FILE COPY

New England Economic Review

Federal Reserve Bank of Boston

July/August 1992

Jane Sneddon Little — *Public-Private Cost Shifts in Nursing Home Care*

Geoffrey M. B. Tootell — *Purchasing Power Parity within the United States*

Stephen K. McNees — *How Large Are Economic Forecast Errors?*

Richard W. Kopcke — *The Capitalization and Portfolio Risk of Insurance Companies*

FILE COPY

President
Richard F. Syron

Senior Vice President and Director of Research
Alicia H. Munnell

Research Department:
International
Norman S. Fieleke, Vice President
Jane Sneddon Little

Monetary and Financial
Richard W. Kopcke, Vice President
Richard E. Randall, Vice President
Eric S. Rosengren, Vice President
Joe Peek
Katerina Simons
Robert Tannenwald

National Business Conditions
Stephen K. McNees, Vice President
Jeffrey C. Fuhrer
Geoffrey M. B. Tootell

Regional
Lynn E. Browne, Vice President and Deputy Director of Research for Regional Affairs
Katharine L. Bradbury, Assistant Vice President
Yolanda K. Henderson
Christopher J. Mayer

The **New England Economic Review** is produced in the Research Department and is edited by Joan T. Poskanzer. Heather Schofield coordinates the design and production. The views expressed are those of the individual authors, and do not necessarily reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System. The authors will be glad to receive comments on their articles. The Review is available without charge. Requests to be placed on the mailing list or for additional copies should be sent to the Research Library–D, Federal Reserve Bank of Boston, P.O. Box 2076, Boston, MA 02106-2076. Articles may be reprinted if the source is credited. Please send copies of the reprinted materials to the editor.

Contents

Public-Private Cost Shifts in Nursing Home Care

Jane Sneddon Little

This article presents estimates of the degree to which individuals paying privately for nursing home services are being asked to subsidize their state's Medicaid program. Recently published data permit making some first rough estimates of the extent of this activity. Although these crude measures must be interpreted with great care, the data suggest that cost-shifting activity is widespread and sizable.

Because, as the article shows, cost-shifting activity undermines policy makers' efforts to regulate nursing home costs, the article suggests that private patients should not be asked to subsidize Medicaid's long-term care program; all comparably disabled individuals in a given institution should be charged the same rate. The author concludes by explaining why her findings bolster the case for including long-term care coverage in Medicare. 3

Purchasing Power Parity within the United States

Geoffrey M. B. Tootell

Economics has many articles of faith. One of the most dearly held is Purchasing Power Parity, which posits that the price of the same good in different regions should be equivalent when no barriers to arbitrage exist. Because Purchasing Power Parity (PPP) is an important assumption in much of international economic theory, this article examines empirical evidence testing this proposition.

Instead of analyzing international data, this study analyzes PPP between regions of the United States. By comparing regions within a country, it eliminates many of the hypotheses offered to explain the failure of PPP. The results of this study are suggestive. PPP fails to hold within regions of the United States. Instead, the inclusion of nontraded goods in the total consumer price indices for these regions is shown to be the major cause of this failure. When the nontraded components of these indices are removed, PPP holds. Some categories of goods do seem to move in lockstep while others do not, as one would expect, and as PPP predicts. 15

Contents

How Large Are Economic Forecast Errors?

Stephen K. McNees

Opinion about the reliability of economic forecasts ranges widely. Some argue that they are literally worthless, at the same time that most forecasters can point to a sequence of near perfect predictions. How much confidence should one place in economic forecasts? The errors vary with many factors.

A crucial determinant of the size of forecast errors is the forecast period; some periods are very difficult to predict while others are relatively easy. By far the largest errors were the sustained underestimations of the acceleration of inflation in 1973–75 and again in 1978–80. In addition, variations in the difficulty of predicting different variables can be illustrated by examining the forecasts. The results show drastic differences among variables in the forecasters' ability to outperform a naïve standard of comparison. Finally, much of the interest in forecast accuracy stems from the wish to know "Who is the best forecaster?" Even a cursory examination of the information in this article shows that no single forecaster dominates all outliers for all, or even most, of the variables. 25

The Capitalization and Portfolio Risk of Insurance Companies

Richard W. Kopcke

The strategies of financial intermediaries in the United States presumed a stability of interest rates that began to break down in the late 1960s. Not only did rising interest rates during the past two decades tend to depress the value of the assets of all intermediaries, they also fostered competition among intermediaries as all sought new opportunities for profit. In order to cope, many financial institutions assumed new bets by "reaching" for riskier assets offering higher yields or by operating with less capital per dollar of assets. To varying degrees, many insurance companies have adopted these strategies.

Of all the remedies inspired by the recent investigations of the insurance industries, none appears to be more important than raising more capital. Insurers need to reduce their leverage if their contracts are to be as secure as they were supposed to be prior to the late 1960s. This article concludes that many insurers must increase their capital to cope safely with the consequences of an enduring slump in the value of commercial real estate, a substantial decline in corporate profits, or a significant rise in credit market yields. 43

Public-Private Cost Shifts in Nursing Home Care

Jane Sneddon Little

Economist, Federal Reserve Bank of Boston. *The author is grateful to Michael D. Jud for his excellent research assistance. The author would also like to thank Phyllis Torda of Families USA Foundation for her helpful comments.*

A recent article in this *Review* argued that policymakers trying to tame Medicaid spending might want to reexamine how they regulate and pay nursing homes. The article (Little 1992) made four main points. First, differences in Medicaid nursing home reimbursement rates contribute significantly to the large cross-state differences in Medicaid payments per recipient—for all services as well as for long-term care.

Second, Medicaid nursing home reimbursement rates vary much more across states than do per capita personal health care expenditures or average pay for all health care workers. In other words, broadly defined state health care costs do not appear to explain the differences in per diem reimbursement rates.

Third, by contrast, nursing home worker pay, and thus "costs" narrowly defined, do appear to justify much of the variation in nursing home reimbursement rates. However, as the article argued, the direction of causality is not clear. Indeed, it seems likely that relatively generous reimbursement rates permit above-average expenditures which, in turn, justify above-average per diems in an interactive cycle. The ease with which nursing homes can charge private patients more than they receive for comparable Medicaid patients may contribute to this cycle, the article suggested.

Finally, differences in Medicaid payments per recipient do not necessarily reflect relative efficiency. Below-average payments per recipient may indicate efficiency, but they may also reflect below-average quality, or above-average use of cross-subsidies. Although anecdotal evidence suggests that nursing home residents paying privately often pay more than—and thus subsidize—similar residents supported by the public, most states do not publish data on average charges to private patients; thus, it has not been easy to measure the amount of cost-shifting activity by state.

Table 1
Ratio of Estimated Average Private Revenue to Medicaid Reimbursement per Nursing Home Patient Day, 1989

	Private Rate/ Medicaid Rate		Private Rate/ Medicaid Rate
Alabama	1.56	Montana	1.14
Alaska	n.a.	Nebraska	n.a.
Arizona	n.a.	Nevada	2.04
Arkansas	n.a.	New Hampshire	n.a.
California	1.25	New Jersey	1.61
Colorado	1.13	New Mexico	1.26
Connecticut	1.68	New York	.51
Delaware	1.08	North Carolina	1.09
D.C.	1.76	North Dakota	n.a.
Florida	1.16	Ohio	1.06
Georgia	1.82	Oklahoma	n.a.
Hawaii	n.a.	Oregon	.99
Idaho	1.19	Pennsylvania	n.a.
Illinois	1.23	Rhode Island	.41
Indiana	n.a.	South Carolina	1.10
Iowa	n.a.	South Dakota	n.a.
Kansas	n.a.	Tennessee	.97
Kentucky	.88	Texas	1.10
Louisiana	1.11	Utah	1.96
Maine	1.58	Vermont	1.50
Maryland	n.a.	Virginia	1.20
Massachusetts	1.25	Washington	n.a.
Michigan	.93	West Virginia	n.a.
Minnesota	.86	Wisconsin	1.25
Mississippi	n.a.	Wyoming	1.13
Missouri	1.46		

n.a. = not available

Source: Health Care Investment Analysts, Inc. and Arthur Andersen (1991); HCFA State Medicaid Data Disk for FY1989; and National Governors' Association (1989).

Within the last two years, however, Health Care Investment Analysts, Inc. and Arthur Andersen have begun compiling and publishing a wealth of information on nursing home operations.[1] Combining this information with previously available data permits *estimating* average daily revenue per private nursing home resident for most states.

This article presents such estimates and compares them with the average Medicaid per diem in the same state. In most cases, the ratio of estimated private to public reimbursement is well above one. The article then discusses what these ratios suggest about long-term care in this country and in particular states. It ends by arguing that for reasons of equity and effective cost control, nursing home patients supported by private funds should not be asked or, indeed, allowed to subsidize Medicaid's long-term care program. In other words, all comparably disabled individuals in a given institution should be charged the same rate, regardless of the source of their support. These findings also underscore the need to reconsider the way this country finances long-term care.

Private/Public Reimbursement per Nursing Home Day

Table 1 presents the ratio of private to Medicaid reimbursement for all states for which average private revenue per nursing home patient day can be estimated.[2] These ratios suggest that private payors may be subsidizing the Medicaid program in most states. (Alternative interpretations of these ratios are discussed in the next section.) In 27 of the 34 states for which estimates can be made, the ratio is above 1. Within the 27, private daily revenue exceeds the Medicaid reimbursement rate by an average of 36 percent. For all 34 states, estimated private revenue exceeds the Medicaid reimbursement rate by an average of 21 percent.

Table 2 presents the same data but with the states ranked according to the amount of cost-control pressure the state regulators seem to be exerting on the nursing homes. This pressure is measured by the difference between two ratios, the ratio of the state's Medicaid reimbursement rate to the U.S. median reimbursement rate, and the ratio of the state's personal health care expenditures per capita to the U.S.

[1] The 1990 and 1991 volumes of *The Guide to the Nursing Home Industry* contain a great deal of detailed data compiled from publicly available sources, primarily the Medicaid and Medicare cost reports filed by the nursing homes. The earliest year covered is 1987. The author is indebted to Bob Moran of the Pioneer Institute for bringing this valuable source to her attention.

[2] Average revenue per private patient day was estimated according to the following equation:

Average Revenue per Private Patient Day = [Net Patient Revenue − (Medicaid Share of Resident Days * Weighted Average Medicaid Per-Diem)] / [1.00 − Medicaid Share of Resident Days].

The Medicaid per diem is a weighted average of the rates reported for skilled nursing facilities and for intermediate care facilities other than those for the mentally retarded. The weights were the number of Medicaid recipients in each type of facility in 1989. The Medicaid reimbursement rates reported to the National Governors' Association may themselves have been means, weighted means, or medians. While these differences would distort cross-state comparisons, the mean and the median of a large group of numbers converge. Just as an example of the numbers involved, in 1991 Maryland had 105 nursing facilities with over 10,000 beds.

Table 2

Comparisons of Personal Health Care Expenditures per Capita and Nursing Home Reimbursement per Patient, FY1989

State	Medicaid Nursing Home Reimbursement Rate/ U.S. Average (1)	Personal Health Care Expenditures per Capita/ U.S. Average (2)	Difference: Relative Medicaid Nursing Home Reimbursement Rate and Relative Personal Health Care Expenditures (3) = (1) − (2)	Private Rate/ Medicaid Rate[a] (4)	Medicaid Payments per Nursing Home Recipient[b]/ U.S. Average (5)
New York	1.80	1.16	.64	.51	2.42
New Hampshire	1.20	.82	.38	n.a.	1.55
New Jersey	1.18	.92	.26	1.61	1.58
New Mexico	.97	.74	.23	1.26	1.01
Vermont	1.03	.81	.22	1.50	1.08
Wyoming	.93	.72	.21	1.13	.81
Idaho	.90	.71	.19	1.19	.85
Delaware	1.12	.94	.18	1.08	1.63
North Carolina	.93	.76	.17	1.09	.89
Maine	1.03	.90	.13	1.58	1.32
Pennsylvania	1.16	1.05	.11	n.a.	1.24
South Carolina	.79	.70	.09	1.10	.93
Connecticut	1.19	1.11	.08	1.68	1.25
Oregon	1.02	.95	.07	.99	.76
Virginia	.92	.86	.06	1.20	1.00
Utah	.80	.74	.06	1.96	.89
Kentucky	.83	.77	.06	.88	.80
Florida	1.05	1.00	.05	1.16	.92
Minnesota	1.07	1.02	.05	.86	1.19
Indiana	.94	.91	.03	n.a.	.94
Montana	.88	.85	.03	1.14	.98
Massachusetts	1.27	1.25	.02	1.25	1.76
Rhode Island	1.13	1.12	.01	.41	1.11
Mean:	1.05	.90	.15	1.18	1.17
Ohio	.99	1.03	−.04	1.06	.98
Colorado	.94	1.00	−.06	1.13	.68
Wisconsin	.92	1.01	−.09	1.25	.85
California	1.02	1.19	−.17	1.25	.98
Tennessee	.75	.93	−.18	.97	.71
Michigan	.88	1.06	−.18	.93	.78
Arkansas	.60	.80	−.20	n.a.	.69
Georgia	.64	.85	−.21	1.82	.76
Texas	.67	.90	−.23	1.10	.66
Oklahoma	.64	.88	−.24	n.a.	.76
Nevada	.87	1.14	−.27	2.04	1.05
Louisiana	.62	.90	−.28	1.11	.73
North Dakota	.82	1.10	−.28	n.a.	.88
Nebraska	.73	1.01	−.28	n.a.	.78
Alabama	.65	.94	−.29	1.56	.72
Missouri	.77	1.06	−.29	1.46	.79
Iowa	.67	.97	−.30	n.a.	.63
Illinois	.72	1.08	−.36	1.23	.80
South Dakota	.58	.96	−.38	n.a.	.82
Mean:	.76	.99	−.23	1.30	.79

[a]Ranked by the difference between the relative Medicaid nursing home reimbursement rate and the relative personal health care expenditures per capita.

[b]Medicaid recipients of services provided by skilled nursing facilities and intermediate care facilities other than those for the mentally retarded.

Source: Health Care Investment Analysts, Inc. and Arthur Andersen (1991); HCFA State Medicaid Data Disk for FY1989; Families USA Foundation (1990); and National Governors' Association (1989).

average. If a state has comparatively expensive personal health care, its nursing home reimbursement rate presumably ought to reflect these costs. Accordingly, if a state's relative personal health care costs greatly exceed its relative Medicaid reimbursement rate, its regulators are probably pushing the nursing homes to control Medicaid costs. By contrast, a large positive gap would suggest relative generosity on the part of the regulators.

Table 2 indicates that cost-shifting activity appears to be widespread even where regulators have not pursued particularly stringent reimbursement strategies. Nevertheless, the table provides some evidence that cost shifting is somewhat more pronounced in the high-pressure states. The mean ratio of private-to-Medicaid reimbursement is 1.18 for the states where the relative Medicaid reimbursement rate exceeds relative personal health care costs, but it is 1.30 for the states where the Medicaid reimbursement rate looks relatively low.

After presenting alternative interpretations of these ratios, this article will discuss how these measures help to clarify whether or not the long-term care portion of a state's Medicaid program is relatively low cost. The article will then explain why cost shifting matters.

Alternative Interpretations of These Ratios

When the average Medicaid reimbursement rate is less than the average daily revenue per private patient, Medicaid recipients are either 1) receiving lower-quality care in the same homes; or 2) are congregated in less expensive nursing homes; or 3) are on average less disabled (and, thus, less costly to care for)[3] than nursing home residents paying privately; or 4) are in a program subsidized by the private sector. In general, a ratio above 1 probably suggests some combination of these four interpretations.[4]

In most states it is not legal for a nursing home to give lower-quality care to a Medicaid recipient than to a private payor. In addition, once having accepted a patient, a nursing home usually cannot force that person to move if her savings are exhausted and she becomes a Medicaid recipient. In other words, the first rationale is probably not the primary explanation for the ratios.[5]

Nevertheless, some states do allow nursing homes to refuse admittance to individuals already dependent on Medicaid. As a consequence, Medicaid recipients may cluster in less expensive (lower quality?) institutions. In addition, Medicaid recipients may congregate in less expensive homes by happenstance rather than as a result of nursing home policy. For instance, a Medicaid recipient's family might prefer to have their relative in a nearby home that can be reached by public transportation, whereas private payors' families are able to drive to homes in relatively affluent suburbs or rural areas. Accordingly, a ratio above 1 could indicate that some Medicaid recipients are having trouble getting access to average quality long-term care or that Medicaid recipients are choosing less expensive homes.

Like the first, the third possibility—that Medicaid recipients are on average less disabled than nursing home residents paying privately—seems unlikely. After all, the frail elderly tend to become more disabled over time, and nursing homes are known to give some preference in the admissions process to applicants with private support. Moreover, a significant share of Medicaid recipients were originally private patients who, after some years in a nursing home, have exhausted their savings and "spent down" to Medicaid eligibility.[6] In addition, research on publicly supported home and community care programs generally concludes that they do not succeed in delaying Medicaid recipients' admissions to nursing homes. In other words, most Medicaid recipients admitted to nursing homes really need to be there; by the time they are admitted, there is simply no alternative to nursing home placement (Weissert 1991). Finally, as will be discussed further below, the ratio of 0.86 for Minnesota, where by law private patients cannot be charged more than comparable Medicaid patients, suggests that Medicaid recipients

[3] Because the cost of providing care rises with the disability of the patient, some states use a case-mix system in setting Medicaid reimbursement rates. Accordingly, if the average Medicaid recipient were less disabled than the average private patient, the ratio of the private to the Medicaid reimbursement rate could be above 1.

[4] A thorough investigation of these issues would require examining private/Medicaid reimbursement rates institution by institution with adjustments for case mix. Currently available data do not permit such an undertaking.

[5] On the other hand, even North Dakota, one of the three states that regulate private rates, permits nursing homes to charge more for private rooms and other services that Medicaid does not allow.

[6] States have the option of providing Medicaid coverage to "medically needy" people. Under this option, individuals who fit into Medicaid-eligible categories but are poor only because of high health care expenses may "spend down" to meet Medicaid income and asset criteria. They "spend down" by incurring medical or remedial care expenses that reduce their remaining income and liquid assets to a level below that allowed by their state's program.

are on average *more* disabled than their privately supported counterparts—contrary to the third hypothesis.

By contrast, the fourth possibility—that private payors are subsidizing the Medicaid program—is generally accepted to be a common occurrence (Health Care Investment Analysts, Inc. and Arthur Andersen 1991, p. 9). Within individual institutions private patients often pay more than comparable patients supported by Medicaid. Similarly, a given patient usually pays a higher rate when paying privately than he does once his savings are gone and he is forced to turn to Medicaid.

Why have private payors tolerated these cost shifts? For several reasons. Private applicants generally turn to nursing homes as a last resort, when they are desperate. Then they frequently face waiting lists because state regulators usually limit the supply of beds in order to control Medicaid spending. Moreover, these nursing home residents or their families tend to have a strong preference for a particular region; thus, they agree to pay high nursing home charges, without looking very far afield for lower-cost alternatives. Finally, they know that, should worst come to worst and their savings be exhausted, Medicaid will generally provide for their continued care.

Heretofore, the extent of this cost-shifting activity has not been easy to measure. Only three states—Connecticut, Minnesota, and North Dakota[7]—currently regulate nursing home charges to private patients, and, as mentioned previously, most states do not publish average, let alone institution-specific, private rates. Nevertheless, previous research provides limited evidence of such practices (Birnbaum and others 1981, for example), and cross-subsidies are relatively well documented in hospital care. For example, at a Redwood City, CA hospital, commercial inpatient revenue exceeds the cost of care received by these private patients by roughly 50 percent; the hospital is shifting uncompensated Medicaid and Medicare expenses to the private payors in order to cover their operating costs by a slim margin (Marchasin 1991).

That private payors are subsidizing the Medicaid program is generally accepted to be a common occurrence.

But what about the four states—Kentucky, Minnesota, New York, and Rhode Island—where average private revenue appears to be well below the average Medicaid per diem? The possible explanations are most limited in Minnesota, one of the three states that regulate private rates. Minnesota nursing homes cannot charge private patients more than comparable Medicaid patients. Accordingly, the state's relatively low ratio of 0.86 suggests that Minnesota Medicaid recipients may be more disabled than the average private resident[8] or that Medicaid recipients may be living in relatively expensive institutions or areas.

Case mix differences and geography are also likely to play significant roles in other states, such as New York.[9] In that state, Medicaid recipients may be clustered in the metropolitan area where land is

[7] Connecticut regulates but does not necessarily equalize private and Medicaid rates. North Dakota equalizes rates for similar services but permits nursing homes to charge for single rooms and other services that Medicaid does not allow. In Minnesota, private patients cannot be charged more than comparable Medicaid patients.

[8] If Minnesota's Medicaid recipients really are more disabled than the average nursing home resident, and if that result is typical of other states, the implication is that the subsidization of the public sector by the private sector may be even greater than the ratios in Table 1 suggest. If, by contrast, Minnesota is atypical and Medicaid recipients are generally less disabled than the average nursing home resident paying privately, then that situation would help to explain why revenues per private patient day tend to be greater than Medicaid per diems. Perhaps, for example, private payors have the resources to pay for home care and, thus, are able to maintain their independence for longer than a similarly frail person without such resources. In fact, however, this second possibility seems unlikely, since public support for home and community care programs appears to be a (much-needed) supplement to but not a substitute for long-term care. These programs do not appear to delay an individual's entry into a nursing home; most states restrict the supply of nursing home beds to the point where applicants face waiting lists, and most people do not seek to enter a nursing home until there is simply no alternative. In other words, the implications of the Minnesota data should be kept in mind.

[9] A telephone call to the Medical Assistance Division of the New York State Department of Social Services yielded a figure for the 1992 average Medicaid reimbursement rate but no information on the average charge for private patients. Comparing the average private charge cited in a press report (Freudenheim 1992) with the Medicaid reimbursement rate provided on the phone results in a private/public ratio of 1.35. Such a ratio suggests that cost shifting currently is as prevalent in New York as it is nationally rather than well below average, as suggested by the figure in Table 1 for 1989. Although the author does not know how either of these "averages" was calculated, this result raises a red flag and underscores the difficulty of estimating and interpreting average revenue per private patient day from publicly available data.

relatively expensive and nursing homes must compete with big city hospitals for direct care staff.[10] By contrast, private patients may be more heavily represented in upstate nursing homes with below-average operating costs. As for case mix differences, New York is unusual in using RUG (resource utilization group) methods to adjust Medicaid reimbursement to reflect the intensity of care required by each recipient. The results again suggest that the average Medicaid recipient may be more disabled than the average private patient. In addition, New York is reputed to have an unusual number of very expensive, special-purpose facilities, such as those for people with head injuries or Alzheimer's disease.

Nursing home reimbursement ratios are rough indicators and must be interpreted with great care.

In addition, policymakers in both New York and Rhode Island have been concerned about an incipient shortage of nursing home beds. For example, a Rhode Island assessment projects a shortage of 2,250 beds, or 20 percent of the 1989 level, by 1995. Accordingly, policymakers in those states may have offered relatively generous Medicaid reimbursement rates to lure additional providers into the business. Similarly, in 1989 Minnesota had the highest ratio of nursing home beds to its elderly population of any state in the country.

Finally, as reported to the National Governors' Association, Medicaid reimbursement rates are not entirely comparable across states. For example, New York and Kentucky are two of 10 states that include (expensive) prescription drugs in Medicaid's per diem reimbursement rate for the nursing homes. Other states cover some of these services separately.

In sum, these ratios must be interpreted with great care. They are rough indicators and should be assumed to reflect a mix of geographic happenstance, quality/access/case mix issues, and cross-subsidization. Nevertheless, cost-shifting activity appears sufficiently widespread that the ratios shown in Table 1 can be used to flag the states where private subsidization of public responsibilities could be substantial.

What the Ratios Suggest about the States' Medicaid Programs

Cost-shifting efforts complicate the task of interpreting cross-state differences in Medicaid payments per recipient. As already mentioned, relatively low payments per recipient could suggest relative efficiency, but they could also suggest relatively poor quality or an above-average amount of cost-shifting activity. Alternatively, states with relatively high payments per Medicaid recipient may be shifting public costs onto private payors to a below-average extent. The ratios shown in Table 2 help to narrow the likely interpretations.

For example, as Table 2 indicates, among the 14 states where Medicaid payments per recipient of nursing home services appear low relative to the U.S. average (with a ratio of 0.80 or less) *and* to their own personal health care costs, only three—Michigan, Oregon, and Tennessee—do not appear to be shifting costs to the private sector. In these three states, nursing home care appears to be truly low-cost, either because the industry is relatively efficient or because the quality of care is relatively spartan—for everyone. For another three states, Colorado, Louisiana, and Texas, the data suggest that the cost shifts may be quite modest. But the other "low-cost" states appear to be shifting part of the burden supposedly carried by the state's taxpayers onto specific members of the private sector, individuals paying privately for long-term care.

Among the "high-cost" states, New York and Minnesota may seem to be high cost in part because they are *not* shifting public costs to the private sector; in these states, the taxpayers appear on average to be bearing the full cost of the Medicaid long-term care programs, even if these programs are of unusually high quality or below-average efficiency. By contrast, the other high-cost states have expensive Medicaid programs (high-quality/inefficient?) *despite* sizable cost shifts to private payors.

[10] According to the associate commissioner of the Medical Assistance Division of the New York State Department of Social Services, nursing homes in New York currently charge private patients $62,000 per year, on a statewide average basis, and $69,000 in the metropolitan area. In other words, the metropolitan area rate is 11 percent above the state average rate for private patients (Freudenheim 1992).

[11] Many citizens also do not realize that Medicaid beneficiaries generally pay for a significant share of their nursing home care from their own income. Except for a very small personal allowance, a Medicaid recipient's Social Security, pension, or other income is devoted to paying as much as possible of the Medicaid reimbursement rate set by the state; Medicaid pays the balance.

Why Does Cost Shifting Matter?

Cost shifting matters because it is inequitable. It also encourages inflation of nursing home costs.

To start with the issue of equity, when an individual requires nursing home care, the emotional and financial burdens facing that person and his family are substantial. On top of these burdens, a relatively small number of private patients are then asked to finance a significant part (often 20 to 30 percent) of the cost of the Medicaid long-term care program, a program that citizens believe to be broadly tax-funded.[11]

Furthermore, asking a private patient to provide one-third, say, of a Medicaid recipient's care in addition to supporting herself will only hasten the day when the private payor exhausts her savings and spends down to Medicaid eligibility—a painful transition for all involved. Indeed, the realization that cross-subsidies are commonplace could very well encourage resort to the asset protection schemes that are currently so widely advertised.[12] Such schemes have led some members of the press to describe Medicaid's long-term care provisions as a "middle-class entitlement program." While some families may benefit from these divestment efforts, others are forced by cross-subsidies to contribute disproportionately to the nation's long-term care program.

In addition, permitting nursing homes to shift Medicaid costs onto private payors undermines policymakers' efforts to slow the rise in Medicaid expenditures. Let's assume, for example, that a state's rate-setters wish to encourage nursing home operators to control costs and become more efficient. To this end, they raise the Medicaid reimbursement rate by slightly less than the increase in nursing home operating costs during the previous year. However, the nursing home operators remain free to increase the gap between the Medicaid per diem and the private rate and do so. Accordingly, the nursing home managers retain the ability to give their employees a slightly bigger pay increase, or hire an extra aide, or plant an extra tree. These "extras" become incorporated into the costs on which Medicaid rates are based the following year.

Cost shifting matters because it is inequitable. It also encourages inflation of nursing home costs.

To see whether cost-shifting activity appears to contribute to higher-than-expected nursing home operating costs to a statistically significant extent, this study uses simple regression analysis. The dependent variable is state average annual pay for nursing home workers, relative to the U.S. average; wages, which represent roughly half of nursing homes' total operating costs, are a proxy for "costs."[13] The independent variables are the state's relative Medicaid reimbursement rate, the ratio of the average revenue per private patient day to the Medicaid reimbursement rate, and relative average annual pay for individuals working in the retail sector. This last variable represents local wage costs in a sector where wages are set by market forces.[14] In each case the relationship between the independent and the dependent variables was expected to be positive: holding the other explanatory variables unchanged, the higher the relative Medicaid rate, the local retail wage, or the greater the amount of cost-shifting activity, the

[12] Federal and state legislation limit a Medicaid recipient's ability to shift assets, but loopholes exist and should be closed at the federal level. For example, federal legislation limits the "look-back" period to 30 months. For progressive conditions like Alzheimer's, however, such a look-back period is too short.

[13] The regression presented in this article uses average annual pay as a proxy for costs, because total nursing home operating costs are very closely related to total revenue, from which the estimates of revenue per private patient day were derived. In other words, a regression equation using total operating costs and the ratio of private to Medicaid reimbursement appeared to involve a near identity. Nevertheless, a regression in which the ratio of state operating costs to U.S. average costs is the dependent variable, and the independent variables are the relative Medicaid reimbursement rate, relative per capita personal health care costs, and the ratio of private to Medicaid reimbursement, yields results similar to those shown in Table 3. Holding relative Medicaid reimbursement rates and personal health care costs unchanged, a greater amount of cost-shifting activity is positively associated with higher total operating costs to a statistically significant extent. Interestingly, however, in this case, relative personal health care costs do not appear to have a statistically significant link with relative nursing home operating costs.

[14] The retail sector is broadly defined and includes drug stores, fast food restaurants and gas stations. These establishments employ many low-skilled, part-time workers. In other words, nursing homes and retail establishments employ similar types of labor; however, while retail wages are set competitively, market forces do not operate very freely in the nursing home industry.

Table 3
Relationship between the Ratio of Average Private[a] to Average Medicaid Reimbursement Rates and the Average Annual Pay of Nursing Home Workers, 33 States,[b] FY1989

Dependent Variable: Average Annual Pay for Nursing Home Workers	
Independent Variables:	
Constant	−.46** (−4.79)
Medicaid Reimbursement Rate Relative to the U.S. Median	.58** (8.68)
Ratio of Private[a] to Medicaid Reimbursement Rates	.08* (2.32)
Average Annual Pay for Retail Workers	.81** (6.58)
R^2 (adjusted)	.91

Numbers in parentheses are t-statistics.
**Statistically significant at the 1 percent level.
*Statistically significant at the 5 percent level.
[a]Estimated.
[b]All states except Alaska, Arizona, Arkansas, Hawaii, Indiana, Iowa, Kansas, Maryland, Mississippi, Nebraska, New Hampshire, North Dakota, Oklahoma, Pennsylvania, South Dakota, Washington, and West Virginia.
Source: Health Care Investment Analysts, Inc. and Arthur Andersen (1991); HCFA State Medicaid Data Disk for FY1989, Families USA Foundation (1990); National Governors' Association (1989); and U.S. Bureau of Labor Statistics, ES202 Data Tape.

higher the relative average annual pay for nursing home workers was expected to be.[15]

Table 3 presents the results of this very simple regression analysis. The results suggest that additional cost-shifting activity does appear to be associated with relatively high nursing home costs to a statistically significant extent. In other words, cost-shifting activity appears to contribute to the process by which nursing home costs are driven higher than can be justified by the state's relative wage rates. The results of most regression analysis should be taken with a large pinch of salt; these results undoubtedly require a larger pinch than usual, because the ratio of the estimated private to Medicaid reimbursement rate is subject to the many interpretations already discussed. Nevertheless, the results are suggestive.

Policy Implications

The ratios presented in this article could reflect happenstance, access problems, case mix differences, or cost-shifting activity. Whatever the cause, a ratio significantly different from 1 has implications that deserve policymakers' attention. If the explanation for a ratio above 1 is case mix differences, for instance, the state's taxpayers might benefit from better screening and case management procedures. If the problem is access, Medicaid recipients may be left in hospitals where beds are even more expensive than they are in nursing homes. Thus, wherever the private-public ratio is significantly different from 1, policymakers ought to know why.

If, after further investigation, a ratio above 1 turns out to indicate a significant amount of cost-shifting activity, policymakers may want to require nursing homes to charge the same (institution-specific) rate for comparable patients, whether these individuals are supported by their own resources or by the public sector. In other words, policymakers may want to establish an all-payor system comparable to those frequently advocated in the context of acute or hospital care. For the average state, such a one-time change in policy would raise Medicaid expenditures by roughly 5 percent.[16] At least half of the additional expenditure would be reimbursed with federal matching funds. In return for a net 2.5 percent increase in Medicaid expenditures, policymakers would reduce some sizable inequities, slow the pace at which private payors spend down to Medicaid eligibility, improve observers' ability to measure relative efficiency across state Medicaid programs, and enhance the regulators' ongoing ability to control Medicaid costs.

[15] Of course, the direction of causality—from cost-shifting to relative costs or from relative costs to cost-shifting—is not clear. However, if the direction of causality is the reverse of that assumed in this regression, then relatively high nursing home costs not validated by the regulators appear to lead to additional cost-shifting activity. Accordingly, policymakers would probably want to take account of any positive link between relative costs and cost-shifting activity, whatever the direction of causality.

[16] Assumptions: the ratio of the average private charge to the average Medicaid reimbursement rate is 1.21 and this ratio reflects cost shifts rather than geographic or case mix differences; Medicaid accounts for 63 percent of the total resident days and for 44 percent of payments to nursing homes made for Medicaid recipients; Medicaid must pay all of the increase in the Medicaid reimbursement rate required to bring the Medicaid per diem up to the average revenue per patient day; and nursing home care accounts for 28 percent of total Medicaid vendor payments.

At a time when total state Medicaid payments have been rising 19 percent per year, a one-time increase of 2.5 percent does not seem like an exorbitant price to pay for these advantages. But, of course, these are difficult days for state budget makers, who have been under great pressure to cut Medicaid spending, one of the fastest-growing items in most state budgets. Under these circumstances, state policymakers may conclude that they would rather ask moderate-income families to subsidize the long-term care portion of the Medicaid program than 1) cut Medicaid coverage of acute care for poor children and others or 2) raise state taxes.

The public should be aware of the choices being made. One problem with subsidies is that they are largely invisible.

Nevertheless, the United States Congress and most state legislatures have agreed that society has a responsibility to support elderly and disabled individuals who need nursing home care and meet Medicaid income and asset eligibility criteria. If those responsibilities are not being fulfilled in certain states, the public should be aware of the choices being made. One problem with subsidies is that they are largely invisible.

Requiring nursing homes to charge comparable publicly and privately supported patients the same rates would end the inequity of asking individuals paying privately to subsidize a public responsibility. It would also help state regulators to get a better handle on Medicaid costs. However, it would not alleviate other problems inherent in the U.S. approach to long-term care. In particular, because the need for long-term care is uncertain, and the actuality is extraordinarily expensive, most families do not or cannot save enough to pay for more than a brief stay in such an institution.

To start with the first point, only a relatively small fraction of the population ever actually requires long-term care. In 1990 nursing home residents of all ages (but almost 85 percent are aged 75 and above) equaled less than 1 percent of the entire population, less than 5 percent of the population aged 65 and above, and 11 percent of the population aged 75 and over. But for those comparatively few individuals who do require nursing home care, the cost is staggering—with the average charge to private payors approaching $30,000 per year. Accordingly, families who are unfortunate enough to need long-term care and who play by the rules of the game often end with their savings exhausted and their relative a "burden to society." Others may succeed in hiding their assets, but their relative still becomes a burden to society.

When people tend to undersave for or underinsure against a risk that can produce a significant liability for the taxpayers, governments generally intervene: they either create tax incentives for purchasing private insurance, or they establish a social insurance program to pay for the required services directly.[17] The United States has followed both approaches in dealing with acute medical care, for which individuals also tend to underinsure. For workers and their families, the government has created tax incentives that encourage employers to provide health insurance as a fringe benefit. For the elderly and the disabled, who are not expected to work, the government has established Medicare, an extremely popular social insurance program. But, as already discussed, Medicare provides very limited coverage of long-term care.[18]

[17] New York state is introducing a somewhat different approach to this problem. The new program will encourage state residents to buy private insurance covering three years of nursing home care and six years of home care. Those who do so but require additional care will be allowed to become eligible for Medicaid without spending down their assets. Connecticut and Indiana have started, and California is applying for federal approval to start, somewhat similar programs (Freudenheim, May 3, 1992).

In the absence of a federal initiative, this innovative program is clearly a step in the right direction: it will probably save the taxpayers some money and will stimulate a market for (more) affordable long-term care insurance. More people will have an incentive to buy, thereby reducing the problem of adverse selection, and the state will limit the insurers' liability. Nevertheless, this program will primarily benefit upper- and middle-income people, while lower-income individuals may not be able to afford the private insurance (with estimated annual premiums of $1,500 to $2,000 for policies bought at age 65) and will have to spend down to become eligible for Medicaid. Moreover, this approach still does not spread the burden of paying for long-term care as widely as would a federal social insurance program. Should many states follow this approach, the insurance companies will undoubtedly force a reduction in the ratio of private to Medicaid reimbursement rates. These private insurers will not want to subsidize the Medicaid program for very long.

[18] Medicare coverage of nursing home care is currently limited to 100 days and requires prior hospitalization and the need for skilled services. In other words, it covers convalescent or terminal care, not long-term care.

What, then, should be done about individuals' tendency to underinsure against the risk of needing long-term care? One solution would be to expand Medicare coverage to include long-term care and to transfer the long-term care portion of Medicaid to that program. Congress did not provide Medicare coverage for long-term care when the program was first established because members were concerned about incurring an open-ended fiscal liability. Now that the country has had 25 years' experience with Medicaid coverage of long-term care, however, it seems clear that the demand for nursing home services will not surge merely because Medicare coverage has become available. Rather, it seems that individuals generally resist going into a nursing home as long as any viable alternative remains. Nevertheless, predictable demographic changes indicate that the need for long-term care will continue to grow unless health care providers find a way to prevent or alleviate such disabling conditions of aging as Alzheimer's disease and arthritis.

Once government limited an individual's liability for costly nursing home care, the private market for Medigap insurance for long-term care would develop rapidly.

Expanding Medicare to cover even the current use of long-term care would have a major impact on the Medicare program. In 1989 such a change would have meant expanding Medicare expenditures by 50 percent. However, because federal, state, and local taxes already pay for one-half of all long-term care through Medicaid, the net impact on U.S. citizens' tax burden would have amounted to 25 percent of 1989 Medicare expenditures.[19] It is worth remembering, moreover, that such a change would most likely have had little impact on society as a whole, since society pays, one way or another, for required long-term care.[20] The change would simply have spread the burden more evenly across society instead of focusing half of the total cost on the relatively few individuals (less than 1 percent of the population) who now pay for all or some of their nursing home care from private sources.

Given the strained state of the hospital insurance trust fund (the primary source of funding for Medicare)[21] and the negative reaction to and eventual repeal of The Medicare Catastrophic Coverage Act of 1988, paying for an expansion of the Medicare program might present political difficulties. Nevertheless, several options exist. For example, the government could raise FICA rates, or it could raise the maximum income subject to FICA taxes. It could raise federal estate tax rates (which are lower than those in most other industrial countries) and earmark the revenue for the long-term care portion of Medicare. It could tax relatively wealthy Medicare recipients on these benefits.

To limit the fiscal burden and give consumers some incentive to shop for efficient institutions and consider alternatives to institutionalization, the government could also require moderate co-payments for nursing home services. These co-payments could be calculated on a sliding scale or they could be capped.

Once the government had limited an individual's liability for costly nursing home care, the private market for Medigap insurance for long-term care would undoubtedly develop rapidly. At present, the private insurance market for long-term care coverage is severely underdeveloped[22] because of the huge potential liabilities involved and the problem of adverse selection. Only the individuals who know that they are most at risk of needing long-term care are likely to buy the expensive and limited policies currently available. Indeed, the contrast between Medigap health and long-term care insurance coverage is dramatic: 75 percent of the elderly have Medigap health insurance policies; only 3 percent of the elderly have long-term care insurance (Health Care Investment Analysts, Inc. and Arthur Andersen, 1991, pp. 8–9). These figures suggest that the private insurance market could expand rapidly once Medicare entered the long-term care picture.

[19] Presumably, thus, an individual paying the maximum FICA tax (on $48,000) in 1989 would have had to pay $174 more to the hospital insurance trust fund. The supplemental medical insurance premiums and the contributions from general revenue would have been proportionately greater as well.

[20] This assertion assumes that the demand for nursing home care is not very price-elastic. Because most people resist institutionalization, this assumption seems reasonable in the case of nursing home care, but it may well not apply to the demand for home or community care.

[21] The primary sources of funds for Medicare are the hospital insurance trust fund (which, in turn, is largely funded by payroll taxes), general revenue, and Medicare premiums.

[22] At present, private insurance provides only 1 percent of payments to nursing homes.

Conclusions

This article has presented estimates of the extent to which individuals paying privately for nursing home services are being asked to subsidize their state's Medicaid program. Although anecdotal evidence has long indicated that such cross-subsidies are widespread, measuring the extent of this activity has been very difficult because states do not publish data on nursing home charges to private patients. Now, however, recently published data permit researchers to make rough estimates of the ratio of nursing home revenues per private patient day to the Medicaid per diem reimbursement rate across states. Although these ratios undoubtedly reflect differences in access and degree of disability as well as cost shifts, these data suggest that cost-shifting activity may be occurring in over three-fourths of the states for which these calculations can be made. And it is occurring in states where policymakers have been relatively open-handed as well as in states where policymakers have been relatively unaccommodating. In states where the private sector appears to be subsidizing the Medicaid program, the subsidies amount, on average, to one-third of a Medicaid recipient's care.

Combined with previous research, these data help to distinguish between states where the Medicaid long-term care program is actually relatively low-cost and states where the long-term care program only appears to be low-cost because private individuals are subsidizing the public sector. Similarly, some states' Medicaid expenditures may appear relatively high partly because the taxpayers are paying the full cost of the program. Unfortunately, the data still do not permit researchers to distinguish differences in quality from differences in efficiency other than through institution-by-institution inspection.

Having found evidence of widespread cost-shifting activity, this article argued that such behavior is highly inequitable and may be self-defeating from a policymaker's perspective. It clearly hastens the absorption of private payors' resources and either encourages middle-income people to use asset protection schemes or speeds their transition to Medicaid dependence. It also complicates the task of regulating nursing home operating costs and curbing the rapid rise in Medicaid expenditures.

Accordingly, policymakers may want to investigate more thoroughly than is currently possible using publicly available data why a state's average private reimbursement rate differs from its Medicaid rate. Any ratio substantially different from 1 could indicate a situation requiring policy action. If the explanation turns out to be cost shifting, the authorities may want to require nursing homes to charge comparable patients the same amount. Such a change would restore the social contract and help to slow spiraling Medicaid costs. For the average state, the one-time transition cost would be a fraction of recent (albeit painfully large) increases in its annual Medicaid bill.

Expanding Medicare to cover long-term care most likely would have little impact on society as a whole, since society pays, one way or another, for required long-term care.

However, requiring nursing homes to charge all comparable residents the same rate would not address the private sector's tendency to undersave or underinsure against the uncertain but very costly risk of needing long-term care—with the consequence that a significant number of people become dependent on public support. In dealing with this issue in the context of acute care needs, the government has chosen to use tax incentives to encourage employers to provide health insurance as a fringe benefit in the work place. For the elderly and disabled, it has established Medicare, an extremely popular social insurance program. The data in this article bolster the case for including long-term care coverage in Medicare.

References

Birnbaum, Howard, A. James Lee, Christine Bishop and Gail Jensen. 1981. *Public Pricing of Nursing Home Care.* Cambridge, MA: Abt Books.

Families USA Foundation. 1990. *Rising Health Costs in America: 1980–1990–2000.* A Families USA Foundation Report in Cooperation with Citizen Action. Washington, D.C.: October.

Freudenheim, Milt. 1992. "Medicaid Plan Promotes Nursing-Home Insurance." *The New York Times*, May 3, 1:1.

Health Care Investment Analysts, Inc. and Arthur Andersen. 1991. *The Guide to the Nursing Home Industry.* Baltimore, MD.

Little, Jane Sneddon. 1992. "Lessons from Variations in State Medicaid Expenditures." *New England Economic Review*, January/February, pp. 43–66.

Marchasin, Sidney. 1991. "Cost Shifting: How One Hospital Does It." *The Wall Street Journal*, December 9.

National Governors' Association. 1989. "Medicaid-Memo: Medicaid Payment for Nursing Facilities: Skilled Nursing Facilities and Intermediate Care Facilities (1986–1989). Washington, D.C.

U.S. Congress. House of Representatives. Committee on Ways and Means. 1990. *1990 Green Book; Overview of Entitlement Programs; Background Material and Data on Programs within the Jurisdiction of the Committee on Ways and Means* (June 5). Washington, D.C.: Government Printing Office.

Vertrees, James C., Kenneth A. Manton, and Gerald S. Adler. 1989. "Cost Effectiveness of Home and Community-Based Care." *Health Care Financing Review*, vol. 10, Summer.

Weissert, William G. 1991. "A New Policy Agenda for Home Care." *Health Affairs*, Summer, pp. 67–77.

Purchasing Power Parity within the United States

Geoffrey M.B. Tootell

Economist, Federal Reserve Bank of Boston. *The author wishes to thank Adam T. Rosen for valuable research assistance.*

Economics has many articles of faith. One of the most dearly held is Purchasing Power Parity, which posits that the price of the same good in different regions should be equivalent when no barriers to arbitrage exist. If a BMW costs $20,000 in Germany and $40,000 net of transportation costs in the United States, some entrepreneur will start buying BMWs in Germany and sending them to this country. BMW's profits may suffer because of its decreased ability to segment these two markets, but the arbitrager will reap huge gains. Purchasing Power Parity is an important assumption in much of international economic theory, and this article examines empirical evidence in support of this proposition.

To date, empirical support for Purchasing Power Parity (PPP) has been mixed. Dornbusch (1978, 1985), Frenkel (1981), Roll (1979), and Giovannetti (1992), as well as work by Meese and Rogoff (1983), have found varying degrees of evidence that PPP fails; in both the short run and the long run, single-currency prices do not equilibrate. On the other hand, studies using annual data over a very long time period, as in Friedman (1980) and Edison and Klovland (1987), have found some support for PPP. But even though some long-run support for various versions of PPP exists, particularly during fixed exchange rate regimes, recent evidence has not supported this simplest of market mechanisms.

The negative empirical results for PPP have elicited two types of responses. The first stresses the flaws in the actual price indices compared, while the second approach places a renewed emphasis on theories that postulate no long-run, stable relationship between prices in different countries. Because price indices in different countries include different goods, many of them nontraded, and different weights for each good, PPP can fail when relative prices change. As a result, PPP may only appear to fail, because of an index number problem. Alternatively, theories that predict changes in the relationship between these prices have been revitalized, based on the evidence that PPP fails. These

theories include both macro models, represented by traditional Keynesian analysis or the new real business cycle theories, and micro models, emphasizing pricing to segmented markets. The macro models tend to emphasize the relationship between changes in the nominal exchange rate and PPP.[1] The micro models stress the goods market structure, the idea being that a firm such as BMW would, and could, segment the markets in these two countries, violating PPP.

This article injects new evidence into the debate. Instead of examining international data, this study analyzes PPP between regions of the United States. By comparing regions within a country, it eliminates many of the hypotheses offered to explain the failure of PPP. The nominal exchange rate between any two regions is immutably fixed at one, thus avoiding any violations of PPP related to its movements, predicted in either the Keynesian or real business cycle macro models. Since national monetary and fiscal policies are essentially identical across regions, different policy mixes cannot be blamed. The new micro theories of why PPP fails, hypothesizing market segmentation, also become much less believable; the ability of firms to isolate markets is much more limited within a country than between countries, as goods arbitrage is much easier. Finally, price indices within a country are much more consistent than those collected internationally because the same agency collects data for each region, the same types of goods and methodology are used, and the weights on each good are more uniform. Yet, even in the absence of these postulated causes for the failure of PPP, empirical tests of PPP fail to hold between different regions in the United States.

Instead of examining international data, this study analyzes PPP between regions of the United States.

The results of this study are suggestive. Volatility in the nominal exchange rate is not required for PPP to fail, since PPP fails to hold within regions of the United States. Instead, the inclusion of nontraded goods in the total consumer price indices (CPIs) for these regions is shown to be the major cause of this failure. When the nontraded components of these indices are removed, PPP holds. Some categories of goods do seem to move in lockstep while others do not, as one would expect, and as PPP predicts.

The next part of this article defines the real exchange rate and explains why it is so closely related to PPP. Previous explanations for PPP's failure are also examined in more detail. Section II outlines the empirical model; a brief description of recent analysis on the statistical properties of time series is required to fully appreciate its elegance. The third section presents the results, and Section IV discusses the implications. A conclusion follows.

I. The Real Exchange Rate

Purchasing Power Parity,

$$P = eP^*, \tag{1}$$

relates the price of a good in the home country, P, with its own-currency price in another country, eP^*. The nominal exchange rate, e, represents the home-currency value of a unit of the foreign currency. The "strong form" of PPP sets this relationship as an equality. The real exchange rate, RXR, is derived by a simple rewriting of the PPP condition,

$$RXR \equiv \frac{e\ CPI^*}{CPI_{us}}, \tag{2}$$

and represents the domestic-currency price of foreign goods relative to home goods. In the BMW example, the real exchange rate would equal one-half: the dollar value of the BMW in Germany divided by its dollar value in the United States.

The strong form of PPP predicts that the real exchange rate equals one, as these two prices equilibrate. In the short run, the real exchange rate may deviate from this value, but in the long run it cannot stray too far from one if the strong form of PPP is to hold. Thus, tests of PPP, tests of equation (1), are identical to tests of whether the real exchange rate equals one. Using consumer price indices (CPIs), and occasionally wholesale price indices, such tests have shown the real exchange rate to diverge significantly from one, with little tendency to revert back. As a result, equation (1) does not hold by equality.

[1] Specifically, macro theory has analyzed the high correlation between the nominal and real exchange rates.

In reality, no theoretical justification demands that the exchange-rate-adjusted price indices used in these tests should equilibrate. The predominant reason that these indices need not be related in the long run is that the CPIs of different areas or countries include nontraded goods. PPP as a goods arbitrage condition should hold only for those goods that can be traded. Without trade, goods arbitrage has no direct way to equate a price in both markets. One cannot pick up a house and an acre of land in Wyoming and move it to downtown Tokyo, or even to downtown San Francisco, for that matter. Thus, the price of land need not be equal in these two regions.

Dornbusch (1978, 1985) emphasizes another potential cause of the instability of the real exchange rate and the failure of PPP; the weights of different goods need not be identical in the price indices of different countries. The French may drink wine and the Germans beer. If the price of wine rises relative to that of beer, the French CPI will rise relative to that of Germany since France weighs the increase in wine prices more heavily. As a result, relative price changes can produce permanent changes in the real exchange rate without violating the goods arbitrage condition. Changes in tastes, or in these weights, can also change the real exchange rate forever, making it look unstable. Comparing indices that include nontraded goods and are composed of differing weights seems to ensure that tests will show instability in the real exchange rate.

The CPIs in France and Germany can also differ for reasons other than measurement problems. In a world of imperfect substitutes, goods prices need not equate. Macro policies that can affect the nominal exchange rate can also affect the real rate. Since France and Germany may have drastically different and variable macroeconomic policies and are subject to different shocks to their production, the real exchange rate can move. Furthermore, imperfect substitution allows firms the possibility of discriminating by price; if demand for some good is much more elastic in Germany than in France and the firm can control the cross-border trade in its good, then the price of that good in Germany will be lower than it is in France. The real exchange rate can be unstable and differ from one for both macro and micro theoretic reasons.

Although measurement issues and theoretical explanations provide compelling reasons why PPP should fail, good reasons also support the idea that the real exchange rate should be stationary. If tastes do not change, then differences in weights need not cause changes in the indices. Different weights and the existence of tariffs and transportation costs may force the real exchange rate away from one. Yet the real exchange rate can still be stable around a mean other than one. For example, assume that the differential between the U.S. and German dollar prices of a BMW is caused by a 100 percent tariff. Although the average real exchange rate for BMWs is one-half, the real exchange rate is stable around that mean; doubling the dollar price in Germany would double the U.S. dollar price, leaving the RXR the same. This stability around the different mean is known as the "weak form" of PPP, and it is perfect for the analysis of an inflationary world.

The test of the weak form of PPP is simply a test of the stability of the real exchange rate, not a test of its stationarity around one.

Suppose the inflation rate is an evenly distributed 10 percent in the home country and zero in the foreign country; the exchange rate should depreciate by 10 percent to ensure that no real variables have changed, such as relative prices between home-produced and foreign-produced goods. Again, this exchange rate movement results from goods market arbitrage; the differential inflation rates alter the relative prices of foreign and domestic goods at the old nominal exchange rate. For example, if the equilibrium of the relative dollar prices of BMWs and Oldsmobiles is two, the 10 percent inflation of Oldsmobile prices at the same nominal exchange rate increases the dollar prices of American cars relative to BMWs. As a result, the increased demand for foreign goods, and therefore for foreign currency, and the decreased demand for domestic goods and currency at the old nominal exchange rate, drive the value of the domestic currency lower. The numerator in equation (2) rises proportionally to the denominator. Although the real exchange rate does not equal one and so the strong form of PPP fails, the weak form of PPP holds,

where the real exchange rate is stable. Thus, the test of the weak form of PPP is simply a test of the stationarity or stability of the real exchange rate, and not a test of the stationarity of the real exchange rate around one, as in the strong form of PPP. Development of the theory of a weak form of PPP has renewed the justification for tests of PPP.

Examining CPIs from different regions of the same country minimizes the above-mentioned measurement problems faced by international comparisons. Within the same country, similarity of culture across regions reduces the biases caused by relative price changes, since the consumers have very similar tastes and, therefore, weights. In fact, among the traded goods, the prices of exactly the same goods will be sampled in the various regions.[2] Furthermore, the many subcategories of the CPI within the United States permit an attempt to disentangle traded and nontraded goods. Even the existence of a nontraded good within the United States is more tenuous than it is internationally. Since labor is unquestionably mobile between regions in the United States, nontraded goods will be a decreasing percentage of the CPI in the long run. For example, it might not be possible to move an acre of land from Seattle to San Francisco, but someone, or some company, in San Francisco can very easily move to Seattle, equilibrating the prices of nontraded goods. The measurement problems with the international PPP studies are, to a large extent, avoided with an interregional study.

The examination of regions within a single country also removes much of the need to consider various theories to explain the failure of PPP. Identical federal monetary and fiscal policies diminish potential fluctuations of the real exchange rate. The absence of nominal exchange rate movements reduces most of the short-run volatility in real exchange rates seen when looking at international data. Finally, pricing to market is more difficult within, rather than between, countries.

Since many of the theoretical explanations for the instability of the real exchange rate do not apply to interregional comparisons, if real exchange rates are found to be unstable interregionally, then other causes must be investigated. As a result, domestic policy implications become more important. Finding the source of the instability will help produce policies to alleviate any problems caused by a long-run divergence of the regional CPIs. The results in this article shed some light on the sources of the breakdown in interregional PPP and provoke some speculation on possible policy implications.

II. The Empirical Model

Recent time series techniques, found in Dickey and Fuller (1979, 1981), Engle and Granger (1987), Hendry (1986), and Engle and Yoo (1987), for example, are designed to examine such issues as the long-run behavior of the real exchange rate. A brief outline of these procedures will facilitate analysis here.

Many macroeconomic time series, GNP or the money supply for example, tend to be unstable. They grow continuously rather than converging toward some long-run mean value. In an inflationary environment, the CPI also fails to revert to a stable level. A series is said to be "nonstationary" if it does not settle around a stable mean. The properties of the statistics that relate one series to another depend on the stationarity of the variables being examined; as a result, researchers must carefully test the stationarity of each variable to ensure that the proper statistical conclusions can be drawn. Dickey and Fuller have constructed the asymptotic properties of the distributions of such tests. A brief explanation of these tests is necessary, since they are at the heart of the analysis concerning the stability of the real exchange rate and therefore the validity of PPP.

If the value of a variable today, P_t, depends on its value in the past,

$$P_t = \beta\ P_{t-1} + \varepsilon_t, \tag{3}$$

and $\beta = 1$ and ε_t is a mean zero random variable with a finite variance, then P is said to be "difference stationary." In other words, in levels P is a random walk, but once it is differenced so that changes in P are examined, it is stationary around ε's zero mean. P's variance is infinite, because, with $\beta = 1$, the error at time $t - 1$, for example, is permanently remembered through the P_{t-1} term in equation (3); thus the variance of P_t is the sum of the variances of the preceding t errors, which gets infinitely large as time progresses.

The problem with comparing nonstationary series in levels is depicted in Figure 1a. The positive growth in both series over time fools the statistical analysis into believing the two series are closely

[2] When using the GDP deflator, changes in terms of trade are an important concern. Analyzing CPIs within one country eliminates this concern. Furthermore, the sample of goods is identical even if the weights are not; this is not clear in foreign comparisons, in which different goods might actually be included in the different samples, so that the weights are zero in their counterparts.

related. No stable long-run relationship may exist between these two variables, but a regression will falsely conclude that one exists; this mistaken conclusion derives from the "spurious correlation" that haunts data not screened for such trends. As a result, much of the empirical work in macroeconomics today is performed on changes, not levels.

All CPIs grow over time and a simple comparison of their levels, as in a direct test of PPP, would show a strong relationship. The regression of the CPI of Philadelphia on the CPI of Los Angeles has a coefficient near one, suggesting that PPP holds. But this coefficient occurs simply because both CPIs are increasing. Differencing the variables and testing their correlation is one method of avoiding any spurious element of the correlation.

Although two series may be nonstationary, they still may have a stable long-run relationship. Income and consumption are both nonstationary, but they are, in theory, tied to each other. In the short run, deviations from this long-run stable relationship occur, but eventually the two series converge to their long-run connection. When such a relationship exists, the two variables are said to be "cointegrated." Cointegration of income and consumption, for example, may be examined by analyzing the residuals from a regression of consumption on income,

$$Y_t = \gamma + \theta\, C_t + \mu_t, \qquad (4)$$

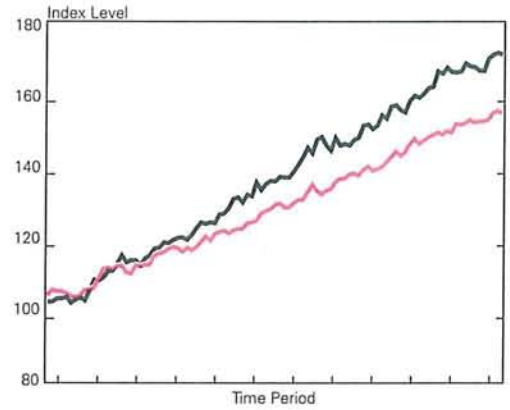
where Y_t is income at time t, C_t is consumption at time t, and both series are difference-stationary. The test for cointegration examines the time series properties of the residuals.[3] If μ is a random walk, then the two series can diverge in the long run. If, on the other hand, μ is stationary, the two variables do not diverge from a long-run stable relationship.[4] Figure 1A graphs two variables without a stable relationship; Figure 1B graphs two variables with one, consumption and income. Note that the two cointegrated series stay together, while the two series that are not cointegrated tend to separate.

A cointegration test is ideal for examining the stability of the real exchange rate. Since this article compares regions of the United States, where the nominal exchange rate is constant and equal to 1, the test of the stability of the real exchange rate is a test of whether the CPIs of two regions are cointegrated. Intuitively this makes sense; the cointegration test on the real exchange rate simply examines whether prices in one part of the country are tied to prices in another. Since these prices could differ because of the existence of nontraded goods, this study also attempts to separate traded and nontraded components of the indices, in order to find the origins of any difficulty.

Figure 1

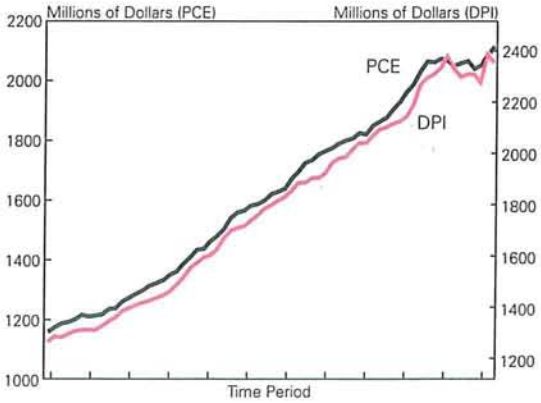
A

A Comparison of Levels of Two Nonstationary Series without a Stable Relationship



B

A Comparison of Two Cointegrated Series: Personal Consumption Expenditures and Disposable Personal Income



[3] Note that the cointegration test also could run income on consumption. The test is not a test of causality, but simply of whether a long-run relationship exists.

[4] In fact, Campbell and Mankiw (1989) find that income and consumption are difference-stationary and cointegrated.

III. The Results

The study uses CPIs for the Philadelphia, Chicago, New York, and Los Angeles metropolitan areas to examine the real exchange rates between these regions. Thus, six real exchange rates can be analyzed. Several problems existed with the data, however. First, each series had to be seasonally adjusted.[5] More importantly, the price index for the non-shelter component of the CPI for each of the six cities went back only to 1980, so this component had to be extended back to the longer sample.[6] Since the total CPI and its shelter component were available for all four cities back to 1956, the extension of the ex-shelter series was relatively simple.[7]

Before examining any cointegration tests, it is necessary to be sure that the variables to be compared are all of the same order of integration. For CPIs, that requires Dickey-Fuller tests to ensure that all the various price indices and subindices are nonstationary in levels and stationary in differences. Table 1 shows the confidence level at which the simple null hypothesis that the price levels of the series are nonstationary can be rejected. The null hypothesis of a random walk cannot be rejected for any variable at the 5 percent level.[8] Performing a test of the hypothesis of stationarity of the first differenced series accepts that they are all stationary. As a result, cointegration tests are performed on the levels of the CPI, as is consistent with previous real exchange rate tests.

Table 1
Results of Dickey-Fuller Tests of the CPIs of Four Metropolitan Areas
Level at which the hypothesis that the series is a random walk can be rejected:

Metropolitan Area	Total	Shelter	Ex-Shelter	Reconstructed (Ex-Shelter)
New York	.998	.218	.405	.999
Philadelphia	.999	.391	.295	.295
Los Angeles	.998	.982	.298	.998
Chicago	.999	.964	.437	.999

Sample period: 1967:12–1992:3.

Table 2
Results of Cointegration Tests of the Six Possible Exchange Rate Combinations of the CPIs of Four Metropolitan Areas

Metropolitan Areas Compared	Level at which the hypothesis that the CPIs are not cointegrated can be rejected:
Chicago—Los Angeles	.04
Philadelphia—Chicago	.16
Philadelphia—New York	.04
New York—Chicago	.35
Los Angeles—New York	.36
Philadelphia—Los Angeles	.06

Sample period: 1967:12–1992:3

Table 2 presents the results of the cointegration tests comparing the total CPIs from all six possible combinations of the four regions. Only two of the six possible real exchange rates, the Chicago-Los Angeles and the Philadelphia-New York, accepted at the 5 percent level that the real exchange rate was stable. The other four appear to be random walks, violating PPP. Even within the same country, negative results are produced, similar to those found in the international literature examining the stability of the real exchange rate. In other words, no long-run relationship exists between the price levels in different regions of the same country.

[5] The BLS does not seasonally adjust CPI data at the metropolitan area level. Each raw series was adjusted using the SAS procedure for x-11. Three years of raw data were used to adjust the first year of sample data. Each succeeding year was adjusted with an additional year of raw data until 10 years of raw data were available. For the remainder of the sample a 10-year rolling base of raw data was used.

[6] The BLS changed its procedure for calculating this subcategory in 1980. The total from 1956 to 1980 was calculated using a different method for shelter. Thus, the only shelter series the BLS presents goes back to 1980. However, the shelter, ex-shelter, and total series are consistent for all four cities in any one year, though they differ from pre-1980 to post-1980. All that matters for this test is that they are the same at each point in time.

[7] The total CPI is simply a weighted average of the shelter and ex-shelter components; thus, the missing ex-shelter subindex can be backed out of the total and shelter indices if the weights used in the construction of the total for each city can be obtained. These weights were available for each city back to 1967. Thus the results cover the period from 1967:12–1992:3. Extrapolating the 1967 weights back to 1956 produces a longer, though less reliable, series.

[8] Adjusted Dickey-Fuller tests were also performed to test for stationarity. The adjusted Dickey-Fuller tests were basically consistent with the unadjusted tests in Table 1. The unadjusted are used because no obvious problems were found with the error terms in the unadjusted tests.

It is important to note that some components of the CPIs for the different regions could be stationary, or cointegrated, while others are not. For example, say an index is the summation of two components, X and Z,

$$T = X + Z, \tag{5}$$

and these two parts have different time series properties: X is stationary and Z is not. T will then be a nonstationary variable even though it has a stationary component. The nonstationarity of the one component will mask the stationarity of the other. Similarly, if two different series, T_i and T_j, have components X_i and X_j that are cointegrated and components Z_i and Z_j that are not, then T_i and T_j will not be cointegrated. Given that the total CPI is the weighted average of the traded and nontraded goods components, it is possible that the source of the violation of PPP is the failure of the nontraded goods prices to be cointegrated.

Labor and capital may be very mobile, but some factors and goods, such as land, are not. Conventional wisdom also says that some prices, for example housing expenses, do seem to differ even in the long run between regions of the country. Certainly the price of land in San Francisco has behaved very differently than has its counterpart in Philadelphia. Although the supply of useful land and housing, as well as factor mobility from high-cost to low-cost regions, would tend to equilibrate these prices, aside from compensating differentials, factor mobility does not appear to have accomplished this feat in the last decade.

Because of the nontradable nature of land, the price indices for each metropolitan area were separated into shelter and ex-shelter components. The shelter component includes housing and renters' costs, which are very sensitive to the differential in real estate prices. This subindex certainly represents a large part of the nontraded goods component of the total CPI. A cointegration test was then performed on each subindex, to see if nontraded shelter is the source of the negative results on the cointegration tests of the total CPIs.

Table 3 presents results for both the shelter and the ex-shelter components for all combinations of the four cities. In four of the six city-by-city comparisons, the expected pattern holds; the two shelter subindices are not cointegrated while the ex-shelter indices are. Only New York violates this pattern; its behavior with Chicago and Los Angeles was unstable in both subcategories. Otherwise all city combinations suggest that the nontraded goods elements contained in the CPI strongly tilt the total CPI results toward a rejection of PPP. The shelter component, representing about one-quarter of the total CPI, appears to be a major cause of the instability of the real exchange rates between these areas. Conversely, the real exchange rate between different areas of the United States for traded goods was generally stable.

Table 3
Results of Cointegration Tests of Subindexes of the CPIs of Four Metropolitan Areas

Metropolitan Areas Compared	Level at which the hypothesis that the subindexes are not cointegrated can be rejected: Ex-Shelter	Shelter
Chicago—Los Angeles	.0002	.49
Philadelphia—Chicago	.001	.91
Philadelphia—New York	.03	.31
New York—Chicago	.22	.24
Los Angeles—New York	.19	.99
Philadelphia—Los Angeles	.005	.15

Sample period: 1967:12 to 1992:3.

IV. Implications

These simple results must be interpreted with care, since the methodology used here is not powerful. Dickey-Fuller tests assume that the variable is a random walk. As Hakkio (1986) discusses, the power of these tests is not great, and the results are only indicative. It is likely that the null of a random walk will be accepted when it is false, making it harder to prove that PPP holds. A longer sample is needed with low-power tests. Unfortunately, the importance of PPP as a theoretical device is diminished if it holds only over centuries.

Even given the low power of these tests, however, the finding of unstable real exchange rates between regions in the United States is surprising, and the cause of this instability may shed light on the origins of the negative results from international data. Looking within a single country makes it more likely that cointegration of the price levels will occur,

since some of the traditional explanations for failure in the international tests are largely avoided in a regional study. The major reason for the finding of an unstable real exchange rate here appears to be the presence of nontraded goods in the CPI.

International Implications

In the international tests, the failure to factor out the nontraded goods component makes it impossible to fully test PPP. Factor mobility is even lower internationally than regionally, leaving very little reason to believe that nontraded goods prices should equilibrate, and more reason to believe nontraded goods prices will determine the results of the real exchange rate tests. If the international tests are to mean anything substantive, if it is important to determine whether or not the real exchange rate is a random walk, then better attempts must be made to control for the nontraded goods in the international indices. Within the United States, it has been shown that the distinction between traded and nontraded goods in the price indices determined the stability of the real exchange rate.

These results are also suggestive because they show that the nominal exchange rate is not responsible for either the short-run or the long-run relative instability of the real exchange rate. Although in international economics the correlation between the nominal and the real exchange rates is high, this study using regional price indices shows that nominal exchange rate volatility need not be an important corollary to real exchange rate instability. Thus, the exchange rate regime is not vital to the finding of instability in the real exchange rate, and arguments for fixed exchange rate regimes built on this evidence are tenuous.

If international tests are to mean anything substantive, then better attempts must be made to control for the nontraded goods in the international indices.

Domestic Implications

The instability in the real exchange rates between regions of the United States is also relevant to the issue of regional shocks and regional disequilibria. If all shocks were national, the real exchange rates between these regions would be stable; it is the interregional disturbances that create real exchange rate movements. The empirical findings in this article support the conclusion that the nontraded goods of a region bear the brunt of the effects of regional disturbances, whether positive or negative. If, for example, the demand for region A's product declines while the demand for region B's increases, either real wages for producers in region A fall or unemployment increases. Blanchard and Katz (1992) find that real wages tend to be rigid, with unemployment ensuing. As a result, labor moves to the high-wage, high-job region B. The decline in employment and income in region A decreases the demand for the nontraded good and, assuming an elasticity of supply for this nontraded good that is less than infinity, the price will fall. The price of the nontraded good in region B rises, again depending on its elasticity of supply. Essentially the losses from the disturbance fall on the immobile, nontraded goods. In equilibrium, unemployment disappears, and the real exchange rate of nontraded goods permanently changes.

Yet is this an equilibrium, since nontraded goods prices are lower in region A and capital and labor are mobile? Why would firms not move from region B to region A and produce at lower costs, either because the nontraded good is a factor of production or because labor has to be paid less, nominally?[9] Unless the goods produced in region B need some other fixed local endowment for their production, then the price of nontradables across regions should also equate in the long run. The long-run instability of the real exchange rate is evidence either for a nontraded, immobile factor needed in the production of the goods from each region, or for economies of scale, so that the growing areas tend to have lower costs and higher productivity. The failure of the nontradable goods prices to be cointegrated could support either increasing returns to scale or specialization due to different regional endowments.

[9] Although it is assumed that the real wage paid to the worker is fixed, regional prices decline with the decline in the price of nontraded goods. Thus, nominal wages could fall in region A, or fail to grow as fast as in region B.

If, however, real wages are flexible, then real exchange rate instability can support a conclusion of labor immobility. With immobile labor, the initial unemployment in region A eventually forces down the real wage in that region. Incomes fall and the price of nontradables declines. Without perfect mobility of labor, it is unavoidable that the price of nontradables will diverge among regions. Still, capital could move to the region with the lower wages and nontradable goods prices, so a reason why capital does not move is also required, in order to explain long-run real exchange rate movements. Many plausible explanations for interregional labor immobility can be found, but few clear reasons exist for capital immobility. One possible reason capital may not move to a low-wage region is that labor is not homogeneous between the two areas; in this case, region A labor will not be as productive as region B labor with region B capital. This heterogeneity would also explain why labor would not move. Alternatively, firms may simply set nominal wages equal to the value of labor's product for the firm. As a result, unemployment need not decrease wages in region A and capital would have no reason to move into that area. Either story predicts that the long-run real exchange rate for nontradables would be unstable.

The policy implications of this story are more interesting still. Traditional analysis would suggest that labor mobility determines whether two regions are an optimal currency area. Labor and capital mobility should force both wages and prices of nontraded goods to equilibrate between regions. Since labor is not mobile, and wages are sticky, a nominal exchange rate depreciation would cure the short-run unemployment problem. The price of tradables would rise relative to nontradables, but the region avoids the transitory unemployment. In this case, the real exchange rate data could be suggesting that the best social policy is to break the United States into different currency regions!

These results show that the nominal exchange rate is not responsible for either the short-run or the long-run relative instability of the real exchange rate.

V. Conclusion

The results of these simple tests of PPP are suggestive. Real exchange rates are not stable when measured by total consumer price indices of the regions in the United States; PPP fails within the United States. Even without the additional problems often cited as reasons for real exchange rate instability in the international comparisons, PPP fails. One major cause of this failure appears to be the inclusion of nontraded goods in these indices. Nontraded goods are less of a problem in interregional trade because no tariff or nontariff barriers to trade exist; thus, this finding must understate the problems found in international empirical literature.

Before other theories are stressed based on the instability of the real exchange rate, care should be given to eliminating the nontraded elements of whatever index is chosen for study. Furthermore, if one wishes to prove that changes in the relative prices of traded to nontraded goods in different countries are the source of the real exchange rate instability, as postulated in real business cycle theories, then a direct proof of nontraded goods prices as the cause should be undertaken.

One interesting sidelight of this research is its implication for regional performance and policy in the United States. Obviously, conclusions can only be tentative, since the power of the tests is weak and no further analysis has been performed. What causes the prices of nontraded goods in different regions to diverge in the long run? Certainly regional booms and busts, caused by shocks to the productivity of a region's endowments, almost by definition have long-run effects on the real exchange rate, but long-run divergences between diversified regions like the four metropolitan areas in this study are not so easy to explain, even by real business cycle theorists. It will be necessary to discover the causes of real exchange rate instability between regions—whether shocks to regional endowments or problems with labor and capital mobility—in order to solve the problems, such as diverse and persistent regional unemployment rates, associated with these swings in the real exchange rate.

References

Blanchard, Olivier J. and Lawrence Katz. 1991. "Regional Evolutions." Mimeo, Harvard University.

Campbell, John Y. and N. Gregory Mankiw. 1989. "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence." In *NBER Macroeconomics Annual 1989*, National Bureau of Economic Research, pp. 185–216. Cambridge, MA: The MIT Press.

Dickey, David A. and Wayne A. Fuller. 1979. "Distributions of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association*, vol. 74, pp. 427–31.

———. 1981. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica*, vol. 49, no. 4 (July), pp. 1057–72.

Dornbusch, Rudiger. 1978. "Monetary Policy under Exchange Rate Flexibility." In *Managed Exchange Rate Flexibility: The Recent Experience*, Federal Reserve Bank of Boston Conference Series no. 20.

———. 1985. "Purchasing Power Parity." National Bureau of Economic Research, Working Paper no. 1591.

Edison, Hali J. and Jan Tore Klovland. 1987. "A Quantitative Re-assessment of the Purchasing Power Parity Hypothesis: Evidence from Norway and the United Kingdom." *Journal of Applied Econometrics*, vol. 2, pp. 309–33.

Engle, Robert F. and C.W.J. Granger. 1987. "Co-integration and Error Correction: Representation, Estimation, and Testing." *Econometrica*, vol. 55, no. 2 (March), pp. 251–76.

Engle, Robert F. and Byun Sam Yoo. 1987. "Forecasting and Testing in Co-integrated Systems." *Journal of Econometrics*, vol. 35 (May), pp. 143–59.

Frenkel, Jacob A. 1981. "The Collapse of Purchasing Power Parities during the 1970s." *European Economic Review*, vol. 16, pp. 145–65.

Friedman, Milton. 1980. "Prices of Money and Goods across Frontiers: The £ and $ over a Century." *The World Economy*, vol. 2, pp. 497–511.

Giovannetti, G. 1992. "A Survey of Recent Empirical Tests of the Purchasing Power Parity Hypothesis." *Banca Nazionale del Lavoro Quarterly Review*, no. 180 (March), pp. 81–101.

Hakkio, Craig S. 1986. "Does the Exchange Rate Follow a Random Walk? A Monte Carlo Study of Four Tests for a Random Walk." *Journal of International Money and Finance*, vol. 5, pp. 221–29.

Hendry, David F. 1986. "Econometric Modelling with Co-integrated Variables: An Overview." *Oxford Bulletin of Economics and Statistics*, vol. 48, no. 3, pp. 201–12.

Meese, Richard A. and Kenneth Rogoff. 1983. "Empirical Exchange Rate Models of the Seventies: Do They Fit out of Sample?" *Journal of International Economics*, vol. 14, pp. 3–24.

Roll, Richard. 1979. "Violations of Purchasing Power Parity and Their Implications for Efficient International Commodity Markets." In Marshal Sarnat and G. P. Szego, eds., *International Finance and Trade*, vol. 1. Cambridge, MA.: Ballinger Publishing Co.

How Large Are Economic Forecast Errors?

Stephen K. McNees

Vice President and Economist, Federal Reserve Bank of Boston. *The author gratefully acknowledges the research assistance of Kim Gilbo and Delia Sawhney. He is also indebted to the forecasters who so kindly provided their forecasts.*

Virtually everyone follows some forecaster's views, analyzing each pronouncement and eagerly awaiting the next. Opinion about the reliability of economic forecasts ranges widely, however—some argue that they are literally worthless, even though most forecasters typically can point to a sequence of predictions that virtually replicates the eventual outcome. How much confidence should one place in economic forecasts?

The answer would seem straightforward: To measure a forecast's reliability, one need simply compare it with what "actually" occurs. The diversity of opinion on reliability indicates the answer is not so simple. Two problems arise immediately, one philosophical and one practical. The philosophical problem is one of induction: Forecast accuracy cannot be measured until what actually happened is known, but the main interest typically lies in the accuracy of current forecasts for which, necessarily, no actual outcome is available. Despite many attempts to make headway with this problem, some form of assumption must be made that the future will resemble the present. Neither logic nor econometrics can provide assurance that this assumption will hold. In fact, the future is almost certain to differ at least somewhat from previous experience. Nevertheless, no alternative exists to blithely assuming that the reliability of today's forecasts will resemble the reliability of previous forecasts—that some forecaster (or model) has captured the essential lasting features of past and future behavior.

The practical problem in measuring the accuracy of past forecasts is that so many different forecasts are available—and, in some cases, so many different measures of what actually happened—that millions of different errors can be calculated, and this varied experience can be summarized in many different ways. The problem, in other words, is not the paucity of measures of reliability but their multiplicity or, more precisely, their variety. The errors vary with many factors, including (1) the economic series or variable predicted, (2) the forecaster, (3) the

time period being forecast, (4) the horizon of the forecast, and (5) the choice of "actual" data to measure what really happened.

Much attention focuses on the first three factors—the economic variable, the forecaster, and the forecast period. To illustrate the importance of the fourth and fifth factors, consider the accuracy of one prominent forecaster's predictions over the last 10 years of real GNP growth in the current quarter. The top panel of Table 1 describes the accuracy of the predictions as measured against the first official estimate of real growth ("preliminary actual data," or "advance" actual data); the bottom panel, the accuracy of the predictions when measured against the final revised estimate of real growth (prior to the benchmark revision). The first column shows the accuracy of forecasts made late in the first month of each quarter, just after the preliminary estimate of the prior quarter became available; these are called "early-quarter" forecasts. The second column shows "mid-quarter" forecasts, those made in the middle month of each quarter. The final column shows the errors of the forecasts made in the last month of the quarter, or "late-quarter" forecasts. These forecasts are customarily the expectations against which the press and financial market participants judge the preliminary GNP data release.

Table 1
Accuracy of Current Quarter Forecasts of Annual Growth Rate of Real GNP, 1981:III to 1991:III
Percentage Points unless Otherwise Specified

Relative to PRELIMINARY Actual Data	Early (First month of quarter)	Mid (Second month of quarter)	Late (Third month of quarter)
RANGE	−5.2 to 4.7	−3.2 to 2.8	−2.1 to 1.9
>1	59%	41%	29%
>2	15%	17%	2%
>3	10%	5%	0%
MAE	1.4	1.0	.8
RMSE	1.8	1.4	.9
MEAN	−.2	−.1	−.1
Relative to REVISED Actual Data			
RANGE	−5.8 to 4.4	−4.0 to 3.8	−4.0 to 4.2
>1	61%	51%	68%
>2	37%	34%	24%
>3	22%	20%	15%
MAE	1.9	1.6	1.5
RMSE	2.4	2.0	1.9
MEAN	−.4	−.3	−.3

Note: Preliminary Actual Data are the first estimates released in the month immediately following each quarter's end and are equivalent to what the U.S. Department of Commerce terms "advance" actual data. Revised Actual Data are the last estimates made prior to the benchmark revision. MAE = Mean Absolute Error, RMSE = Root Mean Squared Error, MEAN = Mean Error.

The table documents two obvious points: (1) The forecasts are much more accurate predictions of the preliminary data, which are based largely on information also available to the forecaster, than they are of the final revised data, which are based on information that does not become available until much later. (2) Forecasts made later in the quarter, when the forecaster has more information, are more accurate than earlier forecasts. Note, however, that the improvement in forecast accuracy is much greater compared to the preliminary than to the revised actual data. For example, 10 percent of the forecasts of real growth made in the first month of the quarter were off the mark by more than 3 percentage points, while none of the forecasts made during the last month of the quarter missed the preliminary estimate by more than 2.1 percentage points. The elimination of the large outliers, through the incorporation of incoming high-frequency data, cuts the root mean square error (RMSE) in half between the first and third months. In contrast, relative to the revised actuals, the proportion of errors exceeding 3 percentage points falls only from 22 percent of the forecasts made in the first month of the quarter to 15 percent of the forecasts made in the last month of the quarter; the proportion of forecast errors exceeding 1 percentage point was actually somewhat larger for the forecasts made in the last month of the quarter. The RMSE falls only by about 20 percent over the quarter. Thus, while the incoming high-frequency data shed a lot of light on what the preliminary estimate of real GNP will be, they provide relatively little new information on what the final revised number will be.

Table 2 presents comparable information for forecasts of the current-quarter rate of growth of the consumer price index (CPI). Note first that little difference can be seen in the accuracy of the predictions whether compared to the preliminary or the revised data. Unlike real GNP, where additional information is collected to improve the estimates, the CPI is based on a survey conducted each month, which cannot be repeated; all revisions come solely from changing the seasonal adjustment factors. Note also that the timing of the forecast is even more

Table 2
Accuracy of Current Quarter Forecasts of Annual Growth Rate of CPI, 1980:I to 1992:I
Percentage Points unless Otherwise Specified

Relative to PRELIMINARY Actual Data	Early (First month of quarter)	Mid (Second month of quarter)	Late (Third month of quarter)
RANGE	−5.0 to 4.8	−2.7 to 2.2	−3.5 to 1.2
>1	35%	20%	4%
>2	22%	8%	2%
>3	12%	0%	2%
MAE	1.2	.7	.3
RMSE	1.8	1.0	.6
MEAN	.1	.1	−.0
Relative to REVISED Actual Data			
RANGE	−5.2 to 4.0	−2.9 to 2.8	−1.7 to 2.1
>1	39%	22%	14%
>2	22%	10%	2%
>3	10%	0%	0%
MAE	1.2	.8	.5
RMSE	1.7	1.1	.7
MEAN	.1	.1	−.0

Note: See Table 1.

important for the CPI than for real GNP; this reflects the fact that CPI data are collected and released monthly so that by the time the late-quarter forecast is made, forecasters know the actual outcome for two of the three months of the quarter.

Forecasters have often been accused of bias. However, none of these forecasts shows a systematic tendency to either overestimate or underestimate the actual outcome. The mean errors are essentially zero, whatever the forecast horizon and whichever actual data are used.

Should forecast accuracy be assessed relative to the preliminary or to the revised actuals? The answer depends entirely on the purposes of the forecast. If the objective is to understand what influences behavior at the time—for example, if one is interested in the reaction of investors in financial markets—the preliminary data are the obvious choice, as the revised data are not available until much later. However, if the objective is to measure how close the forecast comes to what actually occurred—what nonfinancial decisionmakers, modelbuilders, and policymakers presumably would want to know—it is equally clear that the revised data, based on the most complete information set, provide a better estimate of reality.

This is particularly true for comparative evaluations: If forecaster A provides the most accurate predictions of what was initially thought to have happened (preliminary data), but forecaster B provides the best forecasts of what turns out to have actually occurred, once all the facts are in, it would seem odd to call A the better forecaster of the economy, even though forecaster A clearly is a superior forecaster of the social accountants who produce GNP estimates. Fortunately, the distinction between preliminary and revised data becomes less important for forecasts of longer time spans, such as one-year-ahead forecasts, and for variables other than the National Income and Product Accounts and the monetary aggregates, such as the CPI and the unemployment rate. For example, prices in financial markets (stock prices and interest and exchange rates) are measured precisely and thus are not subject to revision.

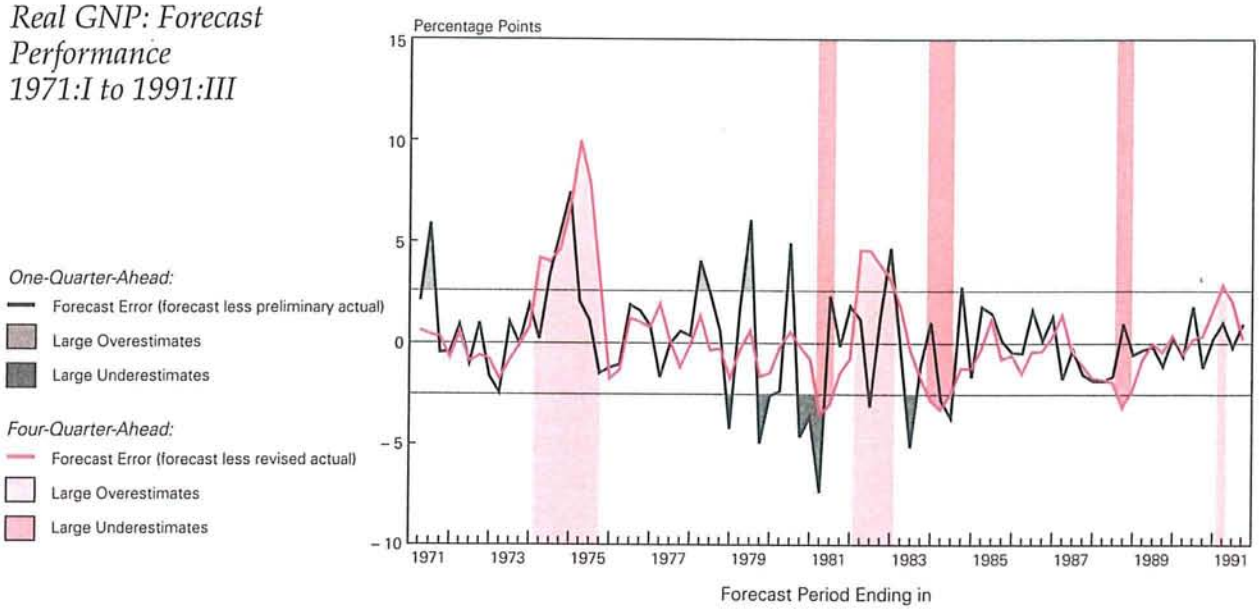
Variations in Forecast Accuracy over Time

A crucial determinant of the size of forecast errors is the forecast period; some periods are very difficult to predict while others are relatively easy. Figure 1 shows the errors of one-quarter-ahead and four-quarter-ahead forecasts, made by one prominent forecaster, of growth in real GNP from 1971:I to 1991:III. The errors for the different time spans follow different patterns: The four-quarter-ahead forecasts are dominated by the overestimates of the two major recessions, 1974–75 and 1981–82, and the underestimates of the early recoveries from the 1980 and 1981–82 recessions. The only other errors in the four-quarter-ahead forecasts that exceeded 2½ percentage points were a 3.2 percentage point underestimate of the rate of real growth in the year after the October 1987 stock market crash and a 2.9 percentage point overestimate for the 1990:I to 1991:I period, which included the 1990–91 recession.

The one-quarter-ahead forecasts are not so clearly linked to business cycle turning points, even though the largest errors were the overestimates in 1974 and the underestimates of the early recovery from the 1980 recession. In addition, large errors occurred in 1978, 1979, 1983, and 1984. But because the one-quarter-ahead errors, although large, were

Figure 1

Real GNP: Forecast Performance 1971:I to 1991:III



offsetting, the errors of forecasts covering multi-quarter time spans were not especially great.

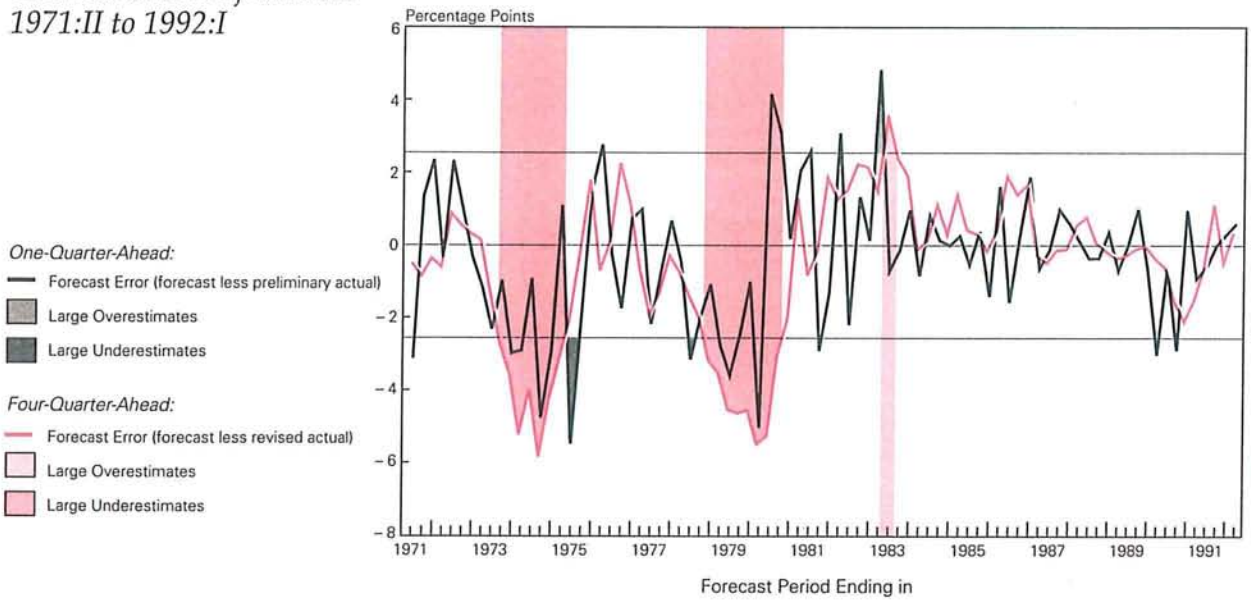
Forecasters' reputations probably reached the nadir in 1979–80, when for six quarters in a row virtually all one-quarter-ahead forecasts were in the wrong direction—when forecasts expected positive real growth, it was negative and vice versa. And in the only quarter (1980:II) when everyone's forecast was of the correct algebraic sign, the size of the decline was vastly underestimated.

Figure 2 shows corresponding information for CPI forecast errors. By far the largest errors were the sustained underestimations of the acceleration of inflation in 1973–75 and again in 1978–80. From these experiences forecasters gained the reputation of systematically underestimating inflation. These shortfalls were followed by large overestimates of the rate of inflation in 1983, which was undoubtedly associated with the underestimation of the severity of the 1981–82 recession. Since 1983, the record of forecasting the CPI has been much improved. The one-quarter-ahead forecast errors have exceeded 2 percentage points only in 1990:I and in 1990:III, when the forecasts were made just prior to the sharp increase in oil prices associated with Iraq's invasion of Kuwait. These errors resulted in the 2.1 percentage point underestimate of the inflation rate for the year 1990, the first error that large since the overestimates in 1983. The fact that CPI forecasting errors have declined in absolute terms does not necessarily indicate that forecasting ability has improved, however. The variability of the inflation rate has also been much smaller in the last 10 years. Relative to the 1970s, the 1980s have been an easy time to forecast inflation.

Large variations in forecast accuracy over time have several important implications. First, in terms of comparing different forecasters, it is critically important to compare identical forecast periods. The best forecaster's errors in the 1970s would be far larger, in absolute terms, than an inferior forecaster's errors in the 1980s. More fundamentally, the fact that accuracy varies over time poses a challenge to the constancy assumption needed to make inferences about future periods. Is it possible to know whether the current "easy" period will last or whether we will revert to the hectic 1970s? In the former case, only recent experience would be relevant for estimating the accuracy of current forecasts. But in the latter case, recent experience would be deceptive; it will be important to look at a longer sweep of history to remind us of how much uncertainty there can be.

Figure 2

CPI: Forecast Performance 1971:II to 1992:I



Has Forecast Accuracy Improved?

The figures clearly suggest forecast accuracy has improved over the past 20 years. Since the four-quarter period ending in 1984:I, no four-quarter real GNP forecast error has exceeded 3¼ percentage points and only two have exceeded 2½ percentage points. The record for inflation forecasts has been more impressive: Since the four-quarter period ending in 1983:IV, no four-quarter-ahead CPI forecast error has exceeded 2¼ percentage points and only one (1989:I to 1990:I) has exceeded 2 percentage points.

These facts undoubtedly overstate the degree of improvement that has been achieved. History shows a close association between business cycle turning points and the size of forecast errors. Much of the improvement merely reflects the fact that no turning point occurred for the 92 months between November 1982 and July 1990. Forecast errors did increase during the 1990–91 recession, when real growth was overestimated by nearly 3 percentage points and inflation underestimated by about 2 percentage points. Even errors this large, far larger than average, pale in comparison with those from earlier recessions.

In order to try to distinguish genuine improvement from a string of good luck, it is helpful to examine a longer time period. Table 3 summarizes the longest consistent forecasting record available—the forecasts of real GNP growth in the following year made each November since 1952 by the Research Seminar in Quantitative Economics (RSQE) at the University of Michigan. The distribution of errors has been fairly stable over time: About half of the errors were less than 1 percentage point, ranging only from a low of 40 percent in the 1970s to 60 percent in the 1960s and 1980s; about one-fifth of the errors exceeded 2 percentage points, ranging only from a low of 10 percent in the 1980s to a high of nearly 30 percent in the 1950s. In absolute terms, the largest errors, underestimates of the first years of expansions, occurred in the 1950s. Errors were far smaller in the relatively tranquil 1960s but rose somewhat in the turbulent 1970s; errors in the 1980s were about the same as the 1960s. The 1990s are off to a poor start: The errors for 1990 and 1991 are both larger than the average for the entire period, nearly double the average error in the 1980s.

A long-term trend toward greater accuracy is more apparent when the errors are judged relative to standards, in order to account for varying degrees of

Table 3
Accuracy of RSQE Forecasts of Real GNP, 1953 to 1991
Percentage Points unless Otherwise Specified

Years	MEAN (1)	MAE (2)	MAE/N4 (3)	RMSE (4)	RMSE/SD Actual (5)
All	−.1	1.3	.51	2.0	.70
1953–71	−.8	1.4	.62	2.2	.84
1972–91	.5	1.2	.43	1.6	.57
1950s	−1.5	2.1	.59	3.2	.90
1960s	−.7	1.0	.71	1.4	.85
1970s	.6	1.4	.39	1.9	.55
1980s	.2	.9	.44	1.3	.51

Note: MEAN = Mean Error, MAE = Mean Absolute Error, RMSE = Root Mean Squared Error, N4 = naïve "same as four-year average" forecast, SD Actual = standard deviation of actual real growth in forecast period.

Source: Forecasts: Research Seminar in Quantitative Economics, University of Michigan, *The Economic Outlook for 1992*, Table 1, p. 4.

difficulty over time. Column (3) in the table compares the MAE of the RSQE forecast with that of a naïve rule of thumb that predicts real growth each year to be equal to its average rate in the four previous years. (This rule is more accurate than the simple rule that predicts next year's growth will be the same as this year's growth.) The RSQE errors were 40 to 30 percent smaller than those of the naïve rule in the 1950s and 1960s, respectively, and improved to a level nearly 60 percent smaller in the 1970s and 1980s. Column (5) shows that the RMSE of the Michigan forecast has declined steadily relative to the standard deviation of real GNP in each forecast period. The standard deviation of real GNP is a direct measure of the difficulty of forecasting in each period. Alternatively, it can be viewed as the RMSE of a forecaster who knew in advance the average actual growth rate in the forecast period but knew nothing about the yearly deviations from that true average. The Michigan forecasts have improved steadily relative to that hypothetical straw man.

Thus, forecast accuracy seems to have improved, whether viewed from the perspective of several decades or by comparing the recent performance with the rather dismal record in the 1970s and early 1980s. Continuing improvement is not inevitable; the performance in the 1990–91 recession was distinctly worse than average. Future improvement (deterioration) depends on whether forecasting techniques improve more rapidly (slowly) than changes occur in the structure of the economy.

Variations in Accuracy among Variables

It is commonly asserted that particular economic variables are "unpredictable." Because it is easy to find someone who will gladly predict anything, such statements are intended to refer to the accuracy of predictions and not the difficulty of making some prediction, no matter how reliable. It is obvious that some variables can be predicted more accurately than others, but not at all obvious how to compare errors in forecasts of different variables. Is a $10 billion error in GNP better or worse than a 50 basis point error in interest rates? Is a 1 percentage point error for the CPI the same as a 1 percentage point error for the unemployment rate? Clearly, forecast errors for different variables cannot simply be added up. Some kind of standardization is required if a comparison of different variables is even to be attempted.

Although perfection is the goal of forecasting, we know that the future is unknown and we do not expect forecasts to eliminate all uncertainty. A forecast is useful if it can reduce uncertainty. But to measure a reduction presumes some estimate of the level of uncertainty that prevailed initially. Forecast evaluation cannot be done in absolute terms but only relative to some standard, because no unique estimate of the level of uncertainty exists, no totally obvious standard of comparison. The only sensible standard of comparison is some alternative forecasting technique. Traditionally, forecasts have been evaluated relative to simple rule-of-thumb forecasts, such as no change or same change (as in some past period). A no-change standard of comparison is a sensible, even a surprisingly stringent, standard of

A forecast is useful if it can reduce uncertainty, but to measure a reduction presumes some estimate of the initial level of uncertainty.

Table 4
Mean Absolute Errors of Forecasts Relative to Naïve Straw Man

Variable (Straw Man) Forecast	High (1)	Low (Ratio) (2)	Median (3)	>1.2 (4)	>1.1 (5)	>1 (Percent) (6)	>.9 (7)	>.8 (8)
Short-term interest rate								
Next half-year	1.48	.97	1.28	67	92	92	100	100
Next year	1.67	.95	1.20	50	67	92	100	100
Long-term interest rate								
Next half-year	1.59	1.04	1.28	83	83	100	100	100
Next year	1.57	.89	1.20	50	75	83	92	100
Unemployment rate								
Next half-year	2.26	.57	.84	3	14	28	31	55
Next year	2.71	.63	.97	14	24	31	62	76
CPI growth rate								
Next half-year	.98	.59	.72	0	0	0	7	21
Following half-year	1.02	.56	.68	0	0	7	10	14
Next year	1.11	.38	.54	0	3	3	10	14
GNP (Lag*)								
Next half-year	2.05	.63	.86	21	31	34	48	72
Following half-year	1.21	.56	.75	3	3	7	17	34
Next year	1.54	.56	.69	3	3	7	14	31
GNP (Lead**)								
Next half-year	3.25	1.00	1.30	69	86	93	100	100
Following half-year	1.49	.74	.96	17	28	38	66	76
Next year	2.09	.78	.99	28	38	48	69	97

Note: Short-term interest, long-term interest, and unemployment rates are relative to a no-change straw man. CPI and GNP growth rates are relative to a same-change straw man.

*Lag: Last observed half-year growth rate prior to forecast.

**Lead: Next half-year growth rate after forecast.

Source: Twelve individual forecasters' interest rate forecasts, 1982–91; other variables, 29 individual forecasts, 1986–91, as published in *The Wall Street Journal.*

comparison for several variables—primarily ratios of two variables, such as unemployment rates, profit rates, foreign exchange rates, and interest rates. Most economic variables, however, grow exponentially over time. For these variables, a same-change standard is a more stringent and sensible basis of comparison.

Variations in the difficulty of predicting different variables can be illustrated by examining the forecasts published twice a year in *The Wall Street Journal* from a survey conducted by Tom Herman. Interest rate forecasts for the next six months have been collected since 1982, and for the next year since 1984; forecasts of real GNP, the CPI, and the unemployment rate have been collected since 1986. Although 68 different individuals have submitted at least one forecast, more than half (36) of these have participated in fewer than 10 of the surveys, and only three have participated in all surveys. We have already seen that forecast accuracy varies over time, so that the infrequent forecasters would benefit from skipping difficult periods and suffer if they missed the easy periods. In order to try to control for these missing forecasts, each forecaster's performance is compared, not to those of the other forecasters but to a straw man—a no-change forecast for interest rates and the unemployment rate, and a same-change forecast for the CPI and the real GNP growth rate. Difficult (easy) periods presumably would also be more (less) difficult for the straw man, so that individuals' performance relative to the straw man would be affected less by missing forecasts or gaps.

The results, summarized in Table 4, show drastic differences among variables in the forecasters' ability to outperform the straw man. At one extreme, none of the forecasters could predict the long-term interest

rate a half-year into the future as well as the simple assumption that the rate would not change; 83 percent (10 of the 12 forecasters) were more than 20 percent *less* accurate than the naïve straw man. Only one forecaster, a different individual for the half-year and the full-year horizons, could predict short-term interest rates more accurately than the straw man, and neither forecaster was more than 5 percent more accurate.

At the other extreme, everyone could predict the CPI better than the simple straw man forecast, which predicted that future changes will be the same as the most recent change. Only 14 percent (four of 29 forecasters) were unable to beat the straw man by more than 20 percent in forecasting CPI growth over the next year.

The real GNP growth and unemployment rates are more difficult to estimate than the CPI but not as difficult as interest rates. Only about one-third of the forecasters were unable to outperform the no-change straw man for the unemployment rate. Nearly half of the forecasters could beat the straw man by more than 20 percent for the half-year horizon, and almost 25 percent of the forecasters were over 20 percent more accurate in the year-ahead forecast.

Real GNP forecasts are compared to two straw men. The first, GNP lag, is the simple idea that real GNP will continue to grow at the same rate as it grew in the last observed half-year. One-third of the forecasters could not improve upon this forecast of the next half-year, while all but one could improve upon this forecast of the following half-year and of the entire year after the forecast is made.

The forecasts were made during the first few days of January and July, a few weeks before the initial estimate of actual growth in the prior quarter was released. Although they did not yet know the preliminary official estimate of the previous quarter, the forecasters had a considerable amount of information on that quarter. A second straw man—GNP lead—compares the forecasts with the preliminary estimate of real GNP growth in the half-year before the forecasts, which is released a few weeks after the forecasts were made. Only a few forecasters slightly outperformed this straw man for the first half-year period, but a majority were more accurate in forecasting real growth in the subsequent half-year and in the full year after the forecast.

This contrast reinforces the earlier observation concerning the importance of forecast release dates. It also illustrates the importance of the choice of a straw man as a standard of comparison. Although the no-change and same-change standards applied here seem reasonable, other standards could alter the results. These results are not sensitive, however, to the summary error measure or the actual data employed. Similar results hold for the RMSE instead of the MAE, or for revised actual data in place of the preliminary actual data used in the table.

Variations in Forecast Accuracy among Forecasters

Much of the interest in forecast accuracy stems from the wish to know "Who is the best forecaster?" Appendix A presents the mean absolute errors of nine different forecasters for 24 different variables over the period from 1986:I through 1991:III, corresponding to the period when the National Income and Product Accounts were based to 1982, and prior to the December 1991 benchmark revision to a 1987 base year.[1]

Even a cursory examination of the information in Appendix A shows that no single forecaster dominates all outliers for all, or even most, of the variables.[2] In light of the importance of the time within the quarter when the forecast was made, consider only the early-quarter forecasts, those made in the first month. For most variables, the most accurate forecaster varies depending on the horizon of the forecast. Even for the few exceptions (gross domestic final sales, housing starts, state and local government purchases, and the unemployment rate), three different forecasters were "the best." One of the two remaining forecasters was best in predicting the GNP deflator up through seven quarters ahead. However, different forecasters have different interests; to deem one of these forecasters the best, based on a few variables, runs the risk of misleading those forecast users whose primary interest is in some other variable.

Suppose attention is confined to the concept of the inflation rate; Appendix A shows one forecaster who excels for the CPI measure while a different forecaster excels for the GNP deflator. Assume a forecast user cares only about one specific variable and one specific horizon. Appendix A can be used to determine which forecaster has been the most accurate for that particular variable and horizon, but this

[1] Additional summary error measures (RMSEs, Theil coefficients, and mean errors) are available on request from the author.

[2] Further information on the participating forecasting organizations is provided in Appendix B.

does not imply that this particular forecaster will continue to be the most accurate in the future. The reason is that the *differences* in accuracy are typically fairly small; the "best" forecaster's errors were, on average, less than 10 percent smaller than those of the second best forecaster. These differences are of doubtful economic or statistical significance.

The fact that the accuracy of the most prominent group of forecasters is similar does not imply that *all* forecasters are equally accurate. A few of the individuals whose performance was summarized in Table 4 commonly made errors that were large multiples of the simple straw man used as a standard of comparison. It is as easy to make poor forecasts as it is difficult to consistently make the best forecasts.

Conclusion

With so much variability in forecasting accuracy, it is easier to disprove any generalization than to offer a valid one. Nevertheless, it seems clear that a major factor in forecast accuracy is the time period to be forecast. Errors were enormous in the severe 1973–75 and 1981–82 recessions, much smaller in the 1980 and 1990–91 recessions, and generally quite minimal apart from business cycle turning points. Because turning points also tend to be periods when simple rule-of-thumb forecasts fare poorly, the moral for the forecast user seems to be not to ignore the forecasts but rather to think carefully about plausible outcomes far from the consensus view.

Clearly, accuracy also varies among variables. For good theoretical reasons, it is difficult to forecast a financial variable where genuinely unique knowledge presents an opportunity to profit. These reasons do not hold as forcefully for standard nonfinancial variables—real GNP, inflation, and unemployment rates—where the opportunities for profit are less apparent. Nevertheless, some nonfinancial variables are also extremely difficult to predict. A prominent example is the change in business inventories, where forecasts are often inferior to a no-change rule of thumb.

The interplay between forecast accuracy and the length or span of the forecast is also important. Forecast accuracy obviously tends to improve as the horizon of the forecast declines. But, at least for real GNP, the improvement is relatively slow over time until the forecast period actually starts, when some actual high-frequency data can be incorporated into the forecast. At the same time, longer time spans are often easier to forecast, as aberrations in the economy and/or noise in the measurement procedures "average out." The variability of four-quarter or eight-quarter cumulative changes is generally smaller than that of quarterly changes.

Finally, the importance of the forecaster, as a determinant of accuracy, is often exaggerated, perhaps by the forecasters themselves. Some forecasters have much to fear from a clear statement of the accuracy of their forecasts. But the vast majority of prominent forecasters, including those who have invited public scrutiny of their performance, have much to gain from disclosure of how accurate their forecasts have been. First, although it may be disappointing to learn that others' performances have been similar, it must be comforting to learn that others cannot document a clearly superior performance. Second, and more importantly, there has been much disillusionment with macroeconomic forecasting. Some of this is justified, but some of it may reflect forecasters' failure to educate forecast users in how much (little) confidence to place in their forecasts. In forecasting, an explanation of how much (little) the forecaster knows can be more useful to the user than a single best guess of what the future will be. Only with some understanding of how large forecast errors are likely to be does the forecaster's message become valuable.

Appendix A

Mean Absolute Errors 1986:I to 1991:III

Forecaster	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
	Forecast Horizon (Quarters)							
	Change in Business Inventories (Billions of Current Dollars)							
Early Quarter								
DRI	19.7	26.0	31.2	28.5	29.8	31.2	27.7	25.4
GSU	15.8	22.3	28.0	27.5	30.2	37.5	32.9	26.7
LHMA	17.3	21.3	28.3	27.5	29.2	32.8	29.3	28.5
RSQE	17.9	19.7	25.4	26.5	28.3	31.4	30.1	—
WEFA	21.2	20.9	26.3	30.5	34.9	34.8	31.3	31.9
Mid Quarter								
DRI	16.3	23.7	31.2	29.8	31.6	32.6	32.8	27.8
LHMA	15.9	23.2	27.8	29.3	30.9	33.1	31.9	27.5
WEFA	19.4	21.9	23.9	28.7	33.3	34.6	33.4	30.6
Late Quarter								
DRI	15.5	24.6	31.6	29.7	32.6	35.2	36.0	29.0
LHMA	13.2	21.5	24.5	28.2	31.8	33.3	32.9	26.0
	Real Change in Business Inventories (Billions of 1982 Dollars)							
Early Quarter								
DRI	16.8	19.3	22.9	22.0	21.5	23.5	17.9	17.9
GSU	13.2	17.3	20.8	20.2	22.1	27.7	23.6	18.7
LHMA	13.4	16.6	20.6	20.4	22.2	26.4	21.0	23.0
RSQE	14.1	14.3	18.9	20.4	20.4	23.8	20.0	—
WEFA	17.8	17.2	20.8	23.1	24.0	26.9	21.0	23.0
Mid Quarter								
DRI	12.2	17.8	21.4	21.7	22.0	24.0	19.6	19.3
LHMA	13.7	17.3	20.4	21.8	22.2	26.4	22.0	19.9
UCLA	13.1	22.4	28.3	29.2	27.9	24.6	—	—
WEFA	14.9	16.4	18.9	21.2	22.5	26.2	21.7	22.4
Late Quarter								
DRI	12.2	20.0	22.5	22.3	22.7	26.1	21.5	20.7
LHMA	11.9	16.2	19.6	20.6	21.5	26.1	23.5	19.9
SPF	13.3	17.0	20.7	21.2	21.7	—	—	—
	Total Civilian Employment—Household Survey (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	.6	.5	.5	.6	.6	.6	.5	.4
GSU	.4	.6	.7	.8	.8	.9	.8	.8
LHMA	1.0	.8	.9	.9	.9	1.0	.9	.9
WEFA	.9	.7	.8	.7	.7	.7	.7	.7
Mid Quarter								
KEDI	2.2	1.8	1.6	1.4	1.4	1.4	1.3	—
LHMA	.6	.7	.7	.7	.8	.9	.8	.8
UCLA	.5	.6	.6	.7	.8	.7	—	—
WEFA	.6	.6	.6	.7	.7	.7	.7	.6
Late Quarter								
DRI	.4	.5	.5	.5	.6	.6	.5	.4
LHMA	.5	.7	.7	.8	.8	.9	.9	.8

Note: — = more than two forecasts not available.

Appendix A

Mean Absolute Errors 1986:I to 1991:III, continued

Forecaster	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
	Forecast Horizon (Quarters)							
	Civilian Unemployment Rate							
Early Quarter								
DRI	.1	.2	.3	.5	.6	.7	.7	.7
GSU	.1	.3	.4	.5	.6	.7	.8	.9
LHMA	.2	.3	.4	.6	.7	.9	1.0	1.1
RSQE	.1	.3	.4	.6	.6	.7	.7	—
WEFA	.2	.3	.4	.5	.6	.7	.8	.9
Mid Quarter								
DRI	.1	.2	.3	.5	.6	.7	.7	.6
KEDI	.1	.3	.4	.5	.6	.7	.8	—
LHMA	.1	.2	.3	.5	.6	.7	.8	.9
UCLA	.1	.3	.4	.6	.7	.7	—	—
WEFA	.2	.3	.4	.5	.6	.7	.8	.9
Late Quarter								
DRI	.1	.2	.3	.4	.5	.7	.7	.7
LHMA	.1	.2	.3	.5	.6	.8	.9	.9
SPF	.1	.3	.4	.5	.6	—	—	—
	Consumer Price Index (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	.9	.9	.8	.7	.6	.5	.4	.3
GSU	1.0	.9	.8	.6	.5	.4	.2	.2
LHMA	.8	.9	.8	.7	.5	.5	.4	.4
RSQE	1.2	1.1	1.0	.8	.8	.7	.6	1.3
WEFA	1.0	1.0	.9	.8	.6	.5	.4	.3
Mid Quarter								
DRI	.4	.7	.7	.6	.5	.5	.4	.3
KEDI	1.9	1.1	.9	.7	.7	.7	.7	—
LHMA	.5	.7	.7	.6	.6	.5	.4	.4
UCLA	.7	.7	.7	.6	.5	.5	—	—
WEFA	.6	.8	.8	.8	.6	.5	.4	.3
Late Quarter								
DRI	.2	.6	.6	.6	.5	.5	.4	.3
LHMA	.3	.6	.7	.6	.5	.4	.4	.4
	Federal Government Purchases, Nominal (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	7.4	5.4	4.9	4.1	3.5	3.2	2.6	1.9
GSU	7.5	6.3	5.1	3.3	3.5	3.0	2.8	2.8
LHMA	8.3	6.5	4.5	3.8	3.5	3.2	2.8	2.3
RSQE	6.9	5.4	4.6	4.5	4.0	3.7	3.2	—
WEFA	8.1	5.6	4.6	3.5	3.2	2.6	2.2	1.9
Mid Quarter								
DRI	7.6	5.1	4.2	3.8	3.3	3.2	2.6	2.0
LHMA	7.6	6.2	4.5	3.9	3.6	3.3	2.9	2.4
WEFA	7.1	4.7	3.3	2.8	2.8	2.5	1.7	1.5
Late Quarter								
DRI	6.7	4.9	3.9	3.8	3.5	3.3	2.8	2.1
LHMA	7.2	5.7	4.2	3.8	3.6	3.2	2.9	2.5

Appendix A

Mean Absolute Errors 1986:I to 1991:III, continued

Forecaster	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
	Forecast Horizon (Quarters)							
	Federal Government Purchases, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	9.0	4.7	5.4	4.6	3.7	2.9	2.1	1.8
GSU	1.4	5.5	4.4	3.2	2.5	2.1	2.1	1.9
LHMA	10.1	5.3	4.7	3.7	3.3	2.9	2.6	1.8
WEFA	9.0	4.5	4.7	3.4	2.8	2.2	1.8	1.5
Mid Quarter								
DRI	8.6	4.8	4.7	4.3	3.5	3.0	2.2	2.0
LHMA	8.8	5.3	4.3	3.7	3.3	2.8	2.5	1.9
WEFA	8.9	4.1	3.6	3.1	2.4	1.9	1.5	1.4
Late Quarter								
DRI	8.5	4.5	4.4	4.1	3.5	2.9	2.2	1.9
LHMA	9.2	5.3	4.5	3.7	3.2	2.6	2.5	2.1
	Federal Surplus (Billions of Current Dollars)							
Early Quarter								
DRI	20.1	29.1	33.0	34.3	32.8	30.9	30.9	34.8
GSU	22.3	27.5	28.1	31.5	31.8	39.2	39.1	44.4
LHMA	28.0	29.8	31.0	37.9	39.8	42.5	43.8	41.8
WEFA	22.4	27.6	25.9	34.7	35.4	39.6	40.2	42.8
Mid Quarter								
DRI	14.8	26.2	34.2	33.2	33.0	31.0	33.2	35.0
KEDI	31.9	28.1	27.2	29.1	35.8	41.6	40.9	—
LHMA	21.7	34.8	32.8	34.9	40.7	42.5	41.7	37.9
WEFA	17.8	28.7	28.5	31.1	33.3	35.4	39.0	43.2
Late Quarter								
DRI	13.8	24.4	30.2	33.0	31.4	29.6	32.2	31.9
LHMA	21.0	30.2	30.9	35.0	40.0	40.9	42.3	39.7
	Final Sales, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	1.6	1.2	.9	.9	.9	.9	.9	.8
GSU	1.8	1.3	1.0	1.1	1.1	1.3	1.3	1.2
LHMA	1.5	1.1	1.0	1.0	1.0	1.0	1.0	1.0
RSQE	1.6	1.1	1.1	1.1	1.0	.9	.9	—
WEFA	1.3	1.1	1.0	.9	.9	1.0	.9	.8
Mid Quarter								
DRI	1.5	1.2	.9	.9	.9	.9	.8	.8
LHMA	1.5	1.2	.9	.9	.9	1.0	1.0	.9
WEFA	1.4	1.1	1.0	.9	.9	.9	.9	.8
Late Quarter								
DRI	1.5	1.2	.8	.8	.8	.8	.8	.8
LHMA	1.4	1.2	.8	.9	.9	1.0	1.0	1.0
SPF	1.1	.9	.8	.9	1.0	—	—	—

Note: — = more than two forecasts not available.

Appendix A

Mean Absolute Errors 1986:I to 1991:III, continued

Forecaster	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
	Forecast Horizon (Quarters)							
	Gross Domestic Purchases, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	2.1	1.6	1.5	1.6	1.5	1.4	1.2	1.1
GSU	1.8	1.8	1.6	1.5	1.5	1.6	1.5	1.5
LHMA	1.9	1.6	1.4	1.4	1.4	1.5	1.4	1.3
RSQE	2.0	1.9	1.7	1.6	1.5	1.3	1.1	—
WEFA	2.1	1.5	1.4	1.3	1.3	1.2	1.1	1.0
Mid Quarter								
DRI	1.7	1.6	1.4	1.5	1.5	1.5	1.3	1.2
LHMA	1.6	1.6	1.4	1.4	1.3	1.4	1.3	1.2
WEFA	1.9	1.6	1.5	1.4	1.3	1.2	1.1	1.0
Late Quarter								
DRI	1.7	1.6	1.4	1.4	1.4	1.4	1.3	1.2
LHMA	1.4	1.3	1.2	1.3	1.4	1.4	1.4	1.3
	Gross Domestic Final Sales, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	1.8	1.3	1.3	1.3	1.2	1.2	1.0	1.0
GSU	1.8	1.2	1.1	1.2	1.2	1.4	1.3	1.3
LHMA	1.6	1.1	1.1	1.1	1.1	1.2	1.1	1.1
RSQE	2.0	1.4	1.2	1.2	1.1	1.0	.9	—
WEFA	1.6	1.0	.9	.9	.9	.9	.9	.8
Mid Quarter								
DRI	1.6	1.3	1.3	1.3	1.2	1.2	1.1	1.0
LHMA	1.6	1.1	1.0	1.1	1.1	1.1	1.1	1.0
WEFA	1.6	1.1	1.0	1.0	1.0	.9	.9	.8
Late Quarter								
DRI	1.4	1.1	1.2	1.2	1.1	1.2	1.1	1.0
LHMA	1.1	1.0	.9	1.1	1.1	1.2	1.1	1.0
	Gross Domestic Final Private Sales, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	1.8	1.7	1.5	1.6	1.4	1.3	1.2	1.0
GSU	2.1	1.7	1.6	1.7	1.6	1.7	1.7	1.6
LHMA	2.1	1.6	1.5	1.6	1.5	1.5	1.5	1.4
WEFA	1.8	1.2	1.2	1.2	1.3	1.3	1.2	1.1
Mid Quarter								
DRI	1.6	1.7	1.5	1.5	1.4	1.4	1.2	1.1
LHMA	2.0	1.6	1.4	1.5	1.5	1.5	1.4	1.3
WEFA	1.7	1.3	1.3	1.3	1.3	1.3	1.1	1.1
Late Quarter								
DRI	1.5	1.5	1.4	1.5	1.4	1.4	1.2	1.1
LHMA	1.6	1.4	1.4	1.5	1.6	1.6	1.5	1.4

Appendix A

Mean Absolute Errors 1986:I to 1991:III, continued

Forecaster	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
	Forecast Horizon (Quarters)							
	Gross National Product, Nominal (Percentage Points, Annual Rates of Growth)							
Early Quarter								
BMARK	1.8	1.8	1.7	1.8	—	—	—	—
DRI	1.7	1.3	1.3	1.2	1.2	1.1	1.1	1.1
GSU	1.6	1.7	1.6	1.6	1.6	1.5	1.4	1.3
LHMA	1.8	1.5	1.5	1.5	1.5	1.5	1.4	1.4
RSQE	2.1	1.6	1.5	1.4	1.3	1.1	1.0	—
WEFA	2.1	1.5	1.5	1.4	1.3	1.2	1.1	1.0
Mid Quarter								
DRI	1.4	1.3	1.3	1.2	1.2	1.2	1.1	1.1
KEDI	2.5	2.3	2.4	2.3	2.1	1.8	1.6	—
LHMA	1.5	1.5	1.5	1.4	1.4	1.4	1.4	1.3
UCLA	1.5	1.4	1.4	1.4	1.4	1.2	—	—
WEFA	1.6	1.5	1.5	1.4	1.3	1.2	1.1	1.0
Late Quarter								
DRI	1.3	1.2	1.2	1.1	1.2	1.2	1.1	1.1
LHMA	1.2	1.3	1.3	1.3	1.4	1.5	1.4	1.4
SPF	1.5	1.3	1.4	1.4	1.3	—	—	—
	Gross National Product, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
BMARK	1.9	1.7	1.7	1.6	—	—	—	—
DRI	1.7	1.2	1.2	1.2	1.2	1.1	1.0	1.0
GSU	1.6	1.6	1.5	1.5	1.5	1.5	1.4	1.4
LHMA	1.9	1.3	1.3	1.3	1.3	1.4	1.3	1.3
RSQE	1.6	1.6	1.6	1.5	1.4	1.3	1.1	—
WEFA	2.1	1.5	1.5	1.4	1.3	1.2	1.2	1.0
Mid Quarter								
DRI	1.5	1.3	1.2	1.2	1.2	1.1	1.0	1.0
KEDI	1.8	1.7	1.6	1.6	1.6	1.4	1.2	—
LHMA	1.6	1.4	1.3	1.3	1.3	1.3	1.2	1.1
UCLA	1.6	1.6	1.6	1.5	1.4	1.1	—	—
WEFA	1.7	1.6	1.5	1.4	1.3	1.2	1.2	1.0
Late Quarter								
DRI	1.5	1.3	1.2	1.1	1.2	1.1	1.0	1.0
LHMA	1.4	1.2	1.2	1.2	1.3	1.3	1.2	1.2
SPF	1.4	1.2	1.2	1.3	1.2	—	—	—
	Housing Starts (Millions of Units)							
Early Quarter								
DRI	.1	.1	.2	.2	.2	.3	.3	.3
GSU	.1	.1	.1	.2	.2	.2	.2	.2
LHMA	.1	.1	.1	.2	.2	.2	.3	.3
Mid Quarter								
DRI	.0	.1	.1	.2	.2	.3	.3	.3
KEDI	.1	.2	.2	.3	.4	.4	.5	—
LHMA	.0	.1	.1	.2	.2	.2	.3	.3
UCLA	.1	.1	.1	.2	.2	.3	—	—
Late Quarter								
DRI	.0	.1	.1	.2	.2	.2	.3	.3
LHMA	.0	.1	.1	.2	.2	.2	.2	.3
SPF	.1	.1	.1	.2	.2	—	—	—

Note: — = more than two forecasts not available.

Appendix A

Mean Absolute Errors 1986:I to 1991:III, continued

Forecaster	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
	Forecast Horizon (Quarters)							
	Implicit GNP Price Deflator (Percentage Points, Annual Rates of Growth)							
Early Quarter								
BMARK	1.1	.9	.7	.6	—	—	—	—
DRI	1.2	.9	.7	.6	.4	.3	.3	.2
GSU	1.1	.8	.6	.5	.4	.4	.4	.5
LHMA	.8	.7	.5	.4	.4	.3	.3	.3
RSQE	1.4	1.1	1.0	.8	.7	.7	.6	—
WEFA	1.1	.8	.6	.5	.5	.4	.3	.4
Mid Quarter								
DRI	.8	.7	.6	.4	.4	.3	.2	.1
KEDI	2.1	1.4	1.4	1.1	.8	.7	.7	—
LHMA	.8	.7	.5	.4	.3	.2	.2	.3
UCLA	.9	.6	.5	.4	.3	.2	—	—
WEFA	1.1	.8	.5	.5	.5	.5	.4	.5
Late Quarter								
DRI	.8	.7	.7	.5	.4	.3	.3	.2
LHMA	.9	.7	.5	.5	.4	.3	.3	.2
SPF	1.0	.7	.5	.5	.5	—	—	—
	Investment in Residential Structures, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	7.5	6.4	6.2	6.0	5.7	5.3	5.1	4.6
GSU	5.0	6.0	5.7	6.5	5.9	5.5	5.3	5.0
LHMA	6.7	5.9	5.4	5.5	5.5	5.5	5.5	5.7
RSQE	7.3	6.7	7.0	8.1	8.1	7.9	7.9	—
WEFA	7.0	6.4	6.1	6.2	6.1	5.9	5.6	5.2
Mid Quarter								
DRI	5.9	5.5	5.7	5.4	5.6	5.5	5.5	5.0
LHMA	6.3	5.5	5.2	5.0	5.1	5.0	5.1	5.2
WEFA	5.8	5.2	4.8	5.3	5.7	5.7	5.4	5.3
Late Quarter								
DRI	5.5	5.2	5.5	5.8	6.2	5.9	5.4	4.8
LHMA	6.5	5.5	5.5	5.8	5.9	5.6	5.4	5.3
	Net Exports of Goods and Services (Billions of Current Dollars)							
Early Quarter								
DRI	10.7	14.8	23.6	26.6	28.3	29.3	25.5	25.1
GSU	11.7	15.6	20.0	23.2	23.7	26.7	27.4	27.5
LHMA	10.8	14.5	16.3	22.4	28.6	34.3	38.8	41.1
RSQE	14.0	21.1	25.6	30.0	36.5	44.3	48.6	—
WEFA	10.0	14.0	19.3	25.1	31.5	38.3	40.4	40.8
Mid Quarter								
DRI	9.8	17.8	24.3	29.6	31.3	31.4	29.7	27.5
LHMA	9.8	13.7	28.2	24.1	30.4	36.5	42.8	45.3
WEFA	10.1	16.8	21.4	28.5	35.0	40.7	43.7	45.0
Late Quarter								
DRI	11.1	17.9	22.6	28.9	31.9	31.1	29.8	27.1
LHMA	10.1	16.3	18.4	20.8	27.0	32.9	39.8	42.2

Appendix A

Mean Absolute Errors 1986:I to 1991:III, continued

Forecaster	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
	Forecast Horizon (Quarters)							
	Real Net Exports of Goods and Services (Billions of 1982 Dollars)							
Early Quarter								
DRI	12.6	17.1	21.9	24.0	27.8	28.0	30.0	29.7
GSU	13.2	17.3	16.7	15.3	17.5	19.5	20.1	19.0
LHMA	12.4	16.7	14.1	15.5	21.6	22.1	24.5	25.9
RSQE	14.2	15.7	14.3	14.9	21.4	22.0	20.3	—
WEFA	11.0	12.4	15.6	14.7	18.8	22.9	20.6	19.0
Mid Quarter								
DRI	11.5	16.5	19.1	22.4	24.6	27.6	28.4	27.4
LHMA	11.5	15.2	13.8	16.5	22.0	22.5	25.7	25.7
WEFA	11.6	14.5	16.6	19.2	22.1	27.4	27.7	24.8
Late Quarter								
DRI	12.1	16.6	18.0	22.9	25.6	28.9	30.0	28.9
LHMA	11.1	16.0	14.0	14.7	21.8	24.6	24.4	25.4
	Nonresidential Fixed Investment, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	6.1	4.8	4.1	3.7	3.6	3.3	3.0	2.4
GSU	7.5	5.8	4.6	4.3	4.3	4.2	3.6	3.7
LHMA	8.0	5.1	4.4	3.6	3.5	3.3	2.9	2.6
RSQE	7.7	5.9	4.7	4.3	4.2	3.7	3.5	—
WEFA	7.3	4.8	3.8	3.2	2.9	2.7	2.6	2.3
Mid Quarter								
DRI	5.8	4.8	3.9	3.6	3.5	3.4	3.0	2.6
LHMA	7.7	5.1	4.0	3.4	3.5	3.1	2.9	2.7
WEFA	6.3	4.7	3.6	3.2	2.9	2.6	2.2	2.0
Late Quarter								
DRI	6.2	4.2	3.5	3.5	3.4	3.4	3.1	2.6
LHMA	7.4	4.7	3.9	3.3	3.5	3.1	2.8	2.8
	Personal Consumption Expenditures, Durable Goods, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	7.4	5.2	4.6	4.4	3.7	3.7	3.5	3.3
GSU	8.8	5.4	4.5	4.3	3.9	4.4	4.1	3.7
LHMA	7.1	4.6	3.8	3.6	3.2	3.6	3.3	3.3
RSQE	8.1	5.8	4.7	4.2	3.5	3.1	2.8	—
WEFA	7.3	4.9	3.7	3.5	2.7	2.8	3.0	2.7
Mid Quarter								
DRI	7.0	4.5	3.9	4.1	3.5	3.7	3.5	3.3
LHMA	7.7	5.5	4.1	3.7	3.0	3.5	3.2	3.0
WEFA	7.1	4.9	4.0	3.5	2.8	2.9	2.9	2.6
Late Quarter								
DRI	4.3	4.1	3.9	4.1	3.6	3.6	3.4	3.2
LHMA	4.9	4.2	3.6	3.5	3.1	3.5	3.3	3.0

Note: — = more than two forecasts not available.

Appendix A

Mean Absolute Errors 1986:I to 1991:III, continued

Forecaster	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
	Forecast Horizon (Quarters)							
	Personal Consumption Expenditures, Nondurable Goods and Services, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	1.3	1.0	.9	.7	.6	.5	.5	.5
GSU	1.2	1.0	.9	.8	.8	.8	.7	.7
LHMA	1.4	1.2	1.0	.9	.9	.9	.8	.7
RSQE	1.5	1.3	.9	.7	.7	.7	.6	—
WEFA	1.1	.9	.7	.7	.7	.6	.6	.6
Mid Quarter								
DRI	1.2	1.1	1.0	.7	.6	.7	.5	.5
LHMA	1.2	1.0	.9	.9	.8	.8	.7	.7
WEFA	1.2	1.0	.8	.7	.6	.7	.6	.6
Late Quarter								
DRI	1.2	1.2	.9	.7	.6	.6	.5	.5
LHMA	1.2	1.2	.9	.9	.8	.8	.8	.7
	State and Local Government Purchases, Real (Percentage Points, Annual Rates of Growth)							
Early Quarter								
DRI	2.5	2.2	1.9	1.8	1.9	1.9	1.9	1.8
GSU	2.2	1.7	1.4	1.4	1.3	1.3	1.3	1.3
LHMA	1.8	1.5	1.2	1.0	1.0	1.0	1.0	1.0
WEFA	1.6	1.2	.9	.9	.9	.9	.8	.8
Mid Quarter								
DRI	2.0	1.8	1.7	1.6	1.7	1.6	1.7	1.6
LHMA	1.8	1.4	1.3	1.0	.9	.9	.9	.9
WEFA	1.7	1.2	1.0	.9	.9	.8	.8	.8
Late Quarter								
DRI	2.0	1.8	1.7	1.7	1.7	1.7	1.7	1.6
LHMA	1.8	1.4	1.2	1.0	.9	.9	.8	.9
	90-Day Treasury Bill Rate							
Early Quarter								
DRI	.1	.4	.8	1.0	1.1	1.2	1.3	1.5
LHMA	.2	.5	.8	1.0	1.1	1.3	1.4	1.6
RSQE	.1	.4	.8	1.1	1.3	1.4	1.5	—
WEFA	.2	.5	.7	.8	.9	.9	.9	.9
Mid Quarter								
DRI	.1	.3	.7	.9	1.0	1.1	1.3	1.5
KEDI	.3	.6	.8	.9	1.1	1.3	1.4	1.4
LHMA	.1	.4	.7	.8	1.0	1.1	1.2	1.4
WEFA	.1	.4	.7	.8	.8	.8	.8	.9
Late Quarter								
DRI	.0	.3	.6	.8	1.0	1.2	1.4	1.5
LHMA	.0	.3	.6	.8	.9	1.1	1.2	1.3

Appendix B

Summary Information on Forecasting Organizations Studied

Forecasting Organization (Abbreviated Title), Contact for Further Information	Approximate Number of Macroeconomic Variables Forecast	Typical Forecast Horizon, Quarters	Frequency of Release, per Year	Date Forecast First Issued Regularly
1) Benchmark Forecast (BMARK), George Washington University, Frederick Joutz (202) 994–4899	30	8	4	1976
2) Data Resources, Inc. (DRI), Roger Brinner (617) 863–5100	1,200	10 to 12	12	1969
3) Georgia State University (GSU), Economic Forecasting Project, Donald Ratajczak (404) 651–3282	540	8	4	1973
4) Kent Economic and Development Institute, Inc. (KEDI), Vladimir Simunek (216) 678–8215	1,700	10	12	1974
5) Laurence H. Meyer & Associates, Ltd. (LHMA), Larry Meyer (314) 721–4747	450	7 to 11	12	1983
6) Research Seminar in Quantitative Economics (RSQE), University of Michigan, Saul Hymans (313) 764–3299	200	8	8	1969
7) Survey of Professional Forecasters (SPF), Federal Reserve Bank of Philadelphia, formerly ASA/NBER, Dean Croushore (215) 574–3809	20	5	4	1968
8) University of California at Los Angeles (UCLA), School of Business, David Hensley (310) 825–1623	1,000	8 to 12	4	1968
9) Wharton Econometric Forecasting Associates, Inc. (WEFA), Kurt Karl (215) 660–6357	1,000	12	12	1963

The Capitalization and Portfolio Risk of Insurance Companies

Richard W. Kopcke

Vice President and Economist, Federal Reserve Bank of Boston. *The author thanks Betsy L. Morgan, Inge Schaefer-Schmidbauer, and David Zanger for research assistance as well as James Bugenhagen and Robert Klein of the National Association of Insurance Commissioners for data and information.*

Insurance companies, by their nature, bear risks. Some of these risks depend on insurers' ability to anticipate the magnitude and timing of the losses that are covered by their policies. Other risks resemble those borne by other financial intermediaries. Because insurers hold portfolios of assets to pay their obligations, they assume the risk that the value of their assets may not exceed that of their contractual liabilities.

Recent failures of insurance companies raise questions about the financial condition of the insurance industries.[1] Many of the specific difficulties confronting insurance companies tend to be unique to each insurer or to its lines of business. In one general respect, however, the same difficulty that confronts thrift institutions, banks, and most other intermediaries also confronts insurers. The financial strategies of financial intermediaries in the United States presumed a stability of interest rates that began to break down in the late 1960s. Not only did rising interest rates during the past two decades tend to depress the value of the assets of all intermediaries, they also fostered competition among intermediaries as all sought new opportunities for profit.

In order to cope, many financial institutions assumed new bets by "reaching" for riskier assets offering higher yields or by operating with less capital per dollar of assets. To varying degrees, many insurance companies have adopted these strategies. Life insurance companies holding one-quarter of their industry's assets have relatively low capitalization, and companies holding more than four-fifths of industry assets have substantial investments in assets that currently are considered risky. Casualty companies representing approximately one-fifth of that industry's assets have comparatively little capital by historical standards, and companies representing three-fifths of industry assets would have low capital if interest rates were to rise substantially in the near future.

Of all the remedies inspired by the recent investigations of the insurance industries, none appears to be more important than raising

more capital. With more capital, the value of insurers' assets would exceed their contractual liabilities by a greater margin. As a consequence of the increasing volatility of the relative yields on assets and the increasing competition among financial intermediaries during the past two decades, insurers need to reduce their leverage if their contracts are to be as secure as they were supposed to be prior to the late 1960s.

Some remedies propose a greater reliance on guaranty associations to protect those who hold insurance contracts from the potential insolvency of their underwriters. These associations essentially allow the customers of insolvent insurers to draw upon the resources of other participating insurers to fulfill a portion of their unsatisfied claims. Consequently, guaranty associations inherently are no stronger than the capital of participating insurers. These associations, alone, cannot compensate for insurers' lack of capital unless they commit the government to indemnifying customers of insurance companies.

Regulatory reforms could do much to control the risks borne by insurers and those holding insurance contracts, but the potential efficacy of these reforms is limited. As financial intermediaries, insurers invest in some assets whose risks and returns are difficult for outsiders to assess. Furthermore, much of the risk borne by an insurance company arises from the blends of both assets and liabilities that constitute the company's balance sheet. Successful regulation could foster an adequate diversification of assets or the proper matching of assets and liabilities; yet, after a point, assessing adequacy and propriety requires the oversight and skills of a resident shadow management.

Of all the remedies inspired by recent investigations of the insurance industries, none appears more important than raising more capital.

The first section of this article discusses the risks inherent in financial intermediation. The second section describes the roles of life and casualty insurance industries in credit markets, discussing some of the changes in their aggregate balance sheets during the last three decades. The third section, using reports submitted to the National Association of Insurance Commissioners for 1990, examines the distribution of assets, capital, and liabilities among life insurance companies and among casualty insurance companies. This article concludes that many insurers must increase their capital to cope safely with the consequences of an enduring slump in the value of commercial real estate, a substantial decline in corporate profits, or a significant rise in credit market yields.

I. Financial Intermediation and Risk

Economic development and capital formation depend on the efficient transfer of resources from those who would save to those who would invest. In the United States, more than three-quarters of the funds transferred to investors in the form of credit market instruments or loans flow through financial intermediaries. On one hand, insurance companies, depository institutions, pension funds, and other intermediaries issue financial claims with features that appeal to savers; on the other hand, these intermediaries accept financial obligations from borrowers on terms that appeal to borrowers. Without this intermediation, each financial contract must accommodate at once the frequently disparate motives of savers and investors. Intermediaries also serve savers and investors by evaluating investors' prospects, monitoring their performance, and providing both savers and investors a dependable access to funds on terms commensurate with their risks and returns.[2]

By design, intermediaries, which transform primary securities issued by investors into assets valued by savers, manage an unmatched book. To compensate for this risk, these intermediaries expect to receive a sufficiently large margin between the effective yields they offer savers and the effective yields they earn on their assets. Savers may be willing to earn a yield below that prevailing in financial markets

[1] See, for example, IDS (1990); U.S. Congress (1990); Stevenson (1990); Laing (1990); American Council of Life Insurance (1990); and Kramer (1990).

[2] See Gurley and Shaw (1955, 1956, 1960); Navin and Sears (1955); Baskin (1988); Jensen and Meckling (1976); Leland and Pyle (1977); Smith and Warner (1979); Diamond and Dybvig (1983); Diamond (1984); Fama (1985); Bernanke and Gertler (1987); Gertler (1988); and Hoshi, Kashyap, and Scharfstein (1989).

or to sacrifice liquidity in order to receive services not offered by primary securities or by mutual funds. Investors who are not recognized in public credit markets may be willing to pay greater yields or to accept terms more stringent than those prevailing in financial markets in order to cultivate a reliable source of funds. The more savers value competitive yields and the more investors can avail themselves of competitive yields, the more intermediaries' expected profit and capacity for bearing risk shrink.

The Role of Capital

The capacity of intermediaries to bear risk also depends on their leverage. With more equity capital and surplus per dollar of assets, intermediaries can honor their contracts despite deeper or longer financial setbacks. In principle, more capital could increase the odds of survival when expected profit margins are low compared to the volatility of profits. Yet, with lower profit margins, intermediaries ordinarily require greater leverage to maintain a competitive return on capital. From the viewpoint of their customers, increasing leverage under these circumstances would compromise safety and soundness.

With increasing leverage, the interests of those who own and manage insurance companies are less likely to coincide with the interests of their customers. Extraordinary losses or competitive pressures encourage insurance companies, like other intermediaries, to acquire assets promising greater yields and risks or to increase the volume of their underwriting relative to their surplus. These strategies increase both the odds that the contracts of weak insurers will not be honored in full and the odds that failing insurers will not recover. These risky strategies often are the most appealing for imperiled intermediaries, because the price of obtaining new capital can appear to be too expensive for the existing owners.

Regulation and Guaranty Associations

Because the interests of those who own and manage financial institutions do not necessarily coincide with the interests of their customers, intermediaries typically are regulated by public agencies. But this reliance on oversight by outsiders also can pose risks. Assessing the specific values of insurers' assets and liabilities or their inherent risks and returns is difficult for customers and regulators alike.[3]

Some proposals for reforming the regulation of domestic financial intermediaries advocate relaxing direct oversight in favor of more reliance on market discipline (caveat emptor). When customers and regulators cannot audit accurately the risks borne by intermediaries, direct oversight, including substantial capital requirements, achieves a higher degree of safety and soundness than alternative approaches.[4] Instead of asking outsiders to discipline intermediaries, regulations could encourage insiders to do so. Relatively high minimum capital requirements financed entirely by common stockholders or by surplus accounts encourage intermediaries to pursue financial strategies that are more sympathetic to the interests of their customers. If the ownership of an intermediary is to be transferred should it fail to meet its capital requirement (while the value of its capital is still considerable), then its owners and managers bear more of the burden of controlling its financial risks.

Because the interests of those who own and manage financial institutions do not necessarily coincide with the interests of their customers, intermediaries typically are regulated by public agencies.

Many insurance contracts are covered to some degree by guaranty associations in most states. Like deposit insurance for thrifts or banks, the strength of these associations depends on the ability of participating insurers to pay the necessary assessments. And, like deposit insurance funds, the failure of these associations may uncover an implicit put written on state or federal governments. In cases when the federal government provides disaster relief or catastrophic insurance coverage, insurers, their customers, and their guaranty associations possess an explicit put option. Sometimes this put is less obvious: insurers may be able to claim tax deductions or tax credits for assessments paid to guaranty associations.

[3] See Randall (1989). Assessing these risks also may be difficult for insiders; see Simons and Cross (1991).

[4] See Kambhu (1990), Jensen and Meckling (1976) and Galai and Masulis (1976).

Because of the ambivalent status of these associations, governments that bear the potential burden of this put option attempt to design regulations that limit the inevitable failure of insurers to isolated, manageable cases.

This put option on the government also has deeper consequences for regulation and economic policy.[5] Even if intermediaries hold well-diversified portfolios of assets, their financial condition is contingent on the stability of the prices of capital assets.

Even if intermediaries hold well-diversified portfolios of assets, their financial condition is contingent on the stability of the prices of capital assets.

For example, if economic policy does not ratify the expectations of investors who install an "excessive" number of factories or develop an "excessive" amount of real estate, then the subsequent collapse in the prices of capital assets could entail extraordinary losses among financial intermediaries. Accordingly, the success of "deposit insurance" ultimately depends on the ability of economic policy and financial regulation to avoid binges and purges, to foster a flow of investments generally consistent with the potential growth of the economy.

Neither regulation nor guaranty associations necessarily promote safety and soundness. At times, regulations limit either the assets intermediaries hold or the variety of liabilities they issue in a fashion that diminishes their efficiency, perhaps reducing their expected returns more than the potential variability of their returns.[6] At other times, intermediaries reporting substantial current returns (by undertaking a risky investment strategy not perceived as such by outsiders) may appeal strongly to customers and may not be examined closely by regulators; these institutions also may be allowed to carry less capital or surplus than their competitors.[7] To the degree customers believe that regulated intermediaries bear a seal of approval, and to the degree that intermediaries are covered by explicit guarantees or by an implicit put option onto the government, financial institutions can become less sound, unless regulators can assess accurately their financial strategies.

Charles Darwin Meets the Winner's Curse

Direct oversight by regulators may be necessary for achieving safety and soundness even if the interests of those who own or manage financial intermediaries coincide reasonably closely with the interests of their customers. Intermediaries and their customers are not exempt from winner's curses: An intermediary that bids on assets or offers products for what it thinks they are worth, "will, in the long run, be taken for a cleaning."[8]

As a consequence of chance and familiar waves in the pattern of economic development, at any time some assets will boast a recent record of consistently high returns with little apparent risk. The appeal of these assets can be considerable for intermediaries that need to restore or maintain a competitive rate of return on their portfolios. The need to offer customers competitive terms also encourages intermediaries to pay relatively high prices for these assets, whose returns appear to be great compared to their risks. For life insurers representing a substantial proportion of the industry's assets, commercial real estate, low-grade bonds, and venture capital investments appeared to be an attractive tonic for revitalizing overall returns on assets in the 1980s.

Many of the assets held by life and casualty insurers offer great returns because they are risky investments. The prices of risky assets that have enjoyed a good run ultimately are set by the most optimistic bidders, those who foresee the least risk. Intermediaries holding these assets appear to be more profitable than their competition, whereas intermediaries that do not emulate these winners may imperil their market share or their independence. This consequence may only be reinforced when customers are encouraged to discipline intermediaries: Those that offer relatively low yields and shun apparently successful strategies receive few rave reviews from analysts or financial advisers.[9] "Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally" (Keynes 1936, p. 158).

[5] See, for example, Keynes (1936) and Minsky (1985).

[6] Regulations designed to make intermediaries more secure by limiting the liabilities they may issue and the assets they may hold might instead make both the economy and intermediaries less stable; see Kopcke and Rosengren (1989).

[7] See, for example, U.S. Department of the Treasury (1991).

[8] This observation is adapted from the words of Capen, Clapp, and Campbell (1971). See also Thaler (1988).

[9] See also Koeppel (1991).

This natural selection tends to reduce diversity within the gene pools of intermediaries, making the financial system less robust. Moreover, to the degree the competitors that appear to be the fittest ultimately are victims of the winner's curse, earning inadequate returns for the risks they are bearing, the likelihood of systemic distress is greater.

II. Insurance Companies as Financial Intermediaries

Insurance companies manage approximately 16 percent of all the financial assets held by intermediaries in the United States (Table 1). The share of assets under their control is nearly as great as the share of assets held by the thrift institutions; only the share held by commercial banks is significantly higher.

Since the 1950s, casualty insurers' share of all financial assets held by intermediaries has remained constant, while the share managed by life companies has fallen by almost one-half. During the early 1950s, life companies alone managed about 21 percent of intermediaries' assets. Currently, their share is under 12 percent. About two-thirds of this decline occurred in the late 1960s and in the 1970s; since then, the share of life insurers has changed little.

The presence of insurance companies traditionally has been greatest in the bond and mortgage markets (Table 2). During the 1960s life insurers held about one-half of the outstanding corporate bonds. Although this share has fallen with the advent of mutual funds and the growth of pension plans, life companies still hold approximately one-third of corporate bonds. Over the past 30 years, life insurers consistently have held approximately 30 percent of commercial mortgages, while their shares of residential mortgages have declined because of the growth of the thrift industry. Casualty insurers hold approximately one-fifth of the outstanding municipal bonds.

Both life and casualty insurers invest more than one-half of their assets in longer-term securities bearing fixed yields (Table 3). Bonds account for almost 50 percent of life insurers' assets, and mortgage loans, four-fifths of which were commercial mortgages in 1990, account for another 20 percent. Together, real estate holdings and corporate equities, mostly the common stock of affiliates, represent less than 8 percent of life insurance assets.

Table 1
Allocation of Financial Assets among Financial Intermediaries
Percent of Total Financial Assets Held by Financial Intermediaries

Financial Intermediaries	1900	1912	1922	1929	1952–1955	1956–1960	1961–1965	1966–1970	1971–1975	1976–1980	1981–1985	1986–1990
Life Insurance Companies	10.1	13.0	12.2	14.4	21.1	20.2	18.0	16.0	13.4	12.1	11.4	11.6
Casualty Insurance Companies	2.9	3.2	4.1	6.2	4.4	4.4	4.3	3.8	3.7	4.1	4.1	4.5
Commercial Banking	64.1	65.5	64.7	52.7	47.2	40.8	37.1	37.5	39.2	37.9	34.8	30.9
Thrift Institutions	19.1	15.2	13.6	14.8	15.4	18.4	20.9	20.5	21.0	22.3	20.3	17.9
Pension Funds	—	—	.1	.4	5.6	8.4	10.8	12.4	13.5	15.3	17.1	17.6
Private	n.a.	n.a.	n.a.	n.a.	3.4	5.4	7.2	8.3	8.9	10.4	11.6	11.3
State and Local Government	n.a.	n.a.	n.a.	n.a.	2.2	3.0	3.6	4.1	4.6	4.8	5.5	6.3
Investment Trusts	—	—	.2	2.6	1.4	2.3	3.3	4.0	3.4	1.8	2.1	6.6
Mutual Funds	n.a.	n.a.	n.a.	n.a.	1.4	2.3	3.3	3.9	2.8	1.6	2.0	5.2
Finance Companies	—	—	—	2.2	3.7	4.3	4.6	4.7	4.7	4.8	4.9	5.1
Securities Brokers and Dealers	3.8	3.1	5.1	6.7	1.2	1.1	1.1	1.2	1.0	1.1	1.5	1.9
Money Market Mutual Funds	n.a.	n.a.	n.a.	n.a.	0	0	0	0	—	.7	3.8	3.8

Note: — = less than 0.05%
n.a. = not available.
Source: All data 1900 to 1929 from Goldsmith (1955) and Goldsmith (1958). All data 1952 to 1990 from the Board of Governors, Federal Reserve System, *Flow of Funds.*

Table 2
Insurance Companies' Holdings of Selected Financial Assets
Percent of Total Value Outstanding of Each Security

	1961–1965	1966–1970	1971–1975	1976–1980	1981–1985	1986–1990
Tax-Exempt Bonds						
Life Companies	4.5	2.7	1.9	2.2	1.8	1.3
Casualty Companies	11.9	11.8	14.5	21.2	19.3	18.7
Corporate Bonds						
Life Companies	50.7	42.8	34.9	34.8	33.7	31.9
Casualty Companies	1.9	3.2	3.3	4.5	3.9	4.6
Corporate Equities						
Life Companies	1.3	1.2	1.8	2.2	2.0	1.4
Casualty Companies	2.3	1.9	2.5	2.6	2.1	1.4
Commercial Mortgages						
Life Companies	30.5	31.3	28.9	30.2	29.4	27.1
Casualty Companies	.3	.3	.2	.3	.6	.8
Multifamily Mortgages						
Life Companies	19.7	26.3	21.5	16.4	11.9	8.9
Home Mortgages						
Life Companies	15.8	11.6	5.9	2.4	1.4	.7

Source: Board of Governors, Federal Reserve System, *Flow of Funds*; A.M. Best Company, *Best's Aggregates and Averages—Life/Health*, various years; and A.M. Best Company, *Best's Aggregates and Averages—Property/Casualty*, various years.

Casualty insurers invest almost 60 percent of their assets in bonds and another 10 percent in equities. Their holdings of mortgage loans and real estate are minimal. The average maturity of bonds in both life and casualty insurers' portfolios exceeds 10 years, and the average maturity of mortgages is approximately one-half that of bonds.

The Correspondence between Assets and Liabilities

The invested assets of insurance companies are financed principally by the premiums they have collected for writing their contracts and by capital or surplus, which represents the contribution of those who own the companies. Most of the assets of insurance companies are held in reserves to pay the claims of those holding their contracts.

Although life insurers anticipate paying most of their claims only after their contracts have been in force for many years, those who own these contracts often possess the option to borrow against their reserves (frequently at favorable rates of interest) or to cancel their contracts for cash. Recently, some life insurers have aggressively sold guaranteed investment contracts (GICs) in addition to their more traditional insurance and annuity products. Because GICs are comparatively short-term liabilities, which appeal to buyers mainly by offering a competitive rate of interest, insurers relying on these contracts reduce the average maturity of their liabilities.

The reserves of casualty companies are held mainly against homeowner, automobile, and commercial policies. Casualty insurers ordinarily expect to pay most of their claims within a few years of writing their contracts. Yet, when casualty companies can replace expiring contracts with new contracts and cover their claims by their flow of premium receipts, they may manage a relatively stable portfolio of assets over many years.

If yields on securities are relatively stable, insurers can comfortably regard their liabilities as being of long duration and invest them in long-term assets. Indeed, when the yields on longer-term securities exceed those on shorter-term securities, insurers can price their contracts most attractively by investing their assets in longer-term securities.

Should all yields rise significantly and remain high, however, established insurers cannot continue to offer competitive terms on existing contracts without diminishing their return on surplus. Casualty insurers, especially, may depend on the flow of premiums to pay claims, should the values of their assets fall at the same time that the magnitude of their losses unexpectedly rises. Under these circumstances, insurers could find themselves relying on comparatively short-term liabilities to finance long-term assets.

Table 3
Balance Sheets of Life Insurance Companies and Property and Casualty Companies
Percent of Total Assets

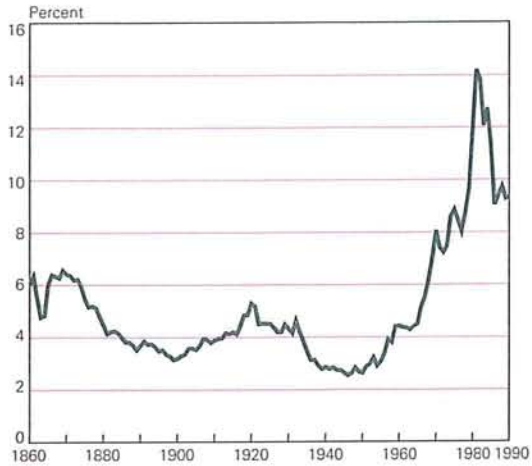
	1961–1965	1966–1970	1971–1975	1976–1980	1981–1985	1986–1990
Life Insurance Companies						
Assets						
Bonds	46.5	41.8	40.4	43.6	41.2	47.7
Government Bonds	8.8	5.9	4.7	6.4	10.1	12.8
U.S. Government	n.a.	n.a.	n.a.	n.a.	6.2	8.0
Special Revenue	n.a.	n.a.	n.a.	n.a.	2.1	3.5
Corporate Bonds	37.7	35.9	35.7	37.2	31.1	34.9
Utilities	n.a.	n.a.	n.a.	n.a.	9.0	7.6
Industrial	n.a.	n.a.	n.a.	n.a.	21.0	26.6
Corporate Stock	5.2	5.6	6.7	6.4	5.7	4.7
Preferred Stock	n.a.	n.a.	n.a.	n.a.	1.9	.8
Common Stock	n.a.	n.a.	n.a.	n.a.	3.9	3.9
Industrial	n.a.	n.a.	n.a.	n.a.	1.5	1.1
Affiliates	n.a.	n.a.	n.a.	n.a.	1.9	2.4
Mortgage Loans	36.1	37.2	32.3	27.4	22.5	19.6
Commercial Mortgages	9.6	11.7	13.9	15.6	15.2	15.4
Real Estate	3.1	3.0	3.0	2.7	2.5	2.3
Policy Loans	4.7	6.4	8.1	8.0	8.1	4.9
Separate Accounts Assets	n.a.	1.2	3.9	5.7	9.6	10.6
Other Assets	4.5	4.9	5.6	6.1	10.3	10.2
Liabilities						
Reserves	81.0	80.3	81.2	81.1	78.4	84.3
Life Insurance	58.0	56.1	53.5	45.0	32.0	24.0
Health and Disability Insurance	.8	1.3	2.0	2.6	3.3	3.0
Annuities and Supplemental Contracts	22.1	22.9	25.8	30.9	26.6	31.3
Separate Accounts	n.a.	n.a.	n.a.	n.a.	9.6	10.5
Other Liabilities	10.5	10.9	11.1	11.6	13.6	7.8
Capital and Surplus	8.5	8.7	7.6	7.1	8.0	7.9
Property and Casualty Companies						
Assets						
Bonds	49.5	50.7	52.3	63.2	56.6	58.8
U.S. Government	16.0	11.0	7.1	10.0	12.0	15.3
State and Municipal	15.4	14.2	15.3	13.3	10.4	9.2
Special Revenue	12.1	14.6	18.4	26.1	24.2	20.8
Industrial	5.3	10.2	10.8	12.8	9.4	12.6
Common Stocks	33.1	30.4	27.0	16.3	12.6	9.1
Preferred Stocks	2.6	2.7	3.5	3.4	3.6	1.9
Other Invested Assets	—	—	.1	.3	.6	.7
Mortgage Loans	.4	.4	.2	.3	.9	1.1
Real Estate	1.5	1.6	1.6	1.3	.6	.2
Other Assets	12.9	14.1	15.2	15.3	25.2	28.0
Liabilities	57.9	65.2	69.3	74.9	74.7	75.1
Losses	26.2	32.2	38.0	45.3	45.2	43.3
Loss Adjustment Expense	n.a.	n.a.	n.a.	n.a.	6.6	7.6
Unearned Premiums	25.2	25.6	23.3	20.5	17.5	16.2
Reinsurance Funds	1.2	1.5	1.2	1.2	1.4	1.4
Other Liabilities	5.4	5.9	6.9	7.9	6.5	5.2
Capital and Surplus	42.1	34.8	30.7	25.1	25.3	24.9

Note: — = less than 0.05%.
n.a. = not available.
For 1961 to 1976, data for the property and casualty companies are on a nonconsolidated basis.
Source: For life insurance companies from 1960 through 1979, American Council of Life Insurance, *Life Insurance Fact Book*, various years. For life insurance companies from 1980 to 1990, A.M. Best Company, *Best's Aggregates and Averages—Life/Health*, various years. For property and casualty insurance companies, A.M. Best Company, *Best's Aggregates and Averages—Property/Casualty*, various years.

Although the history of interest rates during the century ending with the 1960s encouraged insurance companies to invest their reserves in long-term assets, their experience during the subsequent two decades undermined their confidence in this strategy. Between 1860 and 1960, interest rates on bonds were relatively stable (Figure 1). During the past three decades, however, a doubling of yields brought many changes to the insurance industries.

Figure 1

Interest Rate on Corporate Bonds



Note: For 1860 to 1900, the interest rate is for high-grade railroad bonds. For 1901 to 1940, the interest rate is for prime corporate bonds. For 1941 to 1990, the interest rate is for Moody's Aaa corporate bonds.
Source: Homer, Sidney, *A History of Interest Rates*, 1963, various pages and *Economic Report of the President*, 1990, p. 38.

The Performance of the Life Insurance Industry

Since the 1950s, the capitalization of life insurance companies as a whole has varied little, remaining near 8 percent of the value of their assets as reported on their books. But at times during the 1970s and 1980s, the yields on their bonds and mortgages were sufficiently below yields prevailing in credit markets that their capitalization would have been below zero had their assets alone been marked to market.

Although policy lapse rates and loans to policyholders increased during this interval, the vast majority of policyholders left their funds "on deposit" with life insurers through 1985, when the returns on insurers' assets once again compared favorably with the yields prevailing in credit markets. Nonetheless, life insurance companies' share of the flow of funds into intermediaries fell significantly beginning in the late 1960s.

Established insurers coped by promoting new liabilities or new lines of business, while new companies, unburdened by investments bearing low yields, expanded their share of the life insurance, annuity, and pension businesses. Life insurers also acquired new assets promising greater or more flexible returns, often accompanied by more risk. As a result of this experience of the past two decades, life insurers increasingly are promoting their liabilities as investment contracts, and those purchasing these liabilities increasingly value them mainly as financial investments.[10] These innovations may diminish life insurers' ability to bear risk in the future.

The Performance of the Casualty Insurance Industry

During the past three decades, the capitalization of casualty insurance companies fell more than two-fifths. In the early 1960s, the capital and surplus of casualty insurers averaged more than 40 percent of assets. After earning a low rate of return on surplus in both the mid 1960s and the mid 1970s, their capital and surplus fell below one-fourth of assets.

Though the average capitalization of casualty companies as reported on their books has not changed greatly since the 1970s, at times during the 1980s their capital would not have exceeded one-sixth of assets, had their assets alone been marked to market. Customers of casualty insurers cannot cash their policies, so marking only the assets of these insurers to market understates their capital and surplus. Nevertheless, during the 1980s persistent underwriting losses substantially depressed the return on surplus for casualty insurers as a whole. Since 1980, for example, the average return on surplus for casualty insurers has been less than that of banks (10 percent versus 12.8 percent), even though the return on surplus for casualty insurers has been more volatile. This performance may be attributed partly to established insurers' pricing existing and new contracts attractively in order to maintain their flow of premium receipts.

[10] Lautzenheiser and Barks (1991) also stress that life insurers extend generous options to their customers, allowing them to call their funds away from the company, often at the expense of customers who retain their contracts.

III. Financial Characteristics of Insurance Companies in 1990

Within the life and casualty insurance industries, the financial characteristics of each company can differ considerably from those for its industry. Though the aggregate statistics for life insurers show that the industry as a whole has not assumed great risks, companies holding one-quarter of the industry's assets have relatively low capitalization, and companies holding more than four-fifths of assets have substantial investments in risky assets. Casualty companies holding one-fifth of that industry's assets have relatively little capital by historical standards. If interest rates were to rise substantially in the near future, the capitalization of casualty companies holding more than three-fifths of the industry's assets would be less than one-half of recent industry averages.

In retrospect, many insurance companies carried too little capital in the 1970s to cover adequately the risks inherent in their balance sheets. The capitalization of these insurers is now less than that of the 1970s, while their risks have not diminished. By this standard, rather than any minimum acceptable ratio of capital to assets, the capital of many life and casualty insurers appears to be too low given the risks they are bearing.

Life Insurance Companies

Table 4 describes the distribution of assets in 1990 for the 61 largest life insurance groups, representing about 80 percent of the industry's assets. Almost one-quarter of the sample's assets were held by companies with capital and surplus less than 5 percent of assets (column 1). Approximately three-quarters of the sample's assets are held by companies for which capital and surplus is no more than 6 percent of assets. Weighted by assets, the median capitalization of life insurers in this sample is 5.6 percent.

The table also subdivides the sample of life insurers according to their investments in real estate, equity, low-grade bonds, and mortgages. For example, companies with capital to asset ratios below 5 percent hold 24.9 percent of the sample's assets. The entries in the first row of columns 2, 3, and 4 (which sum to 24.9 percent) partition this share according to investments in risky assets: 22.2 percent of assets are held by companies for which capitalization is less than 5 percent *and* for which investments in real estate, equity, low-grade bonds, and mortgages are greater than three times capital and surplus. Similarly, the entries in the first row of columns 5, 6, and 7, columns 8, 9, and 10, or columns 11, 12, and 13 (each group of three columns summing to 24.9 percent) partition the share of assets held by the compa-

Table 4
Allocation of Assets among Life Insurance Companies, 1990
Percent of Total Assets

Life Insurance Companies, Grouped by Capital and Surplus as a Percentage of Assets		Total Risk Assets (percent of capital and surplus)			Risk Assets: Real Estate, Equity and Other Assets (percent of capital and surplus)			Risk Assets: Bonds Below Investment Grade (percent of capital and surplus)			Risk Assets: Mortgages (percent of capital and surplus)		
	Total (1)	<100 (2)	100–300 (3)	>300 (4)	<50 (5)	50–100 (6)	>100 (7)	<50 (8)	50–100 (9)	>100 (10)	<100 (11)	100–300 (12)	>300 (13)
(1) <5	24.9	0	2.7	22.2	7.7	5.7	11.6	7.7	15.2	2.0	2.0	1.5	21.5
(2) 5–6	49.2	1.2	.9	47.0	1.0	9.3	38.8	21.5	23.3	4.4	4.5	1.7	43.0
(3) 7–10	18.2	1.9	1.4	14.8	1.9	9.5	6.7	16.7	0	1.4	3.3	9.2	5.7
(4) >10	7.7	.5	6.4	.8	2.2	5.5	0	7.7	0	0	4.0	3.7	0
Total	100	3.7	11.5	84.8	12.8	30.0	57.1	53.7	38.5	7.8	13.8	16.1	70.1

Note: Risk assets include: real estate, common equity, bonds below investment grade, mortgages, and "other assets," which comprise mostly real estate limited partnerships and venture capital investments. The real estate, equity, low-grade bonds, mortgages, and other assets shown are assets explicitly reported in general accounts, Schedule BA, Part I, and Schedule D. Short-term assets include: cash, bonds with a maturity of less than one year and short-term investments. Separate accounts are not included in either total assets or total liabilities. Data are for the 61 largest life insurance groups, representing about 80 percent of industry assets. Numbers may not add to totals because of rounding.

Source: National Association of Insurance Commissioners (NAIC) Database of Annual Statements.

nies with low capitalization according to their investments in specific assets: 7.7 percent of assets are held by companies for which capitalization is less than 5 percent *and* for which investments in real estate and equity are less than one-half of capital and surplus, for example.

Most of the assets of insurers are held by companies for which capital and surplus is between 5 and 6 percent of assets (Table 4, row 2). Although these companies have assets invested in real estate, equities, and low-grade bonds, these investments generally are not as great as their investments in mortgages, four-fifths of which are commercial loans. Whereas together these companies hold 49.2 percent of the industry's assets, 43.0 percent of industry assets are held by these insurers for which mortgages are at least three times capital and surplus (row 2, last column); only 4.4 percent of assets are held by these insurers for which holdings of low-grade bonds exceed capital and surplus (column 10); but 38.8 percent of assets are held by these insurers for which real estate and equity exceed capital and surplus (column 7).

Real estate, equity, low-grade bonds, and mortgages currently are regarded as risky assets because the potential losses on these assets seem to be too great compared to their returns and because the assets and liabilities of insurers are not adequately diversified to cope with these losses. If the duration of insurance contracts matched that of insurers' assets and the penalties for customers' recalling funds from those contracts were sufficiently large (marking cash values to market, rear-end load charges), then the yields on insurers' liabilities would be linked more closely to the yields on their assets. In these circumstances, the value of insurers' liabilities would tend to vary with the value of their assets, and their capital would be relatively stable. Instead, many insurance contracts are of short duration, and many longer-duration contracts impose negligible penalties on customers who call their funds out of the contracts by means of loans or cancellations. Indeed, the premium that insurers "charge" for writing this call option often is negative.[11] Accordingly, insurers' capital might fall when the yields on their commercial mortgages, for example, fail to meet expectations or match the yields expected from other assets. The consequences of these potential losses are magnified by insurers' relatively high degree of leverage.

More than four-fifths of the assets of the sample of life insurers are held by companies for which real estate, equities, low-grade bonds, and mortgages are more than three times their capital and surplus (Table 4, column 4). Among these companies, risky assets are more than 6.6 times capital and surplus. Should the value of these assets fall by one-tenth, for instance, the capital of these companies would fall by two-thirds. In this event, almost four-fifths of the entire sample's assets would be held by companies for which capital would be less than 4 percent of assets, and almost one-half held by companies with capital less than 2 percent of assets.

That the value of risky assets could fall by one-tenth or more for companies that have invested more than three times their surplus in these assets is not a remote concern. Real estate and mortgages represent more than 80 percent of the risky assets held by these companies (Table 4, columns 4 and 13). About one-half of this real estate and these mortgages have been acquired since 1983. From 1983 to 1986, the Russell-NCREIF index of office property values rose about 15 percent; since then, this index has fallen and is now approximately 20 percent below its value of 1983. Consequently, the potential losses for life insurers that are greatly committed to office properties could amount to one-tenth of their risky assets. Insurers committed to retail or warehouse properties have fared better so far. Between 1983 and 1990, the value of retail properties increased 40 percent and ware-

Table 5
Allocation of Assets among Life Insurance Companies Relying on Guaranteed Investment Contracts (GICs), 1990
Percent of Total Assets

Life Insurance Companies, Grouped by Capital and Surplus as a Percentage of Assets	Total (1)	GICs Relative to Capital and Surplus <50 (2)	50–100 (3)	100–300 (4)	>300 (5)
(1) <5	24.9	9.2	5.4	1.6	8.9
(2) 5–6	49.2	19.5	1.6	5.1	23.0
(3) 7–10	18.2	5.2	0	8.7	4.2
(4) >10	7.7	4.6	.6	1.3	1.2
Total	100	38.5	7.6	16.7	37.3

Note and Source: See Table 4.

[11] Consider the rational behavior of customers who buy the shares of a "mutual fund" that invests in stocks and bonds while always declaring a net asset value of $1. See also footnote 10.

Table 6
Allocation of Assets among Life Insurance Companies That Issue Guaranteed Investment Contracts (GICs) and Hold Risk Assets, 1990
Percent of Total Assets

Life Insurance Companies, Grouped by Risk Assets as a Percentage of Capital and Surplus	Total (1)	GICs Relative to Capital & Surplus: 50–100, GICs Relative to Short-Term Assets <50 (2)	50–100, 50–200 (3)	50–100, >200 (4)	100–300, <50 (5)	100–300, 50–200 (6)	100–300, >200 (7)	>300, <50 (8)	>300, 50–200 (9)	>300, >200 (10)
(1) <100	1.2	0	0	0	0	0	0	0	0	1.2
(2) 100–300	4.2	.4	.6	0	0	1.4	.5	0	0	1.2
(3) >300	56.1	5.4	1.2	0	0	5.1	9.7	0	13.2	21.6
(4) Total	61.5	5.8	1.8	0	0	6.5	10.2	0	13.2	24.0

Note and Source: See Table 4.

house properties appreciated just over 20 percent; since 1990, their values have fallen only about 10 percent.

Life insurers also assume risk by financing their assets with short-term guaranteed investment contracts (GICs). Even if a company were to invest only in high-grade bonds, by relying on GICs for financing, it risks losing capital should interest rates rise. Should the company invest in riskier assets, those holding its GICs might not renew their contracts if the value of these assets were to be questioned. While GICs are the most visible source of short-term financing for life insurers, their permanent life and annuity contracts also grant their customers options to withdraw funds from the company should these contracts become sufficiently unattractive.

Almost four-tenths of the assets of life insurers were held by companies for which outstanding GICs were at least three times their capital in 1990 (Table 5, column 5). If these funds were invested in short-term, high-grade securities, this reliance on GICs would not be an issue. Yet, as much as one-third of the assets of the industry were held by insurers whose GICs were twice as great as their short-term assets (Table 6, row 4, columns 4, 7, 10). Of these companies, insurers representing two-tenths of the industry's assets not only issued GICs exceeding three times their capital and surplus, but also invested three times their capital in real estate, equities, low-grade bonds, and mortgages (row 3, last column).[12]

Casualty Insurance Companies

Table 7 describes the distribution of assets, according to capitalization and return on surplus, for the 60 largest casualty insurance groups, representing about 90 percent of the industry's assets in 1990. Only about 43 percent of the industry's assets were held by companies for which capital and surplus exceeded 20 percent of assets. Only one-third of these, in turn, reported a return on surplus exceeding 9 percent. Four-tenths of the industry's assets were represented by companies for which capital and surplus was less than 20 percent of assets while, at the same time, returns on surplus were less than 9 percent.

In comparison with the standards that prevailed before the late 1970s, much of the casualty insurance business is undercapitalized. Those insurers with capital and surplus amounting to less than 20 percent of assets may be vulnerable either to unexpectedly large underwriting losses or to a substantial increase in interest rates.

For example, if bond yields were to rise 3 percentage points and dividend-price ratios on equity

[12] Commercial mortgages, constituting four-fifths of total mortgages, represent most of these "risky investments." Some analysts contend that the funds raised by selling GICs were invested in commercial mortgages. Although the maturities of the GICs and these mortgages are similar, the value of commercial mortgages is questionable, because of high vacancy rates and low rents. See Shulman (1990) and Borman (1991).

Table 7
Distribution of Assets among Casualty Insurance Companies, 1990
Percent of Total Assets

Casualty Insurance Companies, Grouped by Capital and Surplus as a Percentage of Assets	Actual for 1990				Higher Interest Rate Alternative
	Return on Capital & Surplus				
	<9 (1)	9–15 (2)	>15 (3)	Total (4)	(5)
(1) 1–10	10.8	0	0	10.8	30.2
(2) 11–15	8.4	0	2.2	10.6	33.2
(3) 16–20	20.6	4.0	10.7	35.2	14.0
(4) 21–25	10.7	6.1	1.1	17.9	6.1
(5) 26–30	2.3	.4	5.8	8.5	2.4
(6) 31–35	.4	1.1	.4	1.9	11.8
(7) >35	14.1	1.0	0	15.1	2.3
Total	67.2	12.6	20.2	100.0	100.0

Note: For the calculation of the higher interest rate alternative, see Appendix 1. Data are for the 60 largest casualty insurance groups, representing about 90 percent of industry assets.
Source: National Association of Insurance Commissioners (NAIC) Database of Annual Statements.

were to rise 1 percentage point, the median ratio of capital to assets (weighted by assets) for casualty insurers could fall from 20 percent to 14 percent. Under these circumstances, about three-tenths of the industry's assets would be held by companies with capital and surplus less than 10 percent of assets (Table 7, last column), and almost two-thirds by companies with capital and surplus less than 15 percent of assets.

The capital of casualty insurers is sensitive to changes in yields because the average maturity of their bonds exceeds 10 years and the average maturity of their loss payments is approximately 2.5 years. In essence, with rising interest rates, established insurers sell their bonds at a loss to pay current claims. If these insurers retain their bonds and avoid reporting their capital loss after yields rise, then they will report a substandard rate of return on investments over the next decade. If they also price their new policies very attractively in order to increase their cash flow, they may report substandard underwriting income. Whether or not established insurers sell their bonds after interest rates rise, the consequences ultimately are the same for their ratios of capital and surplus to assets.

IV. Conclusion

Examples of the gambler's ruin extending back more than a century have prominently featured insurance companies.[13] Probability theory has long taught that the risks inherent in forecasting deaths and casualty losses eventually can undermine any insurer whose access to new capital is limited. During the past century, both theory and practice have shown that the inevitable failures among insurers can be infrequent, isolated occurrences when insurers maintain adequate capitalization and manage their underwriting prudently.

Insurers also bear risks arising from their roles as financial intermediaries. These risks, too, entail an inevitable gambler's ruin that may be especially threatening for highly levered insurance companies that have assumed substantial risks in their portfolios of assets and liabilities. The risks arising from intermediation can even be of more concern to insurers and their public regulators than the underwriting risks posed by unusually large claims. Unusually great storm damage in an East Coast city probably will not threaten insurers representing a substantial portion of the casualty industry's assets; the specific risks covered by most insurers generally are very different.[14] But, should many insurers be highly levered and their reserves in similar assets, then an event like rising interest rates or declining real estate values may imperil companies repre-

[13] See, for example, Peirce (1878).

[14] Some underwriting risks arising from unforeseen diseases such as AIDs or from unforeseen liabilities such as environmental pollution can pose more widespread problems for the insurance industries.

senting a substantial proportion of their industry's assets.

Traditionally, both life and casualty insurance companies have invested their policyholders' reserves in long-term securities. This strategy provided businesses with a substantial flow of long-term financing at attractive prices. This strategy also allowed insurers to offer their customers relatively attractive returns on their contracts, because the yields on long-term securities exceeded those of shorter-term securities.

Though this strategy was attractive, it also was risky. The increase in yields during the 1970s and 1980s left insurance companies and their policyholders holding assets offering below-market rates of return. Insurers that no longer offered their customers a competitive rate of return lost business, whereas insurers that continued to offer their policyholders competitive returns, absorbing the losses themselves, diminished both their return on capital and subsequently their capital relative to their assets. Some insurers attempted to increase their return on surplus by acquiring a riskier portfolio of assets or by writing a substantial volume of new contracts in order to invest the proceeds in new long-term securities. Any of these steps increases the odds of insurers' failing to honor their contracts fully because of unexpected underwriting losses or unexpected increases in rates of interest.

In retrospect, 20 years ago insurers carried too little capital to cover adequately their bets against rising interest rates. Today, the capitalization of most insurers is less than that of the 1970s, while the risks inherent in their assets and liabilities have not diminished.

Insurance regulators currently are designing capital requirements that depend on the specific assets held by insurers. "Risky" assets require more capital than "safer" assets. To the degree these requirements are a preliminary step toward increasing the capitalization of those companies managing risky portfolios of assets and liabilities, they will promote safety and soundness within the insurance industries. But, if these requirements are regarded as a remedy in themselves, they may not achieve their goals and they may impede the efficient operation of credit markets.

Capital requirements eventually should depend on the risks and returns inherent in an insurance company's overall balance sheet, not the classifications of specific assets. Assets, by themselves, are neither risky nor safe.[15] An apparently risky asset, when held in a properly diversified portfolio, can increase an insurer's expected rate of return while diminishing the potential variability of its returns. A supposedly safer asset, when held in an undiversified portfolio, can increase risk at the expense of expected returns. Furthermore, even an apparently safe portfolio of assets may pose substantial risks for insurers when the financial characteristics of their assets differ greatly from those of their liabilities. Consequently, recent requirements that encourage insurers to assess their contract offerings and their investment strategies under a variety of economic assumptions suggest a promising method for evaluating the adequacy of their capital.

[15] See, for example, Chirinko and Guill (1992).

Appendix: Calculations for Table 7

Using the NAIC reports for each of the 60 largest casualty groups for 1990, the change in capital and surplus equals the change in the value of the groups' bonds, plus the change in the value of common stock, less the change in the value of the groups' expected loss payments.

The change in the value of the bond portfolio when interest rates increase 3 percentage points equals

$$\Delta B = \left\{ \sum_{i=1}^{m} (C + X(1-X)^{i-1})/(1 + C + .03)^{i} + (1-X)^{m-1}/(1 + C + .03)^{m} - \sum_{i=1}^{m} (C + X(1-X)^{i-1})/(1 + C)^{i} - (1-X)^{m-1}/(1 + C)^{m} \right\} * B.$$

B is the value of bonds held by the group,
M is the average maturity of bonds (from Schedule D of the NAIC Annual Statement),
C is the average coupon payment on bonds (interest income on bonds divided by B), and
X is the rate at which bonds are prepaid (.05).

The change in the value of common stock when dividend-price ratios rise 1 percentage point equals

$$\Delta S/S = -((D/P)^{-1} - ((D/P) + .01)^{-1}) * (D/P).$$

S is the value of common stocks held by the group, and D/P is the dividend-price ratio for those stocks.

The change in the value of loss payments when interest rates increase 3 percentage points equals

$$\Delta R/R = -((1.09)^{-D} - (1.12)^{-D}) * (1.09)^{D}.$$

R is losses and loss adjustment expenses, and
D is the average maturity of loss payments (from Schedule P of the NAIC Annual Statement).

The typical profile of payments for a given year's losses is the average of the profiles of reported payments, beginning with 1980. Then, taking into account the vintages of reserves and the profiles of their remaining payments (calculated from the typical profile), D is the weighted mean of the timing of expected future payments. Because D estimates the average maturity of payments, the foregoing formula (a duration equation using an initial return of 9 percent) tends to overstate the change in the value of these liabilities. This bias, which is small because D is near 2.5, tends to reduce the estimated loss of capital.

Note: This article was prepared originally as a paper for "Financial Crises: Duration, Cure and Prevention," a conference sponsored by the Jerome Levy Economics Institute of Bard College. It adapts and updates the analysis of Kopcke and Randall (1991).

References

A.M. Best Company. 1990. *Best's Agents Guide to Life Insurance Companies 1990*, 17th Annual Edition. Oldwick, NJ: A.M. Best Company.

———. *Best's Aggregates and Averages—Life/Health*, various years.

———. *Best's Aggregates and Averages—Property/Casualty*, various years.

American Council of Life Insurance. 1990. *Report of the ACLI Task Force on Solvency Concerns*. Washington, D.C., September.

———. Various years. *Life Insurance Fact Book*. Washington, D.C.

Baskin, Jonathan Barron. 1988. "The Development of Corporate Financial Markets in Britain and the United States, 1600–1914: Overcoming Asymmetric Information." *Business History Review*, vol. 62, Summer, pp. 199–237.

Bernanke, Ben S. and Mark Gertler. 1987. "Banking and Macroeconomic Equilibrium." In William A. Barnet and Kenneth Singleton, eds., *New Approaches to Monetary Economics*, pp. 89–111. New York: Cambridge University Press.

Board of Governors of the Federal Reserve System. 1990. *Flow of Funds Accounts. Financial Assets and Liabilities, Year-End 1966–1990*. December.

Borman, Thomas H. 1991. "Hitting the Mark on Real Estate Values." *Best's Review*, April, pp. 18–20, 92–93.

Capen, E. C., R. V. Clapp, and W. M. Campbell. 1971. "Competitive Bidding in High-Risk Situations." *Journal of Petroleum Technology*, vol. 23, June, pp. 641–53.

Chirinko, Robert S. and Guill, Gene D. 1992. "Aggregate Shocks, Loan Losses, and Portfolio Concentrations: Lessons for Assessing Depository Institution Risk." In James R. Barth and R. Dan Brumbaugh, Jr., eds., *The Reform of Federal Deposit Insurance: Disciplining the Government and Protecting Taypayers*, pp. 225–45. New York: Harper Business.

Council of Economic Advisers. 1991. *Economic Report of the President*. Washington, D.C.: Government Printing Office.

Diamond, Douglas W. 1984. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies*, vol. 51, pp. 393–414.

Diamond, Douglas W. and Philip H. Dybvig. 1983. "Bank Runs, Deposit Insurance Liquidity." *Journal of Political Economy*, vol. 91, no. 3, pp. 401–19.

Fama, Eugene F. 1985. "What's Different About Banks?" *Journal of Monetary Economics*, vol. 15, January, pp. 29–39.

Galai, Dan and Ronald W. Masulis. 1976. "The Option Pricing Model and the Risk Factor of Stock." *Journal of Financial Economics*, vol. 3, March, pp. 53–81.

Gertler, Mark. 1988. "Financial Structure and Economic Activity: An Overview." National Bureau of Economic Research, working paper no. 2559, April.

Goldsmith, Raymond W. 1955. *A Study of Saving in the United States*. Princeton, NJ: Princeton University Press.

———. 1958. *Financial Intermediaries in the American Economy Since 1900*. Princeton, NJ: Princeton University Press.

Gurley, John G. and Edward S. Shaw. 1955. "Financial Aspects of Economic Development." *The American Economic Review*, vol. 45, September, pp. 515–38.

———. 1956. "Financial Intermediaries in the Saving-Investment Process." *Journal of Finance*, vol. 11, March, pp. 257–76.

———. 1960. *Money in a Theory of Finance*. Washington, D.C.: The Brookings Institution.

Homer, Sidney. 1963. *A History of Interest Rates*. New Brunswick, NJ: Rutgers University Press.

Hoshi, Takeo, Anil Kashyap, and David Scharfstein. 1989. "Bank Monitoring and Investment: Evidence from the Changing Structure of Japanese Corporate Banking Relations." Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series, no. 110, August.

IDS Financial Services. 1990. *Will the U.S. Life Insurance Industry Keep Its Promises?* Minneapolis, MN, March.

Insurance Information Institute. Various years. *Property/Casualty Insurance Fact Book*. New York.

Jensen, Michael C. and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure." *Journal of Financial Economics*, vol. 3, pp. 305–60.

Kambhu, John. 1990. "Concealment of Risk and Regulation of Bank Risk Taking." *Journal of Regulatory Economics*, vol. 2, pp. 397–414.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money*. New York, NY: Harcourt Brace and World, 1965 ed.

Koeppel, Jeffrey A. 1991. "The Insolvency Looking Glass." *Best's Review*, Life/Health Insurance Edition, vol. 92, September, pp. 23–30.

Kopcke, Richard W. and Richard Randall. 1991. "Insurance Companies as Financial Intermediaries: Risk and Return." In Kopcke and Randall, eds., *The Financial Condition and Regulation of Insurance Companies*, pp. 19–52. Federal Reserve Bank of Boston, Conference Series No. 35, June.

Kopcke, Richard W. and Eric S. Rosengren. 1989. "Regulation of Debt and Equity." In Kopcke and Rosengren, eds., *Are the Distinctions between Debt and Equity Disappearing?* pp. 173–206. Federal Reserve Bank of Boston, Conference Series No. 33, October.

Kramer, Orin. 1990. *Rating the Risks: Assessing the Solvency Threat in the Financial Services Industry*, December 31. New York, NY: Insurance Information Institute.

Laing, Jonathan R. 1990. "Flawed Policies." *Barron's*, October 1.

Lautzenheiser, Barbara and Bradly Barks. 1991. "Why Punish Persistency?" *Best's Review*, Life/Health Insurance Edition, vol. 92, September, pp. 18–19, 21, 116.

Leland, Hayne E. and David H. Pyle. 1977. "Information Asymmetries, Financial Structure, and Financial Intermediation." *Journal of Finance*, vol. 32, May, pp. 371–87.

Minsky, Hyman. 1985. *Stabilizing an Unstable Economy*. New Haven, CT: Yale University Press.

National Association of Insurance Commissioners. 1984 to 1989. NAIC Database Tapes of the annual statements, Life/Health data for all companies, 1984 to 1989. Property/Casualty data for all companies, 1984 to 1989.

———. 1985. "The Baldwin-United Bankruptcy: Its Significance for Insurance Regulation." February.

Navin, Thomas R. and Marian V. Sears. 1955. "The Rise of a Market for Industrial Securities, 1887–1902." *Business History Review*, vol. 30, pp. 105–38.

Peirce, Charles Sanders. 1878. "Illustrations of the Logic of Science. Fourth Paper—The Doctrine of Chances." *The Popular Science Monthly*, vol. 12, March, pp. 604–15, reprinted in Justus Buchler, ed., *Philosophical Writings of Peirce*. New York, NY: Dover Publications, 1955, pp. 157–73, esp. pp. 161–2.

Randall, Richard E. 1989. "Can the Market Evaluate Asset Quality Exposure in Banks?" *New England Economic Review*, July/August, pp. 3–24.

Shulman, David. 1990. "Yesterday, Today, and Tomorrow: The Coming Default Crisis in Commercial Mortgages." *Salomon Brothers Bond Market Research, Real Estate*, June 14.

Simons, Katerina and Stephen Cross. 1991. "Do Capital Markets Predict Problems in Large Commercial Banks?" *New England Economic Review*, May/June, pp. 51–56.

Smith, Clifford W. and Jerold B. Warner. 1979. "On Financial Contracting." *Journal of Financial Economics*, vol. 7, June, pp. 117–61.

Stein, Benjamin J. 1991. "Sunk by Junk: First Executive's Woes Put Policyholders at Risk." *Barron's*, April 1, pp. 8, 9, 23–25.

Stevenson, Merril, 1990. "A Question of Balance; A Survey of American Insurance." *The Economist*, October 27, pp. S5–S20.

Thaler, Richard H. 1988. "Anomalies: The Winner's Curse." *Journal of Economic Perspectives*, vol. 2, Winter, pp. 191–202.

U.S. Congress, House of Representatives, Committee on Energy and Commerce. 1990. *Failed Promises, Insurance Company Insolvencies*. Washington, D.C.: Government Printing Office, February.

U.S. Department of the Treasury. 1991. *Modernizing the Financial System*. Washington, D.C., February.

Federal Reserve Bank of Boston
P.O. Box 2076
Boston, Massachusetts 02106-2076

Address Correction Requested

Bulk Rate
U.S. Postage
PAID
Richmond, VA
Permit No. 930