turn, affect people's expectations about the second-period technology shock. Hence, one can imagine shifts in aggregate conditions that cause shifts in both the supply of funds by savers and the demand of borrowers.

To this point, there has been no mention of the role of intermediaries in the credit market. The model's specification is such that intermediaries are not necessary to allocate effectively the economy's resources among borrowers. In fact, borrowers could sell securities, offering prorated shares of the payment schedules,  $\rho(\phi, Y, r)$ . In equilibrium, a saver might own shares of the securities issued by a variety of borrowers. The ability of borrowers to contract directly with the savers arises from the absence of informational frictions in the model.

Although intermediaries are not necessary, one can imagine credit in this economy flowing through institutions that take the funds of savers, in exchange for some promised return, and distribute those funds to borrowers. Borrowers, then, make the payments  $\rho(\phi, Y, r)$  to the intermediaries, who use those payments to pay the savers. If the activity of intermediating is costless and if there is free entry into this activity, then, in equilibrium, intermediaries will have zero profits and the allocation of resources will be the same as in the case of the direct securities market. While the discussion that follows will adopt this interpretation of an intermediated credit market, note that the model applies more generally to the competitive allocation of credit, with or without intermediation.

It is assumed that all market participants—borrowers, savers, and intermediaries—take the expected return,  $r$ , as given in making their economic

---

<sup>10</sup> For instance, if savers care only about consumption in the second period, then they will save all of their resources, regardless of the rate of return.

decisions. Under the intermediary interpretation, intermediaries attract total savings of  $S(r)$ , accept loan applications from borrowers, and adopt a lending standard. A lending standard, here, is simply a rule stating the minimum acceptable probability of success. A type  $\phi$  borrower takes a loan from the intermediary that offers the lowest expected repayment,  $E_Y[\rho(\phi, Y, r)]$ , among those that are willing to accept a type  $\phi$  credit risk.

### 3. EQUILIBRIUM

The equilibrium expected return equates savers' supply of funds to the demand of the borrowers. Hence, equilibrium can be represented by a standard supply and demand diagram, as in Figure 2.<sup>11</sup> The market-clearing return, denoted  $r^*$ , in turn determines the minimum acceptable probability of success. This lending standard,  $\hat{\phi}$ , satisfies

$$\hat{\phi}EY = r^*. \quad (3)$$

Given this cut-off, the aggregate amount of credit extended is  $N[1 - F(r^*/EY)]$ . Competition assures that each loan makes zero expected profits for the lender. This zero-profit condition implies that repayment schedules for borrowers of type  $\phi$  are such that  $\phi E_Y[\rho(\phi, Y, r^*)] = r^*$ . Note that all loans are risky in the sense that a borrower only makes payments when successful (with probability  $\phi$ ). Hence, for all but the best borrower (type  $\phi = 1$ ) there is a positive markup between the average contractual repayment rate,  $E_Y[\rho(\phi, Y, r^*)]$ , and the cost of funds,  $r^*$ . Notice, also, that this markup gets larger as the probability of success,  $\phi$ , gets smaller.

All lending brings with it an expectation of losses. A loss will be said to occur when no payment is made. This amounts to defining default as arising only from the borrower-specific failure to produce, not from variations in payments arising from changes in the aggregate shock. That is, the loss on a failed borrower is equal to the payment that would have been expected had that borrower succeeded. For a type  $\phi$  borrower, then, the expected loss to the lender is weighted by the probability of failure:

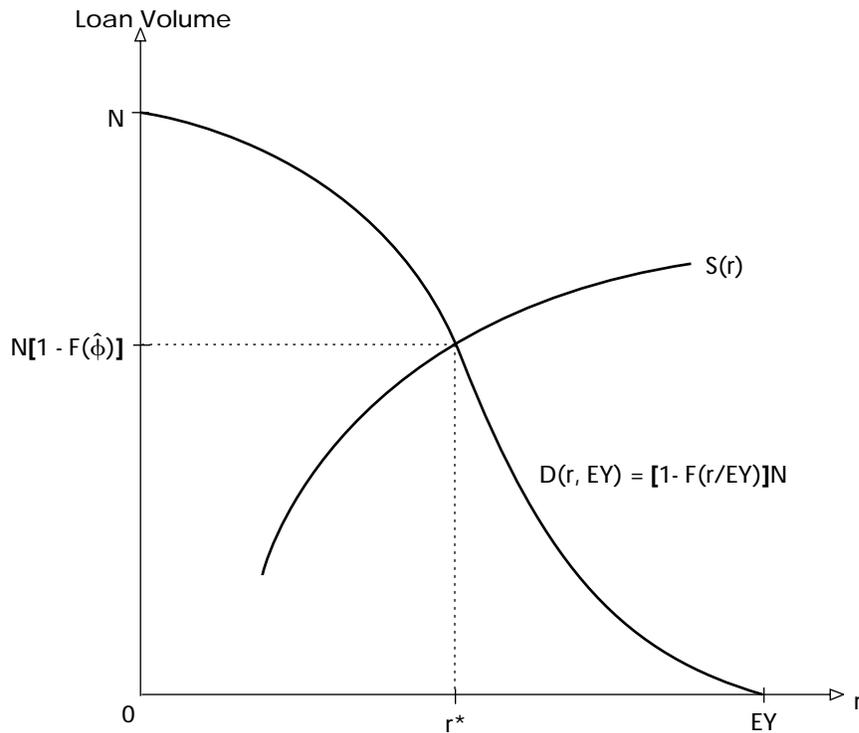
$$L(\phi, r^*) \equiv (1 - \phi)E_Y[\rho(\phi, Y, r^*)] = \frac{(1 - \phi)}{\phi}r^*. \quad (4)$$

Note that  $L(\phi)$  is decreasing in  $\phi$ , as one would expect.

Equation (4) gives a narrow specification of losses. A loan to a given borrower suffers a loss only if that borrower fails to produce a positive output. A broader specification of losses might include shortfalls of payments caused by bad realizations of the aggregate shock. For instance, one might define a loss as occurring any time the realized payment is less than  $E_Y[\rho(\phi, Y, r^*)]$ .

---

<sup>11</sup> Note, however, that in Figure 2, price is on the horizontal axis.

**Figure 2 Supply and Demand in the Credit Market**

This broader specification would complicate the analysis without changing the essential fact that expected losses rise as the probability of success falls.

The narrower specification of losses,  $L(\phi)$ , treats as losses only those shortfalls resulting from borrower-specific performance. With the narrow specification, changes in the equilibrium lending standard will affect expected aggregate losses primarily through the change in the riskiness of loans made. Since this interaction is the intended focus of this article, the narrow specification is sufficient. Expected aggregate losses, per loan made, are given by

$$\bar{L}(r^*) \equiv \frac{1}{[1 - F(r^*/EY)]} \int_{r^*/EY}^1 L(\phi, r^*) f(\phi) d\phi. \quad (5)$$

The right-hand side of equation (5) is the expected value of the function  $L(\phi, r^*)$ , given that  $\phi$  is greater than the threshold value  $\hat{\phi}$ . For each type,  $L(\phi, r^*)f(\phi)$  gives the expected losses weighted by that type's weight in the

population's distribution of types. The integral, then, can be viewed as adding up these weighted expected losses across all borrowers that receive loans. Dividing by the number of borrowers receiving credit gives average losses.

The competitive equilibrium described above achieves an efficient allocation of funds.<sup>12</sup> Given the expected technology shock, the market extends credit to businesses from the top down, until the supply of funds has been exhausted. In other words, it is impossible to find two businesses, one funded and one unfunded, such that the unfunded firm has a higher probability of success. There is no "credit rationing" in the sense that the term is often used. All borrowers who are willing and able to pay the required return (in expected value) receive loans.

#### 4. COMPARATIVE STATICS

As presented above, the key exogenous variable in the model is  $EY$ , the expected value of the aggregate shock. Accordingly, all of the endogenous variables determined in equilibrium are functions of  $EY$ . Expectations, of course, are not truly exogenous variables. Rather, economic decisionmakers form expectations by observing current conditions and making assumptions about the relationship between current conditions and future conditions. Recall that the model's description specifies a two-period time horizon. The market for funds allocates credit in the first period, based on expectations of conditions in the second period. The expectation  $EY$ , then, may be a function of some condition observed in the first period. Since that condition would be exogenous to the decisions made by borrowers and lenders, including the additional variable would contribute little to the current analysis. This section, therefore, simply treats  $EY$  as an exogenous variable and examines how the endogenous variables respond to changes in  $EY$ .<sup>13</sup>

In terms of Figure 2, an increase in  $EY$  brings about an increase in demand, as represented by an outward shift of the demand curve. For any required return, the set of borrowers who can profitably meet that requirement grows as  $EY$  rises. Specifically, the marginal borrower is that for whom  $\phi EY = r$ ; this is the borrower who, if promising payments with expected value  $r$ , has expected earnings net of payments just equal to zero. An increase in  $EY$  lowers the marginal type. Hence, as expected aggregate conditions improve, riskier borrowers become acceptable for any given required return.

---

<sup>12</sup> A deeper discussion of the efficiency of equilibrium would have to include a more explicit treatment of the economic decisions made by savers; as long as those decisions are such that an equilibrium exists and there are no externalities, equilibrium will be efficient (Pareto-optimal).

<sup>13</sup> More precisely, the distribution of  $Y$  is assumed to be subject to exogenous changes. Market participants are aware of the true distribution when they make their first-period decisions and, hence, form expectations rationally.

Of course, since more businesses are able to borrow profitably at any given return, market forces will cause the equilibrium required return,  $r^*$ , to rise if, as in Figure 2, the supply of funds depends on the return to saving. The overall change in the lending standard,  $\hat{\phi}$ , is

$$\frac{d\hat{\phi}}{dEY} = \frac{d}{dEY} \left[ \frac{r^*}{EY} \right] = \frac{dr^*}{dEY} \frac{1}{EY} - \frac{r^*}{(EY)^2}. \quad (6)$$

The right-hand side of this expression includes a direct and an indirect effect. The direct effect is given by the second term; for a fixed required return,  $r^*$ ,  $\hat{\phi}$  falls as  $EY$  rises. The first term is the indirect effect that comes from the change in  $r^*$ . If the increase in demand causes  $r^*$  to rise, then the rising cost of funds will dampen the decline in the lending standard. If the supply of funds is not perfectly inelastic, the equilibrium lending standard falls, and credit is extended to a wider range of borrowers as  $EY$  rises. If supply is perfectly elastic (for instance, if this credit market is a small portion of broader financial markets), then the first term on the right side of equation (6) is zero and only the direct effect on  $\hat{\phi}$  is present.

Since all loans have some risk of loss, the expansion of lending when expected conditions improve causes expected losses on loans to rise. Expected losses, however, also rise as a percent of total loans. That is,  $\bar{L}(r^*)$  rises as  $EY$  rises. This can occur for two reasons. First, as seen in equation (4), the expected loss on a loan to any given borrower rises if  $r^*$  rises. Second, expansion of lending comes at the bottom of the range of acceptable credit risks. Hence, added borrowers bring greater-than-average expected losses.

Closely related to loan losses is the average markup of repayments by businesses that do not fail over the cost of funds,  $r$ . This markup for a type  $\phi$  borrower is  $E[\rho(\phi, Y, r^*)]/r^*$ . Since, in equilibrium,  $\phi E[\rho(\phi, Y, r^*)] = r^*$  for all acceptable borrowers, the markup for a type  $\phi$  borrower is simply  $1/\phi$ . Therefore, for any given borrower, the average markup of repayment (conditional on successful production) over the lender's opportunity cost of funds is independent of aggregate conditions. Accordingly, the average markup across borrowers rises as the lending standard,  $\hat{\tau}$ , falls with a rising  $EY$ . Additional borrowers have a higher risk of failure and, therefore, must make greater average payments, conditional on success.

## 5. DISCUSSION

Can the model presented above yield insights into cycles in lending standards? Recall that in a fully dynamic model, aggregate conditions would evolve according to a process that would allow market participants to form the expectation,  $EY$ , using observations of current and, perhaps, past conditions. More precisely,  $EY$  would be a function of at least the most recently observed productive (or profit) experience of successful business firms.

If there is persistence in the aggregate technology shock that determines  $Y$ , then expected output is high when current output is high. Hence, loan growth and falling standards occur during good economic times. Since a lower standard means the extension of credit to borrowers with a higher risk of default, it seems as though movements in the lending standard can have the effect of widening the swings in the economy, especially in downturns. If, when expectations were high, a downturn occurs because of a low realization of  $Y$ , output falls not just because of the low aggregate shock, but also because of the higher number of failures of marginal borrowers. On the other hand, if the economy experiences a string of good aggregate performance, it is possible to have a period of rising losses while total loans and economic activity continue to grow. This possibility might be reminiscent of much of the expansion of the U.S. economy in the 1980s.

In the model outlined above, changes in lending and lending standards come entirely from the demand side of the market. It may be reasonable to suppose that changes in aggregate conditions also affect the supply of funds. In particular, if a high current value of the output of successful firms means that savers' total resources are correspondingly high, then the supply of funds may expand together with the demand. This would tend to reinforce the effect of aggregate shocks on the expansion of credit while making the effect on expected returns ambiguous.

Popular discussions of changes in lending standards often implicitly treat these changes as originating from the supply side. When standards rise, for instance, there are often references and anecdotes concerning borrowers who are being rationed out of the market. Such references seem to suggest that demand has not changed, but that borrowers who are willing and able to take loans under prevailing market conditions are being denied credit. Casual evidence, however, is difficult to interpret, particularly when credit flows from suppliers to demanders through intermediaries. One can interpret the model presented above as one of an intermediated market. Under this interpretation, for the market to achieve its equilibrium allocation it is not necessary for borrowers to be aware of aggregate conditions. If every borrower submits applications to one or more banks, then the banks, knowing the aggregate condition, will make credit decisions according to the equilibrium described above. Hence, the perceptions of borrowers who get screened out as standards rise (as  $EY$  falls) can be misleading.

Another aspect of popular discussions of variations in lending standards is the degree of competition among intermediaries. As in late 1994 and early 1995, observers often point to rising competition as a cause of weakening bank standards. The model presented in this article, as a model of perfect competition, cannot capture rising competition. Presumably, if competitiveness is to vary, competition must be imperfect. Imperfect competition might arise from technological or legal barriers to entry. Procyclical competitiveness might arise

if lenders in an imperfectly competitive market find it more difficult to refrain from aggressive competition during expansions. This sort of varying competition would reinforce the countercyclical lending standards of this article's model.

The model clearly does not imply that a goal of supervision and regulation of financial intermediaries should be to encourage "smooth" lending standards that do not vary with aggregate conditions in the economy. Rather, the model implies that in times of strong and improving economic activity, businesses (and households) whose individual characteristics make them look risky may become acceptable borrowers. A policy aimed at smoothing credit standards would limit the ability of participants in the economy to adapt to changing economic conditions and to take productive risks when those risks are warranted.

Of course, the existence of regulatory oversight of bank activities seems to presume a bias toward excessive risk taking on the part of banks. The source of that presumption is deposit insurance. It is not clear why the adverse incentive created by deposit insurance should be greatest during periods of strong economic performance and expanding credit. If anything, deposit insurance should have its greatest effect on incentives at the other end of the cycle in credit market conditions. The argument in the literature on this topic is that insured lenders may have an incentive to "bet the bank" when they are in weak financial condition. This suggests that regulatory scrutiny of bank lending behavior should be greatest during a period of low returns (profitability) for banks.

## 6. CONCLUSION

Lending standards are usually thought of in terms of requirements placed on the characteristics of individual loans and borrowers. A central point of this article is that there is a natural tendency for standards to vary inversely with the level of activity in the credit markets. There is, of course, a sense in which lending standards do not vary at all in this article's model. The marginal borrower is always that borrower who can just afford repayment terms that just cover (in expected value) the opportunity cost of savers' funds. That the quality of the marginal borrower varies is the result of the interaction of individual and aggregate conditions in determining the payoff to extending credit. As (expected) aggregate conditions improve, a borrower who did not look creditworthy yesterday may now be deserving of credit.

The frictionless model examined herein is missing many of the ingredients that are perhaps important to the character of modern credit markets and institutions. Even with further complications, however, the fundamental role of the credit market would be the same. Given perceptions of the general condition of the economy, the credit market sets a threshold for acceptable risks. If the market functions well, subject to whatever imperfections may be present in

the economic environment, then an improvement in participants' perceptions of aggregate conditions will lower the threshold.

The concerns about credit quality that are often expressed in times of expanding credit are typically driven by very current news. A more "global" perspective would recognize the interaction of individual and aggregate conditions. The model examined in this article suggests that a policy goal of credit standards that are constant over time is not only unwarranted, but could be counterproductive.

---

---

## REFERENCES

- Bernanke, Ben S., and Cara S. Lown. "The Credit Crunch," *Brookings Papers on Economic Activity*, 2:1991, pp. 205–47.
- Connor, John. "Comptroller Unveils Credit Committee, Citing Slippage in Banks' Loan Standards," *The Wall Street Journal*, April 10, 1995, p. 2.
- Cunningham, Donald F., and John T. Rose. "Bank Risk Taking and Changes in Financial Condition of Banks' Small and Midsized Commercial Customers, 1978–1988 and 1988–1991," *Journal of Financial Services Research*, vol. 8 (December 1994), pp. 301–10.
- Darin, Robert M., and John R. Walter. "Were Bank Examiners Too Strict with New England and California Banks?" Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 80 (Fall 1994), pp. 25–47.
- Dunaief, Daniel. "Bankers Agree: Credit Standards Have Been Sliding," *American Banker*, April 11, 1995, pp. 1 and 24.
- Kareken, John H. "Deposit Insurance Reform or Deregulation Is the Cart, Not the Horse," Federal Reserve Bank of Minneapolis *Quarterly Review*, vol. 7 (Spring 1983), pp. 1–9.
- Lacker, Jeffrey M., and John A. Weinberg. "Coalition Proof Equilibrium in a Private Information Credit Economy," *Economic Theory*, vol. 3 (Spring 1993), pp. 279–96.
- Mathews, Gordon. "Worries Grow that Downturn in Economy Could Punish Banks for Easing Credit," *American Banker*, March 27, 1995, pp. 1 and 28.
- Mueller, Henry P., and Leif Olson. "Credit and the Business Cycle," in Herbert V. Prochnow, ed., *Bank Credit*. New York: Harper and Row, 1981, pp. 78–91.
- Owens, Raymond E., and Stacey L. Schreft. "Identifying Credit Crunches," *Contemporary Economic Policy*, vol. 13 (April 1995), pp. 63–76.

- Rajan, Raghuram G. "Why Bank Credit Policies Fluctuate: A Theory and Some Evidence," *Quarterly Journal of Economics*, vol. 109 (May 1994), pp. 399–441.
- Randall, Richard E. "Safeguarding the Banking System in an Environment of Financial Cycles: An Overview," *New England Economic Review*, March/April 1994, pp. 1–13.
- Schreft, Stacey L., and Raymond E. Owens. "Survey Evidence of Tighter Credit Conditions: What Does it Mean?" Federal Reserve Bank of Richmond *Economic Review*, vol. 77 (March/April 1991), pp. 29–34.
- Stevenson, Bruce G. "Research Report . . . Capital Flows and Loan Losses in Commercial Banking," *Journal of Commercial Lending*, vol. 77 (September 1994), pp. 18–26.

# Errors in Variables and Lending Discrimination

---

Jed L. DeVaro and Jeffrey M. Lacker

**D**o banks discriminate against minority loan applicants? One approach to answering this question is to estimate a model of bank lending decisions in which the probability of being denied a loan is a function of a set of creditworthiness variables and a dummy variable for the applicant's race ( $z = 1$  for minorities,  $z = 0$  for whites). A positive coefficient on the race dummy is taken as evidence that minority applicants are less likely to be granted loans than white applicants with similar qualifications. This approach is employed in many empirical studies of lending discrimination (Schill and Wachter 1994; Munnell et al. 1992), in U.S. Department of Justice lending discrimination suits (Seiberg 1994), and in regulatory examination procedures (Bauer and Cromwell 1994; Cummins 1994).

One weakness of this approach is that an estimate of the discrimination coefficient may be biased when measures of creditworthiness are fallible. In such situations, distinguishing racial discrimination from unmeasured racial disparities in creditworthiness can be difficult. If true creditworthiness is lower on average for minority applicants, the model may indicate that race adversely affects the probability of denial, even if race plays no direct causal role.

There are good reasons to believe that measures of creditworthiness are fallible. First, regulatory field examiners report difficulty finding matched pairs of loan files to corroborate discrimination identified by regression models. An applicant's file often yields a picture of creditworthiness different from the one given by model variables. Second, including more borrower financial characteristics generally reduces discrimination estimates, sometimes to zero (Schill and Wachter 1994). Third, studies of default data find that minority borrowers are more likely than white borrowers to default, even after controlling for income,

---

■ The authors thank Mary Finn and Peter Ireland for helpful conversations during this research and Tom Humphrey, Tony Kuprianov, and Stacey Schreft for helpful comments on an earlier draft. The authors are solely responsible for the contents of this article. The views expressed do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

wealth, and other borrower characteristics related to creditworthiness (Berkovec et al. 1994). This finding suggests that there are race-related discrepancies between the true determinants of creditworthiness and the measures available to econometricians.

Our objective is to develop a method for assessing the sensitivity of lending discrimination estimates to measurement error. In particular, we study the classical errors-in-variables model, in which the components of a vector  $x$  of observed measures of creditworthiness are, one for one, fallible measures of those in a vector of true qualifications  $x^*$ .<sup>1</sup> The implications of errors in variables in the standard linear regression model are well known (Klepper and Leamer 1984; Goldberger 1984).<sup>2</sup> We briefly review these implications in Section 1. Models of lending discrimination generally specify a nonlinear regression model, such as the logit model, because the dependent variable is dichotomous ( $y = 1$  if the loan application is denied;  $y = 0$  if it is accepted). In this article we extend the results for the linear case to cover the nonlinear logit regression model widely used in lending discrimination studies.

Linear errors-in-variables models are underidentified because variation in true qualifications cannot be distinguished from error variance. Assuming that the errors are normally distributed with known parameters, however, the linear model is just-identified, allowing estimation of model parameters depending on the assumed error-variance parameters. Assuming zero error variance yields the standard linear regression model as a special case. By estimating under a range of error-variance assumptions, one can trace out the potential effect of measurement error on model parameter estimates. Note that since the error-variance assumptions make the model just-identified, no one assumption about the error-variance parameters is more likely than any other; that is, estimates of model parameters under alternative error-variance assumptions are all equally consistent with the data. Also note that in the case of normally distributed regressors in the linear model, parameter estimates for alternative error-variance

---

<sup>1</sup> The classical errors-in-variables model is not the only one in which observed variables, taken together, are fallible measures of true creditworthiness. Alternatives include “multiple-indicator” models in which observed variables are fallible measures of a single index of creditworthiness, and “omitted-variable” models in which some determinants of creditworthiness are unobservable. All are alike in that a component of the true model is unobserved by the econometrician; thus, all are latent-variable models. Because errors in variables is one of the simplest and most widely studied models of fallible regressors, it is a useful starting point in examining fallibility in empirical models of lending discrimination.

<sup>2</sup> Interest in the errors-in-variables problem has surged since 1970. As Hausman and colleagues (1995) stated, “During the formative period of econometrics in the 1930’s, considerable attention was given to the errors-in-variable[s] problem. However, with the subsequent emphasis on aggregate time series research, the errors-in-variables problem decreased in importance in most econometric research. In the past decade as econometric research on micro data has increased dramatically, the errors-in-variables problem has once again moved to the forefront of econometric research” (p. 206).

assumptions can be obtained through an algebraic correction to the ordinary least squares estimates.

In Section 2 we examine the logit model under errors in variables and show how estimators depend on assumptions about error variance. Adjusting estimators for error variance is no longer an algebraic correction as it is in the linear setup; the model must be reestimated for each error-variance assumption. For the case in which the independent variables are continuous-valued, we show how to estimate the logit model under various assumptions about error variance. Because of the nonlinearity, the logit model is in some cases identified without error-variance assumptions. In practice, however, the logit model is quite close to underidentified, and little information can be obtained from the data about error-variance parameters. Therefore, we advocate estimating models under a range of error-variance assumptions to check the sensitivity of estimates to measurement error.

In Section 3 we demonstrate our method using artificial data. We show how estimates of a discrimination parameter can be biased when a relatively modest amount of measurement error is present. The magnitude of the bias depends on the model's fundamental parameters. By estimating the model under different assumptions about measurement error variance, we can gauge the sensitivity of the estimators to errors in variables. Section 4 concludes and offers directions for further research.

Bauer and Cromwell (1994) have also studied the properties of logit regression models of lending discrimination, focusing on the small-sample properties of a misspecified model using simulated data. They found that tests for lending discrimination were sensitive to sample size. Our work focuses on the effect of errors in variables on the large-sample properties of otherwise correctly specified logit models of lending discrimination.

## 1. ERRORS IN VARIABLES

The implications of errors in variables are easiest to see in a linear setup such as the following simple model of salary discrimination.<sup>3</sup> Suppose that an earnings variable ( $y$ ) is determined according to the following equations:

$$y = \beta x^* + \alpha z + v, \tag{1a}$$

$$x^* = x_0 + \mu z + u, \tag{1b}$$

---

<sup>3</sup>The exposition in this section is based on Goldberger (1984). This model of salary discrimination has a close parallel in the permanent income theory. Friedman (1957) discusses how racial differences in unobserved permanent income (the counterpart of qualifications in the salary model and creditworthiness in the lending model) bias estimates of racial differences in the consumption function intercept.

$$x = x^* + e, \quad (1c)$$

where the scalar  $x^*$  = true qualification,  $x$  = measured qualification, and  $z$  is a race dummy ( $z = 1$  for minorities,  $z = 0$  for whites). We take  $v$ ,  $u$ , and  $e$  to be mutually independent random variables with zero means and variances  $\sigma_v^2$ ,  $\sigma_u^2$ , and  $\sigma_e^2$ , all independent of  $z$ . The earnings variable in (1a) is a stochastic function of the true qualifications and race. The parameter  $\alpha$  represents the independent effect of race on salary, and  $\alpha < 0$  represents discrimination against minorities. If better-qualified applicants obtain higher salaries, then  $\beta > 0$ . In (1b) qualification is allowed to be correlated with race; the expectation of  $x^*$  is  $x_0$  for whites and  $x_0 + \mu$  for minorities. The empirically relevant case has  $\mu < 0$ . Observed qualification in (1c) is contaminated by measurement error  $e$ .

Consider a regression of  $y$  on the observed variables  $x$  and  $z$ . This estimates

$$E[y \mid x, z] = bx + az.$$

Since the variances and covariances are the same for both white and minority applicants, we can use conditional covariances to calculate the regression slopes. We focus on relationships in a population and thus ignore sampling variability. The least squares estimators are

$$b = \text{cov}(x, y \mid z) / \text{var}(x \mid z) = \text{cov}(x^*, y \mid z) / \text{var}(x \mid z) = (1 - \delta)\beta$$

and

$$\begin{aligned} a &= E[y \mid z = 1] - E[y \mid z = 0] - b\{E[x \mid z = 1] - E[x \mid z = 0]\} \\ &= \alpha + \beta\mu - b\mu \\ &= \alpha + \delta\beta\mu, \end{aligned}$$

where

$$\delta \equiv \sigma_e^2 / (\sigma_u^2 + \sigma_e^2).$$

When there is measurement error ( $\sigma_e^2 > 0$ ), the regression estimator of  $\beta$  is biased toward zero. To see why, substitute for  $x^*$  in (1a) using (1c) to obtain  $y = \beta x + \alpha z + (v - \beta e)$ . The “error”  $v - \beta e$  in the regression of  $y$  on  $x$  and  $z$  is correlated with  $x$  via (1c). Thus a key assumption of the classical linear regression model is violated, and the coefficients are no longer unbiased.

In our case ( $\beta > 0$ ,  $\mu < 0$ ), the estimator of  $\alpha$  is biased downward as well. Bias creeps in because  $z$  is informative about  $x^*$ , given  $x$ ;

$$E[x^* \mid x, z] = (1 - \delta)x + \delta(x_0 + \mu z).$$

Given observed qualification  $x$ , race can help “predict” true qualification  $x^*$ . Race can then help “explain” earnings, even in the absence of discrimination ( $\alpha = 0$ ), because race is correlated with true qualifications.

The model (1) is underidentified (Kapteyn and Wansbeek 1983). A regression of  $x$  on  $z$  recovers the nuisance parameters  $x_0$  and  $\mu$ , along with

$v(x | z) = \sigma_u^2 + \sigma_e^2$ . Other population moments provide us with  $a$  and  $b$ , but these are not sufficient to identify  $\alpha$ ,  $\beta$ , and  $\delta$ . No sample can provide us with enough information to divide  $v(x | z)$  between the variance in true qualifications  $\sigma_u^2$  and the variance in measurement error  $\sigma_e^2$ . Under the assumptions  $\beta > 0$  and  $\mu < 0$ , any value of  $\alpha > a$ , including the no-discrimination case  $\alpha = 0$ , is consistent with the data for some  $\beta$  and  $\sigma_e^2$ .

If  $\sigma_e^2$  were known independently, then we would know  $\delta = \sigma_e^2/(\sigma_u^2 + \sigma_e^2)$  and could calculate the unbiased estimators  $\hat{\alpha}$  and  $\hat{\beta}$  by correcting the ordinary least squares estimators as follows:

$$\hat{\beta} = b/(1 - \delta) \quad (2a)$$

$$\hat{\alpha} = a - \delta b\mu/(1 - \delta). \quad (2b)$$

One could use (2) to study the implications of alternative assumptions about the variance of measurement error; different values of  $\sigma_e^2$  would trace out different estimates of  $\alpha$ .

In (1) the direction of bias in  $a$  is known when the sign of  $\beta\mu$  is known. Matters are different when  $x$  is a vector of characteristics affecting qualifications. Consider a multivariate model:

$$y = \beta'x^* + \alpha z + v, \quad (3a)$$

$$x^* = x_0 + \mu z + u, \quad (3b)$$

$$x = x^* + e, \quad (3c)$$

where  $x^*$  and  $x$  are now  $k \times 1$  random vectors and  $\beta$ ,  $\mu$ , and  $x_0$  are  $k \times 1$  parameter vectors. We take  $u$  and  $e$  to be normally distributed random vectors, independent of  $v$ ,  $z$ , and each other, with zero means and covariance matrices  $\Sigma^*$  and  $D$ . The classical assumption is that measurement errors are mutually independent, so  $D$  is diagonal.

The least squares estimators are now

$$b = (\Sigma^* + D)^{-1}\Sigma^*\beta \quad (4a)$$

and

$$a = \alpha + (\beta - b)'\mu. \quad (4b)$$

The direction of bias is now uncertain, even under the usual assumption that measurement errors are independent ( $D$  is diagonal). To see why, suppose that  $k = 2$ ,  $\Sigma^*$  has  $\rho$  as the off-diagonal element, and  $\Sigma^* + D$  has ones on the diagonal (a normalization of units). Then (4b) becomes

$$a = \alpha + [(D_{11}\beta_1 - \rho D_{22}\beta_2)\mu_1 + (D_{22}\beta_2 - \rho D_{11}\beta_1)\mu_2]/(1 - \rho^2).$$

The bias in  $a$  could be positive or negative, depending on parameter values. For example, suppose only one component of  $x$  is subject to measurement

error, say,  $x_1$  ( $D_{11} > 0$  and  $D_{22} = 0$ ). By itself this would bias  $b_1$  downward, resulting in an upward bias in  $a$ . But  $b_2 = \rho\beta_1 D_{11}/(1 - \rho^2) + \beta_2$  is now biased as well, and this would induce downward bias in  $a$  if  $\rho\beta_1\mu > 0$ . The overall direction of bias is indeterminate (Rao 1973; Hashimoto and Kochin 1980). But again, if the measurement error parameters  $D$  were known, then the least squares estimators  $a$  and  $b$  could be corrected by a simple transformation of (4) (using  $\Sigma^* = \Sigma - D$ , where  $\Sigma = v(x | z)$ ). Each alternative measurement error assumption would imply a different estimator.<sup>4</sup>

## 2. ERRORS IN VARIABLES IN A LOGIT MODEL OF DISCRIMINATION

In model (3) the dependent variable is a linear function of the explanatory variables. In models of lending decisions the dependent variable is dichotomous:  $y = 1$  if the applicant is denied a loan, and  $y = 0$  if the applicant is accepted. In this case the linear formulation in (3) is unattractive (Maddala 1983). A common alternative is the logit model, shown here without errors in variables:

$$\Pr[y = 1 | x, z] = G(\beta'x + \alpha z), \quad (5a)$$

$$G(t) = \frac{1}{1 + e^{-t}}, \quad (5b)$$

where  $x$  is a vector of characteristics influencing creditworthiness. The empirically relevant case has  $\beta < 0$ , so applicants who are more creditworthy are less likely to be denied loans. A value of  $\alpha > 0$  would indicate discrimination against minorities: a minority applicant is approximately  $\alpha(1 - G)$  times more likely than an identical white applicant to be denied a loan.<sup>5</sup>

The parameters  $\alpha$  and  $\beta$  can be estimated by the method of maximum likelihood. The log likelihood function for a sample of  $n$  observations  $\{y_i, x_i, z_i, i = 1, \dots, n\}$  is

$$\log L = \sum_{i=1}^n \log \Pr(y_i, x_i, z_i) = \sum_{i=1}^n \log \Pr(y_i | x_i, z_i) + \sum_{i=1}^n \log \Pr(x_i, z_i), \quad (6)$$

where

$$\Pr(y_i | x_i, z_i) = G(\beta'x_i + \alpha z_i)^{y_i} [1 - G(\beta'x_i + \alpha z_i)]^{(1-y_i)}.$$

Estimators are found by choosing parameter values that maximize  $\log L$ . The likelihood depends on the parameters of the conditional distribution in (5) as

<sup>4</sup> Klepper and Leamer (1984) and Klepper (1988b) show how to find bounds and other diagnostics for the linear errors-in-variables model.

<sup>5</sup> The elasticity of  $G$  with respect to  $z$  is  $\alpha G'/G = \alpha(1 + e^{-t})e^{-t}/(1 + e^{-t})^2 = \alpha e^{-t}/(1 + e^{-t}) = \alpha(1 - G)$ , where  $G$  is evaluated at  $\beta'x + \alpha z$ .

well as on the “nuisance parameters” governing the unconditional distribution of  $(x, z)$ . Since the nuisance parameters appear only in the second sum in (6), while  $\alpha$  and  $\beta$  appear only in the first sum,  $\alpha$  and  $\beta$  can be estimated in this case without estimating the nuisance parameters.

Under errors in variables, (5a) is replaced with

$$\Pr[y = 1 \mid x^*, z] = G(\beta'x^* + \alpha z), \quad (7)$$

where  $x^*$  is the vector of true characteristics. The resulting log likelihood function is

$$\begin{aligned} \log L &= \sum_{i=1}^n \log \Pr(y_i, x_i, z_i) \\ &= \sum_{i=1}^n \log \int \Pr(y_i \mid x_i^*, z_i) \Pr(x_i \mid x_i^*) \Pr(x_i^*, z_i) dx_i^*. \end{aligned} \quad (8)$$

The likelihood function now depends on  $\Pr(x \mid x^*)$ , the probability that  $x$  is observed if the vector of true characteristics is  $x^*$ . Since  $x - x^*$  is the vector of measurement errors,  $\Pr(x \mid x^*)$  is the probability distribution governing the measurement error. In the linear model (3) the least squares estimators could be corrected algebraically for measurement error of known variance. In the logit model, however, there is no simple way to adjust maximum likelihood estimators for errors in variables, since the regression function is nonlinear. Instead, we must estimate  $\alpha$  and  $\beta$  for each distinct assumption about  $\Pr(x \mid x^*)$ .

Unlike the one in (6), the log likelihood function in (8) is not separable in the nuisance parameters of the distribution  $\Pr(x^*, z)$ . Even if we posit an error distribution  $\Pr(x \mid x^*)$ , estimating  $\alpha$  and  $\beta$  requires estimating the parameters of  $\Pr(x^*, z)$  as well. The estimation of these nuisance parameters will be sidestepped here by maximizing the conditional likelihood function

$$\begin{aligned} \log \tilde{L} &= \sum_{i=1}^n \log \Pr(y_i \mid x_i, z_i) \\ &= \sum_{i=1}^n \log \int \Pr(y_i \mid x_i^*, z_i) \Pr(x_i^* \mid x_i, z_i) dx_i^*. \end{aligned} \quad (9)$$

We will assume that  $\Pr(x^* \mid x, z)$ , the distribution of true characteristics conditional on observed characteristics and race, is known.

Our model is completed by adding specific assumptions about the distributions  $\Pr(x \mid x^*)$  and  $\Pr(x^* \mid z)$ , which will allow us to derive  $\Pr(x^* \mid x, z)$ . We will maintain the assumptions embodied in (3b) and (3c):

$$x^* = x_0 + \mu z + u, \quad (10a)$$

$$x = x^* + e, \quad (10b)$$

where  $\beta$ ,  $\mu$ , and  $x_0$  are  $k \times 1$  parameter vectors and where  $u$  and  $e$  are normally distributed random vectors, independent of  $v$ ,  $z$ , and each other, with zero means and covariance matrices  $\Sigma^*$  and  $D$ . Given  $x$  and  $z$ ,  $x^*$  is then normally distributed with mean vector  $m^*$  and covariance matrix  $S^*$ , where

$$m^* = D\Sigma^{-1}\mu z + (I - D\Sigma^{-1})x, \quad (11a)$$

$$S^* = (I - D\Sigma^{-1})D. \quad (11b)$$

With this result in hand, we find that, conditional on  $x$  and  $z$ , the argument of  $G$  is normally distributed with mean  $\beta' m^* + \alpha z$  and variance  $\beta' S^* \beta$ . Therefore, the likelihood in (9) can be written as

$$\Pr(y \mid x, z) = \int G(m + \sigma s)(2\pi)^{-1/2} \exp(-s^2/2) ds, \quad (12)$$

where

$$m = \beta'(I - D\Sigma^{-1})x + (\alpha + \beta'D\Sigma^{-1}\mu)z,$$

$$\sigma = [\beta'(I - D\Sigma^{-1})D\beta]^{1/2}.$$

When  $D = 0$ ,  $m$  collapses to  $\beta'x + \alpha z$  and  $\sigma = 0$ , which is the error-free model.<sup>6</sup>

Because of the nonlinearity of  $G$ , the logit model can potentially be identified without error-variance assumptions, unlike the linear model in Section 1. Thus, in principle, the error-variance parameters could be estimated rather than imposed. In practice, however, the model is so close to linear that the error-variance parameters cannot be estimated; even large samples are uninformative about  $D$ . We therefore recommend estimating the model under a range of alternative error-variance assumptions.

To summarize the procedure, first calculate least squares estimators for the parameters  $x_0$ ,  $\mu$ , and  $\Sigma$ . These parameters are treated as fixed and combined with an assumed  $D$  to obtain the distribution  $\Pr(x^* \mid x, z)$ , which is used in (12) and (9) to obtain maximum likelihood estimates of  $\alpha$  and  $\beta$ . This procedure treats the error variance  $D$  as known, just as the error-free model treats  $D$

---

<sup>6</sup> The joint normality of  $x$  and  $x^*$  given  $z$  implies that given  $x$  and  $z$ ,  $x^*$  is normal with parameters that can be derived algebraically from the parameters of  $\Pr(x \mid x^*)$  and  $\Pr(x^* \mid z)$ . Other distributional assumptions on  $x$  and  $x^*$  are far less convenient. For example, when  $x^*$  takes on discrete values, a more general approach is required to derive  $\Pr(x^* \mid x, z)$ . Given a distribution of the observables  $\Pr(x, z)$ , recover  $\Pr(x^* \mid z)$  using  $\Pr(x \mid z) = \int \Pr(x \mid x^*)\Pr(x^* \mid z)dx^*$ , and then use Bayes's rule to obtain  $\Pr(x^* \mid x, z) = \Pr(x \mid x^*)\Pr(x^* \mid z)/\Pr(x \mid z)$ . The first of these steps involves inverting a very large matrix.

as identically zero. Estimates of  $\alpha$  can then be traced out under alternative assumptions on  $D$ .<sup>7</sup>

Our procedure will misstate the uncertainty about parameter estimates, even conditioning on  $D$ . By implicitly assuming that the estimated parameters  $x_0$ ,  $\mu$ , and  $\Sigma$  are known, we are neglecting their sampling variability. These parameters appear in (12) and thus influence estimates of  $\alpha$  and  $\beta$ . Our procedure therefore misstates their sampling variability as well. When  $D = 0$ , the nuisance parameters disappear from (12), and this problem does not arise.<sup>8</sup>

### 3. EXAMPLES

In the examples in this section, we apply our procedure in a logit model of discrimination to show how the technique is capable of detecting the sensitivity of parameter estimates to errors in variables. We find it convenient to use artificially generated data sets to illustrate our results. Artificial data allow us to isolate important features of the errors-in-variables model for a wide array of cases. Observations are randomly generated under a given, true error variance, and the model is then estimated under various hypothesized error variances.

In the simplest case there is only one explanatory variable besides race ( $k = 1$ ). We assume  $\alpha = 0$ ,  $\beta = -1$ ,  $\mu = -2$ , and  $\Sigma = 1$ . (We focus on the no-discrimination case,  $\alpha = 0$ , solely for convenience.) In this case, if  $a$  is significantly different from zero, then it is also significantly greater than  $\alpha$ , and the usual t-statistic on  $a$  will also show whether  $a$  is significantly biased. The sample was assumed to be half white ( $z = 0$ ) and half minority ( $z = 1$ ). Using these values and an assumed true error variance  $D$ , we generated 10,000 random observations on  $x^*$ ,  $x$ , and  $y$  using equations (7) and (10). We then estimated the model using maximum likelihood, assuming that the true values of  $\mu$  and  $\Sigma$  were known and making an assumption about  $\bar{D}$  (not necessarily the same as  $D$ ). The results are displayed in Table 1. The sample size of 10,000 was chosen to reduce sampling variance.

For the estimates shown in Panel A of Table 1, the true variance of the measurement error is  $D = 0.1$ . This represents one-tenth of the total variance in observed  $x$ , a relatively modest amount. The first line reports estimation under the (incorrect) assumption that the error variance is zero. As expected, the estimate  $b$  is biased toward zero. Consequently,  $a$  is biased upward, toward showing discrimination, and is significant.

---

<sup>7</sup>In related work, Klepper (1988a) extended the diagnostic results of Klepper and Leamer (1984) and Klepper (1988b) to a linear regression model with dichotomous independent variables. These earlier approaches attempted to characterize the set of parameters that maximize the likelihood function. Levine (1986) extended the results of Klepper and Leamer (1984) to the probit model.

<sup>8</sup>Specifically, the hessian of the log likelihood function is then block diagonal across  $(\alpha, \beta)$  and  $(x_0, \mu, \Sigma)$ .

**Table 1 Coefficient Estimates for  
Alternative Error-Variance Assumptions,  $k = 1$**   
 $\mu = -2, \Sigma = 1, n = 10,000.$

	<i>a</i>	<i>b</i>
A. True parameters $\alpha = 0, \beta = -1$ , and $D = 0.1$ :		
Assumed $\tilde{D}$		
0.0	0.1446 (2.4380)	-0.9208 (-32.3477)
0.05	0.0482 (0.7780)	-0.9775 (-31.8322)
0.1	-0.0607 (-0.9308)	-1.0418 (-31.2626)
B. True parameters $\alpha = 0.1, \beta = -0.9$ , and $D = 0.0$ :		
Assumed $\tilde{D}$		
0.0	0.1609 (2.7101)	-0.9260 (-32.4378)
0.05	0.0640 (1.0315)	-0.9832 (-31.9159)
0.1	-0.0456 (-0.6986)	-1.0480 (-31.3393)

Notes: t-statistics are shown in parentheses beneath the coefficient estimates. For each panel, we drew a set of 10,000 random realizations for  $(y, x)$ : 5,000 with  $z = 0$  and 5,000 with  $z = 1$ . Within each panel, estimation was performed on the same data set with different assumptions about the error variance  $\tilde{D}$ .

The last two lines in Panel A show estimates assuming positive error variance. For larger values of  $\tilde{D}$ ,  $b$  is closer to one and  $a$  is closer to zero, the true value. The discrimination parameter is not significantly different from zero when estimated assuming  $D$  is 0.05 or 0.1. In this case, then, our procedure successfully detects the sensitivity of parameter estimates to errors in variables.

In Panel B we examine the case in which no measurement error is present and the true discrimination parameter is positive. The (correct) assumption of no measurement error now yields estimates that are unbiased; they differ from the true parameters only because of sampling error. Imposing the (incorrect) assumption of positive measurement error variance “undoes” a nonexistent bias, resulting in  $a$  near zero and a larger negative  $b$ .

Table 2 shows how the magnitude of the bias varies with the correlation between components of  $x$  when  $k = 2$ .  $\Sigma$  has diagonal elements equal to one and off-diagonal elements equal to a scalar  $\rho$ , where  $-1 < \rho < 1$ .  $D$  has diagonal elements all equal to 0.1; the independent variables other than race suffer from measurement error of the same variance. We maintain  $\alpha = 0$ ,

**Table 2 Coefficient Estimates for Alternative Correlation and Error-Variance Assumptions,  $k = 2$**

$\alpha = 0, \beta = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \mu = \begin{bmatrix} -2 \\ -2 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$  and  $D = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, \tilde{D} = \begin{bmatrix} \tilde{d} & 0 \\ 0 & \tilde{d} \end{bmatrix}, n = 10,000.$

	$a$	$b_1$	$b_2$
A. $\rho = 0$ :			
Assumed $\tilde{d}$			
0.0	0.4299 (4.2028)	-0.8340 (-25.2057)	-0.8663 (-25.7966)
0.1	0.0394 (0.3483)	-0.9557 (-24.3926)	-0.9924 (-24.9389)
B. $\rho = 0.5$ :			
Assumed $\tilde{d}$			
0.0	0.2975 (3.4439)	-0.8797 (-22.8422)	-0.8705 (-22.5769)
0.1	0.0419 (0.4506)	-0.9726 (-20.2974)	-0.9597 (-20.0418)
C. $\rho = -0.5$ :			
Assumed $\tilde{d}$			
0.0	0.7672 (5.5997)	-0.7816 (-21.8801)	-0.7714 (-21.5103)
0.1	-0.0457 (-0.2720)	-1.0084 (-21.4374)	-0.9969 (-21.1531)

Notes: t-statistics are shown in parentheses beneath the coefficient estimates. For each panel, we drew a set of 10,000 random realizations for  $(y, x)$ : 5,000 with  $z = 0$  and 5,000 with  $z = 1$ . Within each panel, estimation was performed on the same data set.

$\beta = (-1, -1)$ , and  $\mu = (-2, -2)$ . Panel A shows that when the components of  $x$  are uncorrelated, the bias is larger than in the comparable  $k = 1$  model: 0.43 versus 0.14. When the components of  $x$  are positively correlated ( $\rho = 0.5$ ), the bias is smaller by almost a third but is still significant. When the components of  $x$  are negatively correlated ( $\rho = -0.5$ ), the bias is substantially larger. Thus the bias in  $a$  varies negatively with  $\rho$ , just as the linear case suggested. A positive value of  $\rho$  implies that measurement error in  $x_1$  biases the coefficient on  $x_2$  away from zero, counteracting the effect of measurement error in  $x_2$ . Although  $b_i$  is biased toward zero by measurement error in  $x_i$ , the bias is somewhat offset by the effects of measurement error in other components of  $x$ .

When  $k = 1$ , the direction of bias is determined entirely by the sign of  $\beta\mu$ . When  $k > 1$ , the direction of bias depends on  $\Sigma$  and  $D$ , even when  $\beta'\mu$  can be signed. Table 3 illustrates this fact for  $k = 2$ , showing a set of parameters for which  $a$  is biased against finding discrimination. Both  $x_1$  and  $x_2$  are plagued by measurement error, but with a strong positive correlation between the two,

**Table 3 Coefficient Estimates for  
Alternative Error-Variance Assumptions,  $k = 2$**

$$\alpha = 0, \beta = \begin{bmatrix} -0.1 \\ -1 \end{bmatrix}, \mu = \begin{bmatrix} -2 \\ -0.1 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0.75 \\ 0.75 & 1 \end{bmatrix} \text{ and}$$

$$D = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, \tilde{D} = \begin{bmatrix} \tilde{d} & 0 \\ 0 & \tilde{d} \end{bmatrix}, n = 10,000.$$

	$a$	$b_1$	$b_2$
Assumed $\tilde{d}$			
0.0	-0.2445 (-3.4442)	-0.2352 (-7.0602)	-0.7703 (-21.9616)
0.1	0.0312 (0.2888)	-0.0887 (-1.6430)	-0.9962 (-17.3444)

Notes: t-statistics are shown in parentheses beneath the coefficient estimates. For each panel, we drew a set of 10,000 random realizations for  $(y, x)$ : 5,000 with  $z = 0$  and 5,000 with  $z = 1$ . Within each panel, estimation was performed on the same data set.

each has a dampening effect on the bias in the coefficient of the other variable. The net bias in  $b_2$  is toward zero, but  $b_1$  is biased away from zero. Since  $x_1$  is more strongly correlated with  $z$ , the net effect is a negative bias in  $a$ . With the correct error-variance assumption, the model detects the lack of discrimination.

In Table 4 we display results for a model with  $k = 10$ , a size that is more like that of the data sets encountered in actual practice. With  $\rho = 0$ , we see in Panel A that with more correlates plagued by measurement error, the bias in  $a$  is larger. With  $\rho = 0.5$ , the various measurement errors partially offset each other, but  $a$  remains significantly biased. Once again, our technique faithfully compensates for known measurement error.

#### 4. SUMMARY

We have described a method for estimating logit models of discrimination under a range of assumptions about the magnitude of errors in variables. Using artificially generated data, we showed how the bias in the discrimination coefficient varies with measurement error and other basic model parameters. Our method successfully corrects for known measurement error, and can gauge the sensitivity of parameter estimates to errors in variables. Our method can be applied to the studies of lending discrimination cited in the introduction. It can also be applied to the empirical models employed in lending discrimination suits and regulatory examinations. Since the stakes are high in such applications, the models ought to be routinely tested for sensitivity to errors in variables.

Further extensions of our method would be worthwhile. Although we allow for errors only in continuous-valued independent variables, studies of lending

**Table 4 Race Coefficient Estimates for Alternative Correlation and Error-Variance Assumptions,  $k = 10$**

$\alpha = 0$ ,  $\beta$  is a  $k \times 1$  vector of  $-1$ s,  $\mu$  is a  $k \times 1$  vector of  $1$ s,  $\Sigma$  is a  $k \times k$  matrix with  $1$ s on the diagonal and off-diagonal elements equal to  $\rho$ ,  $D$  is a  $k \times k$  matrix with  $0.1$ s on the diagonal and off-diagonal elements equal to  $0$ ,  $\tilde{D}$  is a  $k \times k$  matrix with elements  $\tilde{d}$  on the diagonal and off-diagonal elements equal to  $0$ , and  $n = 10,000$ .

		<i>a</i>
A. True parameter $\rho = 0$ :		
Assumed $\tilde{d}$		
0.0		1.0033 (3.3154)
0.1		-0.0339 (-0.1006)
B. True parameter $\rho = 0.5$ :		
Assumed $\tilde{d}$		
0.0		0.2266 (3.4658)
0.1		0.0645 (0.5988)

Notes: t-statistics are shown in parentheses beneath the coefficient estimate. For each panel, we drew a set of 10,000 random realizations for  $(y, x)$ : 5,000 with  $z = 0$  and 5,000 with  $z = 1$ . Within each panel, estimation was performed on the same data set.

discrimination often include discrete variables that are likely to be fallible as well. It would be worthwhile to allow for errors in the discrete variables, as Klepper (1988a) does for the linear regression model. In addition, it would be useful to allow for uncertainty about the nuisance distributional parameters that our method treats as known.

---

## REFERENCES

Bauer, Paul W., and Brian A. Cromwell. "A Monte Carlo Examination of Bias Tests in Mortgage Lending," Federal Reserve Bank of Cleveland *Economic Review*, vol. 30 (July/August/September 1994), pp. 27–44.

Berkovec, James, Glenn Canner, Stuart Gabriel, and Timothy Hannan. "Race, Redlining, and Residential Mortgage Loan Performance," *Journal of Real Estate Finance and Economics*, vol. 9 (November 1994), pp. 263–94.

- Cummins, Claudia. "Fed Using New Statistical Tool to Detect Bias," *American Banker*, June 8, 1994.
- Friedman, Milton. *A Theory of the Consumption Function*. Princeton, N.J.: Princeton University Press, 1957.
- Goldberger, Arthur S. "Reverse Regression and Salary Discrimination," *Journal of Human Resources*, vol. 19 (Summer 1984), pp. 293–318.
- Hashimoto, Masanori, and Levis Kochin. "A Bias in the Statistical Estimation of the Effects of Discrimination," *Economic Inquiry*, vol. 18 (July 1980), pp. 478–86.
- Hausman, J. A., W. K. Newey, and J. L. Powell. "Nonlinear Errors in Variables: Estimation of Some Engel Curves," *Journal of Econometrics*, vol. 65 (January 1995), pp. 205–33.
- Kapteyn, Arie, and Tom Wansbeek. "Identification in the Linear Errors in Variables Model," *Econometrica*, vol. 51 (November 1983), pp. 1847–49.
- Klepper, Steven. "Bounding the Effects of Measurement Error in Regressions Involving Dichotomous Variables," *Journal of Econometrics*, vol. 37 (March 1988a), pp. 343–59.
- \_\_\_\_\_. "Regressor Diagnostics for the Classical Errors-in-Variables Model," *Journal of Econometrics*, vol. 37 (February 1988b), pp. 225–50.
- \_\_\_\_\_, and Edward E. Leamer. "Consistent Sets of Estimates for Regressions with Errors in All Variables," *Econometrica*, vol. 52 (January 1984), pp. 163–83.
- Levine, David K. "Reverse Regressions for Latent-Variable Models," *Journal of Econometrics*, vol. 32 (July 1986), pp. 291–92.
- Maddala, G. S. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.
- Munnell, Alicia H., Lynn E. Browne, James McEneaney, and Geoffrey M. B. Tootell. "Mortgage Lending in Boston: Interpreting the HMDA Data," Working Paper Series No. 92. Boston: Federal Reserve Bank of Boston, 1992.
- Rao, Potluri. "Some Notes on the Errors-in-Variables Model," *American Statistician*, vol. 27 (December 1973), pp. 217–28.
- Schill, Michael H., and Susan M. Wachter. "Borrower and Neighborhood Racial and Income Characteristics and Financial Institution Mortgage Application Screening," *Journal of Real Estate Finance and Economics*, vol. 9 (November 1994), pp. 223–39.
- Seiberg, Jaret. "When Justice Department Fights Bias by the Numbers, They're His Numbers," *American Banker*, September 14, 1994.

# Some Key Empirical Determinants of Short-Term Nominal Interest Rates

---

Yash P. Mehra

**M**any previous studies cannot account for the high level of interest rates in the early 1980s. For example, Clarida and Friedman (1983, 1984) demonstrate that relative to the predictions of either a structural or an astructural model, interest rates in the early 1980s were too high. This article suggests that the customary empirical measures that gauge the short-run impact of monetary policy on interest rates have become increasingly noisy in the 1980s and that this factor may have been partly responsible for deterioration in the predictive ability of these interest rate equations.

In most previous studies, the short-run impact of monetary policy on rates has been captured by detrended measures of the real money supply defined either as real M1 or real M2. Such empirical measures of the money supply do not provide a consistent basis of comparison over time when deposit rate ceilings are removed and new liquid financial claims are introduced. These and other financial developments of the 1980s also have altered the underlying relationships between the public's demand for these assets and their traditional economic determinants, including the nominal rate (Hetzel and Mehra 1989; Feinman and Porter 1992). As a result, many of the published reduced forms for nominal rates that gauge the short-run impact of monetary policy by the

---

■ The views expressed are those of the author and do not necessarily represent those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

real money supply perform poorly in predicting the actual behavior of rates in the 1980s.<sup>1</sup>

In this article, I present an interest rate equation in which the short-run impact of monetary policy on the real component of rates is captured by changes in the real federal funds rate rather than in the real money supply.<sup>2</sup> In addition, I distinguish between the short- and long-run empirical determinants of rates using cointegration and error-correction methodology. In previous short-rate studies, stationarity properties of the data have largely been ignored, thereby muddling the important distinction between the short- and long-run determinants of rates.<sup>3</sup>

The empirical work presented here focuses on the behavior of one-year Treasury bill rates and finds that inflation is the main long-run economic determinant of the level of the nominal rate. Conversely, several customary empirical measures of fiscal policy and a measure capturing foreign capital inflows are not significant when included in the long-run (cointegrating) interest-inflation regression. Also, in the short run, changes in the nominal rate depend largely upon changes in inflation, real output, and the real federal funds rate. Changes in fiscal policy measures and foreign capital inflows do not affect the nominal rate even in the short run.

The short-run interest rate equation estimated here does not exhibit any simultaneous equation bias. The real funds rate is therefore exogenous in the

---

<sup>1</sup> Other studies that have addressed this issue are those of Peek and Wilcox (1987) and Hendershott and Peek (1992). Peek and Wilcox continue to employ the real money supply measure and attribute the high level of rates in the early 1980s to a less accommodative monetary policy. They, however, use dummy variables to capture differences in the average tightness of policy during different Fed chairman regimes. Hendershott and Peek, on the other hand, abandon the money supply measure and use instead innovations in the slope of the term structure (the ratio of the 6- to 60-month Treasury rates) to measure the short-run impact of monetary policy on the real component of short-term rates. Given this proxy, monetary policy is highly relevant in explaining the behavior of rates in the early 1980s.

<sup>2</sup> Unlike the slope of the term structure used in Hendershott and Peek (1992), the nominal federal funds rate has been the main instrument of monetary policy during most of the sample period studied here. I use the real funds rate because in the short run the Fed influences the nominal rate mainly by affecting its real component. Moreover, the real funds rate is correlated with the nominal funds rate in the short run. The simple correlation between these two variables is 0.46. Recently, Goodfriend (1993) has used the federal funds rate to measure the short-run impact of monetary policy on the real component of the long rate.

<sup>3</sup> This distinction is important in describing the effect of monetary policy on the nominal rate. Many analysts believe that monetary policy can control the level of the nominal rate in the long run only through its control over inflation, even though in the short run it could have substantial effects. This distinction between short- and long-run effects can be made more precise using cointegration and error-correction methodology. Thus, the statement that monetary policy determines the level of the nominal rate in the long run through its control over inflation can be interpreted to mean that measures of monetary policy are not at the source of long-run, stochastic movements in the level of the nominal rate, but inflation is. That is, the nominal rate is cointegrated with inflation, but not with measures of monetary policy. However, short-run stationary movements in the nominal rate could still be correlated with measures of monetary policy, indicating that in the short run, monetary policy also influences the nominal rate.

short-run equation, which means that the real funds rate is not correlated with contemporaneous shocks to the short rate. Thus, the Federal Reserve does influence the market rate in the short run. In addition, the nominal rate equation that captures the impact of monetary policy by the real funds rate explains reasonably well the nominal rate in the 1980s, although it does not completely solve the puzzle of high rates during the early 1980s. The equation, for example, significantly underpredicts the level of the short rate in 1981. Finally, the results with other measures of short- and medium-term nominal rates indicate that in the short run, nominal yields on the short end of the U.S. Treasury term structure are determined largely by movements in inflation, the real funds rate, and real GDP. The short-run influences of these economic factors, however, decline with the term to maturity, suggesting that the yield curve at the short end is affected most by the outlook for Fed policy and the state of the economy.

The plan of this article is as follows. Section 1 briefly describes the model and the method used in estimating the nominal rate equation. Section 2 presents empirical results, and Section 3 contains concluding observations.

## 1. THE MODEL AND THE METHOD

### An Economic Specification of the Interest Rate Equation

The nominal interest rate equation that underlies the empirical work presented here is based on a variant of a loanable funds model employed by Sargent (1969), among others. According to that model, the nominal rate depends upon anticipated inflation, changes in the real money supply and income, the deficit, and the level of income.<sup>4</sup> The version examined here measures the impact of monetary policy on the nominal rate by changes in the real federal funds rate rather than in the real money supply. In addition, I examine the role of other fiscal policy measures such as government purchases, net taxes, and foreign capital inflows in determining the nominal rate. Since this model has already been described in an earlier paper (Mehra 1994), I report below just the relevant economic specifications that are used to investigate the behavior of the short rate.

The short-run interest rate equation is estimated using cointegration and error-correction modeling. If the nominal rate and empirical measures of its potential economic determinants are nonstationary, then tests for cointegration provide inferences about the existence of a long-run, equilibrium relationship between the nominal rate and its potential determinants. The error-correction equation then explains short-run changes in the nominal rate.

---

<sup>4</sup> The nominal rate responds positively to anticipated inflation, the deficit, and real growth and responds negatively to increases in the real money supply. A rise in the level of real income, however, generates a larger volume of savings and hence depresses the equilibrium real rate (Sargent 1969).

The nominal rate equation estimated here has two parts: a long-run part and a short-run part. The long-run part that specifies the potential long-run determinants of the level of the nominal rate is given in (1):

$$R_t = a_0 + a_1 \dot{p}_t^e + a_2 RFR_t + a_3 FP_t - a_4 \ln ry_t + a_5 \Delta \ln ry_t + U_t, \quad (1)$$

where  $R$  is the nominal interest rate,  $\dot{p}^e$  is anticipated inflation,  $RFR$  is the real federal funds rate,  $FP$  is a fiscal policy variable,  $\ln ry$  is the logarithm of real income, and  $U$  is the disturbance term. Equation (1) describes the long-run response of the nominal rate to anticipated inflation, the real funds rate, a fiscal policy variable, changes in real income, and the level of real income. The coefficients  $a_i, i = 1$  to 5, measure the long-run responses in the sense that they are the sums of coefficients that appear on current and past values of the relevant economic determinants. The term  $a_1 \dot{p}_t^e$  in (1) captures the inflation premium in the nominal rate, whereas the remaining terms capture the influence of other variables on the equilibrium real component of the short rate. If the nominal rate and anticipated inflation variables are nonstationary but cointegrated as in Engle and Granger (1987), then the other remaining long-run impact coefficients in (1) may all be zero.

Equation (1) may not do well in explaining short-run movements in the nominal rate for a number of reasons. First, it ignores the short-run effects of economic factors. Some economic factors, including those measuring monetary policy actions, may be important in explaining short-run changes in the nominal rate, even though they may have no long-run effects. Second, it completely ignores short-run dynamics. Hence, in order to explain short-run changes in the nominal rate, consider the following error-correction model of the nominal rate:

$$\begin{aligned} \Delta R_t = & c_0 + c_1 \Delta \dot{p}_t^e + c_2 \Delta RFR_t + c_3 \Delta FP_t + c_4 \Delta \ln ry_t \\ & + c_5 \Delta^2 \ln ry_t + \sum_{s=1}^n c_{6s} \Delta R_{t-s} + c_7 U_{t-1} + \epsilon_{1t}, \end{aligned} \quad (2)$$

where  $U_{t-1}$  is the lagged residual from the long-run nominal equation (1),  $\Delta^2$  is the second-difference operator, and other variables are as defined earlier. Equation (2) is the short-run interest rate equation, and the coefficients  $c_i, i = 1$  to 5, capture the short-run responses of the interest rate to economic determinants suggested here.

### **Estimation Issues: OLS Works If Empirical Measures of Economic Determinants Are Nonstationary or Exogenous**

Both the long- and short-run equations (1) and (2) contain contemporaneous values of economic fundamentals. Those values are likely to be correlated with contemporary shocks to the nominal rate. In particular, if the Federal Reserve contemporaneously adjusts its short-run funds objective with respect

to movements in short rates, then the real federal funds rate is likely to be correlated with the disturbance term in the regression.<sup>5</sup> Hence, these equations cannot be consistently estimated by ordinary least squares unless some special assumptions hold.

The long-run equation (1) can be consistently estimated by ordinary least squares if empirical measures of the economic factors included in (1) are non-stationary but cointegrated as in Engle and Granger (1987). Tests of hypotheses on coefficients that appear in (1) can then be carried out by estimating Stock and Watson's (1993) dynamic OLS regressions. I, therefore, test first for nonstationarity and cointegration. The empirical work here examines the stationarity properties of the data using unit root and mean stationarity tests. The test for cointegration used is the one proposed in Johansen and Juselius (1990).<sup>6</sup>

The economic variables that appear in the short-run equation (2) are stationary. This equation can be consistently estimated by ordinary least squares if contemporaneous right-hand side explanatory variables are uncorrelated with the disturbance term. That condition can be tested by performing the test for exogeneity given in Hausman (1978). To implement the test, consider the following VAR representation of these contemporaneous right-hand side explanatory variables. (For simplicity, I am ignoring fiscal policy and some other variables.)

$$\begin{aligned} \Delta \dot{p}_t^e = & d_0 + \sum_{s=1}^k d_{1s} \Delta \dot{p}_{t-s}^e + \sum_{s=1}^k d_{2s} \Delta RFR_{t-s} \\ & + \sum_{s=1}^k d_{3s} \Delta \ln ry_{t-s} + \sum_{s=1}^k d_{4s} \Delta R_{t-s} + \epsilon_{2t} \end{aligned} \quad (3)$$

$$\begin{aligned} \Delta \ln ry_t = & e_0 + \sum_{s=1}^k e_{1s} \Delta \dot{p}_{t-s}^e + \sum_{s=1}^k e_{2s} \Delta RFR_{t-s} \\ & + \sum_{s=1}^k e_{3s} \Delta \ln ry_{t-s} + \sum_{s=1}^k e_{4s} \Delta R_{t-s} + \epsilon_{3t} \end{aligned} \quad (4)$$

---

<sup>5</sup> To illustrate, consider a scenario in which the incoming new data indicates that in the current quarter real growth or inflation is going to be higher than what the market expected based on the past information. If the market believes such information, short rates could rise because accelerations in real growth or inflation are generally associated with higher rates. If the Fed also reacts contemporaneously to such new information and the resulting rise in short rates, then changes in the funds rate would be correlated with the disturbance term. Such correlation will be absent, however, if the Fed does not react or reacts with a lag.

<sup>6</sup> These tests are described in detail in Mehra (1994).

$$\begin{aligned} \Delta RFR_t = & f_0 + \sum_{s=1}^k f_{1s} \Delta \hat{p}_{t-s}^e + \sum_{s=1}^k f_{2s} \Delta RFR_{t-s} \\ & + \sum_{s=1}^k f_{3s} \Delta \ln ry_{t-s} + \sum_{s=1}^k f_{4s} \Delta R_{t-s} + \epsilon_{4t} \end{aligned} \quad (5)$$

This VAR includes only past values of the economic factors that appear in the economic model used here and hence can be consistently estimated by ordinary least squares. Then consider an expanded version of equation (2) given below:

$$\begin{aligned} \Delta R_t = & d_0 + d_1 \Delta \hat{p}_t^e + d_2 \Delta RFR_t + d_3 \Delta \ln ry_t + \sum_{s=1}^n d_4 \Delta R_{t-s} \\ & + d_5 U_{t-1} + d_6 \hat{\epsilon}_{2t} + d_7 \hat{\epsilon}_{3t} + d_8 \hat{\epsilon}_{4t} + \epsilon_{1t}, \end{aligned} \quad (6)$$

where  $\hat{\epsilon}_2$ ,  $\hat{\epsilon}_3$ , and  $\hat{\epsilon}_4$  are residuals from the VAR. If  $d_6 = d_7 = d_8 = 0$ , then  $\Delta \hat{p}_t^e$ ,  $\Delta RFR_t$ , and  $\Delta \ln ry$  are uncorrelated with the disturbance term and hence exogenous in this equation. The hypothesis  $d_6 = d_7 = d_8 = 0$  can be tested using the F-test.<sup>7</sup>

### Data and Definition of Variables

The empirical work uses quarterly data from 1955:1 to 1994:3. The short-term nominal rate,  $R1$ , is the nominal yield on one-year U.S. Treasury bills. I consider two proxies for anticipated inflation. The first one uses actual inflation as measured by the behavior of the consumer price index ( $\hat{p}$ ). The second one uses one-year-ahead inflation rates from the Livingston survey ( $\hat{p}^e$ ). The real federal funds rate,  $RFR$ , is the nominal federal funds rate minus the actual, annualized quarterly inflation rate. Nominal interest rate data are observations from the last month of the quarter, and inflation ( $\hat{p}$ ) is calculated as the change in the log form from the last month of the previous quarter price level to that of the current. In some specifications in which the real money supply is used to measure the impact of monetary policy actions on the real component of the nominal rate, the measure of money used is M2 scaled by the real GDP deflator. Real income,  $ry$ , is real GDP. The real deficit scaled by real GDP ( $DEF/y$ ),<sup>8</sup> real government purchases ( $fg$ ), and real government tax (net of transfers) receipts ( $tx$ ) are alternatively used to measure the impact of fiscal policy on the real component of the short rate. I also consider the impact of

<sup>7</sup> The hypothesis that the right-hand side contemporary regressors in (2) are independent of the disturbance term also can be tested by comparing ordinary least squares and instrumental variables estimates of the equation. Under the null hypothesis that there is no simultaneity equation bias, OLS estimates ( $\hat{\beta}_{OLS}$ ) should not be statistically different from IV estimates ( $\hat{\beta}_{IV}$ ). Hausman (1978) shows that the statistic that tests the null hypothesis  $\hat{\beta}_{OLS} = \hat{\beta}_{IV}$  is distributed Chi-squared with a degree of freedom parameter equal to the number of parameters estimated in the equation. See Maddala (1988) for a simple description of these test procedures.

<sup>8</sup> This specification reflects the assumption that in a growing economy higher deficits result in higher rates only if the deficit rises relative to GDP.

foreign capital inflows measured as the ratio of U.S. Treasury securities held by foreigners to the total of U.S. securities held by domestic and foreign residents ( $fh$ ).<sup>9</sup>

## 2. ESTIMATION RESULTS

### On the Long-Run Determinants of the Nominal Rate

I first present test results that help determine which economic determinants suggested in the long-run equation (1) are relevant.

Table 1 presents test results for determining whether empirical measures of potential determinants such as  $R1$ ,  $\dot{p}$ ,  $\dot{p}^e$ ,  $DEF/y$ ,  $\ln fg$ ,  $\ln tx$ ,  $fh$ ,  $\ln rM2$ ,  $\ln ry$ , and  $RFR$  have a unit root or are mean stationary. As can be seen, the t-statistic ( $t_{\hat{\rho}}$ ) that tests the null hypothesis that a particular variable has a unit root is small for all these series. On the other hand, the test statistic ( $\hat{n}_u$ ) that tests the null hypothesis that a particular variable is mean stationary is large for all these variables with the exception of  $RFR$ . These results indicate that  $R1$ ,  $\dot{p}$ ,  $\dot{p}^e$ ,  $DEF/y$ ,  $\ln fg$ ,  $fh$ ,  $\ln tx$ ,  $\ln rM2$ , and  $\ln ry$  have a unit root and thus are nonstationary in levels.<sup>10</sup> The results are inconclusive for the real funds rate  $RFR$ . Together, these results indicate that most empirical measures of the potential determinants suggested here are nonstationary and hence could be the source of long-run stochastic movements in the nominal rate.

Table 2 presents test statistics for determining whether the short-term nominal rate ( $R1$ ) is cointegrated with any of these nonstationary measures of inflation, fiscal and monetary policies, and foreign capital inflows. Trace and maximum eigenvalue statistics, which test the null hypothesis that there is no cointegrating vector, are large for systems  $(R1, \dot{p})$ ,  $(R1, \dot{p}^e)$ ,  $(R1, \ln ry)$ ,  $(R1, DEF/y)$ , and  $(R1, \ln rM2)$ , but are very small for systems  $(R1, RFR)$ ,  $(R1, \ln fg)$ ,  $(R1, \ln tx)$ , and  $(R1, fh)$ . These results indicate that the nominal rate is cointegrated with inflation (actual or anticipated), the real money supply, the level of income, and the deficit, but not with the real federal funds rate, government purchases, net taxes, and foreign capital inflows. The evidence continues to favor the presence of at least one cointegrating vector even in expanded systems that include inflation and fiscal and monetary policy variables

<sup>9</sup> The data on the Livingston survey are provided by the Philadelphia Fed. The data used in measuring capital inflows are from the Federal Reserve Board's flow of funds data. All other data series are from the Citibank database.

<sup>10</sup> The t-statistic that tests the null hypothesis that first differences of a series have a unit root takes values  $-5.2$ ,  $-4.9$ ,  $-6.4$ ,  $-15.9$ ,  $-4.2$ ,  $-6.2$ ,  $-5.2$ ,  $-5.6$ ,  $-4.7$ , and  $-7.4$  for  $\Delta R1$ ,  $\Delta \dot{p}$ ,  $\Delta \dot{p}^e$ ,  $\Delta RFR$ ,  $\Delta \ln rM2$ ,  $\Delta DEF/y$ ,  $\Delta \ln fg$ ,  $\Delta \ln tx$ ,  $\Delta fh$ , and  $\Delta \ln ry$ , respectively. These t-values are large, indicating that first differences of these series are stationary. The 5 percent critical value taken from Fuller (1976) is  $-2.9$ .

**Table 1 Tests for Unit Roots and Mean Stationarity**

Series X	Panel A Tests for Unit Tools			Panel B Tests for Mean Stationarity
	$\hat{\rho}$	$t_{\hat{\rho}}$	$k$	$\hat{n}_u$
<i>R1</i>	0.94	-1.93	2	0.84*
$\dot{p}$	0.84	-2.83	7	0.49*
$\dot{p}^e$	0.98	-1.80	2	0.98*
<i>RFR</i>	0.85	-2.51	2	0.37
$\ln rM2$	0.99	-1.87	1	1.80*
$\ln ry$	0.97	-1.80	1	0.33
<i>DEF/y</i>	0.92	-2.54	1	1.42*
$\ln fg$	0.98	-1.63	3	0.91*
$\ln tx$	0.96	-1.54	1	1.54*
<i>fh</i>	0.98	-1.58	5	1.19*

\* Significant at the 5 percent level.

Notes: *R1* is the one-year Treasury bill rate;  $\dot{p}$  is the annualized quarterly inflation rate measured by the consumer price index;  $\dot{p}^e$  is the Livingston survey measure of one-year-ahead expected inflation; *RFR* is the real federal funds rate; *rM2* is the real money supply; *ry* is real GDP; *DEF/y* is the ratio of federal government deficits to nominal GDP; *fg* is real federal government purchases; *tx* is real federal government tax (net of transfers) receipts; and *fh* is the ratio of U.S. Treasury securities held by foreigners to total of U.S. Treasury securities held by domestic and foreign residents. The sample period studied is 1955:1 to 1994:3. The values for  $\rho$  and t-statistics ( $t_{\hat{\rho}}$ ) for  $\rho = 1$  in Panel A above are from the Augmented Dickey-Fuller regressions of the form

$$X_t = a_0 + \rho X_{t-1} + \sum_{s=1}^k a_s \Delta X_{t-s}, \quad (a)$$

where  $X$  is the pertinent series. The number of lagged first differences ( $k$ ) included in these regressions are chosen using the procedure given in Hall (1990). The procedure starts with some upper bound on  $k$ , say  $k$  max, chosen a priori (eight quarters here). Estimate (a) above with  $k$  set at  $k$  max. If the last included lag is significant, select  $k = k$  max. If not, reduce the order of the autoregression by one until the coefficient on the last included lag is significant. The test statistic  $\hat{n}_u$  in Panel B above is the statistic that tests the null hypothesis that the pertinent series is mean stationary. The 5 percent critical value for  $\hat{n}_u$  given in Kwiatkowski et al. (1992) is 0.463.

[see systems (*R1*,  $\dot{p}$ , *DEF/y*,  $\ln rM2$ ,  $\ln ry$ ), (*R1*,  $\dot{p}^e$ , *DEF/y*,  $\ln rM2$ ,  $\ln ry$ ), (*R1*,  $\dot{p}$ , *DEF/y*,  $\ln ry$ ), and (*R1*,  $\dot{p}^e$ , *DEF/y*,  $\ln ry$ )].

Panels A and B in Table 3 help determine which variables included in the cointegrating regression<sup>11</sup> are statistically significant. It presents the dynamic OLS estimates of the potential cointegrating regressions with and without the real money supply. As can be seen, inflation (actual or anticipated) is the only variable that enters significantly in these cointegrating regressions. Other

<sup>11</sup> In this article, I focus on a single cointegrating regression that is normalized on the short-term nominal rate. The analysis here thus ignores the possibility that in larger systems there may be multiple cointegrating vectors.

**Table 2 Cointegration Test Results**

System	Trace Test	Maximum Eigenvalue Test	$k$
$(R1, \dot{p})$	22.2*	17.9*	8
$(R1, \dot{p}^e)$	22.1*	18.9*	2
$(R1, RFR)$	17.6	14.3	8
$(R1, \ln rM2)$	20.3*	16.9*	8
$(R1, \ln fg)$	10.1	5.3	2
$(R1, \ln tx)$	12.6	8.5	4
$(R1, DEF/y)$	30.7*	27.5*	8
$(R1, fh)$	14.7	10.7	8
$(R1, \ln ry)$	48.6*	43.4*	2
$(R1, \dot{p}, DEF/y, \ln ry)$	97.1*	42.9*	4
$(R1, \dot{p}^e, DEF/y, \ln ry)$	86.1*	34.9*	4
$(R1, \dot{p}, DEF/y, \ln rM2, \ln ry)$	148.8*	65.2*	2
$(R1, \dot{p}^e, DEF/y, \ln rM2, \ln ry)$	146.4*	64.9*	2

\* Significant at the 5 percent level.

Notes: Trace and maximum eigenvalue tests are tests of the null hypothesis that there is no cointegrating vector in the system. The lag length in the relevant VAR system is  $k$  and is chosen using the likelihood ratio test given in Sims (1980). In particular, the VAR model initially was estimated with  $k$  set equal to a maximum number of eight quarters. This unrestricted model was then tested against a restricted model, where  $k$  is reduced by one, using the likelihood ratio test. The lag length finally selected is the one that results in the rejection of the restricted model.

nonstationary variables such as the real deficit and the real money supply are not significant. Real GDP is significant in some regressions and not in others (see Panels A and B, Table 3). These results thus indicate that inflation is the main long-run economic determinant of the short-term nominal rate.

Panel C in Table 3 presents the cointegrating regressions that include only the inflation variable. The cointegrating regression is estimated with and without the restriction that the nominal rate adjusts one for one with inflation in the long run. The  $\chi^2$  statistic that tests the validity of the full Fisher-effect restriction is not large, indicating that this restriction is consistent with data. These results also indicate that the real rate of interest on one-year Treasury bills is mean stationary. The estimate of this mean falls in a 2.3 to 3.2 percent range (see the constant term in Panel C regressions).

### On the Short-Run Determinants of the Nominal Rate

The short-run equation (2) is estimated here jointly with its long-run part and hence includes levels as well as first differences of the relevant economic

**Table 3 Cointegrating Regressions; Dynamic OLS**

(Leads, Lags)	Panel A: With Real Money Supply
(-4, 4)	$R1_t = 0.7\dot{p}_t - 17.6 \ln rM2_t + 2.05 \ln ry_t + 0.22(DEF/y)_t$ (3.9) (0.9) (1.0) (0.5)
(-4, 4)	$R1_t = 1.1\dot{p}_t^e + 6.1 \ln rM2_t + 6.8 \ln ry_t + 0.02(DEF/y)_t$ (5.9) (0.4) (0.4) (0.1)
	Panel B: Without Real Money Supply
(-4, 4)	$R1_t = 0.9\dot{p}_t - 1.0 \ln ry_t + 0.18(DEF/y)_t$ (21.8) (2.2) (1.5)
(-4, 4)	$R1_t = 1.1\dot{p}_t^e - 1.5 \ln ry_t - 0.10(DEF/y)_t$ (23.1) (12.9) (0.7)
	Panel C: With Inflation Only
(-4, 4)	$R1 = 3.2 + 0.8\dot{p}_t; R1 = 2.3 + 1.0\dot{p}_t; \chi^2(1) = 2.8(0.10)$
(-4, 4)	$R1 = 2.5 + 1.1\dot{p}_t^e; R1 = 2.6 + 1.0\dot{p}_t^e; \chi^2(1) = 0.32(0.57)$

Notes: All regressions are estimated by the dynamic OLS procedure given in Stock and Watson (1993), using leads and lags of first differences of the relevant right-hand side explanatory variables. Parentheses contain t-values corrected for the presence of moving average serial correlation.  $\chi^2(1)$  is the  $\chi^2$  statistic with one degree of freedom (significance levels in parentheses); it tests the hypothesis that the coefficient on  $\dot{p}$  or  $\dot{p}^e$  is unity.

determinants.<sup>12</sup> A preliminary specification search indicated that in the short run, changes in the nominal rate depend largely upon contemporaneous changes in inflation, the real funds rate, and real GDP. The lagged level of the funds rate is also significant. The empirical measures of fiscal policy and foreign capital inflows, however, did not enter the short-run equation. (I formally test these restrictions later on.)

Table 4 presents ordinary least squares as well as instrumental variables estimates of the pertinent short-run equation. The instruments chosen are basically the lagged values of the right-hand side explanatory variables that appear

<sup>12</sup> The short-run equation (2) of the text includes a one-period lagged value of the residual from the long-run equilibrium equation. In joint estimation, the lagged residual is replaced by lagged levels of the variables that enter the long-run equilibrium equation. To see it, assume for the sake of explanation that the long-run cointegrating regression is given in (b) below:

$$R1_t = a_0 + a_1\dot{p}_t^e + U_t. \quad (b)$$

If we solve (b) for  $U_{t-1}$  and substitute for  $U_{t-1}$  into (2), then the short-run equation (2) will include the lagged level of the nominal rate ( $R1$ ) and the inflation rate ( $\dot{p}^e$ ).

**Table 4 Short-Run Nominal Interest Rate Equations, 1957:1 to 1994:3**

Explanatory Variables	Instrumental Variables		Ordinary Least Squares	
	(A.1)	(A.2)	(B.1)	(B.2)
constant	-0.22(1.4)	1.0 (2.9)	0.06( 0.8)	0.6 (3.8)
$\Delta \dot{p}_t$	0.71(8.2)		0.72(12.8)	
$\Delta \dot{p}_t^e$		0.9 (3.2)		0.8 (3.3)
$\Delta RFR_t$	0.61(6.5)	0.11(1.2)	0.64(11.3)	0.17(3.5)
$\Delta \ln ry_t$	0.12(2.6)	-0.02(0.3)	0.06( 3.9)	0.06(2.1)
$R1_{t-1}$	-0.37(3.4)	-0.55(4.9)	-0.33( 4.1)	-0.53(6.1)
$\dot{p}_{t-1}$	0.37(3.4)		0.33( 4.1)	
$\dot{p}_{t-1}^e$		0.55(4.9)		0.53(6.1)
$RFR_{t-1}$	0.34(2.7)	0.28(3.3)	0.30( 3.3)	0.30(4.7)
$\Delta R1_{t-1}$	-0.23(6.2)		-0.21( 5.0)	
$\Delta R1_{t-2}$	0.07(1.2)		0.04( 0.9)	
$\Delta R1_{t-3}$	-0.04(0.5)		-0.06( 0.9)	
$\Delta R1_{t-4}$	0.15(2.9)		0.14( 2.9)	
SER	0.547	0.98	0.52	0.94
Q(36)	33.8	27.7	37.1	38.9
Q(8)	6.9	5.9	9.2	11.9
Q(4)	2.8	3.9	2.2	8.5
Sargan's $\chi^2$	12.5(0.19)	6.5(0.36)		
F	1.4(0.24)	1.0(0.37)		

\* Significant at the 5 percent level.

Notes: See notes in Table 1 for definition of variables. Parentheses following coefficients contain t-values. SER is the standard error of estimate, and Q(36), Q(8), and Q(4) are the Ljung-Box Q-statistics based on 36, 8, and 4 autocorrelations of the residuals.

Sargan's  $\chi^2$  tests the independence of the instruments and the disturbance term (parentheses contain the significance level of the test). F is the F-statistic that tests the null hypothesis that the residuals from the reduced-form regressions of real growth, inflation, and the real federal funds rate are not jointly significant when included in the interest rate equation (parentheses contain the significance level of the test). See footnote 13 of the text for a description of instruments used.

in (2).<sup>13,14</sup> The null hypothesis that contemporary values of changes in inflation, the real funds rate, and real growth are jointly exogenous with respect to

<sup>13</sup> When actual inflation is the proxy for anticipated inflation, the instruments used for estimating the short-run equation are a constant, one-period lagged value of the short rate, inflation, and the real funds rate, two-period lagged values of the change in the short rate, and four-period lagged values of changes in inflation, real income, and the real funds rate. When the Livingston survey is the proxy for anticipated inflation, I use similar instruments, but I treat the Livingston survey as exogenous in the short equation. Hence, first differences and one-period lagged values of the Livingston survey ( $\Delta \dot{p}_t^e$ ,  $\dot{p}_{t-1}^e$ ) are used as instruments.

<sup>14</sup> Two considerations are important in the choice of instruments. First, the instruments chosen should be uncorrelated with the disturbance term. Second, they should be highly correlated

short-run impact coefficients is not rejected (the relevant Hausman F-statistics reported in Table 4 are small).<sup>15</sup> Sargan's (1964)  $\chi^2$  specification test statistic presented in Table 4 is also small, indicating that the instruments chosen are independent of the disturbance term. A casual look at the estimates reported in Table 4 indicates that instrumental variables estimates of short-run impact coefficients are not strikingly different from OLS estimates.<sup>16</sup> These results suggest that in the short-run equation, changes in inflation, the funds rate, and real income are not correlated with contemporary shocks to the short rate. Hence, the estimates of short-run impact coefficients reported here are consistent.<sup>17</sup>

The short-run equation is reported for the full sample 1955:1 to 1994:3.<sup>18</sup> As can be seen, the relevant explanatory variables have coefficients that are

---

with the contemporaneous values of endogenous variables. In the short-run equation estimated here, lagged endogenous variables are valid as instruments if the equation does not exhibit any serial correlation. The Ljung-Box Q-statistics reported in Table 4 indicate that serial correlation is not a problem in the regressions reported there. As regards the second point, lagged endogenous variables are good instruments because they are likely to be highly correlated with the contemporaneous endogenous variables. However, these variables may not be strictly exogenous in the sense that they are completely independent of past shocks to economic variables, including the nominal rate. Thus, changes in the real funds rate may be uncorrelated with the disturbance term in the regression but may not be independent of the past behavior of economic fundamentals, including the short rate.

<sup>15</sup> The individual t-statistics that appear on the residuals from the reduced-form regressions of inflation, the funds rate, and real growth (the t-statistics for  $d_6 = 0$ ,  $d_7 = 0$ , or  $d_8 = 0$  in (6)) are not large in the short-run interest rate equation either. In contrast, the null hypothesis that in the real funds rate equation contemporaneous values of changes in inflation, real GDP, and the nominal rate are jointly exogenous is usually rejected by the F-test.

<sup>16</sup> The null hypothesis that instrumental variables estimates of the short-run equation (2) of the text are not jointly different from OLS estimates is not rejected by the Hausman test described in footnote 7. For equations A.1 and A.2 of Table 4, the relevant  $\chi^2$  statistics are 2.1 and 2.6, respectively. Both these statistics are small.

<sup>17</sup> If one begins with a structural equation in which the nominal rate depends upon contemporaneous anticipated values of economic variables, then lagged values are valid as instruments for unobservables if the order of lag in instruments chosen exceed the order of lag in the serial correlation of the disturbance term. The empirical work here does not begin with any particular structural equation. It does, however, assume that the disturbance term in the short-run equation does not have any serial correlation. Nevertheless, in order to check the robustness of results to the order of lag in instruments, I also reestimated the short-run equations using successively more than one-period lagged values of the right-hand side explanatory variables as instruments, going as far back as five- through eight-period lags. The point-estimates of the short-run impact coefficients that appear on inflation, real GDP, and the funds rate move around somewhat, but they continue to have expected signs and are generally significant. As expected, standard errors of the estimated coefficients increase as the order of lag in instruments chosen increases.

<sup>18</sup> The instrumental variables estimates of the short-run equations for the subsample 1955:1 to 1979:3 are given below, and those estimates look very similar to the ones for the whole period.

$$\begin{aligned} \Delta R1_t = & 0.06 + 0.76\Delta\dot{p}_t + 0.67\Delta RFR_t + 0.00\Delta \ln ry_t - 0.16\Delta R1_{t-1} - 0.15\Delta R1_{t-2} \\ & (8.7) \quad (6.8) \quad (0.0) \quad (2.2) \quad (2.2) \\ & - 0.13R1_{t-1} + 0.13\dot{p}_{t-1} + 0.12RFR_{t-1} \\ & (1.7) \quad (1.7) \quad (1.5) \end{aligned}$$

of expected signs and statistically significant. Thus, the nominal rate responds positively to short-run increases in inflation, real GDP, and the real funds rate. The coefficients that appear on contemporary values of these variables range from 0.7 to 0.9 for inflation, 0.1 to 0.6 for the real funds rate, and 0.0 to 0.12 for real growth. Thus, a one percentage point rise in inflation raises the short rate between 70 and 90 basis points, whereas a similar increase in the real funds rate raises it by 10 to 60 basis points in the short run. A one percentage point rise in the growth rate of real GDP raises the short rate by about 12 basis points in the regression that uses actual inflation data.<sup>19</sup>

The short-run equations reported in Table 4 embody the long-run relationship among the levels of economic determinants. The null hypothesis that coefficients appearing on one-period lagged levels of inflation and the nominal rate sum to zero is not rejected, indicating that the nominal rate adjusts one for one with inflation in the long run. The coefficient that appears on the lagged level of the real funds rate is large and remains statistically significant, indicating that (stationary) movements in the real funds rate have substantial short-run effects on the real component of the short rate.

The short-run equations reported in Table 4 do not include any fiscal policy measure. Nor do they allow for the effect of foreign capital inflows. At this point I formally test the hypothesis that these variables have no significant effects on the nominal rate. Table 5 presents Lagrange multiplier tests for omitted variables.<sup>20</sup> Those test results indicate that the real deficit, real government

$$\begin{aligned} \Delta R1_t = & 1.0 + 1.0\Delta p_t^e + 0.29\Delta RFR_t - 0.06\Delta \ln ry_t + 0.17\Delta R1_{t-1} \\ & (3.4) \quad (2.0) \quad (0.1) \quad (1.4) \\ & - 0.54R1_{t-1} + 0.54p_{t-1}^e + 0.38RFR_{t-1} \\ & (3.3) \quad (3.3) \quad (2.5) \end{aligned}$$

Parentheses contain t-values.

<sup>19</sup> The empirical work here uses actual values of the real funds rate, in contrast with previous studies in which the short-run impact of monetary policy on rates is captured by employing innovations in the pertinent money supply or term-structure measure. The latter approach reflects two assumptions. First, anticipated values of these variables affect the short rate by altering its expected-inflation component. In the short run, only unanticipated changes have an effect on the real component of the short rate. Second, it is possible to decompose the pertinent monetary policy measure into its anticipated and unanticipated components. Some of these assumptions are questionable. Nevertheless, I also estimate the short-run equation using residuals from the short-run real funds rate equation like (5) as a measure of unanticipated monetary policy actions. The coefficient that appears on this measure of policy remains positive and is statistically significant. The estimated coefficient is 0.46 (t-value = 8.6) when the actual inflation data is used, whereas it is 0.20 (t-value = 3.2) when the expected inflation measure is used. These results indicate that monetary policy effects when using the actual funds rate are not spurious.

<sup>20</sup> A Lagrange multiplier test for a set of  $p$ -omitted variables is constructed by regressing the model's residuals on both the set of original regressors and on the set of omitted variables. If the omitted variables do not belong in the equation, then multiplying the  $R^2$  statistic from this regression by the number of observations will produce a statistic distributed as  $\chi^2$  with  $p$  degrees of freedom (Engle 1984; Breusch and Pagan 1980).

**Table 5**  $\chi^2$  Tests for Omitted Variables

Candidate Variable $X$	Lags (0 to $k$ )	Equation B.1	Equation B.2
$\Delta(DEF/y)$	(0,0)	0.0 (0.96)	1.5 (0.21)
	(0,1)	0.7 (0.68)	1.5 (0.47)
	(0,4)	6.1 (0.30)	4.6 (0.47)
$\Delta \ln fg$	(0,0)	0.0 (0.80)	0.9 (0.33)
	(0,1)	1.8 (0.40)	1.2 (0.55)
	(0,4)	3.9 (0.54)	1.8 (0.87)
$\Delta \ln tx$	(0,0)	0.2 (0.28)	2.2 (0.14)
	(0,1)	2.4 (0.30)	4.7 (0.10)
	(0,4)	4.2 (0.51)	7.3 (0.20)
$\Delta fh$	(0,0)	1.2 (0.27)	0.1 (0.71)
	(0,1)	3.2 (0.20)	0.4 (0.81)
	(0,4)	6.1 (0.30)	1.6 (0.89)

Notes: See notes in Table 1 for definition of variables. The statistics reported are the Wald test of the null hypothesis that the pertinent variable is not an omitted variable from the relevant regression. If this statistic is large for some variable, then the variable should be included in the regression. Parentheses contain significance levels of the test.

purchases, net taxes, and foreign capital inflows do not enter the short-run interest equation. Overall, these results assign no significant role to fiscal policy measures and foreign capital inflows in explaining short-run movements in the nominal rate.

### Predicting the Behavior of the Nominal Rate in the 1980s

I now examine whether the short-run equation estimated here can predict the actual behavior of the nominal rate during the 1980s. The predicted values used are the out-of-sample, one-year-ahead dynamic forecasts that cover the subperiod from 1979 to 1993. I focus on the equation that uses actual inflation.

Table 6 presents predicted values generated using the interest rate regression presented in Table 4. Actual values and prediction errors are also reported there. As can be seen, this equation predicts reasonably well the actual behavior of the nominal rate during this period. The mean error is small, only 3 basis points, and the root mean squared error is 0.4 percentage points. This regression outperforms a purely eight-order autoregressive model of the short rate. For the time series model, the mean prediction error is 51 basis points and the root mean squared error is 1.3 percentage points.<sup>21</sup>

<sup>21</sup> The interest rate regression with the real funds rate also outperforms the version in which changes in the real funds rate are replaced by changes in the real money supply. The mean error is 7 basis points and the root mean squared error is 1.22 percentage points.

**Table 6 Predictive Performance**

Year	Actual ( <i>A</i> )	Predicted ( <i>P</i> )	Error
Panel A: Actual and Predicted One-Year Treasury Bill Rate, 1979 to 1993			
1979	10.7	11.2	-0.5
1980	12.6	12.2	0.4
1981	14.5	13.5	1.0*
1982	11.9	12.2	-0.3
1983	9.7	9.9	-0.2
1984	10.9	10.8	0.1
1985	8.3	8.9	-0.6
1986	6.3	6.7	-0.3
1987	6.9	6.4	0.5
1988	7.8	7.9	-0.1
1989	8.5	9.0	-0.6
1990	7.8	7.5	0.3
1991	5.7	5.4	0.3
1992	3.9	-0.1	0.3
1993	3.5	3.2	0.2
Mean Error			0.03[-0.051]
RMSE			0.43[ 1.31 ]
Panel B: $A_t = d_0 + d_1 P_t + e_t$			
Model	$d_0$	$d_1$	$F$
Interest Rate Equation (A.1, Table 4)	0.05	1.0	0.1
AR(8)	0.48	0.89	3.1

\* The predictor error is twice the root mean squared error.

Notes: The predicted values reported above are generated using the regression A.1 of Table 4. AR(8) is an eight-order autoregressive process for explaining changes in the one-year Treasury bill rate. RMSE is the root mean squared error. Brackets contain the mean error and the RMSE value generated using the time series model. F is the F-statistic that tests the null hypothesis that  $d_0 = 0$  and  $d_1 = 1$ .

I evaluate further the predictive performance of the interest rate equation with the real funds rate from 1979:1 to 1993:4 by estimating regressions of the form

$$A_t = d_0 + d_1 P_t,$$

where  $A$  is the quarterly value of the short rate and  $P$  is the value predicted by the short-rate regression. If  $d_0 = 0$  and  $d_1 = 1$ , then regression forecasts are unbiased. As can be seen from Table 6, the coefficients  $d_0$  and  $d_1$  take values

0.05 and 1.0, respectively, for the interest rate equation (A.1 of Table 4) and 0.48 and 0.89, respectively, for the time series model. The hypothesis  $d_0 = 0$  and  $d_1 = 1$  is rejected for the time series model, but not for the interest rate equation.

Despite this overall good predictive performance, the interest rate equation with the real funds rate does not completely solve the puzzle of high rates during the early 1980s. The equation significantly underpredicts the level of the nominal rate in 1981. It predicts very well, however, the declines in nominal rates that have occurred since 1990.<sup>22</sup>

### **The Short End of the Term Structure Is Dominated by the Outlook for Inflation, Fed Policy, and the State of the Economy**

The empirical analysis of one-year Treasury bills summarized in Tables 4 and 5 indicate that in the short run, changes in the nominal rate depend largely upon changes in inflation, real GDP, and the real funds rate. I now argue that the same economic factors pretty much determine the behavior of the short end of the U.S. Treasury term structure. Table 7 presents short-run coefficients that appear on these economic determinants when alternative measures of short-to medium-term interest rates are used in the regression (2). As can be seen, those estimates indicate that changes in inflation, real GDP, and the real funds rate influence most the short end of the term structure. The short-run impact coefficients that appear on these variables steadily decline in size with the term to maturity. These results indicate that the yield curve at the short end of the term structure is dominated by the outlook for inflation, Fed policy, and the state of the economy.

## **3. CONCLUDING OBSERVATIONS**

It is a widely held view both in financial press and academic circles that the Federal Reserve influences short-term nominal interest rates. The empirical work presented here provides one perspective on the potential role of the Federal Reserve in determining short-term rates. The results indicate that it is inflation, not the real federal funds rate, that is at the source of long-run stochastic movements in the level of the nominal rate. Therefore, only through its control over inflation can the Federal Reserve exercise control over the level of the nominal rate in the long run. In other words, the Federal Reserve cannot permanently lower the nominal rate by affecting its real component.

---

<sup>22</sup> In contrast, the interest rate regression with the real money supply performs very poorly in predicting the behavior of the nominal rate during the early 1980s. Nor does it predict well the declines in nominal rates that have occurred since 1990.

**Table 7 Term Structure Effects**

Dependent Variable	Coefficients (t-values) on Contemporary Values of		
	$\Delta \dot{p}_t$	$\Delta RFR_t$	$\Delta \ln ry_t$
R3M	0.73 (8.9)	0.57 (7.5)	0.14 (3.5)
R1	0.71 (8.2)	0.62 (6.4)	0.11 (2.6)
R3	0.51 (6.9)	0.43 (4.5)	0.07 (1.4)
R5	0.41 (5.9)	0.33 (3.6)	0.05 (1.2)
R10	0.29 (5.0)	0.22 (2.8)	0.03 (0.9)

Notes: R3M is the three-month Treasury bill rate; R1, R3, R5, and R10 are the nominal yields on one-year, three-year, five-year, and ten-year Treasury bills. The regression equation used is equation A.1 reported in Table 4. In all of these regressions, the long-run coefficient that appears on the level of the inflation rate ( $\dot{p}$ ) is constrained to be unity, so that in the long run the nominal rate adjusts one for one with inflation. The regressions are estimated by instrumental variables over the sample period 1957:1 to 1994:3. Parentheses above contain t-values.

The results also indicate that in the short run, changes in the real funds rate have considerable effects on the nominal rate. Moreover, the real funds rate variable is exogenous in the short-run interest rate equations, indicating that short-run changes in the real federal funds rate do not respond to contemporaneous movements in the nominal rate. This finding, however, is quite consistent with the possibility that in reduced-form regressions like (5), changes in the real funds rate are highly correlated with lagged values of economic variables, including the nominal rate. Together, these results indicate that in the short run, the Federal Reserve influences the market as well as may be influenced by it.

Fiscal policy measures such as the deficit, government purchases, net taxes, and foreign capital inflows do not affect the short rate, once one controls for the effects of inflation, the real funds rate, and real growth. Overall, the results indicate that in the short run, the behavior of short-term nominal rates is dominated by the outlook for inflation, the funds rate, and the state of the economy.

The interest rate equation predicts reasonably well the nominal rate during the 1980s. It does not, however, completely solve the puzzle of high rates during the early 1980s, particularly in 1981. The analysis here indicates that a significant part of the rise in rates in the early 1980s can be attributed to the behavior of inflation and the real federal funds rate. Since over long periods these variables are endogenously determined, a complete explanation of the behavior of short rates must include an explanation of the behavior of these two variables.

---

---

**REFERENCES**

- Breusch, T. S., and A. R. Pagan. "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics," *Review of Economic Studies*, vol. 47 (January 1980), pp. 239–53.
- Clarida, Richard H., and Benjamin M. Friedman. "The Behavior of U.S. Short-Term Interest Rates Since October, 1979," *Journal of Finance*, vol. 39 (July 1984), pp. 671–82.
- \_\_\_\_\_. "Why Have Short-Term Interest Rates Been So High?" *Brookings Papers on Economic Activity*, 2:1983, pp. 553–78.
- Engle, Robert F. "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics," in Zvi Griliches and Michael D. Intriligator, eds., *Handbook of Econometrics*. Amsterdam, North-Holland, 1984.
- \_\_\_\_\_, and C. W. J. Granger. "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, vol. 55 (March 1987), pp. 251–76.
- Feinman, Joshua, and Richard D. Porter. "The Continuing Weakness in M2," Finance and Economic Discussion Paper #209. Washington: Board of Governors of the Federal Reserve System, September 1992.
- Fuller, W. A. *Introduction to Statistical Time Series*. New York: Wiley, 1976.
- Goodfriend, Marvin. "Interest Rate Policy and the Inflation Scare Problem: 1979–1992," Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 79 (Winter 1993), pp. 1–24.
- Hall, Alastair. "Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection," *Journal of Business and Economic Statistics*, vol. 12 (October 1994), pp. 461–70.
- Hausman, J. A. "Specification Tests in Econometrics," *Econometrica*, vol. 46 (November 1978), pp. 1251–71.
- Hendershott, Patric H., and Joe Peek. "Treasury Bill Rates in the 1970s and 1980s," *Journal of Money, Credit, and Banking*, vol. 24 (May 1992), pp. 195–214.
- Hetzel, Robert L., and Yash P. Mehra. "The Behavior of Money Demand in the 1980s," *Journal of Money, Credit, and Banking*, vol. 21 (November 1989), pp. 455–63.
- Johansen, Soren, and Katarina Juselius. "Maximum Likelihood Estimation and Inference on Cointegration—With Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics*, vol. 52 (May 1990), pp. 169–210.

- Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?" *Journal of Econometrics*, vol. 54 (October-December 1992), pp. 159–78.
- Maddala, G. S. *Introduction to Econometrics*. New York: Macmillan, 1988.
- Mehra, Yash P. "An Error-Correction Model of the Long-Term Bond Rate," Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 80 (Fall 1994), pp. 49–68.
- Peek, Joe and James A. Wilcox. "Monetary Policy Regimes and the Reduced Form for Interest Rates," *Journal of Money, Credit, and Banking*, vol. 19 (August 1987), pp. 273–91.
- Sargan, J. D. "Wages and Prices in the United Kingdom: A Study in Econometric Methodology" in P. E. Hart, G. Mills, and J. K. Whitaker, eds., *Econometric Analysis for National Economic Planning*. London: Butterworth Co., 1964.
- Sargent, Thomas J. "Commodity Price Expectations and the Interest Rate," *Quarterly Journal of Economics*, vol. 83 (February 1969), pp. 127–40.
- Sims, Christopher A. "Macroeconomics and Reality," *Econometrica*, vol. 48 (January 1980), pp. 1–49.
- Stock, James H., and Mark W. Watson. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, vol. 61 (July 1993), pp. 783–820.



# Quantitative Theory and Econometrics

---

Robert G. King

Quantitative theory uses simple, abstract economic models together with a small amount of economic data to highlight major economic mechanisms. To illustrate the methods of quantitative theory, we review studies of the production function by Paul Douglas, Robert Solow, and Edward Prescott. Consideration of these studies takes an important research area from its earliest days through contemporary real business cycle analysis. In these quantitative theoretical studies, economic models are employed in two ways. First, they are used to organize economic data in a new and suggestive manner. Second, models are combined with economic data to display successes and failures of particular theoretical mechanisms. Each of these features is present in each of the three studies, but to varying degrees, as we shall see.

These quantitative theoretical investigations changed how economists thought about the aggregate production function, i.e., about an equation describing how the total output of many firms is related to the total quantities of inputs, in particular labor and capital inputs. Douglas taught economists that the production function could be an important applied tool, as well as a theoretical device, by skillfully combining indexes of output with indexes of capital and labor input. Solow taught economists that the production function could not be used to explain long-term growth, absent a residual factor that he labeled technical progress. Prescott taught economists that Solow's residual was sufficiently strongly procyclical that it might serve as a source of economic

---

■ The author is A.W. Robertson Professor of Economics at the University of Virginia, consultant to the research department of the Federal Reserve Bank of Richmond, and a research associate of the National Bureau of Economic Research. This article was originally prepared as an invited lecture on economic theory at the July 1992 meeting of the Australasian Econometric Society held at Monash University, Melbourne, Australia. Comments from Mary Finn, Michael Dotsey, Tom Humphrey, Peter Ireland, and Sergio Rebelo substantially improved this article from its original lecture form. The views expressed are those of the author and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

fluctuations. More specifically, he showed that a real business cycle model driven by Solow's residuals produced fluctuations in consumption, investment, and output that broadly resembled actual U.S. business cycle experience.

In working through three key studies by Douglas, Solow, and Prescott, we focus on their design, their interrelationship, and the way in which they illustrate how economists learn from studies in quantitative theory. This learning process is of considerable importance to ongoing developments in macroeconomics, since the quantitative theory approach is now the dominant research paradigm being used by economists incorporating rational expectations and dynamic choice into small-scale macroeconomic models.

Quantitative theory is thus necessarily akin to applied econometric research, but its methods are very different, at least at first appearance. Indeed, practitioners of quantitative theory—notably Prescott (1986) and Kydland and Prescott (1991)—have repeatedly clashed with practitioners of econometrics. Essentially, advocates of quantitative theory have suggested that little is learned from econometric investigations, while proponents of econometrics have suggested that little tested knowledge of business cycle mechanisms is uncovered by studies in quantitative economic theory.

This article reviews and critically evaluates recent developments in quantitative theory and econometrics. To define quantitative theory more precisely, Section 1 begins by considering alternative styles of economic theory. Subsequently, Section 2 considers the three examples of quantitative theory in the area of the production function, reviewing the work of Douglas, Solow, and Prescott. With these examples in hand, Section 3 then considers how economists learn from exercises in quantitative theory.

One notable difference between the practice of quantitative theory and of econometrics is the manner in which the behavioral parameters of economic models are selected. In quantitative theoretical models of business cycles, for example, most behavioral parameters are chosen from sources other than the time series fluctuations in the macroeconomic data that are to be explained in the investigation. This practice has come to be called calibration. In modern macroeconometrics, the textbook procedure is to estimate parameters from the time series that are under study. Thus, this clash of methodologies is frequently described as “calibration versus estimation.”

After considering how a methodological controversy between quantitative theory and econometrics inevitably grew out of the rational expectations revolution in Section 4 and describing the rise of quantitative theory as a methodology in Section 5, this article then argues that the ongoing controversy cannot really be about “calibration versus estimation.” It demonstrates that classic calibration studies estimate some of their key parameters and classic estimation studies are frequently forced to restrict some of their parameters so as to yield manageable computational problems, i.e., to calibrate them. Instead, in Section 6, the article argues that the key practical issue is styles of “model evaluation,” i.e., about

the manner in which economists determine the dimensions along which models succeed or fail.

In terms of the practice of model evaluation, there are two key differences between standard practice in quantitative theory and econometrics. One key difference is indeed whether there are discernible differences between the activities of parameter selection and model evaluation. In quantitative theory, parameter selection is typically undertaken as an initial activity, with model evaluation being a separate secondary stage. By contrast, in the dominant dynamic macro-econometric approach, that of Hansen and Sargent (1981), parameter selection and model evaluation are undertaken in an essentially simultaneous manner: most parameters are selected to maximize the overall fit of the dynamic model, and a measure of this fit is also used as the primary diagnostic for evaluation of the theory. Another key difference lies in the breadth of model implications utilized, as well as the manner in which they are explored and evaluated. Quantitative theorists look at a narrow set of model implications; they conduct an informal evaluation of the discrepancies between these implications and analogous features of a real-world economy. Econometricians typically look at a broad set of implications and use specific statistical methods to evaluate these discrepancies.

By and large, this article takes the perspective of the quantitative theorist. It argues that there is a great benefit to choosing parameters in an initial stage of an investigation, so that other researchers can readily understand and criticize the attributes of the data that give rise to such parameter estimates. It also argues that there is a substantial benefit to limiting the scope of inquiry in model evaluation, i.e., to focusing on a set of model implications taken to display central and novel features of the operation of a theoretical model economy. This limitation of focus seems appropriate to the current stage of research in macroeconomics, where we are still working with macroeconomic models that are extreme simplifications of macroeconomic reality.

Yet quantitative theory is not without its difficulties. To illustrate three of its limitations, Section 7 of the article reconsiders the standard real business cycle model, which is sometimes described as capturing a dominant component of postwar U.S. business cycles (for example, by Kydland and Prescott [1991] and Plosser [1989]). The first limitation is one stressed by Eichenbaum (1991): since it ignores uncertainty in estimated parameters, a study in quantitative theory cannot give any indication of the statistical confidence that should be placed in its findings. The second limitation is that quantitative theory may direct one's attention to model implications that do not provide much information about the endogenous mechanisms contained in the model. In the discussion of these two limitations, the focus is on a "variance ratio" that has been used, by Kydland and Prescott (1991) among others, to suggest that a real business cycle arising from technology shocks accounts for three-quarters of postwar U.S. business cycle fluctuations in output. In discussing the practical

importance of the first limitation, Eichenbaum concluded that there is “enormous” uncertainty about this variance ratio, which he suggested arises because of estimation uncertainty about the values of parameters of the exogenous driving process for technology. In terms of the second limitation, the article shows that a naive model—in which output is driven only by production function residuals without any endogenous response of factors of production—performs nearly as well as the standard quantitative theoretical model according to the “variance ratio.” The third limitation is that the essential focus of quantitative theory on a small number of model implications may easily mean that it misses crucial failures (or successes) of an economic model. This point is made by Watson’s (1993) recent work that showed that the standard real business cycle model badly misses capturing the “typical spectral shape of growth rates” for real macroeconomic variables, including real output. That is, by focusing on only a small number of low-order autocovariances, prior investigations such as those of Kydland and Prescott (1982) and King, Plosser, and Rebelo (1988) simply overlooked the fact that there is an important predictable output growth at business cycle frequencies.

However, while there are shortcomings in the methodology of quantitative theory, its practice has grown at the expense of econometrics for a good reason: it provides a workable vehicle for the systematic development of macroeconomic models. In particular, it is a method that can be used to make systematic progress in the current circumstances of macroeconomics, when the models being developed are still relatively incomplete descriptions of the economy. Notably, macroeconomists have used quantitative theory in recent years to learn how the business cycle implications of the basic neoclassical model are altered by a wide range of economic factors, including fiscal policies, international trade, monopolistic competition, financial market frictions, and gradual adjustment of wages and prices.

The main challenge for econometric theory is thus to design procedures that can be used to make similar progress in the development of macroeconomic models. One particular aspect of this challenge is that the econometric methods must be suitable for situations in which we know before looking at the data that the model or models under study are badly incomplete, as we will know in most situations for some time to come. Section 8 of the article discusses a general framework of model-building activity within which quantitative theory and traditional macroeconometric approaches are each included. On this basis, it then considers some initial efforts aimed at developing econometric methods to capture the strong points of the quantitative theory approach while providing the key additional benefits associated with econometric work. Chief among these benefits are (1) the potential for replication of the outcomes of an empirical evaluation of a model or models and (2) an explicit statement of the statistical reliability of the results of such an evaluation.

In addition to providing challenges to econometrics, Section 9 of the article shows how the methods of quantitative theory also provide new opportunities for applied econometrics, using Friedman's (1957) permanent income theory of consumption as a basis for constructing two more detailed examples. The first of these illustrates how an applied econometrician may use the approach of quantitative theory to find a powerful estimator of a parameter of interest. The second of these illustrates how quantitative theory can aid in the design of informative descriptive empirical investigations.

In macroeconometric analysis, issues of identification have long played a central role in theoretical and applied work, since most macroeconomists believe that business fluctuations are the result of a myriad of causal factors. Quantitative theories, by contrast, typically are designed to highlight the role of basic mechanisms and typically identify individual causal factors. Section 10 considers the challenges that issues of identification raise for the approach of quantitative theory and the recent econometric developments that share its model evaluation strategy. It suggests that the natural way of proceeding is to compare the predictions of a model or models to characteristics of economic data that are isolated with a symmetric empirical identification.

The final section of the article offers a brief summary as well as some concluding comments on the relationship between quantitative theory and econometrics in the future of macroeconomic research.

## **1. STYLES OF ECONOMIC THEORY**

The role of economic theory is to articulate the mechanisms by which economic causes are translated into economic consequences. By requiring that theorizing is conducted in a formal mathematical way, economists have assured a rigor of argument that would be difficult to attain in any other manner. Minimally, the process of undertaking a mathematical proof lays bare the essential linkages between assumptions and conclusions. Further, and importantly, mathematical model-building also has forced economists to make sharp abstractions: as model economies become more complex, there is a rapidly rising cost to establishing formal propositions. Articulation of key mechanisms and abstraction from less important ones are essential functions of theory in any discipline, and the speed at which economic analysis has adopted the mathematical paradigm has led it to advance at a much greater rate than its sister disciplines in the social sciences.

If one reviews the history of economics over the course of this century, the accomplishments of formal economic theory have been major. Our profession developed a comprehensive theory of consumer and producer choice, first working out static models with known circumstances and then extending it to dynamics, uncertainty, and incomplete information. Using these developments, it established core propositions about the nature and efficiency of general equilibrium with interacting consumers and producers. Taken together,

the accomplishments of formal economic theory have had profound effects on applied fields, not only in the macroeconomic research that will be the focal point of this article but also in international economics, public finance, and many other areas.

The developments in economic theory have been nothing short of remarkable, matched within the social sciences perhaps only by the rise of *econometrics*, in which statistical methods applicable to economic analysis have been developed. For macroeconomics, the major accomplishment of econometrics has been the development of statistical procedures for the estimation of parameters and testing of hypotheses in a context where a vector of economic variables is dynamically interrelated. For example, macroeconomists now think about the measurement of business cycles and the testing of business cycle theories using an entirely different statistical conceptual framework from that available to Mitchell (1927) and his contemporaries.<sup>1</sup>

When economists discuss economic theory, most of us naturally focus on *formal theory*, i.e., the construction of a model economy—which naturally is a simplified version of the real world—and the establishment of general propositions about its operation. Yet, there is another important kind of economic theory, which is the use of much more simplified model economies to organize economic facts in ways that change the focus of applied research and the development of formal theory. *Quantitative theory*, in the terminology of Kydland and Prescott (1991), involves taking a more detailed stand on how economic causes are translated into economic consequences. Quantitative theory, of course, embodies all the simplifications of abstract models of formal theory. In addition, it involves making (1) judgments about the quantitative importance of various economic mechanisms and (2) decisions about how to selectively compare the implications of a model to features of real-world economies. By its very nature, quantitative theory thus stands as an intermediate activity to formal theory and the application of econometric methods to evaluation of economic models.

A decade ago, many economists thought of quantitative theory as simply the natural first step in a progression of research activities from formal theory to econometrics, but there has been a hardening of viewpoints in recent years. Some argue that standard econometric methods are not necessary or are, in fact, unhelpful; quantitative theory is sufficient. Others argue that one can learn little from quantitative theory and that the only source of knowledge about important economic mechanisms is obtained through econometrics. For those of us that honor the traditions of both quantitative theory and econometrics, not only did the onset of this controversy come as a surprise, but its depth and persistence

---

<sup>1</sup> In particular, we now think of an observed time series as the outcome of a stochastic process, while Mitchell and his contemporaries struggled with how to best handle the evident serial dependence that was so inconsistent with the statistical theory that they had at hand.

also were unexpected. Accordingly, the twin objectives of this paper are, first, to explore why the events of recent years have led to tensions between practitioners of quantitative theory and econometrics and, second, to suggest dimensions along which the recent controversy can lead to better methods and practice.

## 2. EXAMPLES OF QUANTITATIVE THEORY

This section discusses three related research topics that take quantitative theory from its earliest stages to the present day. The topics all concern the production function, i.e., the link between output and factor inputs.<sup>2</sup>

### The Production Function and Distribution Theory

The production function is a powerful tool of economic analysis, which every first-year graduate student learns to manipulate. Indeed, the first example that most economists encounter is the functional form of Cobb and Douglas (1928), which is also the first example studied here. For contemporary economists, it is difficult to imagine that there once was a time when the notion of the production function was controversial. But, 50 years after his pioneering investigation, Paul Douglas (1976) reminisced:

Critics of the production function analysis such as Horst Mendershausen and his mentor, Ragnar Frisch, . . . urged that so few observations were involved that any mathematical relationship was purely accidental and not causal. They sincerely believed that the analysis should be abandoned and, in the words of Mendershausen, that all past work should be torn up and consigned to the wastepaper basket. This was also the general sentiment among senior American economists, and nowhere was it held more strongly than among my senior colleagues at the University of Chicago. I must admit that I was discouraged by this criticism and thought of giving up the effort, but there was something which told me I should hold on. (P. 905)

The design of the investigation by Douglas was as follows. First, he enlisted the assistance of a mathematician, Cobb, to develop a production function with specified properties.<sup>3</sup> Second, he constructed indexes of physical capital and labor input in U.S. manufacturing for 1899–1922. Third, Cobb and Douglas estimated the production function

---

<sup>2</sup> A replication diskette available from the author contains computer programs used to produce the four figures that summarize the Douglas, Solow, and Prescott studies in this section, as well as to produce additional figures discussed below. That diskette also contains detailed background information on the data.

<sup>3</sup> But the mathematics was not the major element of the investigation. Indeed, it was not even novel, as Tom Humphrey has pointed out to the author, having been previously derived in the realm of pure theory by Wickcell (1934, pp. 101–28).

$$Y_t = A N_t^\alpha K_t^{1-\alpha}. \quad (1)$$

In this specification,  $Y_t$  is the date  $t$  index of manufacturing output,  $N_t$  is the date  $t$  index of employed workers, and  $K_t$  is the date  $t$  index of the capital stock. The least squares estimates for 1899–1922 were  $\hat{A} = 1.01$  and  $\hat{\alpha} = 0.73$ . Fourth, Cobb and Douglas performed a variety of checks of the implications of their specification. These included comparing their estimated  $\hat{\alpha}$  to measures of labor’s share of income, which earlier work had shown to be reasonably constant through time. They also examined the extent to which the production function held for deviations from trend rather than levels. Finally, they examined the relationship between the model’s implied marginal product of labor ( $\alpha Y/N$ ) and a measure of real wages that Douglas (1926) had constructed in earlier work.

The results of the Cobb-Douglas quantitative theoretical investigation are displayed in Figure 1. Panel A provides a plot of the data on output, labor, and capital from 1899 to 1922. All series are benchmarked at 100 in 1899, and it is notable that capital grows dramatically over the sample period. Panel B displays the fitted production function,  $\hat{Y} = \hat{A}N^{\hat{\alpha}}K^{1-\hat{\alpha}}$ , graphed as a dashed line and manufacturing output,  $Y$ , graphed as a solid line. As organized by the production function, variations in the factors  $N$  and  $K$  clearly capture the upward trend in output.<sup>4</sup>

With the Cobb-Douglas study, the production function moved from the realm of pure theory—where its properties had been discussed by Clark (1889) and others—to that of quantitative theory. In the hands of Cobb and Douglas, the production function displayed an ability to link (1) measures of physical output to measures of factor inputs (capital and labor) and (2) measures of real wages to measures of average products. It thus became an engine of analysis for applied research.

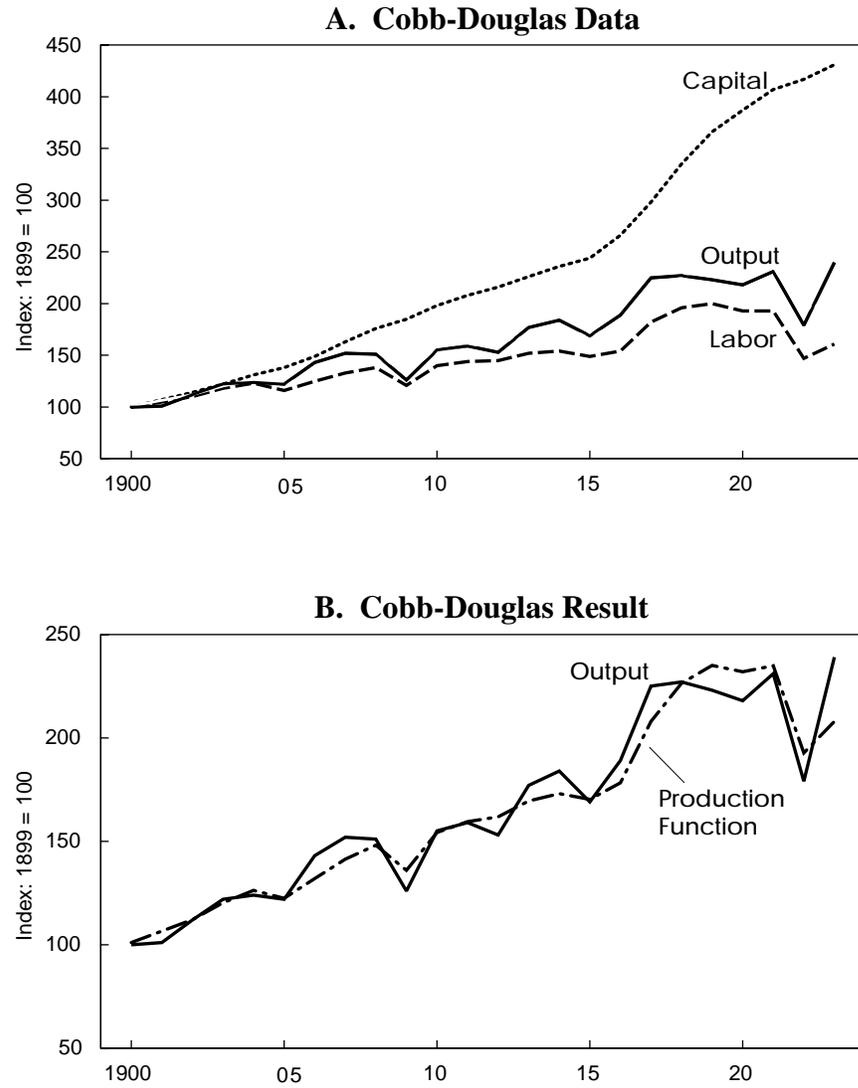
But the authors took care to indicate that their quantitative production function was not to be viewed as an exact model of economic activity for three reasons. Cobb and Douglas recognized that they had neglected technical progress, but they were uncertain about its magnitude or measurability. They also viewed their measure of labor input only as a first step, taken because there was relatively poor data on hours per worker. Finally, and importantly, they found it unsurprising that errors in the production function were related

---

<sup>4</sup> The Cobb-Douglas methodology was also applied to cross-sections of industries in follow-up investigations (as reviewed in Douglas [1948]). Many of these subsequent cross-section investigations used the high-quality New South Wales and Victoria data from Australia. These cross-section studies provided further buttressing of Douglas’s perspective that the production function was a useful applied tool, although from a modern perspective they impose too much homogeneity of production technique across industries.

**Figure 1**

---



to business cycles. They pointed out that during slack years the production function overpredicted output because reductions in hours and capital utilization were not measured appropriately. Correspondingly, years of prosperity were underpredicted.

### The Production Function and Technical Progress

Based on the Cobb and Douglas (1928) investigations, it was reasonable to think that (1) movements in capital and labor accounted for movements in output, both secularly and over shorter time periods and (2) wages were equated with marginal products computed from the Cobb-Douglas production function.

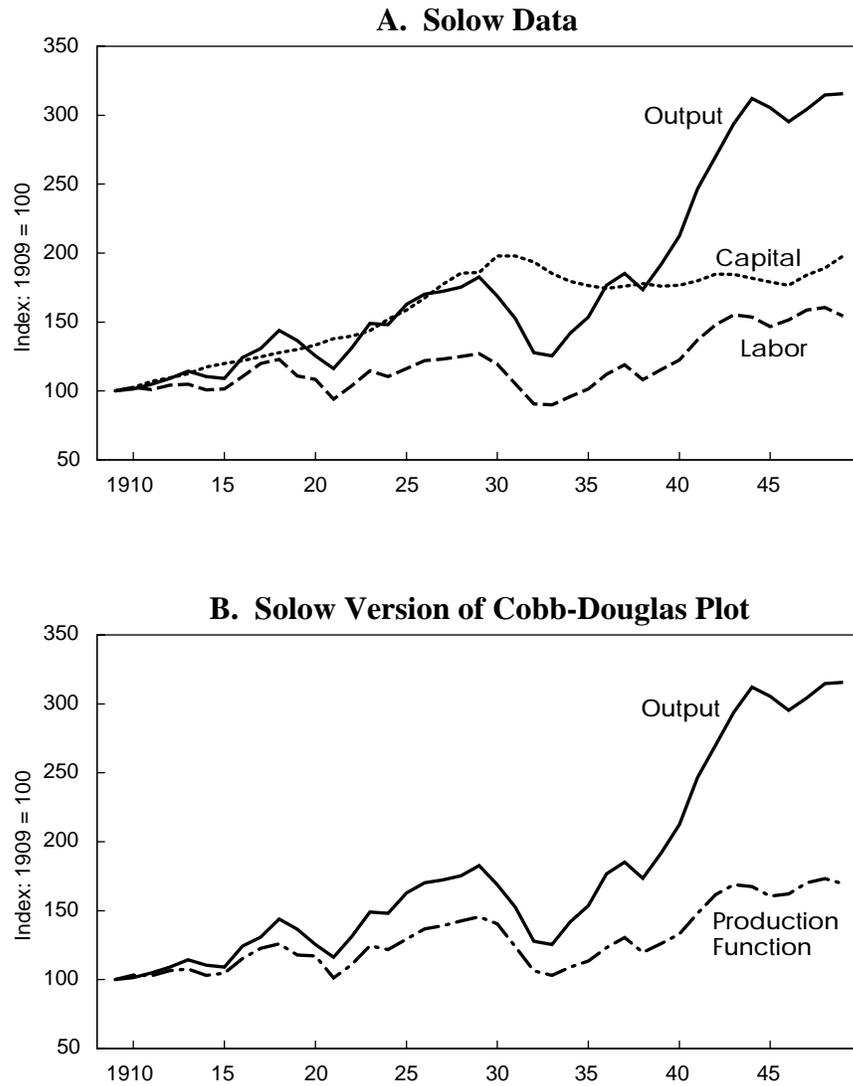
Solow's (1957) exercise in quantitative theory provided a sharp contradiction of the first conclusion from the Cobb-Douglas studies, namely, that output movements were largely determined by movements in factor inputs. Taking the marginal productivity theory of wages to be true, Solow used the implied value of labor's share to decompose the growth in output per man-hour into components attributable to capital per man-hour and a residual:

$$y_t - n_t = s_k(k_t - n_t) + a_t, \quad (2)$$

where  $y$  is the growth rate of output;  $n$  is the growth rate of labor input;  $k$  is the growth rate of capital input; and  $a$  is the "Solow residual," taken to be a measure of growth in total factor productivity. The share is given by the competitive theory, i.e.,  $s_k$  is the share of capital income. Solow allowed the share weight in (2) to vary at each date, but this feature leads to essentially the same outcomes as simply imposing the average of the  $s_k$ . Hence, a constant value of  $s_k$  or  $\alpha = (1 - s_k)$  is used in the construction of the figures to maintain comparability with the work of Cobb and Douglas. Solow emphasized the trend (low frequency) implications of the production function by looking at the long-term average contribution of growth in capital per worker ( $k - n$ ) to growth in output per worker ( $y - n$ ). Over his 50-year sample period, output roughly doubled. But only one-eighth of the increase was due to capital; the remaining seven-eighths were due to technical progress.

Another way of looking at this decomposition is to consider the trends in the series. Figure 2 displays the nature of Solow's quantitative theoretical investigation in a manner comparable to Figure 1's presentation of the Cobb-Douglas data and results. Panel A is a plot of the aggregate data on output, capital, and labor during 1909–49. For comparability with Figure 1, the indexes are all scaled to 100 in the earliest year. The striking difference between Panel A of the two figures is that capital grows more slowly than output in Solow's data while it grows more rapidly than output in the Cobb-Douglas data. In turn, Panel B of Figure 2 displays the production function  $\hat{Y}_t = A N_t^\alpha K_t^{1-\alpha}$ , with  $\alpha$  chosen to match data on the average labor share and  $A$  chosen so that the fit is correct in the initial period. In contrast to Panel B of Figure 1, the production

**Figure 2**



function, which is the dashed line, does not capture much of the trend variation in output,  $Y$ .<sup>5</sup>

It is likely that these results were unexpected to Solow, who had just completed development of a formal theoretical model that stressed capital deepening in his 1956 classic, “A Contribution to the Theory of Economic Growth.” Instead of displaying the importance of capital deepening, as had the Cobb-Douglas investigations, Solow’s adventure into quantitative theory pointed toward the importance of a new feature, namely, total-factor-augmenting technical progress. There had been hints of this result in other research just before Solow’s, but none of it had the simplicity, transparency, or direct connection to formal economic theory that marked Solow’s work.<sup>6</sup>

### Cyclical Implications of Movements in Productivity

By using the production function as an applied tool in business cycle research, Prescott (1986) stepped into an area that both Douglas and Solow had avoided. Earlier work by Kydland and Prescott (1982) and Long and Plosser (1983) had suggested that business cycle phenomena could derive from an interaction of tastes and technologies, particularly if there were major shocks to technology. But Prescott’s (1986) investigation was notable in its use of Solow’s (1957) procedure to measure the extent of variations in technology. With such a measurement of technology or productivity shocks in hand, Prescott explored how a neoclassical model behaved when driven by these shocks, focusing on the nature of business cycles that arose in the model.

Figure 3 displays the business cycle behavior of postwar U.S. real gross national product and a related productivity measure,

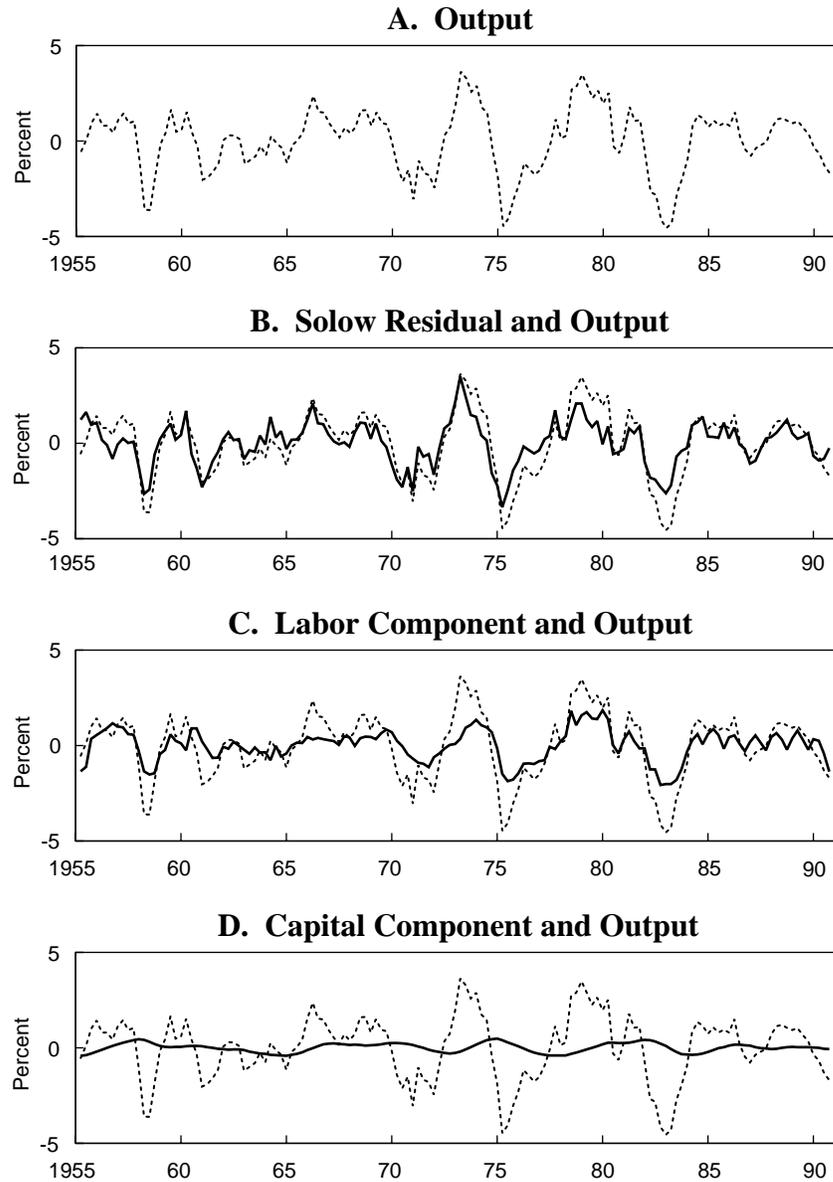
$$\log A_t = \log Y_t - \alpha \log N_t + (1 - \alpha) \log K_t. \quad (3)$$

This figure is constructed by first following a version of the Solow (1957) procedure to extract a productivity residual and then filtering the outcomes with the procedure of Hodrick and Prescott (1980). These “business cycle components” are very close to deviations from a lengthy centered moving

---

<sup>5</sup> It is an interesting question as to why Douglas’s production function results differed so much from Solow’s, particularly since the latter’s have largely been sustained in later work. To begin, Douglas’s work looked only at 23 years, from 1899 to 1922, and the study was restricted to manufacturing. During this interval, there was substantial growth in manufacturing capital (as displayed in Figure 1). Ultimately, it is this growth in capital that underlies the different conclusions from the Solow investigation. The capital series itself was created by Cobb and Douglas using interpolation between U.S. census estimates in 1899, 1904, and 1922.

<sup>6</sup> The construction of Panel B of Figure 2 actually departs somewhat from the procedure used by Solow. That is, the indexes of capital and labor from Panel A are used directly, while Solow corrected capital for utilization by multiplying the capital stock by the unemployment rate. For the trend properties of Figure 2, this difference is of little importance, but for the business cycle issue to which we will now turn, it is likely of greater importance.

**Figure 3**

Notes: Panel A displays the behavior of output relative to trend, measured as discussed in the text. Panels B, C, and D show the decomposition of cyclical output into components attributable to the Solow residual, labor input, and capital input (the solid lines). In each of these panels, cyclical output (the dotted line) is also displayed to provide the reader with a reference point.

average; they also closely correspond to the results if a band-pass filter is applied to each time series.<sup>7</sup> Panel A displays the business cycle variation in output, which has a sample standard deviation of 1.69 percent. This output measure is also the solid line in the remainder of the panels of Figure 3. It is clear from Panel B of this figure that the productivity residual is strongly procyclical: the correlation between the series in Panel B is 0.88. The labor component of output—the filtered version of  $\alpha \log N_t$ —is also strongly procyclical: the correlation between the series in Panel C is 0.85. Finally, Panel D shows there is little cyclical variation in capital input, the filtered series  $(1 - \alpha) \log K_t$ : its correlation with output is  $-0.12$ .

Unlike the earlier studies, real business cycle (RBC) exercises in quantitative theory are explicitly general equilibrium: they take measurements of individual causes and trace the consequences for a number of macroeconomic variables. The simplest RBC model is essentially the growth model of Solow (1956), modified to include optimal choice of consumption and labor input over time. In the Solow model, long-term growth in output per capita must come about mainly from productivity growth. Thus, RBC models may be viewed as exploring the business cycle implications of a feature widely agreed to be important for lower frequency phenomena.

In his quantitative theoretical investigation, Prescott thus assumed that productivity in a model economy was generated by a stochastic process fit to observations on  $\log A_t$ . He then explored how consumption, investment, output, and input would evolve in such an economy.<sup>8</sup> There were two striking outcomes of Prescott's experiment, which have been frequently highlighted by adherents of real business cycle theory.

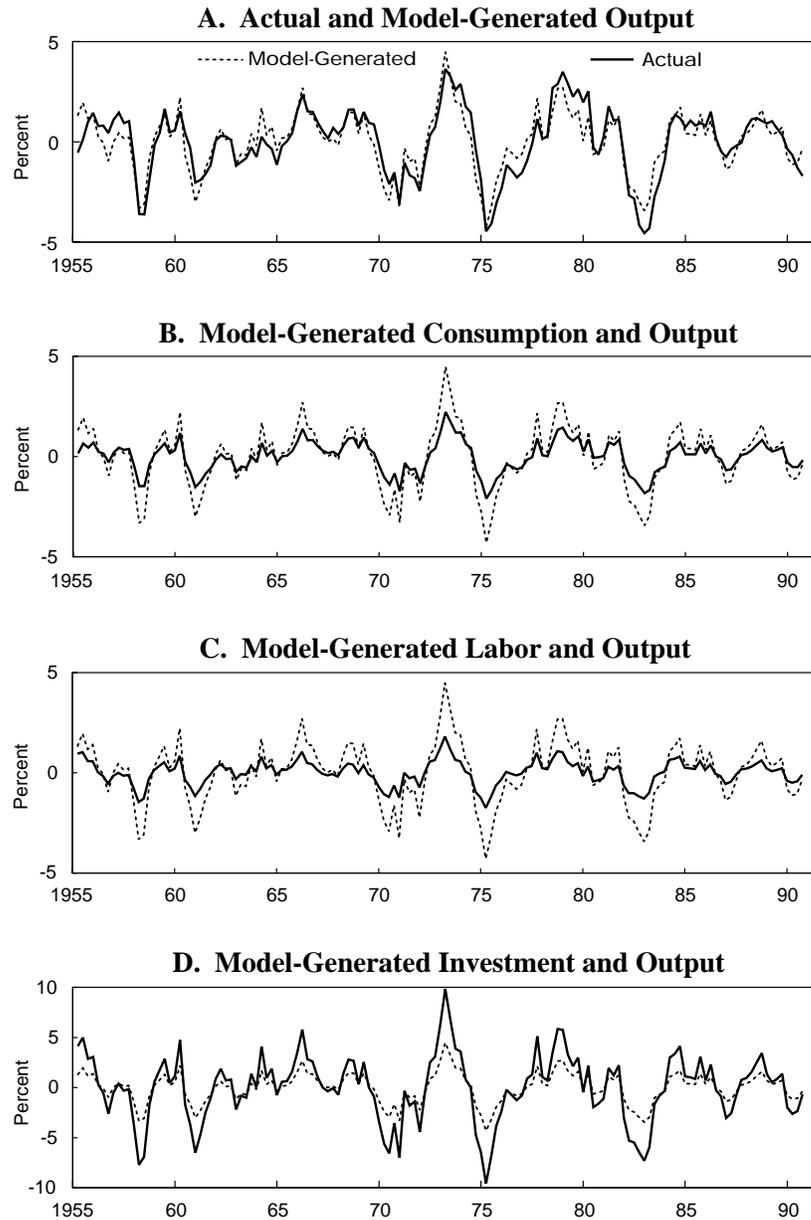
The first, much-stressed finding is that a large fraction of variation in output is explained by such a model, according to a statistical measure that will be explained in more detail below. Figure 4 shows that an RBC model's output (the dotted line) is closely related to actual output (the solid line), albeit with

---

<sup>7</sup> For some additional discussion of the Hodrick and Prescott filter, see King and Rebelo (1993). Recent related work, Baxter and King (1995), showed that Hodrick-Prescott filtered time series on gross national product resembles (1) the deviation from a centered five-year (equal-weight) moving average and (2) a particular approximate high-pass filter, which is designed to eliminate slow-moving components and to leave intact components with periodicities of eight years or less.

In terms of the discussion in Section 7 below, it is useful to have a version of "cyclical variation in productivity" present in Figure 3. However, for Figure 4 below, the original time series on  $\log A_t$  was fed into the dynamic model, not the filtered one. Since optimal dynamic responses imply that "shocks" at one frequency are generally translated to all others, it would be hard to justify using filtered  $\log A_t$  as a forcing process.

<sup>8</sup> To explore the consequences of such "technology shocks," the current discussion will assume that the model economy is driven by the actual sequence of  $\log A_t$ , which is the strategy used in Plosser's (1989) version of this quantitative theoretical experiment. Prescott (1986) instead assumed that the technology shocks,  $\varepsilon_t$ , were drawn from a random-number generator and explored the properties of simulated economies.

**Figure 4**

Notes: Panel A displays U.S. cyclical output (the solid line, also previously reported in Panel A of Figure 3) as well as model cyclical output (the dotted line). The model cyclical output was generated by simulating a basic real business cycle, with the U.S. productivity residual as a driving process. Panels B, C, and D show the comparable simulation results for consumption, labor input, and investment (the solid lines). In each of these panels, model cyclical output (the dotted line) is also displayed to provide the reader with a reference point.

somewhat smaller amplitude.<sup>9</sup> In particular, the variance of output in the model is 2.25 percent with the variance of actual output being 2.87 percent, so that the ratio of variances is 0.78. Further, the correlation between the actual and model output series is 0.90. In this sense, the RBC model captures a major part of economic fluctuations.

The second, much-stressed finding is that the model economy captures other features of observed business cycles, notably the fact that consumption is smoother than output and that investment is more volatile than output, as shown in the additional panels of Figure 4.<sup>10</sup> Moments from the model economy that measure consumption's relative volatility—the ratio of its standard deviation to that of output—are about as large as the comparable ratio for postwar U.S. data.

The design of Prescott's (1986) investigation was solidly in the tradition of Douglas and Solow: a very simple theoretical model was constructed and it was shown to have very surprising empirical properties. Prior to Prescott's study, it was possible to dismiss the RBC interpretation of postwar U.S. economic fluctuations out of hand. One needed only to make one of two arguments that were widely accepted at the time. The first common argument was that there was no evidence of large, procyclical productivity shocks. The second was that equilibrium macroeconomic models, even with productivity shocks, were evidently inconsistent with major features of the U.S. macroeconomic time series, such as the procyclicality of labor input and the high volatility of investment. After Prescott's investigation, as Rogoff (1986) pointed out, it was no longer possible to do this: it was necessary to undertake much more subtle critiques of the RBC interpretation of economic fluctuations.

### 3. HOW WE LEARN FROM QUANTITATIVE THEORY

How, in general, do economists learn from quantitative theory? Taken together, the preceding examples help answer this question.

#### Theory as Abstraction

An essential characteristic of formal economic theory is that it involves abstraction. In constructing a theoretical model, we focus on one set of mechanisms and variables; we neglect factors that are hypothesized to be secondary or are

---

<sup>9</sup> The model economy is the one discussed in King, Plosser, and Rebelo (1988), with parameter choices detailed there. The productivity process is assumed to be a random walk, and the changes in the Solow residual (relative to the mean change) are taken to be productivity shocks. The dynamic model is then used to produce a time series on output, consumption, investment, and labor input; the resulting simulated outcomes are then run through the Hodrick-Prescott filter.

<sup>10</sup> Less realistically, in the specific simple version of the model underlying Figure 4, labor is less volatile than output. However, as Prescott (1986) discussed, there are several ways to increase labor volatility to roughly the same level as output volatility without major changes in the other features of the model.

simply not the focus of the investigation. Quantitative theory similarly involves the process of making sharp abstractions, but it also involves using the theory as an empirical vehicle. In particular, we ask whether an abstraction provides a compelling organization of key facts in an area of economic inquiry.

For Douglas, there were two key questions. First, was there a systematic empirical relationship between factor inputs and outputs of the type suggested by the formal economic theories of Clark and others? Second, was this relationship also consistent with the competitive theory of distribution, i.e., did real wages resemble marginal products constructed from the production function? Prior to Douglas's work, there was little reason to think that indexes of production, employment, and capital would be linked together in the way suggested by the theoretical production function. After Douglas's work, it was hard to think that there were not empirical laws of production to be discovered by economists. Further, Douglas's work also indicated a strong empirical relationship between real wages and average products, i.e., a rough constancy of payments to labor as a fraction of the value of output. Thus, Douglas's work made the theory of production and the theory of competitive determination of factor income payments operational in ways that were crucial for the development of economics. To be sure, as Douglas (1976) pointed out, his simple theory did not work exactly, but it suggested that there was sufficient empirical content to warrant substantial additional work.

One measure of Douglas's success on both fronts is the character of Solow's quantitative theoretical investigation. Solow simply took as *given* that many economists would accept both abstractions as useful organizing principles: he assumed both that the aggregate production function was an organizing principle for data and that the marginal productivity theory of wages was appropriate. But he then proceeded to challenge one of the substantive conclusions of Douglas's analysis, namely, that most of the movement in output can be explained by movements in factors of production. In particular, after reading Solow's article, one finds it hard not to believe that other factors besides physical capital formation are behind the secular rise in wage rates.<sup>11</sup> Solow's reorganization of

---

<sup>11</sup> That is, as it did for Douglas's indexes, input from physical capital would need to grow much faster than final output if it is to be responsible for the secular rise in wage rates. If capital input is proportional to capital stock, the capital stock series that Solow used would have to be very badly measured for this to be the case. However, an important recent exercise in quantitative theory by Greenwood, Hercowitz, and Krusell (1992) suggested that, for the United States over the postwar period, there has been a major mismeasurement of capital input deriving from lack of incorporation of quality change. Remarkably, when Greenwood, Hercowitz, and Krusell corrected investment flows for the amount of quality change suggested by Gordon's (1990) study, capital input did grow substantially faster than output, sufficiently so that it accounted for about two-thirds of growth in output per man-hour. However, one interpretation of Greenwood, Hercowitz, and Krusell's results is that these authors produced a measurement of technical progress, albeit a more direct one than that of Solow. Thus, the net effect of Greenwood, Hercowitz, and Krusell's study is likely to be that it enhances the perceived importance of technical progress.

the production function facts spurred the development of the growth accounting literature—as summarized by Maddison (1981)—and continues to provide major challenges to economists’ thinking about growth and development problems.

Prescott’s (1986) analysis of the role of productivity fluctuations in business cycles builds directly on the prior investigations of Douglas and Solow, yielding two key findings. The first finding is that there is important procyclical variation in Solow residuals. The second finding is that cyclical variations in productivity can explain cyclical variations in other macroeconomic quantities. The first component of this investigation in quantitative theory has had sufficient impact that even new Keynesian macroeconomists like Blanchard and Fischer (1989) list procyclical productivity as one of the key stylized facts of macroeconomics in the opening chapter of their textbook. The second has stimulated a major research program into the causes and consequences of cyclical variations in productivity.

In each of these cases, the investigation in quantitative theory had a surprising outcome: prior to each investigation, there was little reason to think that a specific economic mechanism was of substantial empirical importance. Prior to Douglas’s work, there was little reason to look for empirical connections between outputs and quantities of factor inputs. Prior to Solow’s, there was little reason to think that technical progress was an important contributor to economic growth. Prior to Prescott’s, there was little reason to think that procyclical variations in productivity were important for the business cycle. Each investigation substantially changed the views of economists about the nature of economic mechanisms.

### **Challenges to Formal Theory**

Quantitative theory issues important challenges to formal theory; each of the three examples contains such a challenge.

*Flexible and Tractable Function Forms:* The Cobb-Douglas investigations led to a search for alternative functional forms that could be used in such investigations, but that did not require the researcher to impose all of the restrictions on substitutability, etc., associated with the Cobb-Douglas specification.

*Factor Income Payments and Technical Progress:* Solow’s investigation reinforced the need for theories of distribution of income to different factor inputs in the presence of technical progress. Initially, this was satisfied by the examination of different forms of technical progress. But more recently, there has been much attention given to the implications of technical progress for the competitive paradigm (notably in Romer [1986, 1987]).

*Cyclical Variation in Productivity:* Prescott’s investigation showed dramatically how macroeconomic theories without productivity variation typically have important counterfactual productivity implications (i.e., they imply that output per man-hour is countercyclical). Recent work has explored a range of theories

designed to generate procyclical output per man-hour, including (1) imperfect competition, (2) external effects, and (3) time-varying capacity utilization.

#### **4. HOW MACROECONOMETRICS FAILED (TWICE)**

In recent years, the methods of quantitative theory have become an increasingly used tool in applied research in macroeconomics. This growth results from two very different historical episodes during which macroeconometrics failed, though for very different reasons. In the first of these episodes, econometric practice was subject to *too little* discipline from economic and econometric theory in the development of large-scale macroeconomic models. While the behavioral equations were sometimes motivated by a prevailing economic theory, in practice generous incorporation of lags and dummy variables gave rise to essentially unrestricted empirical specifications. As a result, most models fit the historical data very well, even though they had much more modest success in short-term forecasting. In the second of these episodes, econometric practice was subject to *too much* discipline from economic and econometric theory: rational expectations econometrics produced tightly restricted dynamic macroeconomic models and concluded that no models fit the data very well.

##### **Keynesian Econometrics: The Promise**

The research program of Koopmans and his coworkers at the Cowles Foundation, as reported in Hood and Koopmans (1953), provided a formal structure for Keynesian macroeconometrics, the dynamic simultaneous equations model. This econometric structure utilized economic theory in a very important manner: it stressed that theory could deliver the exclusion restrictions that were necessary for identification of the behavioral parameters. Initially, the Keynesian macroeconomic models were of sufficiently small scale that they could be readily studied by other researchers, like those of Klein (1950) and Klein and Goldberger (1955). Klein's (1950) work is particularly notable in terms of its insistence on developing the relevant behavioral theory for each structural equation. The promise of Keynesian macroeconometrics was that empirical macroeconomic models were to be a research laboratory, the basis for systematic refinement of theoretical and empirical specifications. They were also to be a device for concrete discussion of appropriate policy actions and rules.

##### **Keynesian Econometrics: The Failure**

By the mid-1970s, the working Keynesian macroeconomic models had evolved into extremely large systems; this growth occurred so that their builders could readily answer a very wide range of questions posed by business and government policymakers. In the process of this evolution, they had strayed

far from the promise that early developments had suggested: they could not be readily studied by an individual researcher nor was it possible to determine the components of the model that led to its operating characteristics. For example, in response to most policy and other disturbances, most models displayed outcomes that cycled for many years, but it proved difficult to understand why this was the case. The consequent approach was for individual academic researchers to concentrate on refinement of a particular structural equation—such as the money demand function or the consumption function—and to abstain from analysis of complete models. But this strategy made it difficult to discuss many central issues in macroeconomics, which necessarily involved the operation of a full macroeconomic system.

Then, in the mid-1970s, two major events occurred. There was worldwide “stagflation,” with a coexistence of high inflation and high unemployment. Major macroeconomic models simply got it very wrong in terms of predicting this pattern of events. In addition, Lucas’s (1976) famous critique of econometric policy evaluation highlighted the necessity of producing macroeconomic models with dynamic choice and rational expectations. Taken together with the prior inherent difficulties with macroeconomic models, these two events meant that interest in large-scale macroeconomic models essentially evaporated.

Lucas’s (1976) critique of macroeconomic models had two major components. First, Lucas noted that many behavioral relations—investment, consumption, labor supply, etc.—depended in a central way on expectations when derived from relevant dynamic theory. Typical macroeconomic models either omitted expectations or treated them as essentially static. Second, Lucas argued that expectations would likely be rational—at least with respect to sustained policy changes and possibly for others—and that there were quantitatively major consequences of introducing rational expectations into macroeconomic models. The victory of Lucas’s ideas was swift. For example, in a second-year graduate macro class at Brown in 1975, Poole gave a clear message when reviewing Lucas’s critique in working-paper form: the stage was set for a complete overhaul of econometric models.<sup>12</sup> Better theoretical foundations and rational expectations were to be the centerpieces of the new research.

### **Rational Expectations Econometrics: The Promise**

As a result of the rational expectations revolution, there was a high demand for new methods in two areas: (1) algorithms to solve dynamic rational expectations models, and (2) econometric methods. The high ground was rapidly taken by the linear systems approach best articulated by Hansen and Sargent (1981): dynamic linear rational expectations models could be solved easily and had

---

<sup>12</sup> See also Poole (1976).

heavily constrained vector autoregressions as their reduced form. As Sargent (1981) stressed, the general equilibrium nature of rational expectations models meant that parameters traditionally viewed as important for one “behavioral equation” in a traditional macroeconomic model would also be important for others. For example, parameters of the investment technology would have implications for other quantities (for instance, consumption and labor supply) because changes in investment dynamics would alter the optimal choices of consumption and labor supply by influencing the rational expectations of future wages and interest rates implied by the model. Thus, complicated and time-consuming systems methods of estimation and testing were employed in the Hansen-Sargent program.

The promise of rational expectations econometrics was that “fully articulated” model economies were to be constructed, their reduced forms determined, and the resulting systems compared with unrestricted dynamic models. The “deep parameters” of preferences and technologies were to be estimated via maximum likelihood, and the fully articulated economies were to be used to evaluate alternative macroeconomic policies in a manner consistent with the requirements of Lucas (1976).

### **Rational Expectations Econometrics: The Failure**

With a decade of hindsight, though, it is clear that the Hansen-Sargent program was overambitious, in ways that are perhaps best illustrated by discussing the typical study using this technology in the mid-1980s. The author would (1) construct a rich dynamic macroeconomic model, (2) estimate most of its parameters using maximum likelihood, and (3) perform a likelihood ratio test to evaluate the model, i.e., compare its fit to an unrestricted vector autoregression. Since the estimation of the model had been very time-consuming, the author would have been able to produce only a small number of experiments with alternative specifications of preferences, technologies, and forcing processes.

Further, it would be difficult for the author and the audience to interpret the results of the study. Typically, at least some of the parameter estimates would be very strange, such as implausible discount factors or utility functions lacking concavity properties. Since only limited experimentation with the economic structure was feasible, the author would consequently struggle to explain what features of the data led to these aberrant outcomes. Further, the model would be badly rejected and again the author would have difficulty explaining which features of the macroeconomic time series led to this rejection. Overall, the author would be hard pressed to defend the specific model or, indeed, why he had spent his time conducting the investigation. This experience produced a general reaction that the Hansen-Sargent program had not produced a workable vehicle for systematic development of macroeconomic models.

### **Applied Macroeconomic Research After the Fall**

There were three basic reactions to this double failure of econometrics. First, some researchers sought to use limited-information methods to estimate the parameters of a single behavioral equation in ways consistent with rational expectations (as in McCallum [1976], Kennan [1979], and Hansen and Singleton [1982]). While this work was valuable, it did not aim at the objective of constructing and evaluating complete models of macroeconomic activity. Second, some researchers virtually abandoned the development of dynamic rational expectations models as part of their applied work. This rejectionist approach took hold most strongly in Cambridge, Massachusetts. One of the most sophisticated early applications of rational expectations methods is Blanchard (1983), but this researcher's applied work moved from routinely using the dynamic rational expectations models to the polar alternative, direct behavioral specifications that typically lack any expectational elements.<sup>13</sup> The final approach, quantitative theory, is the topic considered next.

## **5. THE RISE OF QUANTITATIVE THEORY**

In the two seminal theoretical papers on real business cycles, Kydland and Prescott's (1982) "Time to Build and Aggregate Fluctuations" and Long and Plosser's (1983) "Real Business Cycles," each pair of authors faced the following problem. They had constructed rich dynamic macroeconomic models driven by "technology shocks" and wanted to illustrate the implications of their models for the nature of economic fluctuations. Each set of authors sought to use parameters drawn from other sources: data on the input/output structure of industries in Long and Plosser and data on various shares and elasticities in Kydland and Prescott, with the latter pair of authors particularly seeking to utilize information from microeconomic studies.<sup>14</sup> One motivation for this strategy was to make clear that the models were not being "rigged" to generate fluctuations, for example, by fitting parameters to best match the business cycle components of macroeconomic time series.

---

<sup>13</sup> Recently, a small group of researchers has moved to a modified usage and interpretation of the Hansen-Sargent methodology. An early example along these lines is Christiano (1988). In recent work, Leeper and Sims (1994) use maximum likelihood methods to estimate parameters, but supplement likelihood ratio tests with many other model diagnostics. Potentially, promising technical developments like those in Chow (1993) may make it possible to execute the Hansen-Sargent program for interesting models in the future.

<sup>14</sup> In one case, in which they were least sure about the parameter value—a parameter indicating the extent of time nonseparability in utility flows from work effort, which determines the sensitivity of labor supply to temporary wage changes—Kydland and Prescott explored a range of values and traced the consequences for model implications.

### Calibration of Parameters

This process is now called the “calibration” approach to model parameter selection.<sup>15</sup> In line with the definition above, these papers were “quantitative theory”: they provided a strong case for the general mechanisms stressed by the authors. That is, they showed that the theory—restricted in plausible ways—could produce outcomes that appeared empirically relevant. These models were notably interesting precisely because they were evidently “post-Lucas critique” models: expectations were determined rationally about the future productivity. While they did not contain policy rules or policy disturbances, it was clear that these extensions would be feasible, and similar models have since incorporated policy anticipations, particularly on the fiscal side.

In terms of parameter selection, Lucas (1980) set forth a cogent argument for the use of estimates from microeconomic data within his articulation of the quantitative theory approach in his “Methods and Problems in Business Cycle Theory,” albeit in a slightly different context than that so far discussed:

In the case of the equilibrium account of wage and employment determination, parameters describing the degree of intertemporal substitutability do the job (in an empirical sense) of the parameter describing auctioneer behavior in the Phillips curve model. On these parameters, we have a wealth of inexpensively available data from census cohort information, from panel data describing the reactions of individual households to a variety of changing market conditions, and so forth. In principle (and perhaps before too long in practice . . .) these crucial parameters can be estimated independently from individual as well as aggregate data. If so, we will know what the aggregate parameters mean, we will understand them in a sense that disequilibrium adjustment parameters will *never* be understood. This is exactly why we care about the “microeconomic foundations” of aggregate theories. (P. 712)

Thus, one feature of the calibration of a model is that it brings to bear available microeconomic evidence. But Lucas also indicated the value of comparing estimates obtained from microeconomic studies and aggregate time series evidence, so that calibration from such sources is simply one of several useful approaches to parameter selection.

### Evaluating a Calibrated Model

The evaluation of the calibrated models by Kydland and Prescott (1982) and Long and Plosser (1983) was based on whether the models could capture some

---

<sup>15</sup> The calibration approach is now mainly associated with work on real business cycles, but it was simply a common strategy when used by Kydland and Prescott (1982) and Long and Plosser (1983). For example, Blanchard’s (1980) work on dynamic rational expectations models with nominal rigidities also utilized the calibration approach to explore the implications of a very different class of models.

key features of economic fluctuations that the authors were seeking to explain. This evaluation was conducted outside of an explicit econometric framework. In justifying this choice, Kydland and Prescott argued:

We choose not to test our model against the less restrictive vector autoregressive model. This would most likely have resulted in the model being rejected given the measurement problems and the abstract nature of the model. Our approach is to focus on certain statistics for which the noise introduced by approximation and measurement errors is likely to be small relative to the statistic. Failure of the theory to mimic the behavior of the post-war U.S. economy with respect to these stable statistics with high signal to noise ratios would be grounds for its rejection. (P. 1360)

Their argument, then, was essentially that a model needs to be able to capture some first-order features of the macroeconomic data before it can be taken seriously as a theory of business fluctuations. Their conclusions incorporated three further clarifications of the role of quantitative theory. First, they concluded that their model had some success: the “results indicate a surprisingly good fit in light of the model’s simplicity” (p. 1368). Second, they articulated a program of modifications of the theory that were important, many of which have been the subject of their subsequent research activities, including introduction of a variable workweek for capital. Third, they argued that applications of currently prevailing econometric methods were inappropriate for the current stage of model development: “in spite of the considerable recent advances made by Hansen and Sargent, further advances are necessary before formal econometric methods can fruitfully be applied to testing this theory of aggregate fluctuations” (p. 1369).

In Long and Plosser’s (1983) study of the response of a multi-sector macroeconomic model to fluctuations in productivity, the focus was similarly on a subset of the model’s empirical implications. Indeed, since their explicit dynamic equilibrium solution meant that the aggregate and industry allocations of labor input were constant over time, these authors did not focus on the interaction of productivity and labor input, which had been a central concern of Kydland and Prescott. Instead, Long and Plosser highlighted the role of output interrelationships arising from produced inputs for generating sectoral comovement, a topic about which the highly aggregated Kydland and Prescott theory had been silent.

### **Formal Econometric Analysis of RBC Models**

Variants of the basic RBC model were evaluated by Altug (1989) and Christiano (1988) using formal econometric techniques that were closely related to those of Hansen and Sargent (1981). There were three major outcomes of these analyses. First, predictably, the economic models performed poorly: at least some key parameter estimates typically strayed far from the values employed

in calibration studies and the models were decisively rejected as constrained vector autoregressions. Second, the Altug and Christiano studies did not naturally lead to new research, using either the methods of quantitative theory or econometrics, that aimed at resolving specific puzzles that arose in their work. Third, the rejectionist position of Kydland and Prescott hardened.

In particular, Kydland and Prescott called into question all econometric evidence. Prescott (1986) argued that “we do not follow the [econometric] approach and treat the [leisure share] as a free parameter because it would violate the principle that parameters cannot be specific to the phenomena being studied. What sort of a science would economics be if micro studies used one share parameter and aggregate studies another?” (p. 25) More recently, Kydland and Prescott (1991) argued that current econometrics is not faithful to the objectives of its originators (notably Frisch), which they interpret as being the construction of calibrated models. In the conclusion to their paper, they argued “econometrics is by definition quantitative economic theory—that is, economic analysis that provides quantitative answers to clear-cut questions” (p. 176). Thus, macroeconomic research has bifurcated, with a growing number of researchers using calibrated economic models and a very small number using the methods of Hansen and Sargent (1981).

## 6. COMPARING THE METHODOLOGIES

By its very nature, quantitative theory strays onto the turf of theoretical and applied econometrics since it seeks to use economic models to organize economic data. Thus there has been substantial—at times heated—controversy about the methods and conclusions of each area. Some of this controversy is a natural part of the way that economists learn. In this regard, it is frequently the case that applied econometrics poses challenges to quantitative theory: for example, McCallum (1989) marshals the results of various applied econometric studies to challenge RBC interpretations of the Solow residual as a technology shock.

However, controversy over methods can sometimes interfere with accumulation of knowledge. For example, adherents of RBC models have sometimes suggested that any study based on formal econometric methods is unlikely to yield useful information about the business cycle. Econometricians have similarly suggested that one learns little from investigations using the methods of quantitative theory. For this reason, it is important to look critically at the differences that separate studies using the methods of quantitative theory from those that use the more familiar methods of econometrics. As we shall see, a key strength of the quantitative theory approach is that it permits the researcher to focus the evaluation of a model on a specific subset of its empirical implications.

### Parameter Selection and Model Evaluation

To look critically at the methodological differences, we will find it useful to break the activities of econometricians into two general topics: selection of parameters (estimation) and evaluation of economic models (including testing of hypotheses and computation of measures of fit). Theoretical econometricians work to devise procedures for conducting each of these activities. Applied econometricians utilize these procedures in the context of specific economic problems and interpret the results. For concreteness in the discussion below, let the vector of parameters be  $\beta$  and the vector of model implications be  $\mu$ . Solving a model involves constructing a function,

$$\mu = g(\beta), \quad (4)$$

that indicates how model implications are related to parameters. The business of econometrics, then, is to determine ways of estimating the parameters  $\beta$  and evaluating whether the implications  $\mu$  are reliably close to some empirical counterparts  $m$ . In the approach of Hansen and Sargent (1981), the economic model implications computed are a set of coefficients in a reduced-form vector autoregression. There are typically many more of these than there are parameters of the model (elements of  $\beta$ ), so that the theoretical model is heavily overidentified in the sense of Hood and Koopmans (1953). The empirical counterparts  $m$  are the coefficients of an unconstrained vector autoregression. The estimation of the  $\beta$  parameters thus involves choosing the model parameters so as to maximize the fit of a constrained vector autoregression; model evaluation involves a comparison of this fit with that of an unconstrained time series model.

In the studies in quantitative theory reviewed above, there were also analogous parameter selection and model evaluation activities. Taking the first study as an example, in Douglas's production function,  $Y = AN^\alpha K^{1-\alpha}$ , there were the parameters  $A$  and  $\alpha$ , which he estimated via least squares. The model implications that he explored included the time series observations on  $Y$  and the values of  $\alpha$  derived from earlier studies of factor income shares.

Quantitative theory, then, involves selection of values of the parameters  $\beta$  and the comparison of some model implications  $\mu$  with some empirical counterparts  $m$ , which results in an evaluation of the model. Thus, it shares a formal structure with econometric research. This identification of a common structure is important for three reasons. First, while frequently suggested to be dramatic and irreconcilable, the differences in the two methods must be, at least on some level, ones of degree rather than kind. Second, this common structure also indicates why the two methods have been substitutes in research activity. Third, the common structure indicates the potential for approaches that combine the best attributes of quantitative theory and econometrics.

However, it is notable that quantitative theory, as an abstraction of reality, nearly always delivers a probability model that is remarkably detailed in its

implications and, hence, is too simple for direct econometric implementation. In the Douglas-Solow setting, the production function is essentially free of error terms as it is written down: it is equally applicable at seasonal, cyclical, and secular frequencies. But it is clear that Douglas and Solow did not view the production function as exact. Instead, they had strong views about what components of the data were likely to be most informative about aspects of the production function; in his own way, each concentrated on the trend behavior of output and factor inputs. The model driven solely by a productivity shock displays another form of exactness, which is sometimes called a “stochastic singularity.” There are not enough error terms for the joint probability distribution of the model’s outcomes to be nondegenerate. Since the model is dynamic, this does not mean that all the variables are perfectly correlated, but it does mean that there is (1) a singular variance-covariance matrix for the innovations in the model’s reduced form and (2) a unit coherence between variables in the frequency domain.

The key implication of this essential simplicity is that there is always some way to reject econometrically a quantitative theory with certainty. Quantitative theorists have had to come to terms with this simplicity on a number of fronts. First, when we learn from a quantitative theory, it is not because it is literally true but because there is a surprising magnitude of the gap between model implications  $\mu$  and empirical counterparts  $m$ . Sometimes the gap is surprisingly small, as in the work of Douglas and Prescott, and at others it is surprisingly large, as in the work of Solow. Second, when we investigate the implications of a quantitative theory, we must select a subset of implications of particular interest. Thus, for example, Douglas and Solow focused mainly on asking whether the production function explained the trend variation in output. By contrast, Prescott’s concern was with business cycle variations, that is, with deviations from trend.

### Calibration Versus Estimation

The central tension between practitioners of quantitative theory and econometrics is sometimes described as the choice between whether parameters are to be estimated or calibrated. However, returning to the three examples, we see clearly that this cannot be the case. Douglas estimated his parameter  $\alpha$  from observations on  $Y$ ,  $K$ , and  $N$ : he then compared this estimated value to averages of real labor cost as a fraction of manufacturing output, i.e., labor’s share. Solow “calibrated” his labor share parameter from data on labor income as a fraction of total income.<sup>16</sup> Prescott’s construction of an RBC model involved

---

<sup>16</sup> As noted above, Solow’s calibration is at each point in the sample, but this article’s version of it uses the average value of labor’s share over the entire sample. The difference between these two procedures is quantitatively unimportant.

estimating the parameters of the stochastic productivity process from aggregate time series data and calibrating other parameters of the model.

Further, applications of the Hansen-Sargent methodology to dynamic macroeconomic models typically involved “fixing” some of the parameters at specified values, rather than estimating all of them. For example, the discount factor in dynamic models proved notoriously hard to estimate so that it was typically set at a reasonable (calibrated) value. More generally, in such studies, maximizing the likelihood function in even small-scale dynamic models proved to be a painfully slow procedure because of the system nature of the estimation. Many researchers resorted to the procedure of fixing as many parameters as possible to increase the pace of improvement and the time to convergence.

Thus, almost all studies in quantitative theory and macroeconometrics involve a mixture of estimation and calibration as methods of parameter selection. The main issue, consequently, is: How should we evaluate models that we *know* are major simplifications of reality? In this context, it is not enough to compare the likelihoods of a heavily restricted linear time series model (for example, an RBC model with some or all of its parameters estimated) to an unrestricted time series model. For some time, the outcome of this procedure will be known in advance of the test: the probability that the model is true is zero for stochastically singular models and nearly zero for all other models of interest.

### **Model Evaluation in Quantitative Theory**

Implicitly, the three exercises in quantitative theory reviewed above each directed our attention to a small number of features of reality, i.e., a small vector of empirical features  $m$  that are counterparts to a subset of the model’s implications  $\mu = g(\beta)$ . In Douglas’s and Prescott’s case, the quantitative theory was judged informative when there was a small gap between the predictions  $\mu$  and the empirical features  $m$ . In Solow’s case, there was a strikingly large gap for a production function without technical progress, so we were led to reject that model in favor of an alternative that incorporates technical progress.

Looking at the studies in more detail clearly shows that for Douglas, there were two key findings. First, it was instructive that variation in factors of production, when combined into  $\hat{Y}$ , explained so much of the evolution of output  $Y$ . Second, it was striking that the time series of real wages was close to output per unit of labor input multiplied by  $\hat{\alpha}$  (i.e., that the value of labor’s share obtained was close to the estimate  $\hat{\alpha}$  obtained from the time series). This latter finding is an example of Lucas’s reconciliation of aggregate and micro evidence: we understand the parameter  $\alpha$  better because it represents a consistent pattern of behavioral response in micro and aggregate settings. For Solow, it was instructive that the variation in factors of production explained so little of the evolution of output; thus, there was a large amount of growth attributed to technical progress. Kydland and Prescott (1991) noted

that “contrary to what virtually everyone thought, including the authors, . . . technology shocks were found to be an important contributor to business cycle fluctuations in the U.S. postwar period” (p. 176). Prescott (1986) provided a more detailed listing of successes: “Standard theory . . . correctly predicts the amplitude of these fluctuations, their serial correlation properties, and the fact that the investment component of output is about six times as volatile as the consumption component” (p. 11).

For these researchers and for many others, the key idea is that one learns from an investigation in quantitative theory when there are relatively small or large discrepancies between a set of empirical features  $m$  and model implications  $\mu = g(\beta)$ . A key strength of the quantitative theory approach is that it permits the researcher to specify the elements of  $m$ , focusing the evaluation of the model on a specific subset of its empirical implications.

### Limitations of the Quantitative Theory Approach

However, there are also three important limitations of the method of model evaluation in quantitative theory, which may be illustrated by posing some questions about the outcomes of an exercise in quantitative theory.

First, as Singleton (1988) observed, quantitative theory does not provide information about the confidence that one should have in the outcomes of an investigation. For example, when we look at a specific implication  $\mu_1$  and its empirical counterpart  $m_1$ , how likely is it that a small value of  $m_1 - \mu_1$  would occur given the amount of uncertainty that we have about our estimates of  $\beta$  and  $m_1$ ? When we are looking at a vector of discrepancies  $m - \mu$ , the problem is compounded. It is then important to be explicit about the joint uncertainty concerning  $\beta$  and  $\mu$ , and it is also important to specify how to weight these different discrepancies in evaluating the theory.<sup>17</sup>

Second, each of the quantitative theory investigations discussed above spawned a major area of research. In the follow-up research, many investigations adopted the quantitative theory methods of the original investigation. This process of model refinement thus raises an additional question: When can we say with confidence that model A represents an improvement over model B?

---

<sup>17</sup> Particularly in the area of real business cycles, some investigations in quantitative theory follow Kydland and Prescott (1982) by providing “standard deviations” of model moments. These statistics are computed by simulating the calibrated model over a specified sample period (say, 160 quarters) so as to determine the approximate finite sample distribution of the model moments under the assumption that the model is exactly true. Such statistics cannot be used to answer the question posed in the text since they do not take into account the joint distribution of the parameters  $\beta$  and the empirical implications  $m$ . Instead, these statistics indicate how much uncertainty in sample moments is introduced by sample size if the theory were literally true, including use of the exact parameter values.

Third, if a strength of the quantitative theory approach is that one looks at a subset of model implications of the model economy, is this not also a limitation? In particular, when we look at a subset of model implications, this leaves open the possibility that there are also other, interesting implications of the model that are not examined.

Operating independently of econometrics, quantitative theory gives us no ability even to pose, much less answer, these three key questions.

## 7. REEVALUATING THE STANDARD RBC MODEL

The issues of model evaluation raised above are not debating points: absent some systematic procedures for evaluating models, one may regularly miss core features of reality because the “tests” are uninformative. To document this claim, we will look critically at two features of RBC models, which many—beginning with Prescott (1986) and Kydland and Prescott (1991)—have argued are supportive of the theory. These two features are summarized by Prescott (1986) as follows: “Standard theory . . . correctly predicts the amplitude of . . . fluctuations [and] their serial correlation properties” (p. 11).

The first of these features is typically documented in many RBC studies by computing the “variance ratio”:

$$\lambda = \frac{\text{var}(\log Y_m)}{\text{var}(\log Y_d)}. \quad (5)$$

In this expression,  $\text{var}(\log Y_m)$  is the variance of the logarithm of the model’s output measure, and  $\text{var}(\log Y_d)$  is the variance of the corresponding empirical measure.<sup>18</sup> In terms of the series displayed in Figure 4, in particular, the variance of  $\log Y_m$  is 2.25 and the variance of  $\log Y_d$  is 2.87, so that the implied value of  $\lambda = 0.78$ . That is, Kydland and Prescott (1991) would say that the baseline RBC model explains 78 percent of business fluctuations.

The second of these features is typically documented by looking at a list of autocorrelations of model time series and corresponding time series from the U.S. economy. Most typically, researchers focus on a small number of low-order autocorrelations as, for example, in Kydland and Prescott (1982) and King, Plosser, and Rebelo (1988).

### Parameter Uncertainty and the Variance Ratio

In a provocative recent contribution, Eichenbaum (1991) argued that economists know essentially nothing about the value of  $\lambda$  because of parameter uncertainty.

---

<sup>18</sup> In particular, these constructs are 100 times the Hodrick-Prescott filtered logarithms of model and data output, so that they are interpretable as percentage deviations from trend.

He focused attention on uncertainty about the driving process for technology and, in particular, about the estimated values of the two parameters that are taken to describe it by Prescott (1986) and others. The core elements of Eichenbaum's argument were as follows. First, many RBC studies estimate the parameters of a low-order autoregression for the technology-driving process, specifying it as  $\log A_t = \rho \log A_{t-1} + \varepsilon_t$  and estimating the parameters  $(\rho, \sigma_\varepsilon^2)$  by ordinary least squares. Second, following the standard quantitative theory approach, Eichenbaum solved the RBC model using a parameter vector  $\hat{\beta}$  that contains a set of calibrated values  $\beta_1$  and the two estimates,  $\hat{\beta}_2 = [\hat{\rho} \hat{\sigma}_\varepsilon^2]$ . Using this model solution, Eichenbaum determined that the population value of  $\lambda(\hat{\beta})$  is 0.78 (which is identical to the sample estimate taken from Figure 4). Third, Eichenbaum noted that estimation uncertainty concerning  $\hat{\beta}_2$  means there is substantial uncertainty about the implications that the model has for  $\lambda$ : the standard error of  $\lambda$  from these two sources alone is huge, about 0.64. Further, he computed the 95 percent confidence interval for  $\lambda$  as covering the range from 0.05 to 2.<sup>19</sup> Eichenbaum's conclusion was that there is a great deal of uncertainty over sources of business cycles, which is not displayed in the large number of quantitative theory studies that compute the  $\lambda$  measure.

### The Importance of Comparing Models

There is an old saying that "it takes a model to beat a model."<sup>20</sup> Overall, one reason that quantitative theory has grown at the expense of econometrics is that it offers a way to systematically develop models: one starts with a benchmark approach and then seeks to evaluate the quantitative importance of a new wrinkle. However, when that approach is applied to the basic RBC model's variance ratio, it casts some doubt on the standard, optimistic interpretation of that measure.

To see why, recall that the derivation of the productivity residual implies

$$\log Y_{dt} = \log A_{dt} + \alpha \log N_{dt} + (1 - \alpha) \log K_{dt},$$

where the subscript  $d$  indicates that this is the version of the expression applicable to the actual data. In the model economy, the comparable construction is

$$\log Y_{mt} = \log A_{mt} + \alpha \log N_{mt} + (1 - \alpha) \log K_{mt},$$

---

<sup>19</sup> There are some subtle statistical issues associated with the computation of this confidence interval, as Mary Finn and Adrian Pagan have pointed out to the author. In particular, the point estimate in Eichenbaum [1991] for  $\rho$  is 0.986, thus suggesting that the finite sample approximate confidence interval on  $\hat{\rho}$  is both wide and asymmetric because of considerations familiar from the analysis of "near unit root" behavior. Pagan indicates that taking careful account of this would shrink Eichenbaum's confidence interval so that it had a lower bound for  $\lambda$  of 0.40 rather than 0.05.

<sup>20</sup> Both Thomas Sargent and Robert Barro have attributed this saying to Zvi Griliches.

where the subscript  $m$  indicates that this is the model version. In the simulations underlying Figure 4, the model and data versions of the technology shocks are set equal. Thus, the difference of output in the data from that in the model is

$$\log Y_{dt} - \log Y_{mt} = \alpha(\log N_{dt} - \log N_{mt}) + (1 - \alpha)(\log K_{dt} - \log K_{mt}). \quad (6)$$

That is, by construction, deviations between the data and the model arise only when the model's measures of labor input or capital input depart from the actual measures.

In terms of the variance ratio  $\lambda$ , this means that we are giving the model credit for  $\log A_t$  in terms of explaining output. This is a *very* substantial asset: in Panel A of Figure 3, the Solow residual is highly procyclical (strongly correlated with output) and has substantial volatility.

Now, to take an extreme stand, suppose that one's benchmark model was simply that labor and capital were constant over the business cycle:  $\log Y_{mt} = \log A_t$ . This naive model has a value of  $\underline{\lambda} = 0.49$ : the "Solow residual" alone explains about *half* the variability of output. Thus, the marginal contribution of variation in the endogenous mechanisms of the model—labor and capital—to the value of  $\lambda = 0.78$  is only  $\lambda - \underline{\lambda} = 0.29$ . From this standpoint, the success of the basic RBC model looks much less dramatic, and it does so precisely because two models are compared.

This finding also illustrates the concerns, expressed by McCallum (1989) and others, that RBC models may be mistakenly attributing to technology shocks variations in output that arise for other reasons. It is indeed the case that RBC models explain output very well precisely because they utilize output (productivity) shifts as an explanatory variable, rather than inducing large responses to small productivity shifts.

From this standpoint, the major success of RBC models cannot be that they produce high  $\lambda$  values; instead, it is that they provide a good account of the relative cyclical amplitude of consumption and investment. Perhaps paradoxically, this more modest statement of the accomplishments of RBC analysis suggests that studies like those of Prescott (1986) and Plosser (1989) are more rather than less important for research in macroeconomics. That is, these studies suggest that the neoclassical mechanisms governing consumption and investment in RBC economies are likely to be important for any theory of the cycle independent of the initiating mechanisms.

### Looking Broadly at Model Implications

Many RBC studies including those of Prescott (1986) and King, Plosser, and Rebelo (1988) report information on the volatility and serial correlation properties of real output in models and in U.S. data. Typically, the authors of these studies report significant success in the ability of the model to match the empirical time series properties of real output. However, a recent study by Watson

(1993), which is an interesting blend of quantitative theory and econometrics, questions this conclusion in a powerful manner.

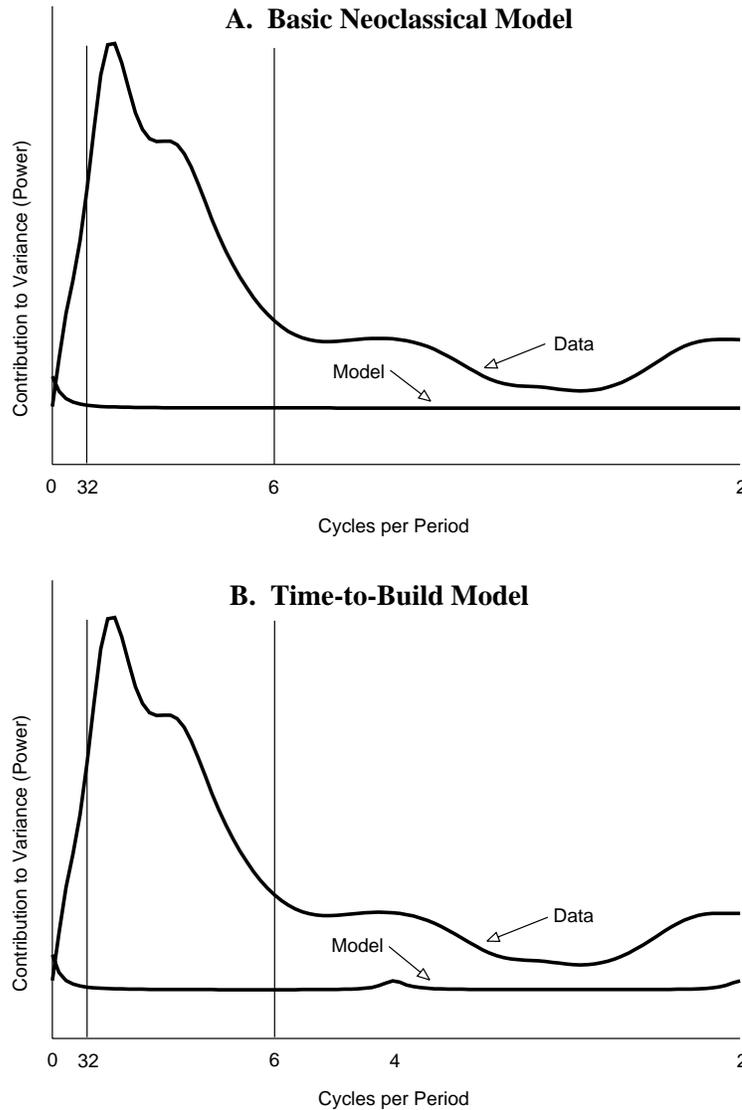
On the econometric side, Watson estimated the time series of productivity shocks, working with two key assumptions about a baseline RBC model drawn from King, Plosser, and Rebelo (1988). First, he assumed that the model correctly specifies the form of the driving process, which is taken to be a random walk with given drift and innovation variance. Second, he assumed that the realizations of productivity shocks are chosen to maximize the fit of the model to the U.S. data. By doing so, he gave the basic RBC model the best possible chance to match the main features of the business cycle.<sup>21</sup> The model did relatively poorly when Watson placed all weight on explaining output variability: the model could explain at most 48 percent of variance in output growth and 57 percent of variance in Hodrick-Prescott filtered data.

On the quantitative theory side, Watson traced this conclusion to a simple fact about the spectrum of output growth in the basic RBC model and in the U.S. data. A version of Watson's result is illustrated in Figure 5, which displays two important pieces of information. To begin to interpret this figure, recall that the power spectrum provides a decomposition of variance by frequency: it is based on dividing up the growth in output into periodic components that are mutually uncorrelated. For example, the fact that the data's power spectrum in Figure 5 has greater height at eight cycles per period than at four cycles per period means that there is greater variation in the part of output growth that involves two-year cycles than one-year cycles. The general shape of the power spectrum in Figure 5 thus indicates that there is a great deal of variability in output growth at the business cycle frequencies defined in the tradition of Burns and Mitchell (1946). These cyclical components, with durations between eight years and eighteen months, lie between the vertical lines in the figure. In addition, since the power spectrum is the Fourier transform of the autocovariance generating function, it displays the same set of information as the *full set* of autocorrelations. In terms of the autocovariances, the key point is that the overall shape of the power spectrum suggests that there is substantial predictability to output growth.<sup>22</sup> Further, the empirical power spectrum's shape for

---

<sup>21</sup> This fitting exercise is one that is easy to undertake because Watson poses it as a simple minimization problem in the frequency domain.

<sup>22</sup> Previously, the empirical power spectrum's shape, or, equivalently, the pattern of autocorrelations of output growth, has been important for the empirical literature on the importance of stochastic trends, as in Cochrane (1988) and Watson (1986). This literature has stressed that it would be necessary to have a relatively lengthy univariate autoregression in order to fit the spectrum's shape well. Such a long autoregression could thus capture the shape with one or more positive coefficients at low lags (to capture the initial increase in the power spectrum as one moves from high to medium frequencies) and then many negative ones (to permit the subsequent decline in the spectrum as one moves from medium to very low frequencies). However, short autoregressions would likely include only the positive coefficients.

**Figure 5**

Notes: Panels A and B contain the estimated power spectrum of the growth rate of quarterly U.S. gross national product. The power spectrum summarizes the decomposition of this growth rate series into periodic components of varying duration. Business cycle components (defined as cycles of length between 6 and 32 quarters) lie between the vertical lines in the figure. Since the total variance of the growth rate is proportional to the area under the power spectrum, this shape of the estimate indicates that a great deal of variability in growth rates occurs at the business cycle frequencies.

Panel A also contains the power spectrum of output growth in a basic real business cycle model if it is driven by random-walk shocks to productivity. Panel B displays the power spectrum for a model that contains the time-to-build investment technology.

output is also a “typical spectral shape for growth rates” of quarterly U.S. time series. King and Watson (1994) showed that a similar shape is displayed by consumption, investment, output, man-hours, money, nominal prices, and nominal wages; it is thus a major stylized fact of business cycles.

Yet, as shown in Panel A of Figure 5, the basic RBC model does not come close to capturing the shape of the power spectrum of output growth, i.e., it does not generate much predictable output growth. When driven by random-walk productivity shocks, the basic model’s spectrum involves no peak at the business cycle frequencies; it has the counterfactual implication that there is greater variability at very low frequencies than at business cycle frequencies.<sup>23</sup> A natural immediate reaction is that the model studied in King, Plosser, and Rebelo (1988) is just too simple: it is a single state variable version of the RBC model in which the physical capital stock is the only propagation mechanism. One might naturally conjecture that this simplicity is central for the discrepancy highlighted in Panel A of Figure 5. However, Panel B of Figure 5 shows that the introduction of time-to-build investment technology as in Kydland and Prescott (1982) does not alter the result much. There is a slight “bump” in the model’s spectrum at four quarters since there is a four-quarter time-to-build, but little change in the general nature of the discrepancies between model and data. Overall, there is an evident challenge of Watson’s (1993) plot: by expressing the deficiencies of the standard RBC theory in a simple and transparent way, this exercise in quantitative theory and econometrics invites development of a new class of models that can capture the “typical spectral shape of growth rates.”<sup>24</sup>

## 8. CHALLENGES TO ECONOMETRICS

The growth of business cycle analysis using the quantitative theory approach has arisen because there is a set of generally agreed-upon procedures that makes feasible an “adaptive modeling strategy.” That is, a researcher can look at a set of existing models, understand how and why they work, determine a new line of inquiry to be pursued, and evaluate new results. The main challenge to econometrics is to devise ways to mimic quantitative theory’s power in systematic model development, while adding the discipline of statistical inference and of replicability.

---

<sup>23</sup> There is also a sense in which the overall height of the spectrum is too low, which may be altered by raising or lowering the variance of the technology shock in the specification. Such a perturbation simply shifts the overall height of the model’s spectrum by the same proportion at every frequency.

<sup>24</sup> Rotemberg and Woodford (1994) demonstrated that a multivariate time series model displays a substantial ability to predict output growth and also argued that the basic RBC model cannot capture this set of facts. Their demonstration was thus the time-domain analogue to Watson’s findings.

Indeed, development of such new econometric methods is essential for modern macroeconomic analysis. At present, it is the last stage of the quantitative theory approach that is frequently most controversial, i.e., the evaluation of new results. This controversy arises for two reasons: (1) lack of generally agreed-upon criteria and (2) lack of information on the statistical significance of new results. On the former front, it seems impossible to produce a uniform practice, and it is perhaps undesirable to try to do so: the criteria by which results are evaluated has always been part of the “art” of econometrics. But on the latter, we can make progress. Indeed, short of such developments, research in modern business cycle analysis is likely to follow the path of the NBER research that utilized the Burns and Mitchell (1946) strategy of quantitative business cycle analysis: it will become increasingly judgmental—leading to difficulties in replication and communication of results—and increasingly isolated.

As suggested earlier, the main challenge for econometric theory is to derive procedures that can be used when we know *ex ante* that the model or models under study are badly incomplete. There are some promising initial efforts under way that will be referred to here as the “Northwestern approach,” whose core features will now be summarized. There are two stages in this procedure: parameter estimation and model evaluation.

### **Parameter Estimation**

The strength of the quantitative theory approach in terms of the selection of parameters is that it is transparent: it is relatively easy to determine which features of the real-world data are important for determining the value of a parameter used in constructing the model economy.

The estimation strategy of the Northwestern approach is similar to quantitative theory in this regard and in sharp contrast to the Hansen and Sargent (1981) approach. Rather than relying on the complete model to select parameters, it advocates using a subset of the model’s empirical implications to estimate parameters. It consequently makes transparent which features of the data are responsible for the resulting parameter estimates.

Essentially, these chosen features of the model are used for the *measurement* of parameters, in a manner broadly consistent with the quantitative theory approach. However, and importantly, the Northwestern approach provides an estimate of the variance-covariance matrix of parameter estimates, so that there is information on the extent of parameter uncertainty. To take a specific example, in building dynamic macroeconomic models incorporating an aggregate production function, many researchers use an estimate of labor’s share to determine the value of  $\alpha$  in the production function, appealing to the work of Solow. This estimate is sometimes described as a first-moment estimate, since it depends simply on the sample average labor’s share. Further, since observations on period-by-period labor’s share are serially correlated,

obtaining an estimate of the amount of uncertainty about  $\alpha$  requires that one adopt a “generalized least squares” procedure such as Hansen’s (1982) generalized method of moments. The procedure is remarkably transparent: the definitional and statistical characteristics of labor income and national income dictate the estimate of the parameter  $\alpha$ .

In a core paper in the Northwestern approach, Christiano and Eichenbaum (1992) demonstrated how a subset of model relations may be used to estimate a *vector* of model parameters with Hansen’s (1982) method in a manner like that used for  $\alpha$  in the discussion above.<sup>25</sup> In this study, some of the parameters are determined from steady-state relations, so that they are naturally first-moment estimators, like the labor share and other long-run parameters which are “calibrated” in Prescott’s (1986) study. Other parameters describe aspects of the model’s exogenous dynamics, including driving processes for productivity and government purchases. These parameter selections necessarily involve consideration of second moments, as in Eichenbaum’s (1991) estimate of the productivity process discussed previously.

More generally, second-moment estimates could also be used for internal elements of the model. For example, an alternative transparent approach to estimating the production function parameter  $\alpha$  is to use the joint behavior of output, capital, and labor as in Cobb and Douglas (1928). With this alternative method, aspects of the trends in output, capital, and labor would be the core features that determined the estimate of  $\alpha$ , as stressed above. However, given the recent focus on technology shocks and the endogenous response of capital and labor to these disturbances, one would presumably not follow Douglas’s ordinary least squares approach. Instead, one would employ a set of instrumental variables suggested by the structure of the model economy under study. For example, in Christiano and Eichenbaum’s setting, one would likely employ government purchases as an instrumental variable for the production function estimation. In this case, the estimate of  $\alpha$  would depend on the comovements of government purchases with output, capital, and labor.

Thus, in the case of  $\alpha$ , a researcher has latitude to determine the core features of the model that are used to estimate the parameter of interest: either first- or second-moment estimators are available. However, with some alternative structural elements of dynamic macroeconomic models, only second-moment estimators are available since the parameters govern intrinsically dynamic elements of the model. For example, absent information on the distribution of investment expenditure over the life of investment projects, there is simply no way to determine the parameters of the “time-to-build” technology of Kydland and Prescott (1982) from steady-state information. However, second-moment

---

<sup>25</sup> In an early appraisal of the relationship between quantitative theory and econometrics, Singleton (1988) forecasted some of the developments in Christiano and Eichenbaum (1992).

estimates of a dynamic investment function could readily recover the necessary time-to-build parameters.

Overall, the parameter selection component of the Northwestern approach is best viewed as an application of the instrumental variables methods of McCallum (1976) and Hansen and Singleton (1982). A key feature, shared with quantitative theory and in contrast to the typical application of the Hansen-Sargent (1981) strategy, is that the complete model is not used for parameter estimation. Rather, since a carefully chosen subset of the model's relations is used, it is relatively easy to understand which features of the data are important for results of parameter estimation. Yet, in contrast to standard techniques in quantitative theory, there are explicit measures of the extent of parameter uncertainty provided by the Northwestern method.

From this perspective, we know from prior theoretical and applied econometric work that instrumental variables estimates of structural parameters from a subset of the model's equations are no panacea. Minimally, one loses efficiency of parameter estimation relative to procedures like those of Hansen and Sargent (1981), if the model is well specified. More generally, there will always be potential for problems with poor instruments and inappropriate selection of the subset of relations that is chosen for estimation. However, at the present stage of macroeconomic research, economists are far from having a well-specified macroeconomic model. It seems best to opt for simple and transparent procedures in the selection of model parameters.

### **Model Evaluation**

A notable feature of quantitative theory is that a selected subset of model implications are compared to their empirical counterparts. This method has its benefits and costs. On the positive side, it permits the researcher to specify a subset of implications that are viewed as first-order for the investigation. This allows an individual researcher to focus on a manageable problem, which is essential for research progress in any discipline. On the negative side, this freedom may mean that the researcher may study model implications that are not too informative about the economic model or models under study. However, once methods are in place that allow for replication and criticism of the results of studies, it is likely that competition among researchers will provide for extensive exploration of the sensitivity of the results of the model evaluation stage of research.

Like quantitative theory, the Northwestern model evaluation approach permits the researcher to focus on a subset of model implications in evaluating small-scale dynamic models. However, since it utilizes standard econometric methodology, it provides diagnostic information about the extent of uncertainty that one has about gaps between model and empirical implications. Any such evaluation of the discrepancies between a model's implications and the corresponding empirical features involves taking a stand on a penalty function

to be applied to the discrepancies, i.e., on a function  $\Delta$  that assigns a scalar loss to the vector of discrepancies  $(m - \mu)$ . If this penalty function is assumed to be quadratic, then one must simply specify a matrix  $L$  that is used to penalize discrepancies between the model's implication and the corresponding empirical features. That is, models may be evaluated using a discrepancy function like

$$\Delta(m, \beta) = [m - g(\beta)]L[m - g(\beta)]', \quad (7)$$

whose statistical properties will depend on the choice of  $L$ .

Christiano and Eichenbaum (1992) took the null hypothesis to be that the model is correctly specified and choose  $L$  to be the inverse of the variance-covariance matrix of  $[m - g(\beta)]$ . This permitted them to perform powerful chi-squared tests of a subset of the model's implications. Essentially, Christiano and Eichenbaum (1992) asked: Can we reject the hypothesis that discrepancies between the model's implication and the corresponding empirical features are due to sampling error?<sup>26</sup> If the answer is yes, then the model is viewed as deficient. If the answer is no, then the model is viewed as having successfully produced the specified empirical features. While their paper focused on implications for selected second moments (variances and covariances of output, etc.), their procedure could readily be applied to other implications such as spectra or impulse responses.

King and Watson (1995) alternatively assumed the null hypothesis that models are not correctly specified and require, instead, that the researcher specify the discrepancy function  $L$ . In contrast to the procedures of Christiano and Eichenbaum (1992), their tests are thus of unknown power. While they discussed how to use this assumed discrepancy function to construct tests of the adequacy of an individual model, most of their attention was directed to devising tests of the relative adequacy of two models. Suppose, for sake of argument, that model A has a smaller discrepancy measure than model B. Essentially, the question that they asked was: Can we reject the hypothesis that a difference in discrepancy measures between two models A and B is due to sampling error? If the answer is yes, then King and Watson would say that model A better captures the specified list of empirical features. While their paper focused on implications for selected impulse responses (comparative dynamics), the procedure could also be applied to other implications such as selected moments or spectra.

---

<sup>26</sup> The phrase *sampling error* in this question is a short-hand for the following. In a time series context with a sample of length  $T$ , features of the macroeconomic data (such as the variance of output, the covariance of output and labor input, etc.) are estimates of a population counterpart and, hence, are subject to sampling uncertainty. Model parameters  $\beta$  estimated by generalized method of moments are consistent estimates, if the model is correctly specified in terms of the equations that are used to obtain these estimates, but these estimates are also subject to some sampling error. Evaluation of discrepancies  $\Delta$  under the null hypothesis that the model is true takes into account the variance-covariance matrix of  $m$  and  $\beta$ .

Thus, the Northwestern approach involves two stages. In the first, the parameters are estimated using a subset of model relations. In the second, a set of features of interest is specified; the magnitude of discrepancies between the model's implications and the corresponding empirical features is evaluated. The approach broadly parallels that of quantitative theory, but it also provides statistical information about the reliability of estimates of parameters and on the congruence of the model economy with the actual one being studied.

It is too early to tell whether the Northwestern approach will prove broadly useful in actual applications. With either strategy for model evaluation, there is the potential for one of the pitfalls discussed in the previous section. For example, by concentrating on a limited set of empirical features (a small set of low-order autocovariances), a researcher would likely miss the main discrepancies between model and data arising from the "typical spectral shape of growth rates." Other approaches—perhaps along the lines of White (1982)—may eventually dominate both. But the Northwestern approach at least provides some econometric procedures that move in the direction of the challenges raised by quantitative theory. The alternative versions of this approach promise to permit systematic development of macroeconomic models, as in quantitative theory, with the additional potential for replication and critique of the parameter selection and model evaluation stages of research.

## 9. OPPORTUNITIES FOR ECONOMETRICS

The approach of quantitative economic theory also provides opportunities for econometrics: these are for learning about (1) the features of the data that are most informative about particular parameters, especially dynamic ones, and (2) how to organize data in exploratory empirical investigations. This section provides one example of each opportunity, drawing on the powerful intuition of the permanent income theory of consumption (Friedman 1957). For our purposes, the core features of Friedman's theory of consumption are as follows. First, consumption depends importantly on a measure of wealth, which Friedman suggested contains expectations of future labor income as its dominant empirical component. Second, Friedman argued that it is the changes in expectations about relatively permanent components of income that exert important wealth effects and that it is these wealth effects that produce the main variations in the level of an individual's consumption path.

### Identifying and Estimating Dynamic Parameters

A common problem in applied econometrics is that a researcher may have little a priori idea about which features of the data are most likely to be informative about certain parameters, particularly parameters that describe certain dynamic responses. Quantitative economic theories can provide an important means of

learning about this dependence: they can indicate whether economic responses change sharply in response to parameter variation, which is necessary for precise estimation.

To illustrate this idea, we return to consideration of productivity and the business cycle. As previously discussed, Eichenbaum (1991) suggested that there is substantial uncertainty concerning the parameters that describe the persistence of productivity shocks, for example, the parameter  $\rho$  in  $\log A_t = \rho \log A_{t-1} + \epsilon$ . In his study, Eichenbaum reported a point estimate of  $\rho = 0.986$  and a standard error of 0.026, so that his Solow residual is estimated to display highly persistent behavior. Further, Eichenbaum concluded that this “substantial uncertainty” about  $\rho$  translates into “enormous uncertainty” about the variance ratio  $\lambda$ .

Looking at these issues from the standpoint of the permanent income theory of consumption, we know that the ratio of the response of consumption to income innovations changes very dramatically for values of  $\rho$  near unity. (For example, the line of argument in Goodfriend [1992] shows that the response of consumption to an income innovation is *unity* when  $\rho = 1$ , since actual and permanent income change by the same amount in this case, and is about *one-half* when  $\rho = 0.95$ ).<sup>27</sup> That is, it matters a great deal for individual economic agents where the parameter  $\rho$  is located. The sensitivity of individual behavior to the value of  $\rho$  carries over to general equilibrium settings like Eichenbaum’s RBC model in a slightly modified form: the response of consumption to the productivity shock  $\epsilon_t$  increases rapidly as  $\rho$  gets near unity. In fact, there are sharp changes in individual behavior well inside the conventional confidence band on Eichenbaum’s productivity estimate of  $\rho$ . The economic mechanism is that stressed by the permanent income theory: wealth effects are much larger when disturbances are more permanent than when they are transitory. Thus, the parameter  $\rho$  cannot move too much without dramatically changing the model’s implication for the comovement of consumption and output. More specifically, the key point of this sensitivity is that a moment condition relating consumption to technology shocks is plausibly a *very* useful basis for sharply estimating the parameter  $\rho$  if the true  $\rho$  is near unity. Presumably, sharper estimates of  $\rho$  would also moderate Eichenbaum’s conclusion that there is enormous uncertainty about the variance ratio  $\lambda$ .

---

<sup>27</sup> The analytics are as follows: If income is a first-order autoregression with parameter  $\rho$ , then the change in permanent income is proportional to the innovation in income with a coefficient:  $r/(r + 1 - \rho)$ . Thus, if  $\rho = 1$ , an innovation in income has a unit effect, and correspondingly, if  $\rho = 0$ , then the effect of a change in income is simply the annuity factor  $r/(r + 1)$ , which is much closer to zero than to one. Even fairly persistent changes in income have much less than a one-for-one effect: with  $r = 0.05$  and  $\rho = 0.95$ , then it follows that the adjustment coefficient is 0.5.

### Organizing Economic Data

Another problem for applied researchers is to determine interesting dimensions along which to organize descriptive empirical investigations. For example, there have been many recent papers devoted to looking at how countries behave differentially in terms of their national business cycles, but relatively few of these investigations provide much information on characteristics of countries that are correlated with differences in national business cycles. There are some interesting and natural decompositions: for example, Baxter and Stockman (1989) asked how national business cycles are related to a country's exchange rate regime and Kouparitsas (1994) asked how national business cycles are linked to industrial structure and the composition of trade. But, beyond these immediate decompositions, how is a researcher to determine revealing ways of organizing data? In the discussion that follows, we show how quantitative theories can provide a vehicle for determining those dimensions along which revealing organizations of economic data may be based.

To continue on this international theme, open economy RBC models such as those of Backus, Kehoe, and Kydland (1992) and Baxter and Crucini (1993) have the implication that there should be a near-perfect cross-country correlation of national consumptions. This striking implication is an artifact of two features of these models. First, they incorporate the desire for consumption smoothing on the part of the representative agent (as in Friedman's permanent income theory). Second, they allow for rich international asset trade, so that agents in different countries can in fact purchase a great deal of risk sharing, i.e., they are complete-markets models.

But cross-national correlations of consumption are in fact lower than cross-national correlations of output.<sup>28</sup> A plausible conjecture is that incomplete asset markets, particularly those for national human capital, are somehow responsible for this gap between theory and reality. Thus, a natural research strategy would be to try to measure the extent of access to the world capital market in various countries and then to look cross-sectionally to see if this measure is related to the extent of a country's correlation with world consumption.

Quantitative theory suggests that this strategy, while natural, would miss a central part of the problem. It thus suggests the importance of an alternative organization of the international economic data. In particular, Baxter and Crucini (1995) investigated a two-country RBC model in which there can be trade in bonds but no trade in contingent claims to physical and human capital. They thus produced a two-country general equilibrium version of the permanent income hypothesis suitable for restricted asset markets. Further, as suggested by Hall's (1978) random-walk theory of consumption, the restriction on asset

---

<sup>28</sup> That is, there is a smaller correlation of Hodrick-Prescott filtered output than of Hodrick-Prescott filtered consumption, as shown in Backus and Kehoe (1992).

markets leads to a random-walk component of each nation's economic activity, which arises because shocks redistribute wealth between countries. When there is low persistence of productivity shocks in a country, Baxter and Crucini (1995) found that the restricted asset two-country model does not behave very differently from its complete-markets counterpart. That is, trade in bonds offers the country a great ability to smooth out income fluctuations; the wealth effects associated with transitory shocks are small influences on consumption. Yet, if the persistence of productivity shocks in a country is high, then there are major departures from the complete-markets model and it is relatively easy for cross-national consumption correlations to be small (or even negative).

In this international business cycle example, quantitative theory thus suggests that a two-way classification of countries is important. One needs to stratify both by a measure of the persistence of shocks and by the natural measure of access to the international capital market. Further, it also makes some detailed predictions about how the results of this two-way classification should occur, if restricted access is indeed an important determinant of the magnitude of international consumption correlations. More generally, quantitative theory can aid the applied researcher in determining those dimensions along which descriptive empirical investigations can usefully be organized.<sup>29</sup>

## 10. THE CHALLENGE OF IDENTIFICATION

To this point, our discussion has focused on evaluating a model in terms of its implications for variability and comovement of prices and quantities; this has come to be the standard practice in quantitative theory in macroeconomics. For example, in Prescott's (1986) analysis of productivity shocks, measured as in Solow (1957), the basic neoclassical model successfully generated outcomes that "looked like" actual business cycles, in terms of the observed cyclical amplitude of consumption, investment, and labor input as well as the observed comovement of these series with aggregate output. In making these comparisons, Prescott measured cyclical amplitude using the standard deviation of an individual series, such as consumption, and cyclical comovement by the correlation of this series with output. That is, he evaluated the basic neoclassical model in terms of its implications for selected second moments of consumption, investment, labor, and output. Further, in Watson's (1993) analysis of the same model reviewed above, the number of second moments examined was expanded to include all of the autocorrelations of these four variables, by plotting the spectra. However, in these discussions, there was relatively little explicit discussion of the link between the driving variable—productivity—and

---

<sup>29</sup> Prescott (1986) previously suggested that theory should guide measurement in macroeconomics. The example in the current section is thus a specific application of this suggestion, applied to the organization of international economic data.

the four endogenously determined variables of the model. In this section, we take up some issues that arise when this alternative strategy is adopted, with specific emphasis on evaluating models of business fluctuations.

### **Identifying Causes of Economic Fluctuations**

*Identification* involves spelling out a detailed set of linkages between causes and consequences. In particular, the researcher takes a stand on those economic variables that are taken to exert a *causal* influence on economic activity; he then studies the consequences that these “exogenous variables” have for the dynamic evolution of the economy.

To implement this strategy empirically, the researcher must detail the measurement of causal variables. This might be as simple as providing a list of economic variables assumed to be causal, such as productivity, money supply, taxes or government deficits, and so on. But it might also be more elaborate, involving some method of extracting causal variables from observed series. For example, researchers examining the effects of government budget deficits on macroeconomic activity have long adjusted the published deficit figures to eliminate components that are clearly dependent on economic activity, creating a “cyclically adjusted” or “full employment” deficit measure. Modern versions of this identification strategy involve treating unobserved variations in productivity as the common stochastic trend in real consumption, investment, and output (as in King, Plosser, Stock, and Watson [1991]) or extracting the unobserved policy-induced component of the money supply as that derived from unpredictable shifts in the federal funds rate (as in Sims [1989]). This stage of the analysis is necessarily open to debate; researchers may adopt alternative identifications of causal variables.

Once these causal variables are determined, then they may be put to three uses. First, one can employ them as the basis for estimation of behavioral parameters, i.e., they may be used as instrumental variables. Second, one can investigate how an actual economy reacts dynamically to shocks in these causal variables, estimating “dynamic multipliers” rather than behavioral parameters. Such estimates may serve as the basis for model evaluation, as discussed further below. Third, one can consider the practical consequences of altering the historical evolution of these causal variables, conducting counterfactual experiments that determine how the evolution of the economy would have been altered in response to changes in the time path of one or more of the causal variables.

### **Identification in Quantitative Theory**

Since they involve simple, abstract models and a small amount of economic data, exercises in quantitative theory typically involve a very strong identification of causal factors. It is useful to begin by considering the identifications implicit in the previously reviewed studies.

For Douglas, the working hypothesis behind the idea of a “practical” production function was that the factors of production, capital and labor, represented causal influences on the quantity of production: if a firm had more of either or both, then output should increase. Estimates of production functions along these lines are now a standard example in introductory econometrics textbooks. Operating prior to the development of econometric tools for studying simultaneous-equations systems, Douglas simply used capital and labor as independent variables to estimate the parameters of the production function. He was not naive: as discussed above, he was concerned that his assumption of a pure causal influence might not be correct, so he gave more weight to the results suggested by the “trends” in capital, labor, and output than he did to results suggested by shorter-term variation. A central objective of his investigation was to be able to make “what if” statements about the practical implications of varying the quantities of these inputs. Thus, he employed his identification for precisely the first and third purposes discussed above. Viewed with more than 50 years of hindsight, many would now be led to question his identification: if there is technical progress, then capital and labor will typically respond to this factor. However, Douglas’s transparent identification led to striking results.

In Solow’s (1957) procedure, particularly as it has been used in RBC modeling, the nature of the identification is also transparent: output is transformed, using data on factor shares and factor inputs, to reveal an unobserved component, productivity, to which a causal role may be assigned.<sup>30</sup> In his investigation, Solow focused mainly on asking the hypothetical question: What if there had been capital growth, but no productivity growth, in the United States?

In the RBC analysis of Prescott (1986), the Solow productivity process was used as the driving variable, with the objective of describing the nature of macroeconomic fluctuations that arise from this causal factor. In line with our third use of an identified causal disturbance, the type of historical simulation produced in Section 2 above provides an implicit answer to the question: How large would fluctuations in macroeconomic activity have been without fluctuations in productivity, as identified by the Solow procedure? The confidence placed in that answer, of course, depends on the extent to which one believes that there is an accurate identification of causal disturbances.

---

<sup>30</sup> There are essentially two equations of Solow’s investigation. In logarithms, the production function is  $y - k = \alpha(n - k) + a$ , and the competitive marginal productivity condition is  $w = (y - n) + \log \alpha + \varepsilon$ , where time-dating of variables is suppressed for simplicity. In these expressions, the notation is the same as in the main text, except that  $w$  is the log real wage and  $\varepsilon$  is a disturbance to the marginal productivity condition. To estimate  $\alpha$ , Solow assumed that the  $\varepsilon$  are a set of mean-zero discrepancies. He made no assumptions about the  $a$  process. Thus, Solow’s identification of  $a$  was conditioned on some additional assumptions about the natures of departures from the competitive theory; it is precisely along these lines that it was challenged in the work of Hall (1988) and others.

### Identification and Moment Implications

Consideration of identification leads one squarely to a set of potential difficulties with the model evaluation strategy used in quantitative RBC theory, reviewed in Section 2 above, and also in the econometric strategies described in Section 8 above. To see the nature of this tension, consider the following simple two-equation model:

$$y - k = \alpha(n - k) + a, \quad (8)$$

$$n - k = \beta a + \theta b. \quad (9)$$

The first of these specifications is the Cobb-Douglas production function in logarithmic form, with  $y = \log Y$ , etc. The second may be viewed as a behavioral rule for setting hours  $n$  as a function of productivity  $a$  and some other causal factor  $b$ , which is taken for simplicity not to affect output except via labor input. In the case in which  $a$  and  $b$  are not correlated, these two expressions imply that the covariance of  $y - k$  and  $n - k$  is:

$$\text{cov}(y - k, n - k) = (1 + \alpha\beta)\beta\sigma_a^2 + \alpha\theta^2\sigma_b^2, \quad (10)$$

where  $\sigma_a^2$  is the variance of  $a$  and  $\sigma_b^2$  is the variance of  $b$ . This covariance is, of course, the numerator of the correlation that is frequently examined to explore the comovement of output and labor input. Expression (10) shows that the covariance depends on how labor and output respond to productivity, i.e., on  $\alpha$  and  $\beta$ ; on how output and labor respond to the other causal factor ( $\theta$ ); and on the extent of variability in the two causal factors ( $\sigma_a^2, \sigma_b^2$ ).

In models with multiple causal factors, moments of endogenous variables like  $y - k$  and  $n - k$  are combinations of behavioral responses determined by the model  $(\alpha, \beta, \theta)$ , with these responses weighted by the variability of the causal factors. If, as most macroeconomists believe, business cycles are the result of a myriad of factors, it follows that covariances and other second moments of endogenous variables will not be the best way to determine how well a model economy describes response to a particular causal factor.

### Implications for Model Evaluation

These difficulties have led to some work on alternative model evaluation strategies; a core reference in this line is Sims (1989). The main idea is to evaluate models based on how well they describe the dynamic responses to changes in identified causal factors.<sup>31</sup> The basic ideas of this research can be discussed within the context of the preceding two-equation model.

---

<sup>31</sup> Sims (1989) studied how several models match up to multiple identified shocks. Single-shock analyses include Rotemberg and Woodford's (1992) analysis of government-purchase disturbances. King and Watson (1995) considered evaluation and comparison of models using individual and multiple shocks.

To begin, consider the relationship between the causal factor  $a$ , productivity, and the endogenous variables  $y - k$  and  $n - k$ . One way to summarize this linkage is to consider the coefficient that relates  $y$  to  $a$ : call this a “response coefficient” and write it as the coefficient  $\pi$  in  $y - k = \pi_{ya} a$ , with the subscripts indicating a linkage from  $a$  to  $y$ . In the model above, this response coefficient is  $\pi_{ya}^m = (1 + \beta)$ , with the superscript indicating that this is the model’s response: productivity exerts an effect on output directly and through the labor input effect  $\beta$ . In the preceding model, the comparable response coefficient that links labor and productivity is  $\pi_{na}^m = \beta$ .<sup>32</sup>

An identified series of empirical productivity disturbances,  $a$ , can be used to estimate comparable objects empirically, producing  $\pi_{ya}^d$  and  $\pi_{na}^d$ , where the superscript indicates that these are “data” versions of the response coefficients. The natural question then is: Are the model response coefficients close to those estimated in the data? Proceeding along the lines of Section 8 above, we can conduct model evaluation and model comparison by treating  $[\pi_{ya}^m \ \pi_{na}^m]$  and  $[\pi_{ya}^d \ \pi_{na}^d]$  as the model features to be explored. The results of model evaluations and model comparisons will certainly be dependent on the plausibility of the initial identification of causal factors. Faced with the inevitable major discrepancies between models and data, a researcher will likely be forced to examine both identification and model structure.

Explicit evaluation of models along these lines subtly changes the question that a researcher is asking. It changes the question from “Do we have a good (or better) model of business cycles?” to “Do we have a good (or better) model of how fluctuations in  $x$  lead to business cycles?” Given that it is essentially certain that business cycles originate from a multitude of causes, it seems essential to ask the latter question as well as the former.

## 11. SUMMARY AND CONCLUSIONS

In studies in quantitative theory, economists develop simple and abstract models that focus attention on key features of actual economies. In this article, quantitative theory is illustrated with examples of work on the production function by Douglas, Solow, and Prescott.

It is typical to view econometrics as providing challenges to quantitative theory. In particular, in the years after each of the aforementioned studies, applied researchers using econometric tools indicated that there were elements missing from each theoretical framework.

---

<sup>32</sup> An alternative procedure would be to focus on a covariance, such as  $\text{cov}(y - k, a)$ , with the difference in the current context simply reflecting whether one scales by the variance of  $a$ . In the dynamic settings that are envisioned as the main focal point of this research, one can alternatively investigate impulse responses at various horizons or cross-correlations between endogenous and exogenous variables. These features summarize the same information but present it in slightly different ways.

However, particularly in terms of the development of dynamic macroeconomic models, there are also important challenges and opportunities that quantitative theory provides to econometrics. To begin, it is not an accident that there has been substantial recent growth in dynamic macroeconomic research using the methods of quantitative theory and little recent work building such models using standard econometric approaches. The article argues that the style of econometrics needed is one that is consistent with the ongoing development of simple dynamic macroeconomic models. It must thus aid in understanding the dimensions along which simple theories capture the interactions in the macroeconomic data and those along which they do not. In this sense, it must become like the methods of quantitative theory. But it can be superior to the current methods of quantitative theory because econometrics can provide discipline to model development by adding precision to statements about the success and failure of competing models.

Quantitative theory, however, is also not without its difficulties. To provide a concrete example of some of these problems, the article uses quantitative theory and some recent econometric work to evaluate recent claims, by Kydland and Prescott (1991) and others, that the basic RBC model explains most of economic fluctuations. This claim is frequently documented by showing that there is a high ratio of the variance of a model's predicted output series to the actual variance of output. (It is also the basis for the sometimes-expressed view that business cycle research is close to a "closed field.") The article's contrary conclusion is that it is hard to argue convincingly that the standard RBC model explains most of economic fluctuations. First, a simple exercise in quantitative theory indicates that most of the variance "explained" by a basic version of the RBC theory comes from the direct effects of productivity residuals, not from the endogenous response of factors of production. Second, recent econometric research indicates that the basic RBC model misses badly the nature of the business cycle variation in output growth (as indicated by comparison of the power spectrum of output growth in the theory and the U.S. data). Thus, the more general conclusion is that business cycle research is far from a closed field. The speed at which a successful blending of the methods of quantitative theory and econometrics is achieved will have a major effect on the pace at which we develop tested knowledge of business cycles.

---

---

**REFERENCES**

- Altug, Sumru. "Time-to-Build and Aggregate Fluctuations: Some New Evidence," *International Economic Review*, vol. 30 (November 1989), pp. 889–920.
- Backus, David, and Patrick Kehoe. "International Evidence on the Historical Properties of Business Cycles," *American Economic Review*, vol. 82 (September 1992), pp. 864–88.
- \_\_\_\_\_, and Finn Kydland. "International Real Business Cycles," *Journal of Political Economy*, vol. 100 (August 1992), pp. 745–75.
- Baxter, Marianne, and Mario J. Crucini. "Business Cycles and the Asset Structure of Foreign Trade," *International Economic Review* (forthcoming 1995).
- \_\_\_\_\_. "Explaining Savings-Investment Correlations," *American Economic Review*, vol. 83 (June 1993), pp. 416–36.
- Baxter, Marianne, and Robert G. King. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," Working Paper 5022. Cambridge, Mass.: National Bureau of Economic Research, February 1995.
- Baxter, Marianne, and Alan C. Stockman. "Business Cycles and the Exchange Rate Regime: Some International Evidence," *Journal of Monetary Economics*, vol. 23 (May 1989), pp. 377–400.
- Blanchard, Olivier J. "The Production and Inventory Behavior of the American Automobile Industry," *Journal of Political Economy*, vol. 91 (June 1983), pp. 365–401.
- \_\_\_\_\_. "The Monetary Mechanism in the Light of Rational Expectations," in Stanley Fischer, ed., *Rational Expectations and Economic Policy*. Chicago: Chicago University Press for National Bureau of Economic Research, 1980, pp. 75–102.
- \_\_\_\_\_, and Stanley Fischer. *Lectures on Macroeconomics*. Cambridge, Mass.: MIT Press, 1989.
- Burns, Arthur M., and Wesley C. Mitchell. *Measuring Business Cycles*. New York: National Bureau of Economic Research, 1946.
- Chow, Gregory. "Statistical Estimation and Testing of a Real Business Cycle Model." Manuscript. Princeton University, March 1993.
- Christiano, Lawrence J. "Why Does Inventory Investment Fluctuate So Much?" *Journal of Monetary Economics*, vol. 21 (March/May 1988), pp. 247–80.

- \_\_\_\_\_, and Martin Eichenbaum. "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," *American Economic Review*, vol. 82 (June 1992), pp. 430–50.
- Clark, John Bates. "The Possibility of a Scientific Law of Wages," *Publications of the American Economic Association*, vol. 4 (March 1889), pp. 39–63.
- Cobb, Charles W., and Paul H. Douglas. "A Theory of Production," *American Economic Review*, vol. 8 (March 1928), pp. 139–65.
- Cochrane, John H. "How Big is the Random Walk in GNP?" *Journal of Political Economy*, vol. 96 (October 1988), pp. 893–920.
- Douglas, Paul H. "The Cobb-Douglas Production Function Once Again: Its History, Its Testing, and Some New Empirical Values," *Journal of Political Economy*, vol. 84 (October 1976), pp. 903–15.
- \_\_\_\_\_. "Are There Laws of Production?" *American Economic Review*, vol. 38 (March 1948), pp. 1–41.
- \_\_\_\_\_. "The Recent Movement of Real Wages and Its Economic Significance," *American Economic Review*, supplement to vol. 16 (March 1926), pp. 17–53.
- Eichenbaum, Martin S. "Real Business Cycles: Wisdom or Whimsy," *Journal of Economic Dynamics and Control*, vol. 15 (October 1991), pp. 607–26.
- Friedman, Milton. *A Theory of the Consumption Function*. Princeton: Princeton University Press, 1957.
- Goodfriend, Marvin S. "Information Aggregation Bias," *American Economic Review*, vol. 82 (June 1992), pp. 508–19.
- Gordon, Robert J. *The Measurement of Durable Goods Prices*. Chicago: University of Chicago Press, 1990.
- Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell. "Macroeconomic Implications of Investment-Specific Technological Change," Discussion Paper 76. Federal Reserve Bank of Minneapolis, Institute for Empirical Macroeconomics, October 1992.
- Hall, Robert E. "The Relationship Between Price and Cost in U.S. Industry," *Journal of Political Economy*, vol. 96 (October 1988), pp. 921–47.
- \_\_\_\_\_. "Stochastic Implications of the Permanent Income Hypothesis," *Journal of Political Economy*, vol. 86 (December 1978), pp. 971–87.
- Hansen, Lars P. "Large Sample Properties of Generalized Methods of Moment Estimators," *Econometrica*, vol. 50 (July 1982), pp. 1029–54.
- \_\_\_\_\_, and Thomas J. Sargent. "Formulating and Estimating Dynamic Linear Rational Expectations Models" and "Linear Rational Expectations Models for Dynamically Interrelated Variables," in Robert E. Lucas, Jr., and Thomas J. Sargent, eds., *Rational Expectations and Econometric Practice*. Minneapolis: University of Minnesota Press, 1981.

- Hansen, Lars P., and Kenneth J. Singleton. "Generalized Instrumental Variables Estimators of Non-Linear Rational Expectations Models," *Econometrica*, vol. 50 (September 1982), pp. 1269–86.
- Hodrick, Robert, and Edward C. Prescott. "Post-War U.S. Business Cycles: A Descriptive Empirical Investigation," Working Paper. Pittsburgh: Carnegie-Mellon University, 1980.
- Hood, William C., and Tjalling C. Koopmans. *Studies in Econometric Method*. New York: John Wiley and Sons, 1953.
- Kennan, John. "The Estimation of Partial Adjustment Models with Rational Expectations," *Econometrica*, vol. 47 (November 1979), pp. 1441–55.
- King, Robert G., Charles I. Plosser, and Sergio T. Rebelo. "Production, Growth and Business Cycles: I. The Basic Neoclassical Model," *Journal of Monetary Economics*, vol. 21 (March/May 1988), pp. 195–232.
- King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson. "Stochastic Trends in Economic Fluctuations," *American Economic Review*, vol. 81 (September 1991), pp. 819–40.
- King, Robert G., and Sergio T. Rebelo. "Low Frequency Filtering and Real Business Cycles," *Journal of Economic Dynamics and Control*, vol. 17 (January 1993), pp. 207–31.
- King, Robert G., and Mark W. Watson. "Money, Interest Rates, Prices and the Business Cycle," *Review of Economics and Statistics* (forthcoming 1996).
- . "On the Econometrics of Comparative Dynamics." Unpublished working paper, University of Virginia, 1995.
- Klein, Lawrence. *Economic Fluctuations in the United States*. New York: John Wiley and Sons, 1950.
- , and Arthur C. Goldberger. *An Econometric Model of the United States, 1929–1952*. Amsterdam: North-Holland Publishing Co., 1955.
- Kouparitsas, Michael. "Industrial Structure and the Nature of National Business Cycles," Working Paper. Charlottesville: University of Virginia, November 1994.
- Kydland, Finn E., and Edward C. Prescott. "The Econometrics of the General Equilibrium Approach to Business Cycles," *Scandinavian Journal of Economics*, vol. 93 (1991), pp. 161–78.
- . "Time to Build and Aggregate Fluctuations," *Econometrica*, vol. 50 (November 1982), pp. 1345–70.
- Leeper, Eric M., and Christopher A. Sims. "Toward a Macroeconomic Model Usable for Policy Analysis," *NBER Macroeconomic Annual*. Cambridge: MIT Press for National Bureau of Economic Research, vol. 9, 1994, pp. 81–118.

- Long, John B., and Charles I. Plosser. "Real Business Cycles," *Journal of Political Economy*, vol. 91 (February 1983), pp. 39–69.
- Lucas, Robert E., Jr. "Methods and Problems in Business Cycle Theory," *Journal of Money, Credit, and Banking*, vol. 12 (November 1980, Part 2), pp. 696–715.
- \_\_\_\_\_. "Econometric Policy Evaluation: A Critique," in Karl Brunner and Allan Meltzer, eds., *The Phillips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy, vol. 1 (1976), pp. 19–46.
- Maddison, Angus. *Capitalist Forces in Economic Development*. New York: Oxford University Press, 1981.
- McCallum, Bennett T. "Real Business Cycles," in Robert J. Barro, ed., *Modern Business Cycle Theory*. Cambridge, Mass.: Harvard University Press, 1989, pp. 16–50.
- \_\_\_\_\_. "Rational Expectations and the Estimation of Econometric Models: An Alternative Procedure," *International Economic Review*, vol. 17 (June 1976), pp. 484–90.
- Mitchell, Wesley C. *Business Cycles: The Problem and Its Setting*. New York: National Bureau of Economic Research, 1927.
- Plosser, Charles I. "Understanding Real Business Cycles," *Journal of Economic Perspectives*, vol. 3 (Summer 1989), pp. 51–78.
- Poole, William W. "Rational Expectations in the Macro Model," *Brookings Papers on Economic Activity*, 2:1976, pp. 463–514.
- Prescott, Edward C. "Theory Ahead of Business Cycle Measurement," in Karl Brunner and Allan Meltzer, eds., *Real Business Cycles, Real Exchange Rates and Actual Policies*, Carnegie-Rochester Conference Series on Public Policy, vol. 25 (Fall 1986), pp. 11–44.
- Rogoff, Kenneth. "Theory Ahead of Business Cycle Measurement: A Comment on Prescott," in Karl Brunner and Allan Meltzer, eds., *Real Business Cycles, Real Exchange Rates and Actual Policies*, Carnegie-Rochester Conference Series on Public Policy, vol. 25 (Fall 1986), pp. 45–47.
- Romer, Paul M. "Crazy Explanations of the Productivity Slow Down," *NBER Macroeconomics Annual*. Cambridge: MIT Press for National Bureau of Economic Research, vol. 2, 1987, pp. 163–201.
- \_\_\_\_\_. "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, vol. 94 (October 1986), pp. 1002–37.
- Rotemberg, Julio, and Michael Woodford. "Is the Business Cycle a Necessary Consequence of Stochastic Growth?" Working Paper. Cambridge: Massachusetts Institute of Technology, 1994.

- \_\_\_\_\_. "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity," *Journal of Political Economy*, vol. 100 (December 1992), pp. 1153–1207.
- Sargent, Thomas J. "Interpreting Economic Time Series," *Journal of Political Economy*, vol. 89 (April 1981), pp. 213–48.
- Sims, Christopher A. "Models and Their Uses," *American Journal of Agricultural Economics*, vol. 71 (May 1989), pp. 489–94.
- Singleton, Kenneth J. "Econometric Issues in the Analysis of Equilibrium Business Cycle Models," *Journal of Monetary Economics*, vol. 21 (March/May 1988), pp. 361–87.
- Solow, Robert M. "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, vol. 39 (August 1957), pp. 312–20.
- \_\_\_\_\_. "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, vol. 70 (February 1956), pp. 65–94.
- Watson, Mark W. "Measures of Fit for Calibrated Models," *Journal of Political Economy*, vol. 101 (December 1993), pp. 1011–41.
- \_\_\_\_\_. "Univariate Detrending with Stochastic Trends," *Journal of Monetary Economics*, vol. 81 (July 1986), pp. 49–75.
- White, Halbert. "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, vol. 50 (February 1982), pp. 1–16.
- Wicksell, Knut. *Lectures on Political Economy, Vol. I: General Theory*. London: George Routledge and Sons, 1934.