Economic Review

Federal Reserve Bank of San Francisco

Fall 1989 Number 4

Table of Contents

Opinions expressed in the Economic Review do not necessarily reflect the views of the management of the Federal Reserve Bank of San Francisco, or of the Board of Governors of the Federal Reserve System.

The Federal Reserve Bank of San Francisco's Economic Review is published quarterly by the Bank's Research Department under the supervision of Jack H. Beebe, Senior Vice President and Director of Research. The publication is edited by Barbara A. Bennett. Design, production, and distribution are handled by the Public Information Department, with the assistance of Karen Rusk and William Rosenthal.

For free copies of this and other Federal Reserve publicatons, write or phone the Public Information Department, Federal Reserve Bank of San Francisco, P.O. Box 7702, San Francisco, California 94120. Phone (415) 974-2163.

Natural Resources and Regional Growth

Ronald H. Schmidt

Senior Economist, Federal Reserve Bank of San Francisco.

The author would like to thank Stephen O. Dean for his assistance. Editorial committee members were Adrian Throop, Carolyn Sherwood-Call, and Gary Zimmerman.

Evidence from the Gross State Product data suggests that natural resource-dependent states out-performed less resource-dependent states over the 1964–86 period. Closer examination, however, suggests that the gains largely were due to an increase in wealth associated with positive resource-price shocks during the period. Interestingly, unlike the Dutch disease problem of the international literature, the non-resource industries of resource-dependent states were the principal beneficiaries of favorable natural resource price shocks.

What is the relationship between a state's dependence on natural resource production and its economic performance? Can differential growth rates of state economies be traced in part to differences in their dependence on agriculture, forestry, mining, and energy?

Past studies have provided mixed evidence on the relationship between natural resources and relative economic performance. The international literature has documented cases where dependence on natural resources has had detrimental side-effects on the economic structure of those countries. Moreover, the boom-bust cycle of mining towns offers evidence of the potentially ephemeral nature of natural resource industries. In contrast, however, other cases explicitly have linked spurts of economic growth to natural resource industry stimulus.

The purpose of this article is to address the linkages between a region's natural resource dependence and its economic performance using information from the Gross State Product (GSP) data prepared by the U.S. Bureau of Economic Analysis. These data, which offer relatively consistent information on the value added by each major industry in each state, allow some cross-state comparisons over the period from 1964 to 1986.

Using descriptive statistics from the data, the relative economic growth of resource-dependent states is compared to that of less dependent states. In general, the results show that resource-dependent states, particularly energy and mining states, had economies that grew more rapidly than average over the period, although they had greater variance around their trend growth rates as well. The evidence suggests, moreover, that the relative gains were more likely a result of strong price movements that generated significant wealth effects, rather than a result of intrinsic advantages associated with natural resource production.

The data also indicate that the "Dutch disease" problem cited in the international literature may not apply to regional economies. The "Dutch disease" refers to the effects of a sharp increase in world natural resource prices on a resource-dependent economy. The resource price increase can cause the resource-dependent economy's currency to appreciate, making its other commodities less competitive on world markets. As a result, a resource price increase can work to the detriment of non-resource industries.

At the regional level, the exchange rate effect is not present because of a common currency, but the increased demand for factors from the resource sector can bid up costs of those factors in the region and make its other outputs less competitive with those of other regions. Results from the state-level data, however, indicate that contrary to the Dutch disease problem, non-resource industries were the principal *beneficiaries* of resource price shocks. The price shocks apparently stimulated non-resource production even more than resource production.

Differences in resource endowments across states are described in the first section of this article, followed in the second section by an examination of the variety of channels through which natural resource dependence might affect the level, growth, and structure of a state's economy. The third and fourth sections present empirical findings. Conclusions are then presented in the fifth section.

I. Differences in Natural Resource Dependence

In this section the degree to which natural resource production varies across states is documented using GSP data. These data, released in 1988 for the first time, provide a measure of the value added annually by each major industrial grouping for the period from 1963 to 1986, and generally provide a better measure of activity than do the income or employment data.[1]

Each state's share of national output, both for the total economy and by resource industry, are shown in Table 1 for 1986. The first column presents each state's share of total national output (the latter being the sum of state GSP across states). The second column gives the share of national natural resource output, comprising agriculture, forestry and lumber, mining, and energy, that is contributed by each state.[2] By comparing these two columns, it is clear that the states' shares of *natural resource output* differ from their contributions to *total output*. Twenty states can be categorized as relatively dependent on natural resource production, in the sense that they contributed a larger proportion of total national natural resource production than would be predicted by their share of national output.

As shown in the table, the distribution of resource production is highly skewed across states and across resources. Rhode Island contributes only 0.05 percent of total natural resource production, while Texas provides 20.78 percent. The top 10 states in each natural resource subcategory, respectively, account for 50.4 percent of U.S. agricultural production, 54.8 percent of the nation's forestry and lumber, 56.9 percent of mineral mining, and 84.7 percent of national energy extraction.[3]

Chart 1
Resource Share of Total Gross State Product

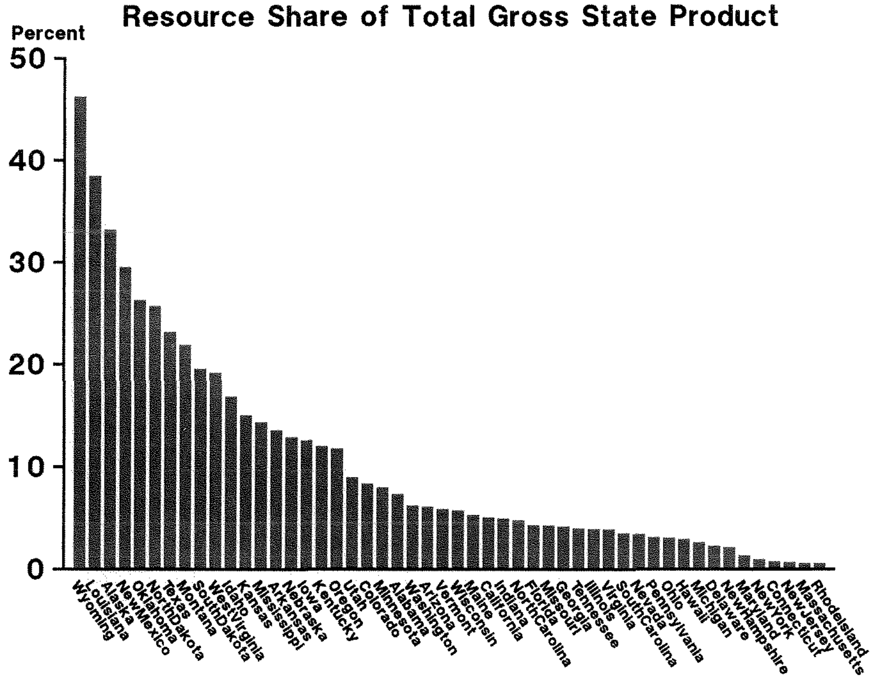


Table 1

Shares of National Natural Resource Production, By State

(Percentage of national total in 1986)

State	All* Industries	All** Resources	Agriculture	Forestry	Mining	Energy
ALABAMA	1.32	1.64	1.49	3.59	0.86	1.44
ALASKA	0.53	3.82	0.03	0.49	0.41	7.62
ARIZONA	1.26	0.81	1.06	1.07	6.17	0.18
ARKANSAS	0.77	1.40	2.30	2.53	0.35	0.56
CALIFORNIA	12.70	8.16	10.42	9.60	8.08	6.14
COLORADO	1.42	1.61	1.66	0.56	3.34	1.65
CONNECTICUT	1.69	0.22	0.37	0.47	0.37	0.05
DELAWARE	0.28	0.12	0.29	0.16	0.03	0.00
FLORIDA	4.16	2.67	4.20	3.21	7.48	1.05
GEORGIA	2.43	1.59	2.35	4.90	6.39	0.03
HAWAII	0.44	0.20	0.51	0.05	0.03	0.00
IDAHO	0.32	0.77	1.41	1.75	1.95	0.01
ILLINOIS	5.03	2.64	4.43	1.48	2.26	1.51
INDIANA	2.08	1.58	2.70	3.03	1.41	0.46
IOWA	1.11	2.39	5.99	1.00	0.90	0.01
KANSAS	1.03	1.94	3.70	0.48	0.53	0.97
KENTUCKY	1.30	2.48	2.23	1.04	1.29	3.04
LOUISIANA	1.87	7.69	0.88	1.79	2.15	14.47
MAINE	0.41	0.27	0.25	1.76	0.04	0.00
MARYLAND	1.79	0.41	0.82	0.50	0.70	0.06
MASSACHUSETTS	2.80	0.28	0.47	0.82	0.93	0.00
MICHIGAN	3.73	1.52	2.14	1.69	2.55	0.93
MINNESOTA	1.86	2.10	4.35	3.02	4.61	0.02
MISSISSIPPI	0.77	1.24	1.24	3.11	0.30	0.94
MISSOURI	1.99	1.19	2.54	1.01	2.32	0.10
MONTANA	0.29	1.03	1.24	0.98	1.63	0.83
NEBRASKA	0.65	1.63	4.18	0.21	0.28	0.05
NEVADA	0.44	0.30	0.14	0.13	6.31	0.02
NEW HAMPSHIRE	0.47	0.13	0.11	0.84	0.26	0.00
NEW JERSEY	3.65	0.40	0.65	1.19	0.85	0.03
NEW MEXICO	0.58	2.03	0.59	0.35	2.53	3.44
NEW YORK	8.46	1.18	1.96	1.95	2.50	0.33
NORTH CAROLINA	2.39	1.58	2.70	4.65	3.17	0.01
NORTH DAKOTA	0.27	1.15	2.02	0.03	0.14	0.77
OHIO	4.28	1.96	2.50	2.59	1.44	1.46
OKLAHOMA	1.26	3.85	2.21	0.36	0.53	6.02
OREGON	0.98	1.61	1.64	10.06	0.52	0.02
PENNSYLVANIA	4.40	2.26	2.48	3.81	2.37	1.78
RHODE ISLAND	0.36	0.05	0.08	0.19	0.06	0.00
SOUTH CAROLINA	1.07	0.44	0.57	1.94	1.15	0.01
SOUTH DAKOTA	0.24	0.81	1.97	0.23	1.21	0.02
TENNESSEE	1.75	0.94	1.55	1.94	1.92	0.20
TEXAS	7.55	20.78	6.35	5.36	5.44	35.96
UTAH	0.58	0.52	0.46	0.33	1.95	0.51
VERMONT	0.21	0.17	0.31	0.44	0.25	0.00
VIRGINIA	2.41	1.32	1.26	2.92	1.55	1.03
WASHINGTON	1.84	1.69	2.66	6.30	1.33	0.07
WEST VIRGINIA	0.59	1.80	0.24	0.62	0.31	3.34
WISCONSIN	1.92	1.89	4.05	3.34	0.89	0.01
WYOMING	0.30	1.72	0.25	0.13	5.96	2.87

* Percentage of total national output (summed over states) contributed by each state.
** Weighted average of agriculture, forestry, mining, and energy shares.

The information in Table 1 does not account for differences in the sizes of states' economies. Such a comparison is presented in Chart 1. Resource industry shares of total state GSP are calculated using averages from the 1964–86 period. As can be seen by comparing the chart, which has shares of *state* output, with Table 1, which reports shares of *national* output, differences in national natural resource production shares have been translated into differences in concentration in resource production. The uneven dispersion of resources across states has resulted in differences in reliance on natural resource industries. Resource dependence, as measured by the share of real GSP accounted for by the resource industries, varies widely across states and across resources. As shown in Chart 1, the average share of real GSP contributed by resource industries (agriculture, forestry and fisheries, mining, and fuel mining) over the period 1964–86 ranged from virtually zero in Rhode Island to nearly 50 percent in Wyoming.

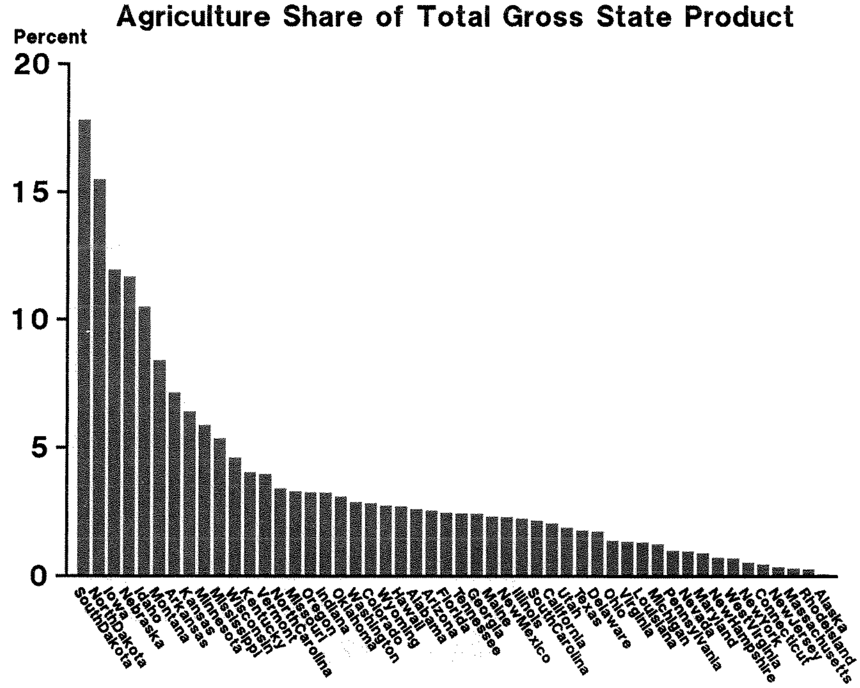
The composition of resource endowments varies significantly across states as well. As shown in charts 2 through 5, the states with the largest shares of GSP in each resource are different across resources, with little overlap between agricultural states and mining states.

The magnitude of dependence on resources varies across resources as well as across states. Among the resource-dependent states, energy stands out as the most dominant single source of resource output. The top six energy-dependent states have between 20 and 40 percent of their gross state product originating in the energy sector. Agriculture also plays a major role in the agriculture-dependent states, with five states reporting an average of 10 to 20 percent of their GSP from agricultural production.

In contrast, mining and forestry play less dominant roles even in the states with the highest concentrations of those activities. Lumber and wood products account for less than five percent of output in all states except Oregon, while mining accounts for less than four percent of total output even in the states with the highest concentrations of mining output. (Mining in this article refers to non-energy mining; coal, oil, and natural gas outputs are combined to form the energy category.)

These figures, however, may understate the importance of the natural resource industries to state economies, particularly in the short run. For example, according to the 1977 California input-output model, agriculture has a multiplier of 3.2 in the economy, suggesting that a 10 percent increase in agricultural output generates a 32 percent increase in aggregate output through the associated increase in demand for inputs, processing, marketing, transporting, and retailing (State of California, 1980.) Because such measures tend to assume that these factors could not be shifted to other uses, however, the multiplier effects tend to overstate the importance of resource industries.[4]

Chart 2
Agriculture Share of Total Gross State Product



II. Natural Resources and Production

These differences in resource endowments can be expected to result in different economic orientations, and thereby to affect the level of activity and the rate of growth of regional economies. In this section, several strands of the literature are summarized to shed some light on the effects that a natural resource orientation can have on different regional economies.

Comparative Advantage

The most obvious effect of differential endowments of natural resources is that states with large shares of a particular resource specialize in the production of that resource. This result displays the basic concept of comparative advantage laid out in the Heckser-Ohlin theorem in the international trade literature. Regions (or countries) with different factor proportions can generate higher total output and consumption by specializing in the production of commodities in which they have a relative abundance of the needed inputs. Total output is maximized when regions specialize in the production of commodities that best reflect the area's relative resource mix (its comparative advantage) and then trade with areas that have a comparative advantage in producing other goods. Consequently, regions with abundant natural resources would be expected to specialize in natural resource-intensive production, using resource outputs to trade for non-resource commodities from other regions.

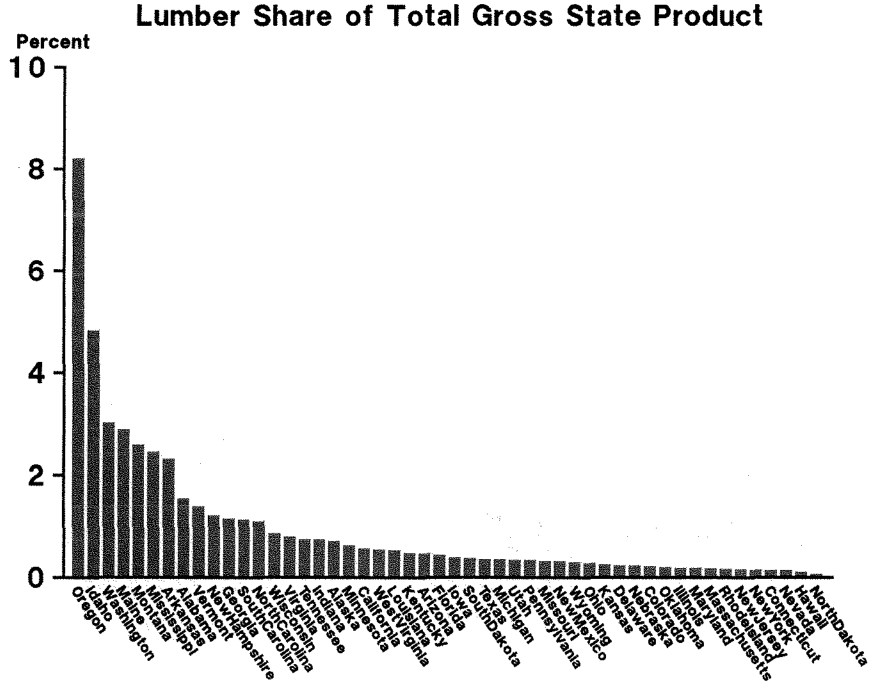
In the regional economics literature, differences in natural resource endowments are important in explaining different regional production processes. If all regions had identical resource endowments, there would be little reason to expect specialization across regions. In fact, differences in resource production and therefore, in industry mix across states, have led to different responses to technology or price shocks. A recent study by Cox and Hill (1988), for example, predicts the differential incidence of exchange-rate shocks at the state level by taking into account industry-mix characteristics and the relative trade-sensitivity of those sectors. Similar research has examined the differential regional effect of oil price shocks. These studies explicitly assume that differences in factor endowments promote different productive processes and outputs, and hence, expose regional economies to different external forces.

Levels vs. Growth Rates

While the principle of comparative advantage predicts that differences in resource endowments would affect the composition of output across regions, natural resources also may be key factors determining both the level and rate of growth of regional economies.

In the case of natural resources, the answer to the question "is more better?" seems obvious at first glance. From the standpoint of an economy, being resource-rich

Chart 3
Lumber Share of Total Gross State Product



increases productivity. Abundant natural resources lower the cost of certain inputs, enhancing the competitiveness of resource-intensive production.

Simple production-function models of an economy suggest that the level of output is a direct function of the resource base. As long as some factor substitution is possible, an abundance of a given resource boosts total output.

The effect of the resource stock on the *rate* of growth is less clear. Natural resources could increase the rate of growth of an economy if the resource stock grew over time, or if the "effective stock" (that is, the actual stock adjusted for changes in that stock's productivity) were to rise over time because of technological efficiency gains. Conversely, natural resources could cause the rate of growth to slow, if increasing scarcity of the natural resources constrained the growth of related sectors.

Consequently, theory suggests that a higher stock of a natural resource should raise the level of activity, but the effect on the rate of growth depends in part on the degree to which natural resource production expands and spurs other industries to develop. As shown in the remainder of this section, theoretical evidence on the extent to which resources spur growth is not conclusive.

Agglomeration Effects vs. Boom Towns

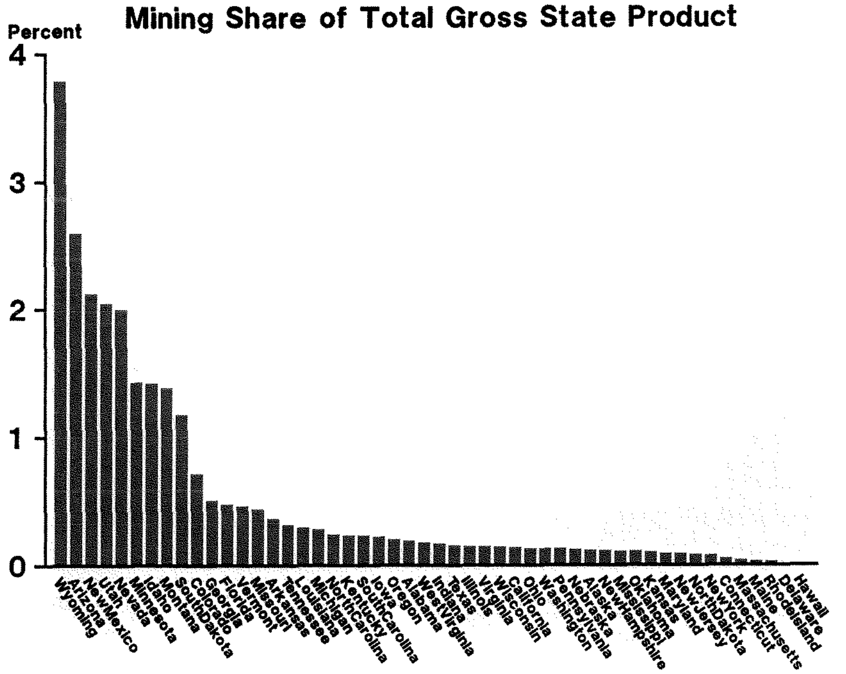
The classic model highlighting the role of natural resources as a stimulus for regional economic growth was developed by North (1955). In North's export-base model, natural resources are viewed as a driving force for economic growth, modeled along the lines of economic development in the United States. Early settlers to a region often are attracted by the economic potential of that area, which typically includes its natural resources. Trappers were attracted to the western areas by available wildlife populations. Farmers were attracted to the Midwest and West by available low-cost land. Miners were attracted to the western states by discoveries of gold and silver. Loggers were attracted to the Pacific Northwest and upper Midwest by available supplies of timber. Discoveries of massive oil reserves brought an influx of people to Texas.

In many parts of the country, therefore, economic growth was either initiated or boosted by the influx of people seeking to use natural resources. Accordingly, exports of natural resource products to other regions and countries became important in the early development of regional economies.

In the export-base literature, the creation of exportable commodities shapes a region's economy and spurs further growth. As production of the resource increases, rising wages and returns to capital encourage the migration of productive factors from other regions (Borts [1960], Borts and Stein [1964], and Schmidt [1985]). Population rises because of immigration, and export earnings and investment allow the capital stock to grow.

In the early stages, this investment is closely tied to the resource sector. Processing and transportation facilities develop, along with services to support the resource industry's employees and production.

Chart 4
Mining Share of Total Gross State Product



Over time, however, other industries not directly tied to resource production develop to take advantage of the growing economic and social infrastructures. Moreover, firms that initially support resource industries diversify into other products. For example, Texas Instruments began as a company manufacturing seismic equipment for oil drilling. As the company grew, it expanded into other electronic instruments. Today, equipment for the oil industry is only a small part of the company's sales.

Natural resources, therefore, can be a primary source of early development, followed by diversification of the economy into other fields not tied to an area's natural resources. But as evidenced by boom towns, natural resources are not always a source of lasting growth. In the case of extractive or nonrenewable natural resources, the extent to which a region diversifies into non-resource production may determine the sustainability of its economic growth. In many cases, natural resource booms have led to temporary growth, followed by decline. This idiosyncratic relationship between growth and natural resources has been described by economic historian Jonathan Hughes in the following way:

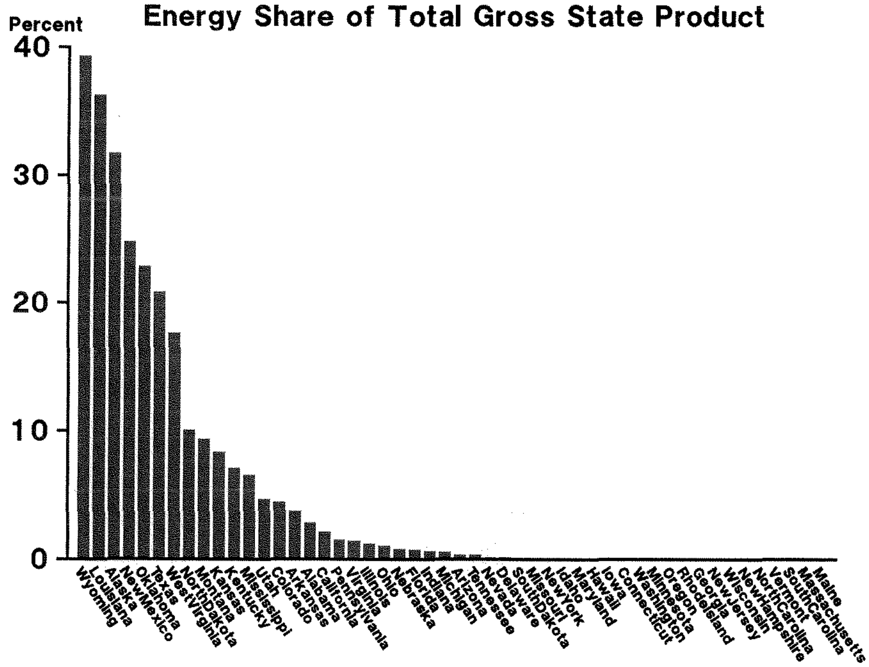
> Apart from agriculture, no doubt the best known cause of increased economic growth in the past came from the discovery and exploitation of natural resources. The ghost towns of the Rockies and the capped oil and gas wells of [the oil states] are witness to the fragile tenure of economic growth from such sources. Exploitation growth via nonreproducible natural resources usually involves only the relatively short-lived creation of fungible wealth that is carried off, leaving a hole in the ground, a stumped-over woodland, or an ocean stripped of one of its main species. In the past, there have been many examples of this purely ephemeral kind of growth. It was caricatured by historian Christopher Lasch once with these lines: 'American capitalism's idea of economic development was to leave the continent a smoking ruin.' Sometimes one must agree with Lasch on this point. Fortunately, the smoking ruin is the exception and not the rule. [Hughes 1985, p.5]

Natural resources, therefore, can be the impetus for sustained growth, but that is not always the case.

Information Effects

Dependence on natural resources also may affect regional economic growth because of the way natural resource price movements and/or discoveries of new deposits tend to change perceptions of the economic opportunities in the affected region. As shown in Plaut and Pluta (1983), Miernyck (1985), Gruben, Martens, and Schmidt (1988), and Gruben and Schmidt (1989), energy price shocks were instrumental in explaining shifts of labor and capital among regions. Those regions that were energy exporters benefited markedly from rising oil prices relative to non-energy producing states. The factor flows were not directed solely to energy industries, but rather to the broader economy. Although part of the expansion and contraction in the energy states can be directly linked to energy-supporting industries, the factor movements appeared to reflect investors' and migrants' expectations of rapid growth in other sectors as well.

Chart 5
Energy Share of Total Gross State Product



As hypothesized by Schmidt and Gruben (1988), this larger effect reflects the imperfect information available to migrants and investors regarding spatial opportunities. Because information is costly and known only with a lag, shocks that have easily recognized impacts convey an unusually large amount of information. The oil price rises in 1973–74 and 1979–80 highlighted investment opportunities in the oil patch states, while the price collapse may have encouraged potential investors to spend their limited resources acquiring information on other regions not likely to be hurt by the collapse.

Because resource price movements often are dramatic and typically are expected to have differential geographic incidence, a state's characterization as natural resource-dependent may give strong signals—whether false or true—regarding the potential opportunities to outside factors. Dramatic price movements tend to focus attention on a resource-rich region, thereby giving investment opportunities a better chance of attracting the needed factors than areas that are less well known.

"Dutch Disease"

The previous aspects of the relationship between natural resources and economic growth have stressed the ways in which natural resource production can attract factors of production from other regions. In this way, natural resources can be a source of non-resource industry growth.

The assumption in those cases is that the income or wealth effects from expanding natural resource production will spill over into non-resource production, thus diversifying the economy. In contrast, the literature in international economics points to the potential for an expanding natural resource sector to crowd out non-resource industries, known in the literature as the "Dutch disease" (Laney [1982]).

This phenomenon has been applied to cases of oil exporting countries, in particular, although the same process affects other markets as well. In this framework, a sudden price shock to a key export industry (such as the oil price increases of 1973–74 and 1979–80) boosts the nominal value of the country's exports, causing an increase in the country's trade surplus. As a result, the country's exchange rate appreciates.

The higher exchange rate then makes imports less expensive and non-petroleum exports more expensive. Consequently, domestic industries that export or compete with imports in the domestic market become less competitive. At the same time, productive factors may flow into the petroleum sector away from other sectors, thus raising factor costs in other sectors as firms bid for potentially scarce labor and capital.

If the price shock persists, the negative effects on the other sectors of the economy are transitional as the economy shifts to a new production mix that reflects the heightened comparative advantage of the natural resource industry. However, if the shock is expected to persist and, instead, proves to be temporary, the cost in terms of misallocated factors, particularly irreversible investment in plant and equipment, can offset the potential gains to the economy of the positive price shock. Moreover, such temporary losses of competitiveness can result in the loss of market share in non-resource industries that may be difficult to recapture. Finally, Dutch disease also can expose a country to greater risk by making its export portfolio less diversified.

Similar forces are at work in a regional context, although there are important differences as well. Resource price shocks can lead to increased production of those resources, which can cause factors to be reallocated from other industries. Moreover, factor prices can be bid upward, increasing the cost of other sectors' output relative to other areas of the country. Thus, even though there is a common currency, a process akin to exchange rate appreciation can occur through a rising relative cost of living.

The most important difference between regions and nations, however, is the greater mobility of factors across regional boundaries. Constraints are not placed on interstate movements of labor or capital, nor are constraints placed on shipments of products. Therefore, whereas the Dutch disease can lead to sharply higher factor prices in the resource-dependent country, factor prices need not rise as sharply in a region where factors can be drawn from other regions relatively easily.

Special Characteristics of Natural Resources

The preceding discussion, while couched in terms of resource industries, is not unique to natural resources. Many of the same relationships between resource industries and regional economies can be found in states that are highly dependent on non-resource-based industries, such as autos and steel.

There are, however, several areas in which natural resource industries differ from other industries. Natural resources lend at least three primary advantages to a developing economy over and above simple comparative advantage: a developed market, low initial technological requirements, and access to capital. First, natural resources often enjoy developed markets, both domestic and international. To an area that is sparsely developed, a natural resource can provide a readily exportable commodity. Unlike many finished commodities, a resource does not require strong local demand during start-up phases of the operation. As long as the region has access to transportation, trade is possible.

A second, related advantage that accrues to natural resource production concerns the relatively low level of skill and capital that is required at initial stages of production. Typically, the existence of abundant resources and access to transportation make the cost of extracting the first units relatively low.[5] Mineral mining can be profitable with a shovel and pan, logs can be obtained with a saw and a river, and oil can be extracted with little more than the technology to drill a water well. Thus, in initial phases, successful development requires risk-taking by the available labor supply and only modest support operations. Production initially does not require a large collection of highly-trained technical and professional workers.

The third advantage of a natural resource is that it initially can produce economic rents that allow the region to import needed capital for other ventures. Because natural resources tend to be scarce worldwide, the price of the resource generally will be sustained above the marginal cost of producing it in areas where it is abundant. Trading with more developed countries allows the region to import capital goods that cannot be developed locally with existing capital. A natural resource, therefore, can serve as a channel through which raw materials are converted into reproducible capital.[6]

At the same time, however, dependence on natural resource production has several disadvantages that can threaten sustained economic growth of more developed economies. To begin with, access to world markets makes the region highly susceptible to fluctuations in those markets. Changes in terms of trade, embargoes, or trade barriers can have a large impact on a resource-dependent economy. Moreover, changes in world supplies, such as the discovery of new reserves elsewhere, or in demand, such as shifts in taste or technology, have immediate impacts on the local economy. Markets are less predictable and controllable than when supply and demand are more insulated geographically.

The downside of resource price volatility clearly has been evident in the oil-dependent states in the 1980s. While the rest of the nation has enjoyed a record peacetime expansion, the oil states have experienced the worst recession in the post-depression era because of the sharp drop in oil prices.

The problem is exacerbated by the close linkage between natural resource production and government policy. In most countries, natural resources are heavily controlled or owned by governments. Oil prices are determined in large part by OPEC's decisions about production. Agricultural production is highly subsidized in most countries. Forestry sales are heavily dependent on decisions concerning access to public stands of timber—decisions that often are influenced by environmental policy considerations. These government controls, therefore, further expose a resource-dependent economy to political uncertainty, as well as to the uncertainty that would normally arise in the market.

The low skill requirements needed in many of the extraction industries also can work to the detriment of long-term regional growth. In many cases, boom periods lead to rapid increases in the wages of low-skilled workers, encouraging a migration of such workers from other areas of the country. These low-skilled workers receive high wages because of temporary output surges and concomitant shortages of labor, rather than as a result of a steady-state measure of their opportunity cost. In periods with slowing demand and production, these workers often become unemployed and typically are less able than skilled workers to find other work at similar pay.

Moreover, the temporarily high returns to low-skilled labor—which may not be perceived to be temporary by the workers—can inhibit investment in human capital. The value of boosting human capital through training and education is obscured by the temporary returns to relatively low levels of such human capital. Over the long run, a drop in the skilled labor pool can slow innovation and productivity growth.

Finally, the theory of exhaustible resource production typically points to declining production after some stage. Optimal production behavior calls for spreading production over time, but even in cases where production capability first must expand, production is expected to decline eventually, both as a share of output and in absolute levels (Pindyck [1978]). Production tends to rise in early stages as discoveries are made and new reserves are developed, but finding large new deposits to replace the reserves that were used in previous periods becomes harder over time. The case of declining U.S. oil production in the face of rising prices during the 1970s offers dramatic evidence of this potential problem.

Summary

Natural resources potentially can play an important role in a region's economy. First, a region heavily endowed with natural resources can be expected to specialize in resource production in conformity with the notion of comparative advantage. Second, other things equal, possession of more resources boosts the potential output of the region.

With respect to the rate of growth, however, the results are mixed. Natural resources have at times been instrumental in providing a catalyst for rapid and sustained growth. In other cases, a resource boom has crowded out other industries to the detriment of the region's longer-term prospects. Moreover, while natural resource industries may provide strong growth at early stages of a region's development, the advantages associated with those industries are more questionable as the economy matures.

III. Relative Performance and the Role of Prices

As indicated in Section II, the relationship between natural resources and differential regional economic growth is complex. In this section, evidence from the GSP data is presented to address two aspects of this relationship. First, gross statistics on the relative economic performance of resource-dependent and non-resource-dependent states are presented. Second, simple regression results are presented to determine whether resource price shocks or resource industries *per se* influence relative performance more.

Relative Performance

For the purpose of this analysis, the 50 states are split into two groups for each natural resource industry based on their resource dependence. Those identified as "resource dependent" are listed in Table 2.[7] As shown in Charts 1–5 earlier, resource dependence is highly skewed across states. In the case of agriculture and energy, the top 10 states are selected as resource dependent. In contrast, because the number of states with significant mining or forestry activity is small, only the top five states are categorized as resource dependent.

The performance of resource-dependent states has differed from that of other states during the 1964–86 period. As shown in Table 3, resource-dependent states ("high") grew more rapidly than did other states ("low"). Over the whole period, the top ten resource states grew at an average annual rate of 3.35 percent, compared to 2.56 percent growth for the other 40 states.[8]

This difference in growth rates was not constant over the whole sample period. Resource industries, particularly the energy sector, have faced significant changes over this period. Accordingly, the sample period was split into three sub-periods—1964 to 1972, 1973 to 1981, and 1982 to 1986—that are divided by the major oil price shocks in 1973–74 and 1979–80.

Although these periods are selected to capture changes in energy industry activity, the shocks were sufficiently traumatic to overall economic activity to make those dates important to the other resource industries. Economic slowdowns early in each of the last two sub-periods, while not necessarily causally related to the oil price shocks, were associated with sharp changes in many of the industries. Higher interest rates dampened the demand for lumber by slowing construction activity, and changes in inflation affected mineral prices and had significant impacts on agricultural land values.

These periods also were characterized by sharp changes in natural resource price behavior. In the early period, natural resource prices were fairly stable. In the middle period, however, energy prices surged. Overall, mineral, lumber, and agricultural prices showed little trend growth over the period, but sharp increases in precious metals

Table 2

Average Share of GSP Contributed by Natural Resource Industries

(1964–86, for the States with the Largest Shares)

ALL RESOURCES*	Avg. Shr.	AGRICULTURE	Avg. Shr.
Wyoming	46.6	South Dakota	17.8
Louisiana	38.9	North Dakota	15.5
Alaska	34.1	Iowa	12.0
New Mexico	30.0	Nebraska	11.7
Oklahoma	26.8	Idaho	10.5
North Dakota	26.3	Montana	8.4
Texas	23.7	Arkansas	7.2
Montana	22.3	Kansas	6.4
South Dakota	20.3	Minnesota	5.9
West Virginia	19.4	Mississippi	5.4
FORESTRY	**Avg. Shr.**	**MINING**	**Avg. Shr.**
Oregon	8.2	Wyoming	3.8
Idaho	4.8	Arizona	2.6
Washington	3.0	New Mexico	2.1
Maine	2.9	Utah	2.1
Montana	2.6	Nevada	2.0
ENERGY	**Avg. Shr.**		
Wyoming	39.5		
Louisiana	36.4		
Alaska	31.8		
New Mexico	24.9		
Oklahoma	23.0		
Texas	21.0		
West Virginia	17.8		
North Dakota	10.2		
Montana	9.5		
Kansas	8.5		

*Combined output from agriculture, forestry, mining, and energy.

prices at the end of the 1970s and sharp agricultural spikes in 1973 and 1980 were important events shaping industry activity.

As indicated in Table 3, resource industries have had major changes in the pace of economic growth over the whole period. In the 1973–81 period, resource states posted growth that exceeded other states' growth by nearly 1.75 percentage points. Conversely, in the 1982–86 period, resource states grew more than four percentage points *slower* than the remaining 40 states.

The more rapid growth in the resource states was accompanied by greater variance. The weighted average root mean square error around trend growth was 0.57 percentage points higher for resource states than for non-resource states. Comparing sub-periods, however, it is clear that this higher variance largely is the result of the last period, during which growth in resource industries slowed dramatically. In the earlier periods, differences in variances were relatively small.

Considerable differences in the relationship between resources and growth are apparent across natural resources. Agricultural states generally grew at the same rate as other states, with faster growth in the 1973–81 period offset by slower growth in the other periods. The variance in growth also was smaller for agricultural states during most time periods.

In the case of forestry, growth rates were higher overall, although the variance was considerably higher in forestry-dependent economies. As in the case of agriculture, the middle period stands out as the period of relative gain, with slower-than-average growth registered in the other periods.

Mineral states have out-performed the rest of the nation in all sub-periods except the last period, and even there, they performed nearly as well. The variance was higher in

Table 3

Average Growth Rates and Variances of State GSP Stratified by Resource Dependence

(Averages weighted by GSP)

	1964–86	1964–72	1973–81	1982–86
All Resource Industries				
High Dependence	3.35	4.00	4.04	0.99
	(4.61)	(2.44)	(2.76)	(2.56)
Low Dependence	2.56	3.03	2.29	5.03
	(4.04)	(2.43)	(3.13)	(1.13)
Agriculture				
High Dependence	2.68	2.76	2.80	3.32
	(3.76)	(2.37)	(2.58)	(2.02)
Low Dependence	2.68	3.21	2.53	4.52
	(4.16)	(2.44)	(3.12)	(1.29)
Forestry				
High Dependence	2.96	2.33	3.96	3.62
	(5.24)	(3.56)	(3.53)	(0.88)
Low Dependence	2.67	3.22	2.49	4.46
	(4.07)	(2.39)	(3.06)	(1.37)
Mineral Mining				
High Dependence	4.33	3.80	5.40	4.13
	(5.30)	(2.77)	(3.60)	(1.90)
Low Dependence	2.62	3.16	2.46	4.44
	(4.09)	(2.43)	(3.06)	(1.33)
Energy				
High Dependence	3.28	3.92	3.93	1.12
	(4.44)	(2.37)	(2.69)	(2.47)
Low Dependence	2.56	3.04	2.29	5.05
	(4.07)	(2.45)	(3.15)	(1.14)

Growth rates are estimated for each state in each sample period, weighted, and summed. Weighted average root mean square errors from the regressions for each state are reported in parentheses.

the mineral-dependent economies, but the growth rates were the highest registered by any resource-dependent category.

Energy-dependent states had the most dramatic variations relative to the rest of the nation over the whole period. Growth during the pre-1982 period exceeded that in the nation, particularly in the 1973–81 period when oil prices rose sharply. The collapse of oil prices in the 1982–86 period caused growth to plunge to one-quarter of the rate of growth experienced in the rest of the country. Furthermore, after enjoying lower-than-average variance in the early periods, the variance rose to twice the national average in the later period. The volatility of the energy states during this period is not surprising given energy's relatively large share of output in these states (see Table 2) and the large swings in prices observed in the period.[9]

The data in Table 3, therefore, suggest that the better-than-average performance of resource industries largely was the result of gains made during the 1973–81 period. Mining and energy states were clear winners overall, while forestry and agriculture also registered some gains.

Separating Price and Share Effects

The results in Table 3 point to strong gains in resource industries during the period of sharp gains in natural resource prices—particularly in the prices of oil and minerals. In this section, an attempt is made to disentangle relative price effects from output effects associated with the growth of resource industries apart from the price shocks.

Ordinary least squares (OLS) and generalized least squares (GLS) regression results are presented in Table 4 for a simple model relating prices and resource shares to relative output growth. The data are estimated in pooled cross-section time series form, using all 50 states and 23 time periods. In the GLS case, the data are corrected for cross-sectional heteroskedasticity. As indicated by the low degree of explanatory power, these estimations fail to capture most of the differences across states in economic growth. The results are useful, however, in the sense that they provide partial correlation statistics for share and price variables, which are better measures than those from simple bivariate correlations between relative growth and prices or shares separately.[10]

Table 4

Estimated Relationships Between Relative State GSP Growth, Natural Resource Shares, and Natural Resource Prices

Dependent variable: Relative State GSP growth*

	OLS Results		GLS Results**	
Variable	Coefficient	t-statistic	Coefficient	t-statistic
Agriculture Share***	−0.0683	−2.92	−0.2629	−2.38
Agricultural price effect****	0.5474	2.44	0.0023	2.04
Lumber Share***	0.0874	1.33	−0.0743	−2.99
Lumber price effect****	0.0424	0.08	0.0069	3.52
Mining Share***	0.1822	1.43	−1.1975	−1.99
Mining price effect****	0.7747	0.77	0.0178	2.96
Energy Share***	−0.0166	−1.68	−0.0320	−1.42
Energy price effect****	0.5391	11.74	1.84E-7	0.15
Relative Per Capita Income in 1963	0.3582	2.49	0.0476	0.50
Adjusted R^2	.12		.06	

* Defined as the difference between annual percentage changes in state GSP less the weighted average annual percentage change in national GSP.

** GLS results from a two-step regression that first corrects for cross-sectional heteroskedasticity using a diagonal covariance matrix.

*** Share variables are defined for each state as the difference between the state's share of GSP in that industry relative to the national average share.

**** The price effect variable interacts the percentage change in the price of the resource (national prices) with the share variable.

The dependent variable is measured as the difference between the annual percentage changes in the level of real GSP for each state from the national average percentage change (formed using total GSP summed across states). Thus, it represents the relative growth of a state's GSP.

Share variables are formed by taking the share of a state's GSP accounted for by the given resource industry and subtracting from it the nation's average resource share. For resource-dependent states, this variable is positive and for those with below average shares, the variable is negative.

Price variables are formed in two steps. First, the real annual percentage change in each resource price is calculated.[11] The price change for a given resource is then multiplied by the state's relative share of that resource described earlier. This specification is necessary because price effects depend on a state's relative dependence on a given resource. Clearly, oil price increases had a positive effect on energy-exporting states and a negative effect on energy-importing states. With this formulation, a positive coefficient indicates that a positive price shock will boost the resource-intensive state and slow growth in non-resource-intensive states.

Finally, to account for the possibility that growth also results from factor movements generated by differences in per capita income, the relative per capita income of persons in each state in 1963 (the beginning of the period) is included. If faster growth simply is the result of a state starting from a low base, this variable will proxy for that effect.

Both OLS and GLS results are reported because together, they convey some sense of the robustness of the relationships. Differences between the two estimates arise from the treatment of variances of state growth rates, which can differ because of a variety of factors, including the size and diversity of the state's economy. The GLS model estimates the coefficients after standardizing the variances of the state, while the OLS model makes no such correction. Both methods yield unbiased coefficients, although the standard errors can be biased in the OLS case.

Results from the regressions differ in the size and significance of the coefficients, but several broad characterizations can be made. First, share variables either do not have a significant influence on relative growth, or where significant at the 95 percent confidence level, tend to have a negative influence on growth. These results suggest that, other things equal, having more natural resource production will not stimulate relative economic growth.

Price variables, on the other hand, were positive in all cases. This result suggests that the sharp movements in resource prices during the period did have an important, positive influence on the relative output growth of resource-dependent economies.

The effect of the starting level of the economy, per capita income in 1963, had a positive effect on relative growth in the OLS case and an insignificant effect in the GLS model. To the limited extent that the starting level mattered, therefore, those states with the strongest economies at the beginning grew faster, making the spread between state incomes larger.

Results from the table, therefore, suggest that the superior performance of the natural-resource dependent states shown in Table 3 may be better interpreted as the result of sharp positive price movements during the sample period, rather than advantages associated with resource production *per se*.[12] To summarize, the results in Table 4 indicate that having a large share of natural resources is detrimental to relative growth prospects, unless the relative price of natural resources rises.

IV. Non-Resource Industries and the Dutch Disease

In the previous section the evidence indicated that natural-resource-based economies out-performed the rest of the nation, although the gains appear to be the result of price effects rather than share effects. This finding allows a direct examination of the applicability of the "Dutch disease" to regional economies. In this section, the data are examined to determine whether the gains in resource-based regional economies led to greater concentration in resource industries—and possibly had detrimental effects on non-resource industries—as would be consistent with Dutch disease.

Table 5 presents average changes in the natural resource share of state GSP over various sub-periods (weighted by GSP) calculated for resource- and non-resource-dependent states. Comparing columns, it can be seen that resource-dependent states had much larger changes in the shares of GSP contributed by the various natural resource industries than did the non-resource-dependent states. This is not surprising given the small shares that those industries contribute in non-resource states.

Comparing resource industries in the resource-dependent states, the largest changes in output shares occurred in the energy sector. Energy's share of output *dropped* in each period, with declines of three and six percentage points in the last two sub-periods in the energy-dependent states. In contrast, mining had less than a 0.4 percentage-point change in share in the states that have the greatest output shares in that industry. Agriculture and forestry had slightly larger changes, but those changes in shares were less than two percentage points between any two subperiods.

As shown in the table, agriculture, forestry, and mining shares dropped sharply during the 1973–81 period, despite positive price shocks. Combined with the earlier information that showed those states doing far better than average during that period, this suggests that non-resource industries in those states were the most important source of growth—causing the resource share to fall because of the faster growth of the other sectors.

A direct comparison of sectoral growth is given in Table 6. Comparing the first two columns, it is clear that resource industries lagged in contributing to growth. In all cases—for both high- and low-resource-dependent states across all categories of resources—the growth rate of the resource industries was below that of the total state economy.

Comparing the two groups of states, the growth rates of resource industries were relatively similar. Only in the case of forestry states was significantly faster growth registered in the resource sector of the high resource-dependent states.

Non-resource industries in resource-dependent states registered the fastest growth in all cases. Consequently, in an apparent refutation of the Dutch disease hypothesis, the non-resource sectors were the prime beneficiary of the resource price shocks.

This conclusion is strengthened by comparing the growth of non-resource industries that are directly tied to resource production (resource processing industries such as refining, pulp and paper, food processing, and stone, clay, and glass) with that of other industries with less direct ties. As shown in the last two columns of the table, non-processing industries ("other") had faster growth than all processing industries except in the case of forest products.

Results from the table, therefore, suggest that the Dutch disease did not afflict regional economies. Price effects boosted the economy, but those prices did not result in increased specialization in resource production and declining competitiveness of other industries.

Differences in relative factor mobility may help to explain this difference between regional and national economies with respect to susceptibility to Dutch disease. In the international case, factor flows are constrained by restrictions on immigration and capital movements. Consequently, relative price shifts that encourage the movement of factors to support the resource industry take factors away from other domestic sectors.

In the regional case, few limits are imposed on factor movements. Labor and capital can flow to areas with potential opportunities. Consequently, increased output by the resource sector does not need to reduce factors available to other industries, since those factors can be imported from other regions. Costs of living can rise as labor is attracted, but the cost increases, in turn, will stimulate additional factor movement, such as the inflow of building materials for additional housing.

Results in Table 6 also highlight the inelastic nature of natural resource production. Resource industries were unable to expand significantly even when sharp positive price movements gave them incentive to do so. Energy states often could not increase output as prices rose because of binding constraints on availability. In Texas, for example, the sharp run-up in prices in 1979–80 slowed the secular trend towards declining production and proven reserves, but production could not rise. As a result, oil wealth tended to be invested in other industries or regions.

Table 5

Changes in Resource Shares During Sub-Periods, By Resource Dependence

(Weighted Averages)

INDUSTRY	HIGH DEPENDENCE	LOW DEPENDENCE
Agriculture		
1963–1972	−1.726	−0.703
1973–1981	−1.561	0.257
1982–1986	0.061	−0.172
Forestry		
1963–1972	1.325	0.492
1972–1981	−1.615	−0.017
1982–1986	0.033	0.075
Mining		
1963–1972	−0.399	−0.021
1973–1981	0.114	−0.332
1982–1986	−0.251	0.000
Energy		
1963–1972	−0.228	−0.431
1973–1981	−6.070	−0.015
1982–1986	−3.024	−0.224

Changes in resource shares reported in the table are the difference in the weighted average share of each group in that industry between the two years listed.

Table 6

Growth Rates of Resource and Non-Resource Sectors High and Low Resource-Intensive States

(1964–86 Annual Growth Rates)

	ALL INDUSTRIES	RESOURCE	NON-RESOURCE* Total	Other	Processing**
Resources					
High Dependence	3.34	0.97	4.30	4.46	2.41
Low Dependence	2.56	1.06	2.64		
Agriculture					
High Dependence	2.68	0.59	3.01	2.99	3.30
Low Dependence	2.68	1.08	2.88		
Forestry					
High Dependence	2.96	1.72	3.10	3.18	1.94
Low Dependence	2.66	1.02	2.88		
Mining					
High Dependence	4.33	1.44	4.90	5.00	3.06
Low Dependence	2.62	1.03	2.82		
Energy					
High Dependence	3.28	0.77	4.23	4.39	2.37
Low Dependence	2.56	1.10	2.64		

* Separate growth rates for "other" and "processing" industries are not reported for the low resource-dependent states because of the limited processing activity in those states.

** Processing industries comprise industries that typically take raw natural resource materials and refine them into intermediate or final goods. Those used in this study are food and kindred products, pulp and paper, stone, clay, and glass, and petroleum refining.

V. Conclusion

The GSP data suggest several relationships between natural resources and relative economic performance. Overall, the experience of the 1964–86 period indicates that resource states grew more rapidly than non-resource states. This faster growth was accompanied by higher volatility in growth, however.

The faster growth and higher volatility of resource-dependent states reflects, in part, the significant volatility in natural resource prices observed during the period, particularly in the 1970s. Sharp price increases boosted resource-dependent economies by providing increased investment in the economy.

Contrary to the Dutch disease, price increases in natural resource industries boosted non-resource industries. Resource industries showed little ability to expand output in the wake of favorable price movements and increases in wealth. These increases in wealth, instead, were invested in non-resource industries. Thus, in the states with large resource industries, these non-resource industries expanded most when positive resource price shocks occurred.

Having a large resource sector, therefore, can be beneficial to a region's growth *when* the industry experiences positive price shocks. If prices fall or remain unchanged, the slow growth (or actual decline) in resource industry output can slow the relative growth of resource-dependent states.

This observation suggests an important area for further study. Why does the additional wealth generated by resource price shocks remain within a resource-dependent region and boost local non-resource industries when investment in such industries outside the region is possible as well? Most theories would argue that non-resource industries in a resource-dependent economy would be harmed, as suggested by the Dutch disease, or at least unaffected in a world of freely-flowing capital. The answer to this question may be associated with the information cost arguments discussed earlier, but a full explanation remains to be uncovered.

Finally, differential effects across states may diminish over time. As noted in this study, resource industries have become less dominant in nearly all states. Output shares have fallen sharply, especially in the resource-dependent states, which should make future regional differences in growth less attributable to natural resource price movements.

NOTES

1. For a discussion of the BEA GSP data, see Giese (1989).

2. The data in the table correspond to shares of national *output,* rather than shares of national *reserves* of those resources. Output and reserves tend to be correlated, but particularly in the case of minerals and energy, there may be some differences in the magnitude of the shares based on the length of time the resource has been extracted.

3. The GSP data do not break forestry separately from the industry data. Instead, BEA reports a total for forestry, fisheries, and agricultural services—an aggregation that is not appropriate for this study. Because nearly all employment in forestry is in the durable goods category "lumber and wood products," that category is used as a proxy for the contribution of forestry to output. Although this procedure understates the role of forestry, it is representative of that impact.

4. Often, characterizations of resource industry importance magnify the effect of the industry by counting the employment of all persons in some way connected to resource processing—some figures for agriculture range as high as 25 percent of the economy. Such a claim would be valid only if (1) those services and production were *not* performed if the state did not have resource production—including retail sales of food—*and* (2) those inputs tied up in the resource chain otherwise would be unemployed.

5. This low initial cost is not always the case, of course, as evidenced by the high cost of developing Alaska's oil fields.

6. This advantage is not limited to developing countries. The Soviet Union, for example, earns much of its hard currency to purchase machinery and supplies from sales of gold and oil to the Western countries.

7. GSP statistics and information used in this article are expressed in real terms unless otherwise noted.

8. Growth rates were estimated by regressing the log of total real GSP on a constant and a time trend for each state. Coefficients on the time trend indicate the growth rate, while the root mean square error from the estimation is used to measure the variance. Averages for the high and low groups are weighted by size of GSP.

9. Evidence on real oil prices is presented by Schmidt (1988). As shown in that article, real oil prices have been trendless over the past 115 years, although prices have tended to be volatile. The price spikes in the 1970s, however, were clear outliers, with unusually large deviations from the historical average.

10. The significance of the coefficients supports the inclusion of resource shares in models of regional performance. See Sherwood-Call (1988) for related work that incorporates farm and oil variables into models explaining deviations of state growth from national performance.

11. Prices for the resources were selected from several sources. Lumber and oil prices are based on the wholesale price indexes for lumber and crude oil, respectively. Agricultural prices are based on the series, "Prices Received," published by the U.S. Department of Agriculture. Mineral prices are the weighted average of iron, copper, lead, and zinc prices (using fixed consumption weights derived from average consumption over the period). In all cases, the prices are deflated by the general wholesale price index.

12. This is particularly true in the case of energy. In the other resource industries, although no trend growth in prices was noted, price increases (the percentage increase) were larger in the positive direction than in the negative direction—that is, price declines were more gradual. Since the price variables were formed using annual percentage changes, the positive price movements helped to explain the better-than-average performance of the resource states.

REFERENCES

Borts, George H. "The Equalization of Returns and Regional Economic Growth," *American Economic Review*, 50, June 1960.

_________ and Jerome L. Stein. *Economic Growth in a Free Market*. New York: Columbia University Press, 1964.

Cox, W. Michael and John K. Hill. "Effects of the Lower Dollar on U.S. Manufacturing: Industry and State Comparisons," *Economic Review*, Federal Reserve Bank of Dallas, March 1988.

Giese, Alenka S. "A Window of Opportunity Opens for Regional Economic Analysis: BEA Releases Gross State Product Data," *FRB Chicago Working Paper Series*, WP-1989-3, Federal Reserve Bank of Chicago, February 1989.

Gruben, William C., Joann E. Martens, and Ronald H. Schmidt. "Interstate Shifts in Nonresidential Construction," *Economic Review*, Federal Reserve Bank of Dallas, July 1988.

Gruben William C. and Ronald H. Schmidt. "State and Local Fiscal Policies, Shocks, and Interregional Migration," Presented at the Western Regional Science Association Meetings, February 1989.

Hughes, Jonathan. "Comments on Long-Run Growth and Development," *Energy and the Southwest Economy*, Proceedings of the 1985 Conference on Energy and the Southwest Economy, Federal Reserve Bank of Dallas, 1985.

Laney, Leroy O. "How Contagious is 'Dutch Disease'?" *Economic Review*, Federal Reserve Bank of Dallas, March 1982.

Miernyk, William H. "Fossil Fuels and Regional Economic Development," *Energy and the Southwest Economy*, Proceedings of the 1985 Conference on Energy and the Southwest Economy, Federal Reserve Bank of Dallas, 1985.

North, Douglass C. "Location Theory and Regional Economic Growth," *Journal of Political Economy*, 63, June 1955.

Pindyck, Robert S. "The Optimal Exploration and Production of Nonrenewable Resources," *Journal of Political Economy*, 86, October 1978.

Plaut, Thomas R. and Joseph E. Pluta. "Business Climate, Taxes and Expenditures, and State Industrial Growth in the United States," *Southern Economic Journal*, 51, July 1983.

Schmidt, Ronald H. "Effects of Energy on Income Growth in the Southwest," *Energy and the Southwest Economy*, Proceedings of the 1985 Conference on Energy and the Southwest Economy, Federal Reserve Bank of Dallas, 1985.

_________. "Hotelling's Rule Repealed? An Examination of Exhaustible Resource Pricing," *Economic Review*, Federal Reserve Bank of San Francisco, Fall 1988.

_________ and William C. Gruben. "Expectations and Regional Growth," Presented at the Regional Science Association Meetings, Toronto, Canada, November 1988.

Sherwood-Call, Carolyn. "Exploring the Relationships between National and Regional Economic Fluctuations," *Economic Review*, Federal Reserve Bank of San Francisco, Summer 1988.

Smith, Donald M. "Regional Growth: Interstate and Intersectoral Factor Allocations," *Review of Economics and Statistics*, 56, 1974.

State of California. *Measuring Economic Impacts: The Application of Input-Output Analysis to California Water Resources Problems*, Bulletin 210, Department of Water Resources, March 1980.

Unemployment-Rate Dynamics: Aggregate-Demand and -Supply Interactions

John P. Judd and Bharat Trehan

Vice President and Associate Director of Research, and Senior Economist, Federal Reserve Bank of San Francisco.

Research assistance was provided by Conrad Gann. We wish to thank the members of the editorial committee, Fred Furlong, Adrian Throop, and Carl Walsh, for valuable comments on earlier drafts of this paper.

The implications for monetary policy of movements in the unemployment rate depend upon the nature of the underlying disturbances that caused those movements. Positive aggregate-demand shocks cause the unemployment rate to fall as inflationary pressures build, whereas positive aggregate-supply shocks are likely to lead to a fall in both the unemployment rate and inflation. In this paper, we employ a recently developed modeling technique to disentangle the effects of aggregate-demand and -supply shocks on the unemployment rate. The technique is agnostic about alternative macroeconomic theories, deriving identifying restrictions from relatively uncontroversial long-run, or steady state, relationships.

The unemployment rate often plays an important role in monetary-policy deliberations, not only because policy makers are concerned about unemployment itself, but also because it is viewed as an important indicator of future inflation. For example, when the unemployment rate declined rapidly to a relatively low level in recent years, a number of Federal Reserve officials became concerned that the economy was developing dangerous inflationary pressures.[1]

One problem in evaluating the policy implications of movements in the unemployment rate (as well as those of other macroeconomic variables) is that these implications often depend on one's *assumptions* about the structure of the economy. Currently there is little agreement among economists concerning the appropriate paradigm; the Keynesian (both the traditional and "new" versions), real-business-cycle, and monetary-misperceptions paradigms all have significant followings among different groups of macroeconomists.[2]

These paradigms differ in the emphasis they place on aggregate-demand versus aggregate-supply shocks in influencing economic activity and labor-market conditions. Real-business-cycle models ascribe a larger role to aggregate-supply shocks, whereas Keynesian and monetary-misperceptions models place greater weight on aggregate-demand shocks. This distinction between demand and supply factors is important because the appropriate monetary-policy response (or lack thereof) to unemployment rate movements depends on the nature of the underlying disturbance. Positive aggregate-demand shocks cause the unemployment rate to fall as inflationary pressures build, and such developments could make a tightening of monetary policy appropriate. By contrast, positive aggregate-supply shocks are likely to lead to a fall in both the unemployment rate and the rate of inflation. Under these circumstances, a tighter monetary policy most likely would be inappropriate.

In this paper, we employ a recently developed modeling technique to disentangle the effects of aggregate-demand and -supply shocks on the unemployment rate (as well as on other important macroeconomic variables.) The technique is agnostic about alternative theories, deriving identifying restrictions from relatively uncontroversial assumptions about long-run, or steady-state, relationships. Given the current lack of agreement about macroeconomic theory, such models have the advantage that they eschew over-identifying restrictions, and choose not to go beyond the minimum number of restrictions necessary to achieve identification.

Our empirical results suggest that for very short horizons of a few quarters, shocks to aggregate demand account for nearly all of the variance of unemployment rate forecast errors. However, at longer horizons of twelve quarters and more, aggregate-supply shocks play a significant role. Moreover, we find that movements in the unemployment rate that are caused by supply shocks (as defined by our model) are *positively* correlated with inflation, whereas those associated with demand shocks are *negatively* correlated with inflation. Thus, decomposing the sources of unemployment rate movements into demand versus supply shocks can be important in designing effective monetary policy.

The paper is organized as follows. Section I reviews the relevant literature on macroeconomic modeling and discusses the rationale for the approach taken in this paper. Section II sets out the theoretical specification of the model. In Section III, we discuss econometric issues that arise in estimating the model and the results of this estimation, as well as their implications for the sources of variation in important macroeconomic variables, including the unemployment rate. Also in this section we analyze the historical evolution of the unemployment rate and the relationship between inflation and the aggregate-demand and -supply components of the unemployment rate. Policy implications and conclusions are discussed in Section IV.

I. Methodological Considerations and Literature Review

Adherents of the main alternative macroeconomic theories—Keynesian, real-business-cycle, and monetary misperceptions—have very different views about the structure of the economy. A major source of controversy concerns the relative importance of demand and supply shocks. The Keynesian and monetary-misperceptions theories stress the role played by aggregate-demand shocks in inducing short-run movements around long-run trends which are independent of those shocks. In contrast, real-business-cycle theories emphasize the role played by technology and labor-supply shocks in producing short-run fluctuations in output around changing equilibrium values which are themselves determined by factors traditionally emphasized in neo-classical growth models.

These alternative macroeconomic theories have different implications for how monetary policy should be conducted. For example, since Keynesians believe that unemployment rate movements mainly are induced by aggregate-demand factors, inflation will rise (fall) when the measured unemployment rate goes below (above) its natural rate. Assuming the monetary authority knows the natural rate of unemployment, Keynesians suggest that the observed rate of unemployment relative to its natural rate can be a major source of information in setting policies to control inflation.

In contrast, real-business-cycle theorists believe that aggregate-supply shocks are the predominant sources of change in macroeconomic variables. Under these circumstances, policy mistakes would be made if the central bank interpreted the unemployment rate as an indicator of aggregate-demand pressures. Further, existing real-business-cycle models generally have modeled business cycles as Pareto-optimal responses to exogenous shocks. Thus, in these models, there is no role for any type of macroeconomic policy aimed at stabilizing the economy.

Macroeconomists have not been able to agree on which theory (or combination of theories) most accurately describes the economy. Each theory implies a different set of identifying restrictions. Thus, a certain degree of agnosticism is warranted in selecting identifying restrictions. This agnostic approach increasingly has shown up in macroeconomic research in recent years. The use of vector autoregressions appears to reflect this view. No identifying restrictions are needed to obtain macroeconomic forecasts. Of course, if these forecasts are to be interpreted in terms of economic theory, identifying restrictions must be added. In early applications, these took the form of assuming a specific recursive structure for the contemporaneous correlations in the data.[3,4]

The Blanchard-Quah Model

Recently, Blanchard and Quah (1989) specified a small vector autoregression of the macroeconomy that achieves identification by imposing relatively uncontroversial constraints on steady-state conditions, thereby avoiding the restrictions associated with alternative theories of the business cycle. Moreover, their model is *exactly* identified, and thus avoids over-identifying restrictions that may raise theoretical controversy. Blanchard and Quah (BQ) assume that supply shocks (those emphasized in real business cycle models) can have permanent effects on the level of real activity, while demand shocks (those emphasized by the Keynesian and monetary-misperceptions models) can have only temporary effects. These assumptions are consistent with each of the three main macro paradigms. Importantly, they are sufficient to identify certain types of VARs incorporating important macroeconomic time series.

Since the BQ approach is used in this paper, albeit on a larger model, it is useful to see how their method works. (A detailed discussion of their method of identification is provided in Appendix A.) BQ specify a VAR with two variables: the rate of growth of real GNP (y), and the level of the unemployment rate (u). Two types of (unobserved) structural disturbances, v and e, are assumed to affect these variables. (As discussed below, we follow BQ in identifying these disturbances with aggregate-demand and aggregate-supply disturbances.) Equations (1) and (2) are moving average representations of y and u in terms of these two disturbances. For simplicity here, we introduce dynamics into the BQ model by including only one lag of v in each equation, although the full BQ model contains several lags.

$$y_t = a_1 e_t + b_1 v_t + c_1 v_{t-1} \tag{1}$$

$$u_t = a_2 e_t + b_2 v_t + c_2 v_{t-1} \tag{2}$$

In order to study the dynamics of this system, it is first necessary to obtain estimates of the various coefficients in equations (1) and (2). This requires placing certain restrictions on e and v. In traditional fashion, BQ assume that e and v are uncorrelated with each other and have unit variance. In addition, e and v are also serially uncorrelated. Given the representations in (1) and (2), these assumptions imply the following identifying restrictions:

$$\sigma_y^2 = a_1^2 + b_1^2 + c_1^2 \tag{3a}$$

$$\sigma_u^2 = a_2^2 + b_2^2 + c_2^2 \tag{3b}$$

$$\sigma_{y_t,\, u_t} = a_1 a_2 + b_1 b_2 + c_1 c_2 \tag{3c}$$

$$\sigma_{y_t,\, u_{t-1}} = c_1 b_2 \tag{3d}$$

$$\sigma_{y_{t-1},\, u_t} = b_1 c_2 \tag{3e}$$

where σ_y^2 is the (observed) variance of y, σ_u^2 is the variance of u, $\sigma_{y_t,\, u_t}$ is the contemporaneous covariance of u and y, and the other variances are defined similarly.

So far, there are five conditions from which we must obtain six coefficients. One more restriction is needed to identify the model. The traditional approach has been to impose a recursive structure on the contemporaneous correlations in the data (Sims [1980]). For example, one might assume that the coefficient $a_1 = 0$; that is, shocks to the unemployment rate do not have a contemporaneous effect on the rate of growth of output. Such an assumption, however, would be theoretically controversial.

BQ avoid having to assume contemporaneous causal orderings by relying on long-run, or steady-state, restrictions. Specifically, they assume that v has no long-run effect on output; that is,

$$b_1 + c_1 = 0. \tag{3f}$$

This restriction, together with the conventional restrictions on variances and covariances, is sufficient to identify the unobserved shocks from observations on y and u. The restriction also leads to a straightforward interpretation of the underlying structural disturbances: v can be interpreted as an aggregate-demand shock since it can have no long-run effect on y, while e can be interpreted as a supply shock since it is permitted permanently to affect y. In other words, the permanent level of real GNP is determined by real factors. Although aggregate-demand shocks can cause real GNP to deviate from this level, it cannot affect the permanent level itself. By construction, neither demand nor supply shocks have a permanent impact on the unemployment rate.

Using this method, BQ found that demand disturbances had a hump-shaped effect on the time path of output, while supply shocks had an effect that increased gradually over time. They also found that demand disturbances accounted for only 35% of the variance of unpredictable changes in real output in the contemporaneous quarter, leaving 65% for supply disturbances, while demand accounted for 13% at a horizon of eight quarters.[5] In contrast, demand disturbances accounted for 100% of the variance of unpredicted changes in the unemployment rate in the current quarter, and for 50% at an horizon of eight quarters.

The Shapiro-Watson Model

One problem with BQ's analysis is that it allows for only two underlying disturbances to the economy. If, as seems plausible, the economy is affected by more than one kind of supply (or demand) shock, their procedure will tend to confound the effects of these different shocks. Based on this reasoning, Shapiro and Watson (1988) used a system that comprised real GNP, total labor hours, inflation and the real interest rate.[6] This set of variables allowed them to account for four different disturbances: two to aggregate supply, which they identified as shocks to labor supply and technology, and two to aggregate demand, which they referred to as IS and LM shocks, but did not identify separately.

Shapiro and Watson (SW) found that aggregate-demand shocks had a smaller impact on real output than BQ did.[7]

Specifically, aggregate-demand shocks accounted for just 28% of the variance of the output forecast error in the contemporaneous quarter, and 20% at an eight-quarter horizon. In addition, they found that labor supply shocks alone accounted for about 45% of the variance of unpredicted changes in output in the contemporaneous quarter. These findings, as well as those of BQ, sharply contradict the Keynesian and monetary-misperceptions views that trend and cycle are neatly separable, with demand shocks playing the dominant role over the business cycle.

II. Model Specification and Identification

In this section, we present the specification of the model estimated in this paper. We begin with a discussion of the variables included in the model, followed by a discussion of the equations that constitute the model, and how we achieve identification.

The model includes five variables: the unemployment rate, real GNP, a nominal rate of interest, a measure of labor supply, and a variable that measures foreign trade. These variables provide broad coverage of important types of activity in the economy, and thus should capture the economic relationships that are important in determining the behavior of the unemployment rate. Movements in the unemployment rate and the interest rate are likely to be highly correlated with two types of underlying aggregate-demand shocks, which can be thought of as being associated with the IS and LM curves of textbook macroeconomic theory. The interest rate should capture shocks both to inflation expectations and real interest rates, while the unemployment rate should reflect aggregate-demand shocks as they affect the level of economic activity. Following previous research, we assume that movements in real GNP are correlated with technology shocks, once we standardize for aggregate-demand shocks.[8]

We use working-age population as our measure of labor supply. This variable is clearly exogenous, and therefore guards against the possibility of confounding labor demand and supply. However, it has the disadvantage of omitting the effects on labor supply of changes in participation rates and average hours worked. One obvious alternative would be to follow SW and use total labor hours as the labor-supply variable. However, our empirical evidence suggests that using labor hours to measure supply causes a serious bias; we are unable completely to separate the demand-induced changes in labor hours from those induced by labor supply. Specifically, when we include labor hours in our model, we find that a positive labor-supply shock leads to a large, sustained *decline* in the unemployment rate, an outcome that suggests a confusion between labor supply and demand. Such confusion could have a profound effect on conclusions concerning the relative importance of supply and demand disturbances in macroeconomic time series. By contrast, as discussed below, the results obtained when population is used as a measure of labor supply are much more plausible. Thus, given this paper's policy-driven focus on the unemployment rate, we have opted for working-age population.

We also extend the BQ and SW models by explicitly incorporating a foreign variable to identify the effects of shocks originating abroad. Given the growing importance of international trade and capital flows to the U.S. economy, it is desirable to incorporate the independent effects of shocks from abroad. While inclusion of the exchange rate appears to be an obvious choice, the move from fixed- to floating-exchange rates in the early 1970s implies a change in the exchange-rate process that precludes sensible estimation results over our 1954–88 period. Instead, we include as a foreign variable the ratio of real exports to real imports.

The Underlying Model

We begin by assuming that the production technology can be described by a neo-classical growth model, so that the long-run level of output is determined by the capital stock and labor supply.[9] The capital stock term can be eliminated by assuming a Cobb-Douglas production function and a constant steady-state capital-output ratio. Thus, the steady-state level of output can be expressed as a function of the steady-state levels of labor supply and technology.

The levels of labor supply and technology may be permanently affected by labor-supply and technology shocks, respectively. The evolution of these variables is described by

$$s_t^* = \alpha^s + s_{t-1}^* + \beta^s(L)\,\mu_t^s \tag{4}$$

$$e_t^* = \alpha^e + e_{t-1}^* + \beta^e(L)\,\mu_t^e \tag{5}$$

where s^* is the log of the steady-state value of labor supply and e^* represents (unobserved) technology. The labor supply and technology shocks, μ^s and μ^e, are uncorrelated, and the lag polynomials $\beta^s(L)$ and $\beta^e(L)$ describe the transitory movements in s^* and e^* as they move to new permanent levels.

Labor supply is not affected, either in the short or long run, by any of the other variables in the system. This assumption follows from our choice of working-age population to represent the influences of labor supply, and yields four of the ten restrictions we need to identify the model.[10]

Both labor-supply and technology shocks can cause short-run movements in output as the level of output adjusts to a new steady-state value. Short-run movements in output also can be the result of aggregate-demand shocks. However, the two types of aggregate-demand shocks are permitted to have only temporary effects on the level of output. These assumptions yield two more identifying restrictions. Foreign shocks cause output to deviate temporarily from its steady-state value, but are not permitted to have a long-run effect on output.[11] This assumption yields one more identifying restriction.

These considerations suggest the following equations for the relationship between observed and equilibrium values:

$$s_t = s_t^* + \chi^s (L) (\mu_t^s) \tag{6}$$

$$y_t = y_t^* + \chi^y (L) (\mu_t^s, \mu_t^e, \mu_t^f, \mu_t^{d_1}, \mu_t^{d_2}) \tag{7}$$

where $\chi^s(L)$ and $\chi^y(L)$ are vectors of lag polynomials (in the indicated variables) that allow for temporary deviations from steady-state levels. Thus, this specification allows the actual level of output to deviate from the level implied by the Cobb-Douglas production function in the short run. As discussed above, y^* itself is a function of s^* and e^*. μ_t^f denotes shocks originating abroad, while $\mu_t^{d_1}$ and $\mu_t^{d_2}$ are the domestic demand shocks.

Statistical tests suggest that output and labor supply are both nonstationary, and thus we take first differences of equations (6) and (7) (see Appendix B). Substituting equations (4) and (5) into the results yields:

$$s_t - s_{t-1} = \alpha^s + \beta^s (L) (\mu_t^s) + (1-L)\chi^s (L) (\mu_t^s) \tag{8}$$

$$y_t - y_{t-1} = \alpha^y + \beta^y (L) (\mu_t^s, \mu_t^e) + (1-L)\chi^y (L) (\mu_t^s, \mu_t^e, \mu_t^f, \mu_t^{d_1}, \mu_t^{d_2}) \tag{9}$$

Consider now the specification of the foreign variable. In addition to disturbances originating abroad, this variable is affected by all the domestic shocks. However, the two aggregate-demand shocks are permitted to affect the foreign variable only temporarily.[12] These assumptions yield two more identifying restrictions. We assume that the long-run evolution of the foreign variable can be described in the same way as output, so it is included in the model in a form similar to (8) and (9). Thus,

$$f_t - f_{t-1} = \alpha^f + \beta^f (L) (\mu_t^s, \mu_t^e, \mu_t^f) + (1-L)\chi^f (L) (\mu_t^s, \mu_t^e, \mu_t^f, \mu_t^{d_1}, \mu_t^{d_2}) \tag{10}$$

Given that the interest rate appears to be non-stationary (see Appendix B), we specify its equation in differenced form:

$$i_t - i_{t-1} = \chi^i (L) (\mu_t^s, \mu_t^e, \mu_t^f, \mu_t^{d_1}, \mu_t^{d_2}) \tag{11}$$

Thus, all the disturbances in the model can have a permanent effect on the nominal interest rate.

There is some ambiguity about how the unemployment rate should be included in the model. On the one hand, there is a large body of theoretical work in macroeconomics to suggest that the unemployment rate is stationary.[13] Tests carried out over long sample periods tend to confirm this.[14] On the other hand, as shown in Appendix B, unit root tests suggest that the unemployment rate is non-stationary over shorter sample periods.

This inability to reject nonstationarity in the unemployment rate over the post-war period poses a problem. Different researchers have dealt with this problem in different ways. BQ for instance, present results both for the case where the unemployment rate is assumed to be stationary and where it contains a deterministic trend. Unfortunately, removal of a linear trend is not sufficient to make the unemployment rate stationary. Evans (1989) allows for an increase in the mean of the unemployment rate beginning in 1974. As indicated in Appendix B, allowing for this shift in the mean appears to make the unemployment rate stationary.

Acceptance of this "solution" to the nonstationarity problem implicitly assumes the existence of some well-defined, exogenous change in the economy that is associated with a change in the mean unemployment rate. While some economists have pointed to the change in participation rates of women and teenagers in the labor force over this period, the issue is by no means resolved.[15] Accordingly, we estimated two alternative versions of the model, one that allows the mean unemployment rate to change in 1974, and one that holds it fixed over the entire 1954–88 period. The results in the two cases are similar, and so we present only those from the specification that allows for a mean shift. (However, we do point out instances below in which the results from the two specifications differ materially.)

Thus, the complete model comprises equations (8)–(11), plus

$$\mu_t = \alpha^u + \chi^u (L) (\mu_t^s, \mu_t^e, \mu_t^f, \mu_t^{d_1}, \mu_t^{d_2}) \tag{12}$$

where α^u is allowed to shift in 1974Q1. Thus, the unemployment rate is affected by all the disturbances in the model. However, because it is entered as a level, none of these disturbances has a permanent effect on it.

In summary, we have identified the model by restricting certain long-run coefficients to equal zero, and by using working-age population, which is strictly exogenous, for our labor-supply variable. As discussed in Appendix A, we require ten identifying restrictions to separate the influences of each of the five shocks—two domestic demand, two domestic supply, and one foreign—on all the variables in the system. The assumption that population is exogenous yields four identifying restrictions. Four additional restrictions come from the assumption that the two aggregate-demand shocks do not have long-run effects on output and the foreign variable. One more restriction comes from the specification that the foreign shock has no long-run effect on U.S. output. This gives us a total of nine restrictions. Following SW, we choose not to identify the two aggregate-demand shocks separately. In this way, we are able to eliminate the need for a potentially controversial tenth restriction.

III. Estimation and Empirical Results

In this section we describe the estimation technique and present our results. The impulse response functions and the variance decompositions presented below provide information about the structure of the economy as estimated by the model. We use this information to analyze the factors that have contributed to the changes in the unemployment rate that occurred over the period from 1955 to 1988. Finally, at the end of this section, we show correlations between our measures of the aggregate-demand and aggregate-supply components of the unemployment rate and the rate of inflation.

Our model includes the log of the unemployment rate and the first differences of the logs of all other variables. Because population is exogenous, we use ordinary least squares to estimate a regression of population growth on six of its own lags. (A lag-length of six is used in all the equations in the model.)

To illustrate the technique used to estimate the remaining equations, we use the real GNP equation.[16] Real GNP is regressed on lags of all the variables in the system plus contemporaneous values of population, the interest rate, the unemployment rate, and the foreign trade variable. We impose the restriction that neither the aggregate-demand shocks nor the foreign shocks has a permanent impact on the level of GNP by taking the difference of the relevant right-hand-side variables one more time and reducing the lag length by one. Thus the first difference of real GNP is regressed on first differences of population, the second differences of the foreign variable and interest rates, and the first difference of the unemployment rate (in addition to lags of first differences of real GNP). Two-stage least-squares is used to estimate the equation because it contains contemporaneous values of the three endogenous variables (that is, of interest rates, unemployment, and the foreign variable). The contemporaneous value of population and lagged values of all variables in the model are used as instruments.

The remaining equations are estimated in a similar manner. Following our discussion above, domestic aggregate-demand variables are restricted to have only a temporary impact on the foreign variable, while no such restriction is placed on the domestic supply variables. No restrictions are placed on the equations for the interest rate and the unemployment rate. As mentioned above, the inclusion of the level of the unemployment rate in the model implies that no shock to the system has a permanent impact on that rate.

The Estimated Structure of the Model

Exhibit IA shows the impulse response functions from the model. The first two columns of the exhibit show the response of the model's four endogenous variables to domestic shocks, while the third column shows the effects of shocks originating abroad. As discussed above, we identify labor-supply and technology shocks separately, but we do not disentangle the two underlying demand shocks. Thus, the impulse response functions in the second column of the exhibit represent responses to a linear combination of the demand shocks.

Positive aggregate-demand shocks reduce unemployment and raise output and interest rates. By construction, the effects on the unemployment rate and GNP are temporary. The effects of aggregate-demand shocks on the unemployment rate die out in about 12 quarters, while those on output last eight to 10 quarters. At first, the ratio of U.S. exports to imports reacts negatively to domestic demand shocks; that is, higher domestic demand leads to a rise in imports relative to exports. The impulse response function then cycles, becoming positive from five to twelve quarters, at which time the effect dampens out.

Positive shocks to technology reduce the unemployment rate. This effect lasts for about 24 quarters before substantially dying out.[17] Shocks to labor supply have insignificant effects on the unemployment rate, causing it to cycle around its original level. Positive shocks to labor supply

Exhibit IA
Impulse Response Functions

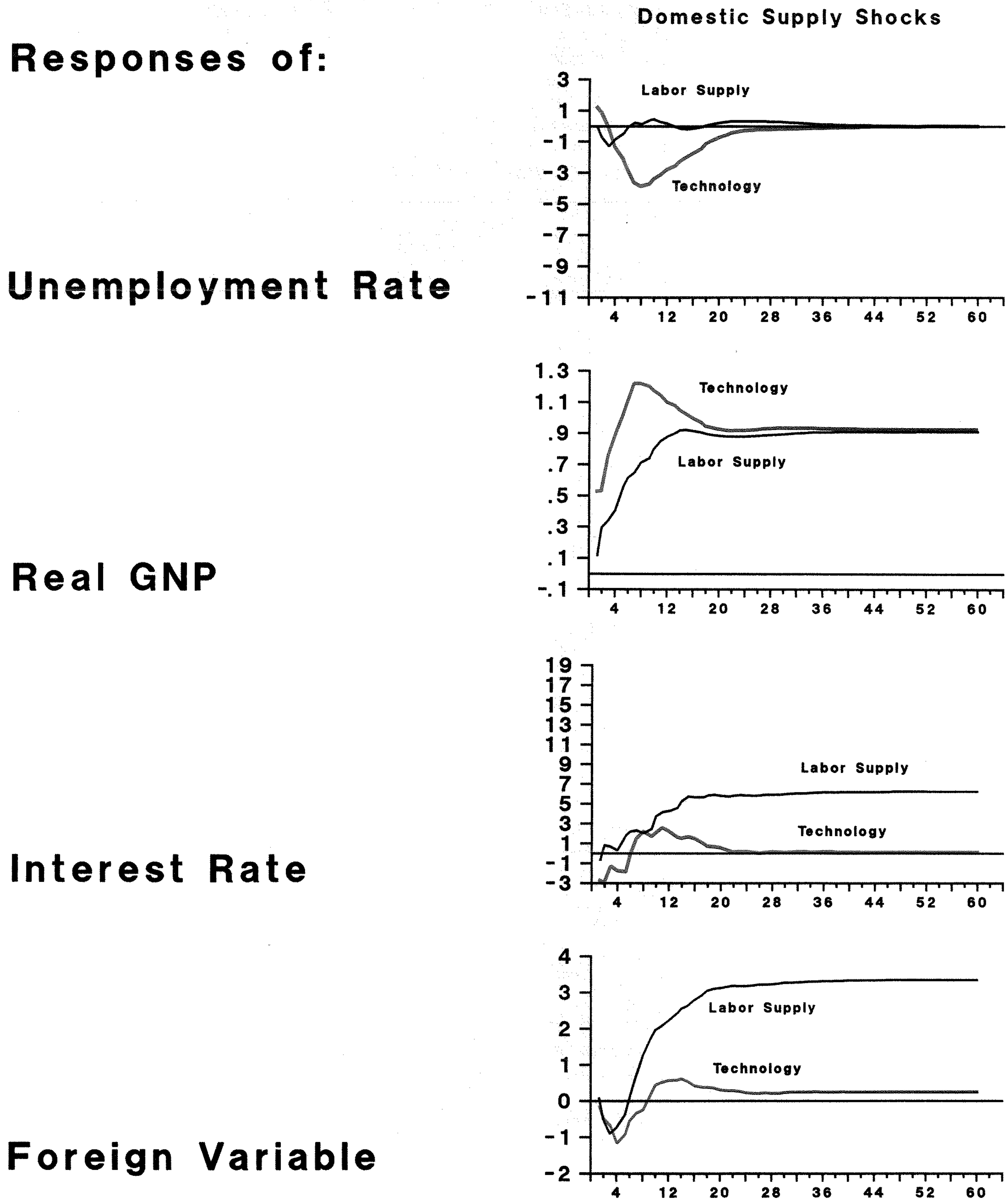
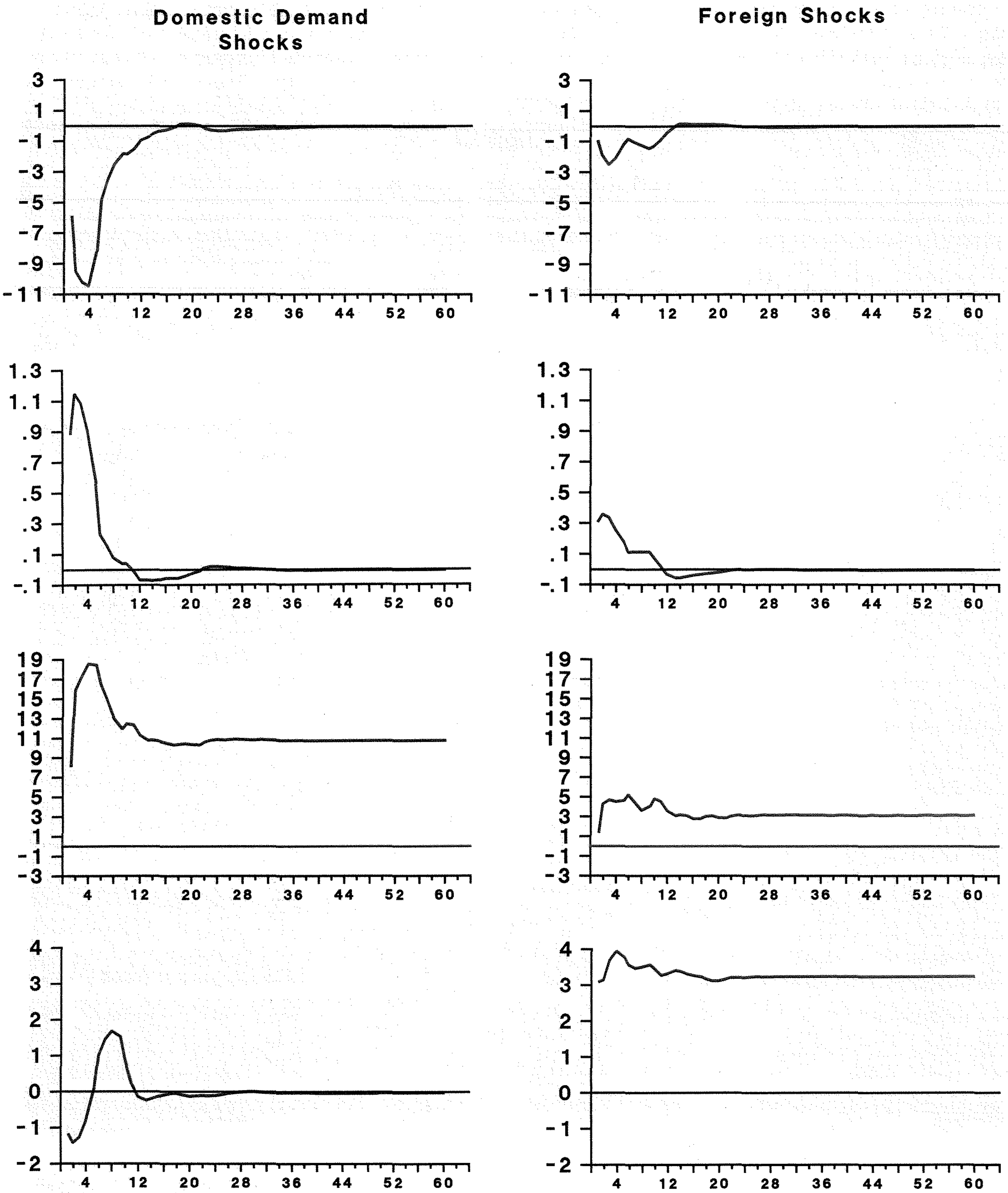


Exhibit IA (continued)



and technology permanently raise output, with the effect of labor-supply shocks building up somewhat more gradually than the effect of technology shocks. Positive shocks to these two variables also permanently raise the level of interest rates, and the ratio of exports to imports.

Positive foreign shocks temporarily raise output and lower the unemployment rate, although the latter effects are relatively small. These shocks also permanently raise the interest rate.

Exhibit IB presents the associated variance decompositions, which show the relative importance of the various kinds of shocks in explaining the errors made in predicting the model's variables. At forecast horizons of up to four quarters, variation in the unemployment rate has been dominated by aggregate-demand shocks. Aggregate-supply shocks begin to play a larger role as the forecast horizon lengthens, reaching 15 percent at eight quarters and 25 percent at 60 quarters. These results suggest that

Exhibit IB
Variance Decompositions

	Shocks				
	Labor Supply (A)	Technology (B)	Domestic Supply (A + B)	Domestic Demand	Foreign
Responses of:					
Unemployment Rate					
Quarters 0	0.0%	6.0%	6.0%	90.0%	3.9%
1	0.7	2.6	3.3	91.4	5.4
2	1.4	1.3	2.7	90.5	6.7
4	1.1	3.5	4.6	90.6	5.7
8	0.9	15.3	16.2	78.2	5.6
12	0.8	21.9	22.7	71.7	5.5
60	1.1	24.7	25.8	68.8	5.3
GNP					
Quarters 0	1.7%	34.7%	36.4%	50.9%	12.7%
1	5.1	28.0	33.1	55.3	11.6
2	6.3	32.7	39.0	50.9	10.1
4	10.7	45.3	56.0	37.1	6.9
8	17.9	60.8	78.7	17.6	3.6
12	24.7	61.7	86.4	11.3	2.3
60	43.0	54.0	97.0	2.5	0.5
Interest Rate					
Quarters 0	0.5%	8.5%	9.0%	87.5%	3.4%
1	0.4	5.4	5.8	85.8	8.3
2	0.3	3.5	3.8	85.8	10.2
4	0.6	2.5	3.1	86.2	10.6
8	1.7	2.3	4.0	84.9	12.0
12	4.9	2.7	7.6	79.9	12.4
60	22.1	0.8	22.9	67.9	9.1
Foreign					
Quarters 0	0.0%	0.1%	0.1%	7.9%	91.9%
1	1.1	1.0	2.1	9.9	88.0
2	2.7	1.7	4.4	8.2	87.4
4	2.4	3.6	6.0	5.2	88.6
8	5.1	2.3	7.4	6.1	86.3
12	12.8	2.2	15.0	4.5	80.5
60	43.2	0.8	44.0	0.7	55.3

unemployment has been substantially affected both by aggregate-demand and -supply shocks during the post-war period.

Aggregate-demand shocks are the most important factor in explaining variation in real GNP in the short run (contemporaneously and at forecast horizons of one and two quarters), accounting for 50 to 55 percent of the variation. Technology shocks also are quite important at these short lags, accounting for from 28 to 35 percent. As the forecast horizon lengthens, technology shocks begin to dominate, as these shocks are permanent, while aggregate-demand shocks are transitory. By the time the lags reach two years, technology shocks dominate demand shocks, with the former factor accounting for 61 percent of the variation and the latter accounting for only 18 percent. Labor-supply shocks begin to become important only after two years. At the frequency of the average business cycle, our results show a larger role for demand shocks relative to supply shocks than does earlier research.

Interest rate variation is dominated at all forecast horizons by domestic demand shocks, although foreign shocks have a noticeable effect in the long run. Domestic supply shocks play only a small role, except at the very long lags. At a lag of 60 quarters, labor supply accounts for 22 percent of the variation in the interest rate, while at shorter lags, the role of this variable is quite small (under five percent).

The foreign variable largely is exogenous with respect to the other four variables in the model—that is, it is determined mainly by its own past behavior—at forecast horizons of up to 12 quarters. At long lags, however, labor supply plays a significant role in the error variance of the foreign variable, reaching 43 percent at 60 quarters. Technology and domestic aggregate demand play only small roles at all forecast horizons.

The effects of the foreign variable on the U.S. economy are relatively modest, as would be expected from the relatively small, albeit growing, role of foreign trade in the U.S. economy. Foreign shocks play a significant role in U.S. real GNP at short lags, accounting for 13 percent of the contemporaneous variation and then declining in importance as the lag lengthens. Foreign shocks also have played a significant role in U.S. interest rate movements, accounting for 10 to 12 percent of the variation in that variable at forecast horizons in the range of two to 12 quarters.

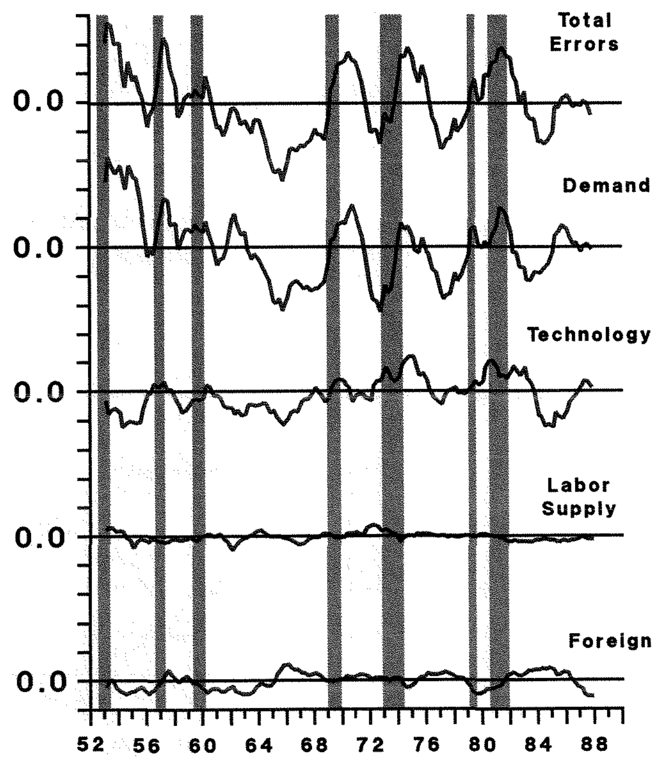
Historical Analysis

We now use our estimated structure to carry out two different exercises that examine the historical evolution of the unemployment rate. First, to understand the factors that have caused movements in the unemployment rate over the course of the business cycle, we look at the sources of our model's forecast errors at a forecast horizon of three years.

The results of this exercise are shown in Exhibit II. By construction, any error in predicting the unemployment rate has to be the result of the unpredicted demand, supply, and/or foreign shocks that took place within the three-year forecast horizon. We obtain the contribution of each kind of disturbance to the forecast error for any particular quarter by multiplying the coefficients in the impulse response functions by the appropriate historical shocks as measured by the model.

The top panel of Exhibit II shows the total error in predicting the unemployment rate twelve quarters ahead over the period from 1955 to 1988. At this forecast horizon, the major errors are closely associated with business cycle swings. The four panels below show the contributions to

Exhibit II
Historical Analysis: Decomposition of 12-Quarter-Ahead Forecast Errors for Unemployment Rate



these forecast errors made by the indicated shocks. Shaded areas represent business cycle downturns.

The most striking feature of this analysis is that aggregate-demand shocks have played by far the largest role in unemployment rate movements over the course of the business cycle in the postwar period. Although technology shocks are important for the average quarterly variability of the unemployment rate over the whole sample, aggregate-demand shocks appear to be more closely related to cyclical swings in the unemployment rate.

Technology shocks do, however, contribute significantly to the longer-run swings in the forecast errors. For example, the well-known productivity surge in the 1960s is picked up in our analysis as a succession of positive technology shocks that led to a lower-than-predicted unemployment rate over most of the decade. Similarly, the slowdown in productivity growth in the early to mid 1970s is picked up as a succession of negative technology shocks.

The 1980s have been marked by large shocks both to aggregate demand and to technology. Not surprisingly, the second panel of Exhibit II shows large, negative aggregate-demand shocks (which pushed up the unemployment rate) during the period from 1980 to 1982, when the Federal Reserve oriented monetary policy around the monetary aggregates to combat the surge in inflation in the late 1970s and early 1980s.

Aggregate-demand shocks then turned positive (thus pushing down the unemployment rate) in 1983 as monetary policy became more accommodative in the face of a continuing recession and falling inflation. In addition, fiscal policy became highly expansionary from 1983 through 1986, with the high-employment deficit jumping sharply in 1983 and reaching a peak in mid 1986. From 1986 through 1988, aggregate demand shocks were relatively small, although on average were slightly negative.

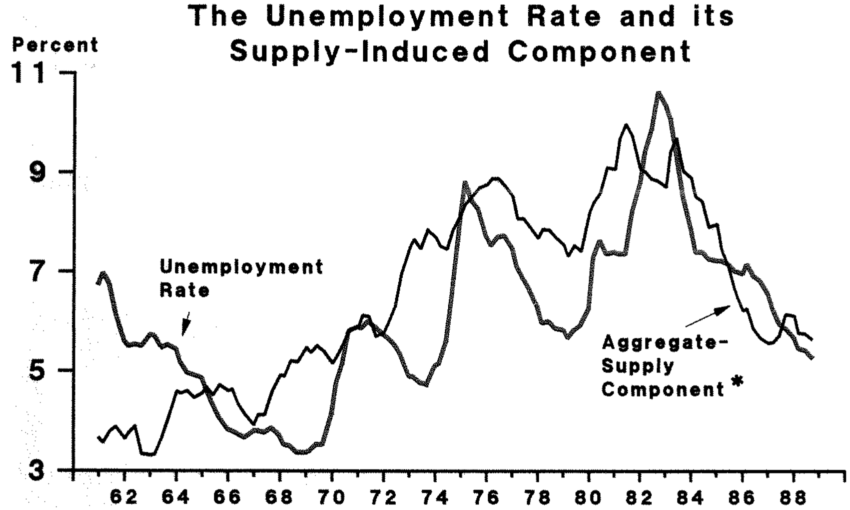
Technology shocks also have been important factors in unemployment rate movements in the 1980s. In fact, they were about as important as aggregate-demand shocks in raising the unemployment rate. This effect was substantial by historical standards and lasted from early 1980 through mid 1984. Technology shocks also accounted for a good part of the unemployment rate decline in 1986 and 1987, when the unemployment rate moved into a range that contributed to the Federal Reserve's concern about future inflation.

What might be responsible for this pattern of technology shocks? Any suggestions would be highly speculative.[18] Several large studies on the sources of productivity change in the U.S., for example, have failed to come up with specific explanations for a substantial portion of that change.[19] Nonetheless, we note that the timing of the negative technology shocks in the early 1970s and the early 1980s is close to the two oil price shocks, suggesting that this factor may have been important. However, as noted elsewhere, inclusion of oil prices causes problems in explaining developments after 1985 (see footnote 6).

Unemployment and Inflation

We turn now to the second exercise of our historical analysis, namely a decomposition of the unemployment rate into its aggregate-demand and -supply components, and a comparison of these components with the inflation

Exhibit III
The Unemployment Rate and its Supply-Induced Component



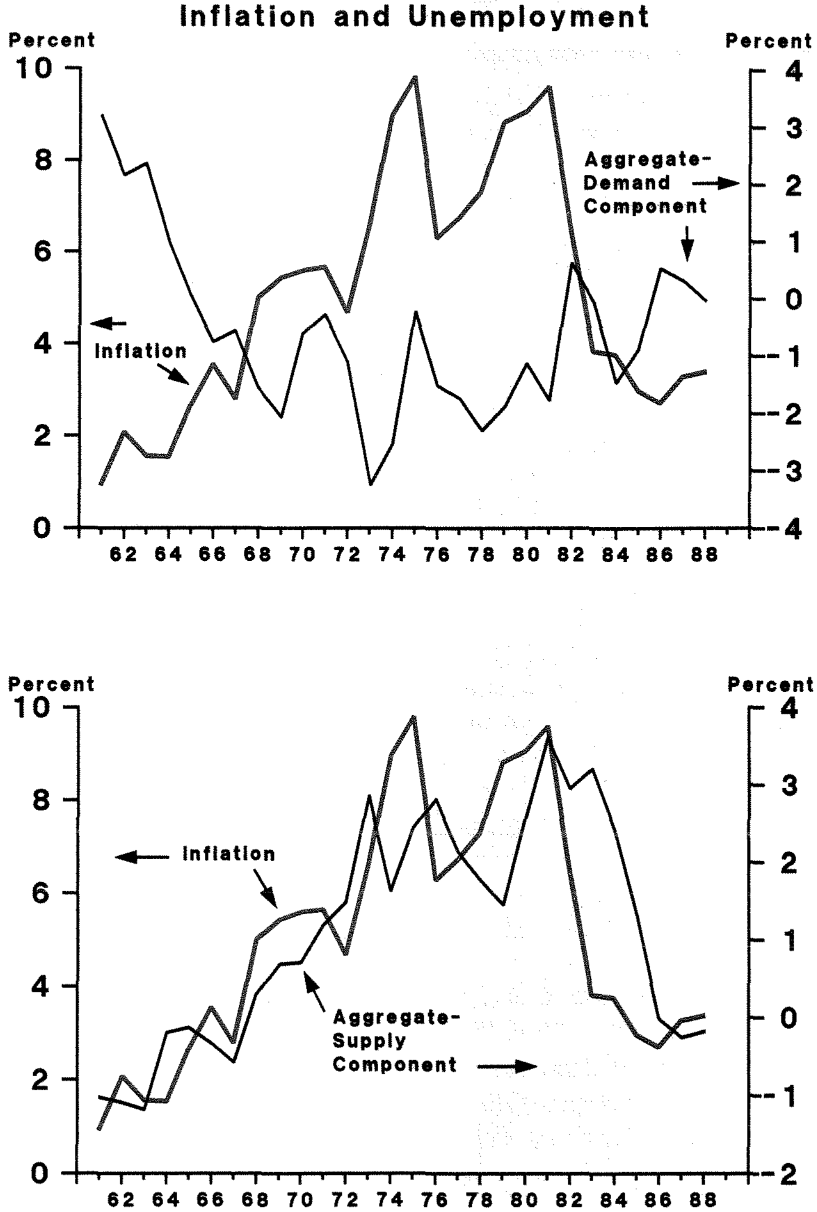
*Includes the estimated mean level of the unemployment rate to allow for comparison with the actual unemployment rate.

rate. In Exhibit III, the actual unemployment rate is plotted against the mean unemployment rate plus the contribution of aggregate-supply factors. To obtain this rate, we subtracted from the unemployment rate both the effects of aggregate-demand-induced changes and the effects of shocks originating abroad.[20] The difference between the two series plotted in Exhibit III represents the effects of aggregate-demand pressures and foreign shocks in the labor market. (Of the two, the latter are not very important.) Demand pressures apparently have reduced the unemployment rate during most of the 1965–1981 period, implying the possibility of an inflationary bias in policy. After 1981, these pressures have been more balanced, sometimes positive and sometimes negative.

According to economic theory, there should be a negative correlation between our measure of the aggregate-demand component of the unemployment rate and the rate of inflation relative to inflation expectations, if our measure is valid. This correlation arises in both the Keynesian Phillips curve and the Lucas-Barro, or monetary-misperceptions, Phillips curve.[21] In the former, an aggregate-demand shock that reduces the unemployment rate leads to higher inflation. In the latter, a positive aggregate-demand shock that raises inflation above inflation expectations (that is, creates an inflation surprise) will lead to a decrease in the unemployment rate.

The expected negative correlation is shown in the top panel of Exhibit IV. (We have used annual averages in

Exhibit IV
Inflation and Unemployment



order to reduce the random fluctuations in the data.) Note that we plot actual inflation rather than the difference between actual and expected inflation. In the following discussion, we implicitly assume a positive correlation between actual and unexpected inflation. The top panel of the exhibit reveals that as the aggregate-demand component of the unemployment rate fell below zero in mid-1960 through 1980, the inflation rate rose, reaching a peak in 1981. Since then, the aggregate-demand component has fluctuated around zero, and inflation has fallen.

The bottom panel of Exhibit IV plots the aggregate-supply component of the unemployment rate and the rate of inflation. As expected, these two series are *positively* correlated. When there is a positive technology shock, for example, inflation falls as prices adjust to a new level, and at the same time the unemployment rate falls as firms' demand for labor rises.

The correlations shown visually in Exhibit IV are presented more rigorously in the first two columns of Exhibit V. The first column presents cross correlations between past, present, and future values of inflation, on the one hand, and the aggregate-demand component of the unemployment rate, on the other. The correlations between the aggregate-demand component of the unemployment rate and inflation are strongly negative, suggesting that our measure of aggregate-demand pressure is functioning as expected.

The second column of the exhibit presents the correlations between our measure of the aggregate-supply component of the unemployment rate and past, present, and future rates of inflation. These correlations are uniformly positive, which appears to validate our concept of the aggregate-supply component of unemployment.

In the third column, we show cross-correlations between the unemployment rate minus its mean rate with past and future values of the inflation rate. The correlations between the aggregate-demand component of the unemployment rate and future inflation are noticeably stronger than those between the (mean-adjusted) unemployment rate and future inflation. Likewise, the positive relationship between past inflation and our measure of the supply-induced changes in the unemployment rate is noticeably stronger than that between past inflation and the unemployment rate.

Notice also that the correlations between past values of inflation and the unemployment rate are positive. Thus, the raw data tend to support the Keynesian Phillips curve, which has causation running from unemployment to future inflation, and refutes the monetary-misperceptions Phillips curve. The latter relationship implies that there should be a *negative* correlation between past inflation surprises and unemployment rates, rather than the positive correlation shown in column three. However, the first column of the table shows that once the aggregate-supply shocks are stripped away, both directions of causation are supported. There is a negative correlation between past inflation and aggregate-demand-induced unemployment (monetary misperceptions) and also between future inflation and unemployment (Keynesian).

Exhibit V

Cross-Correlations Between Inflation and Alternative Measures of Labor Market Pressure

Sample Period: 61:1-88:4

Time[1]	Inflation and Aggregate-Demand Component of Unemployment Rate	Inflation and Aggregate-Supply Component of Unemployment Rate	Inflation and the Unemployment Rate Minus Its Mean
t−12	−.22	.63	.54
t−11	−.22	.67	.57
t−10	−.20	.69	.61
t−9	−.19	.69	.61
t−8	−.20	.71	.61
t−7	−.24	.71	.56
t−6	−.27	.68	.50
t−5	−.31	.67	.43
t−4	−.36	.65	.34
t−3	−.41	.64	.27
t−2	−.50	.62	.17
t−1	−.57	.61	.07
t	−.65	.60	−.03
t+1	−.66	.55	−.10
t+2	−.67	.51	−.16
t+3	−.69	.48	−.20
t+4	−.65	.42	−.21
t+5	−.64	.39	−.22
t+6	−.60	.35	−.22
t+7	−.57	.32	−.21
t+8	−.55	.29	−.21
t+9	−.50	.25	−.21
t+10	−.44	.20	−.19
t+11	−.39	.16	−.17
t+12	−.36	.16	−.14

[1]Exhibit shows correlations between various components of the unemployment rate and past as well as the future values of inflation. For example, the first entry shows the correlation between the current value of the aggregate-demand component of the unemployment rate and the value of inflation 12 quarters earlier.

IV. Policy Implications and Conclusions

The aim of this paper has been to estimate the relative importance of different kinds of disturbances in causing movements in the unemployment rate. Toward that end, we have attempted to keep our model as free as possible of the controversial identifying restrictions that are inherent in the various competing paradigms of the macroeconomy. We find that on average both demand and supply shocks have been important in explaining unemployment rate movements in the postwar period. While demand shocks are relatively more important in causing cyclical swings in the unemployment rate, supply shocks play a significant role in inducing longer-term movements. Our finding that positive supply shocks are correlated with falling unemployment in subsequent periods casts doubt on Phillips-curve analyses, which assume that relative prices and the unemployment rate move independently of each other.

Our historical analysis suggests that supply shocks were important in keeping the unemployment rate low in the 1960s, and relatively high in the early- and mid-1970s. Of particular interest right now is the role played by supply shocks in raising the unemployment rate in the first half of the 1980s, and then lowering it in the second half of the decade. The relatively large role played by supply shocks in the decline in the unemployment rate over the last few years could be one reason the inflation rate has not accelerated as much as past estimates of the unemployment-inflation relationship would have led us to expect.

The analysis is relevant for policy purposes to the extent that policy makers take the unemployment rate into account in determining policy. Policy makers may look at the unemployment rate in order to arrive at an estimate of future inflation. Since movements in the unemployment rate may be the result of either demand or supply factors, looking at the level of the unemployment rate alone (or at the unemployment rate relative to some fixed value) can be misleading in particular episodes; instead, it is necessary first to determine the relative importance of aggregate-demand and -supply forces.

With this in mind, we consider what the model tells us about the conditions that prevailed in 1988 (the last year of our sample period). As Exhibit III indicates, aggregate demand was mildly stimulatory. The unemployment rate averaged 5.5 percent over the year. In the absence of any demand shocks, it would have averaged 5.8 percent. The difference between these two numbers (0.3 percent) provides a measure of aggregate-demand pressures in the economy. A measure of the net impact of supply shocks is obtained as the difference between what the unemployment rate would have been in the absence of demand shocks and what it would have been in the absence of shocks of any kind. Our model implies that in the absence of any shocks to the economy the unemployment rate would have settled at 6.0 percent.[22]

Thus, this difference between the actual 5.5 percent rate in 1988 and the 6.0 percent mean rate is accounted for about equally by demand and supply shocks. Although demand pressures *do* appear to have contributed to labor market tightness in 1988, the degree of pressure probably is not as intense as would be suggested by comparing the prevailing rate with its 6.0 percent mean.

APPENDIX A

Identification

In this appendix we describe the identification problem in terms of the moving average representation of a VAR. Let the vector $X_t = [x_{1t}\ x_{2t} \ldots x_{nt}]$ denote the variables contained in the model, where all the elements are nonstationary, but are not cointegrated. Assume that the structural representation of the model can be written as

$$Z_t = A\,(L)\,e_t \qquad \text{(A1)}$$

where $Z_t = \Delta X_t$, $A(L) = A_0 + A_1L + A_2L^2 + A_3L^3 \ldots$, and the lag operator L is defined by $Le_t = e_{t-1}$. Further, it is assumed that $\Sigma A_i^2 < \infty$, and that the structural disturbance term e is serially uncorrelated.

Let the estimated VAR representation of the model be given by

$$B\,(L)\,Z_t = v_t. \qquad \text{(A2)}$$

Multiplying both sides of (A2) by $C(L) = B(L)^{-1}$ leads to the moving average representation

$$Z_t = C\,(L)\,v_t. \qquad \text{(A3)}$$

where $E(v_t) = 0$, and $E(v_s v_t') = \Omega$ for $t = s$ and is zero otherwise. (A3) is the reduced form representation of (A1), and we have

$$A\,(L)\,e_t = C\,(L)\,v_t. \qquad \text{(A4)}$$

This is satisfied for any v_t such that $v_t = S^*e_t$, and $C(L) = A(L)^*S^{-1}$. Thus, to recover the structural representation from the estimated VAR, we need to obtain the matrix S which links the VAR residuals v_t with the structural disturbances e_t.

The exact form of S will depend upon the structure of the model. Under the usual assumption that the structural disturbances are uncorrelated with each other and that they have unit variance (that is, $E(e_t e_t') = I$), the problem of choosing the appropriate S reduces to choosing the elements of S subject to the condition that S is a square root of Ω (the variance-covariance matrix of the VAR residuals). Since Ω has $n(n+1)/2$ unique elements and S is n^*n, we need $n(n-1)/2$ (that is, $n^2 - [n(n+1)/2]$) additional restrictions in order to identify a unique S. If $n = 2$, for example, Ω contains three unique elements while S contains four. Thus, we need one additional restriction to identify S.

Sims (1980) suggested choosing S such that $s_{ij} = 0$, for all $j > i$, which serves to make the system exactly identified. For a two-variable system, this restriction prevents shocks to the second variable from having any contemporaneous effect on the first variable. Sims' restrictions imply that the underlying structural model is a recursive, simultaneous equations model (with independent error terms), a representation that may sometimes be difficult to reconcile with economic theory. Blanchard and Watson (1986) imposed restrictions on contemporaneous correlations that were explicitly derived from economic theory, and variations of this technique have been implemented by Bernanke (1986) and Walsh (1987), among others.

The technique of identification used in this paper has been suggested recently by Blanchard and Quah. In this technique the restrictions used to identify S can be interpreted as restrictions on the long-run effects of the associated shocks on certain variables. To see how this works, assume that the vector Z_t contains only two elements, so that (A3) becomes

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} c_{11}(L) & c_{12}(L) \\ c_{21}(L) & c_{22}(L) \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}$$

or

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} c_{11}(L) & c_{12}(L) \\ c_{21}(L) & c_{22}(L) \end{bmatrix} \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}^{-1} \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}$$

or

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} c_{11}(L)s_{11} + c_{12}(L)s_{21} & c_{11}(L)s_{12} + c_{12}(L)s_{22} \\ c_{21}(L)s_{11} + c_{22}(L)s_{21} & c_{21}(L)s_{12} + c_{22}(L)s_{22} \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

As discussed above, if e_1 and e_2 are assumed to be independent of each other, only one more restriction is needed to identify S. If it is assumed that e_2 has no long-run effect on x_1 (the first variable in the model), the restriction takes the form

$$c_{11}(1)\, s_{12} + c_{12}(1)\, s_{22} = 0$$

Here, $c_{11}(1)$ is just the sum of the coefficients in the lag polynomial $c_{11}(L)$. Thus, in this case identification is achieved by choosing an S for which a particular weighted sum of the elements of the second column of S is zero. The condition that these weights be the sum of the coefficients of the estimated lag polynomials for the first variable is what ensures that the level of x_1 is independent of e_2.

Shapiro and Watson (1988) show how this restriction can be imposed quite easily in the VAR representation.

APPENDIX B

Data and Preliminary Tests

We use quarterly data over the period 1948Q1–1988Q4 for our estimation.

All data have been obtained from the Citibase data tape. For population, we use noninstitutional population, 16 years and over, after subtracting armed forces. The unemployment rate is the civilian unemployment rate. To make the GNP data comparable, we use real GNP net of federal defense expenditures. The interest rate we use is the six-month commercial paper rate. Data for U.S. exports and imports are from the National Income and Product Accounts.

Tests for Stationarity

We tested for stationarity using the Said-Dickey test, which is recommended by Schwert (1987). The test involves estimating an equation of the form

$$y_t = \alpha + \beta y_{t-1} + \sum_{i=1}^{j} \delta_i \Delta y_{t-i} + e_t.$$

To test whether the y process contains a unit root we have to determine whether $\beta = 1$. However, under the null hypothesis that the process generating y contains a unit root, the ratio of the estimated value of β to its standard error does not have the usual t-distribution. The critical values to be used in this case are tabulated in Fuller (1976).

Schwert (1987) demonstrates that choosing a large value of j (as recommended by Said and Dickey) avoids the problem of falsely rejecting the hypothesis that y contains a unit root. In the table below we present the results for the cases where $j = 8$ and $j = 12$. The table shows that we are unable to reject the null hypothesis of a unit root for population, real GNP, the interest rate, or the foreign variable at the 10% level in either the eight-lag or the 12-lag case.

In the case of the unemployment rate, we present three different sets of results. We are unable to reject the null hypothesis of a unit root in the unemployment rate whether or not we allow for a linear trend. The last column shows the results for the case where we allow for a change in the mean unemployment rate beginning in 1974Q1. For the eight-lag case, the computed test statistic is significant at 5%, while for the 12-lag case the computed value of −2.80 is just below the 5% critical value of −2.89.

Note, however, that these critical values do not allow for a shift in the mean under the alternative hypothesis. It is useful to compare these critical values to those reported in Perron (1988). Perron generalizes the null of a unit root process to allow for a one-time change in the structure of the series, and compares this to the alternative of a stationary series with a discrete change in its mean. (Thus, his null hypothesis is not strictly the same as ours.) It turns out that the critical values vary with the date at which the break occurs. For the case at hand, where the break occurs about two-thirds of the way into the sample, the 5% critical value is −3.33, while the 10% critical value is −3.01.

Table B-1

Unit Root Tests

Sample period 48:1–88:4

					Unemployment Rate		
Variable[1]:	Population[2]	Real GNP[2]	Interest Rate	Foreign	Constant	Trend	Mean Shift
No. of Lags							
8	−2.60	−1.80	−2.09	−2.33	−1.79	−2.68	−3.13
12	−1.85	−1.52	−1.77	−2.54	−1.68	−1.85	−2.80

Notes:

[1]All variables are in logs. Except as otherwise noted, all regressions contain a constant only. The 5% critical value for this case is −2.89, the 10% is −2.58.

[2]Equations contain both a constant and a trend. The 5% critical value is −3.45, the 10% is −3.15.

NOTES

1. See, for example, "Records of Policy Actions of the Federal Open Market Committee," *Federal Reserve Press Release,* for the eight Federal Open Market Committee meetings in 1988.

2. See, for example, Ball, Mankiw, and Romer (1988), Long and Plosser (1983), Lucas (1973), and Greenwald and Stiglitz (1988).

3. Later applications used restrictions derived from economic theory. See Blanchard and Watson (1986), Bernanke (1986), Sims (1986), and Walsh (1987).

4. In another example of theoretical agnosticism, McCallum (1988) has investigated the robustness of nominal-income-targeting rules across different macroeconomic theories.

5. These results refer to the case where no trend is removed from the unemployment rate. Blanchard and Quah also present results for the case where they remove a linear trend from the unemployment rate. Removal of a linear trend tends to increase the relative importance of demand shocks.

6. They also included the price of oil as an exogenous variable, on the grounds that the two recessions during the 1970s were the consequence of the oil price shocks during this period. Inclusion of the oil price variable is problematic, however, since oil prices fell dramatically in 1985 without an obvious effect on real output. Shapiro and Watson estimate their model through the end of 1985 only.

7. SW also present two sets of results: one where there is a deterministic trend in labor hours and one where the trend in hours is stochastic. The results discussed in the text refer to the latter case.

8. See, for example, Blanchard and Quah (1989), Long and Plosser (1983), and Shapiro and Watson (1988).

9. The model outlined here closely follows that in Shapiro and Watson.

10. As described in Appendix A, once we assume that the underlying shocks are uncorrelated and have unit variance, we need $n(n\text{-}1)/2$ additional restrictions to identify a model that contains n variables. Since $n = 5$ here, we need a total of 10 restrictions.

11. This assumption implies symmetric treatment of foreign and domestic aggregate-demand shocks; that is, neither have permanent effects on output. However, the foreign shock also is designed to include the effects of foreign supply disturbances. One drawback of our model is that we are treating foreign and domestic supply shocks asymmetrically; that is, domestic supply shocks can have permanent effects, while foreign supply shocks cannot.

12. The assumption that an aggregate-demand shock induced by monetary policy does not have a long-run effect on the foreign variable is uncontroversial. The assumption that a fiscal-policy shock does not have a long-run effect on real exports and imports is less clear cut. See Krugman (1985) and Mussa (1985) for discussions of these issues and other references.

13. See Phelps (1978). For a contrary view, see Summers and Blanchard (1986).

14. See, for instance, Nelson and Plosser (1982).

15. See, for instance, Gordon (1982) and Congressional Budget Office (1987).

16. The reader interested in more detail is referred to Shapiro and Watson (1988).

17. As noted earlier, we also estimated a model with no mean shift in the unemployment rate, even though under this specification unit root tests suggest that the unemployment rate is non-stationary. The impulse response functions and variance decompositions for this model are nearly identical with those presented in the text, with one exception. The model without a mean shift in the unemployment rate ascribes a larger role to technology shocks and a smaller role to demand shocks in determining the error variance of the unemployment rate. Moreover (consistent with our findings in the unit root tests), the effects of different kinds of shocks on unemployment dissipate more slowly in the model without a mean shift than in the model discussed in the text.

18. The list of real shocks considered by Boschen and Mills (1988), for instance, contains changes in the price of oil, marginal tax rates, real government purchases, working-age population, and real exports.

19. See Jorgenson, et al., (1987).

20. More specifically, to obtain the supply component of the unemployment rate for any given quarter, we subtract the effect of all demand and foreign shocks that occurred as long as 40 quarters ago. The impulse response functions in Exhibit 1A show that this is more than long enough for the effects of these shocks to die out.

21. See Gordon (1982), Barro (1977), and Lucas (1973).

22. Prior to 1974, when we assume a mean shift, this rate is estimated to be 4.8 percent. Also, in the model where the mean is not allowed to shift, the mean rate of unemployment is estimated to be 5.0 percent.

REFERENCES

Ball, Laurence, N. Gregory Mankiw, and David Romer. "The New Keynesian Economics and the Output-Inflation Trade-off," *Brookings Papers on Economic Activity,* 1988:1.

Barro, Robert J. "Unanticipated Money, Output, and the Price Level in the United States," *Journal of Political Economy,* Vol. 86, 1977.

Bernanke, Ben S. "Alternative Explanations for the Money-Income Correlation," Carnegie-Rochester Conference Series on Public Policy, 25: 1986.

Blanchard, Olivier J. and Danny Quah. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *The American Economic Review,* September 1989.

Blanchard, Olivier J. and Mark W. Watson. "Are All Cycles Alike?", in Robert J Gordon, ed., *The American Business Cycle.* Chicago: University of Chicago Press, 1986.

Board of Governors of the Federal Reserve System. "Records of Policy Actions of the Federal Open Market Committee," *Federal Reserve Press Release,* for the eight Federal Open Market Committee meetings in 1988.

Boschen, John F., and Leonard O. Mills. "Tests of the Relation Between Money and Output in the Real Business Cycle Model," *Journal of Monetary Economics,* 22, 1988.

Congressional Budget Office. Congress of the United States, *The Economic and Budget Outlook: An Update.* August, 1987.

Evans, George W., "Output and Unemployment Dynamics in the United States: 1950–85," *Journal of Applied Econometrics,* Vol. 4, 1989.

Fuller, Wayne A. *Introduction to Statistical Time Series.* New York: John Wiley and Sons, 1976.

Gordon, Robert J. "Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment," in Martin N. Bailey, ed., *Workers, Jobs, and Inflation.* Washington D.C.: The Brookings Institution, 1982.

Greenwald, Bruce C. and Joseph E. Stiglitz. "Examining Alternative Macroeconomic Theories," *Brookings Papers on Economic Activity,* 1988:1.

Jorgenson, Dale W., Frank M. Gollop, and Barbara M. Fraumeni. *Productivity and U.S. Economic Growth.* Cambridge, Massachusetts: Harvard University Press: 1987.

King, Robert, Charles Plosser, James Stock, and Mark Watson. "Stochastic Trends and Macroeconomic Fluctuations," unpublished paper, University of Rochester, 1987.

Krugman, Paul R. "Is the Strong Dollar Sustainable?," in *The U.S. Dollar—Recent Developments, Outlook, and Policy Options,* Proceedings of a Symposium Sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 21–23, 1985.

Long, John B. and Charles I. Plosser. "Real Business Cycles," *Journal of Political Economy,* 91, 1983.

Lucas, Robert E. "Some International Evidence of Output-Inflation Tradeoffs," *American Economic Review,* 63, 1973.

McCallum, Bennett T. "Robustness Properties of a Rule for Monetary Policy," Carnegie-Rochester Conference Series on Public Policy, 29, 1988.

Mussa, Michael L. "Commentary on "Is the Strong Dollar Sustainable?," in *The U.S. Dollar—Recent Developments, Outlook, and Policy Options,* Proceedings of a Symposium Sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 21–23, 1985.

Nelson, Charles R. and Charles I. Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics,* 10, 1982.

Perron, Pierre. "Testing for a Unit Root in a Time Series With a Changing Mean," mimeo, Princeton University, 1988.

Phelps, Edmund S. *Microeconomic Foundations of Employment and Inflation.* New York: W. W. Norton, 1970.

Schwert, G. William. "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data," *Journal of Monetary Economics,* 20, 1987.

Shapiro, Matthew D. and Mark W. Watson. "Sources of Business Cycle Fluctuations," *NBER Macroeconomics Annual 1988,* Cambridge, Massachusetts: M.I.T. Press.

Sims, Christopher A. "Macroeconomics and Reality," *Econometrica,* 1980.

__________. "Are Policy Models Usable for Policy Analysis?," *Quarterly Review,* Federal Reserve Bank of Minneapolis, 1986.

Summers, Lawrence and Olivier Blanchard. "Hysteresis and European Unemployment," *NBER Macroeconomics Annual 1986,* Cambridge, Massachusetts: M.I.T. Press.

Walsh Carl E. "Monetary Targeting and Inflation: 1976–1984," *Economic Review,* Federal Reserve Bank of San Francisco, Winter, 1987.

Liquidity Constraints on Consumption: The Real Effects of "Real" Lending Policies

James A. Wilcox

Visiting Scholar, Federal Reserve Bank of San Francisco and Associate Professor, University of California, Berkeley.

I would like to thank Brian Bray, John Duca, Joe Peek, Charles Steindel, David Wilcox, and participants at the Spring 1989 Federal Reserve System Conference on Business Conditions for their views on an earlier version of this article. Especially helpful were the thoughtful and constructive comments of editorial committee members, Ramon Moreno, Bharat Trehan, and Carl Walsh.

This article argues that households are often prevented from consuming as much as their permanent income justifies. The hypothesis is advanced that lending criteria based on payment-to-income ratios often inappropriately constrain borrowing and therefore consumption. The evidence indicates that the variables presumed to proxy for payments and for income, the nominal interest rate and the unemployment rate, respectively, significantly affect consumption growth in the manner suggested by this hypothesis. In contrast, there is little evidence that real interest rates have important effects on consumption.

Personal consumption expenditures typically comprise about two-thirds of total national spending. Not only is consumption the single largest category of spending, but it changes by large amounts. In absolute terms, the variability of consumption expenditures is as large as that of business investment. Moreover, the variabilities of the components of consumption (services, nondurables, and durables) are large. The variabilities of nondurables and of services individually are nearly as large as that of consumer durable expenditures and the variability of the sum of nondurables and services is appreciably larger than that of durables.[1] An understanding of the movements in and determinants of consumption and its components clearly is important for the conduct of monetary policy.

The widely accepted permanent income hypothesis posits that consumption is driven by households' wealth and their expectations of income over the long run. Although actual income may fluctuate, these fluctuations are hypothesized not to affect consumption unless they alter households' expectations of their longer-run average, or permanent, income. Instead, when households are faced with deviations of actual from permanent income, they are presumed to vary borrowing and lending in order to steady consumption.

Considerable recent empirical research, however, based on both macroeconomic and microeconomic data bases, suggests that movements in actual income have sizeable effects on consumption apart from the effect of those movements on permanent income. Likewise, theory suggests that real, after-tax interest rates should affect consumption, but that nominal interest rates should not. The evidence, however, indicates that exactly the opposite is more likely to be true. Consistently, the data point to significant nominal interest rate effects on expenditures for durables, nondurables, and services, and to insignificant real-interest-rate effects.[2] This is surprising indeed.

This article suggests that a single factor helps to explain these two findings—that consumption expenditures are reduced both by higher nominal interest rates and by the shortfall of actual, current real income below its permanent level. These interest-rate and income effects both result from a borrowing constraint which prevents households from obtaining sufficient credit to finance as much consumption as their permanent income justifies. In particular, this borrowing, or liquidity, constraint is hypothesized to emanate from the prevailing lending practice of granting credit subject to virtually-never-changing payment-to-income-ratio ceilings.[3]

The first section presents a brief exposition of the permanent income theory of consumption and its empirical implications. Then it is argued that current credit-granting practices lead to liquidity constraints that are associated with nominal interest rates and real income. Section II reviews the evidence from microeconomic data bases regarding liquidity constraints. It also suggests what these constraints portend for macroeconomic data. Section III econometrically assesses the extent to which nominal interest rates and unemployment rates affect aggregate consumption, apart from their effects on permanent income. The estimates suggest that each has powerful effects on consumption expenditures. The evidence points toward the liquidity constraint that operates through the payment-to-income ratio as the source of these important effects on consumer expenditure. The short- and longer-run implications of these results for national spending and saving, for economic policy, and for financial institutions are discussed in the concluding Section IV.

I. Permanent Income, Consumption, and Constraints

The permanent income hypothesis posits that consumers base desired consumption on permanent income. The theory can be summarized as

$$C = kY^{P} \tag{1}$$

where C is the level of real, per capita consumption, k is the (average and marginal) propensity to consume out of permanent income and Y^{P} is the level of real, per capita permanent income. Permanent income, Y^{P}, is the average, discounted income a consumer expects to receive over the relevant horizon. In logs, (1) becomes

$$log(C) = log(k) + log(Y^{P}) \tag{2}$$

Note that actual income does not appear. The permanent income hypothesis states that actual income affects consumption only to the extent that if affects permanent income. Consumers are presumed to have access to capital markets, and therefore are not constrained by cash flow or current income. Borrowing and lending are viewed as shock absorbers for temporary fluctuations in income, making it possible for households to maintain consumption in the face of changes in actual income that are perceived to be temporary or are anticipated. Access to capital markets also permits consumption to change when permanent income changes in advance of actual income.

Permanent income represents a forecast, not a measured quantity. When consumers use all the information that is available at a given time to form estimates of permanent income, those estimates will change from period to period only as new information is received. Hence, changes in estimates of permanent income will be unpredictable. This is common to optimal forecasts; over time, the change in what is expected to happen over any given future period is random.[4] This is embodied in

$$log(Y^{P}) = log(Y^{P})_{-1} + \mu \tag{3}$$

which shows that today's forecast of future income differs from last period's forecast of income over the same future period by an unforecastable amount, μ. No information available prior to the current period would help predict the change in permanent (or forecasted) income. Otherwise, it already would have been incorporated in last period's estimate of permanent income.

Taking first differences of (2) and using (3) generates

$$\Delta log(C) = \Delta log(k) + \mu \tag{4}$$

On the assumption that k is constant over time, the growth rate of consumption, $\Delta log(C)$, should be random. No information available prior to the current period should reliably predict changes in consumption growth.

This model, however, is correct only if expected, real, after-tax interest rates are constant. When they are not constant, theory predicts that households defer more consumption when the reward for doing so is higher. This means that, ceteris paribus, higher interest rates last period reduce last-period's consumption relative to current consumption, thereby raising the growth rate of consumption:

$$\Delta log(C) = \gamma + \delta r^{e}_{-1} + \mu \tag{5}$$

Liquidity-Constrained Consumption

One potential weakness of this permanent-income formulation is that it assumes perfect capital markets, in which households can borrow and invest in order to smooth consumption across time periods. If capital markets are not perfect, however, households' desired spending patterns may be "liquidity constrained" in significant ways.[5] Thus, the presence of liquidity constraints could make the permanent income hypothesis an inferior explanation for aggregate consumption behavior, particularly since changes in borrowing flows are an empirically important factor in consumption patterns.[6]

The alternative hypothesis advanced here is that liquidity constraints are indeed binding for a significant portion of households and that the aggregate amount of liquidity constraint is associated with unemployment and nominal interest rates. As either rises, liquidity constraints both bind more tightly on previously constrained households and begin to bind on more households. Each aspect drives consumption further below the unconstrained value that the permanent income hypothesis predicts.

Consumers subject to liquidity constraints are assumed to behave differently from those who are not. Their consumption generally will respond vigorously to changes in current income or other sources of cash flow, even if those changes are anticipated, since, by definition, constrained households want to consume more but are prevented from doing so by restrictions on their ability to borrow. Consumers not subject to borrowing constraints, in contrast, would be expected to react little to anticipated, or to past income changes since their estimates of permanent income and their consumption plans already have adjusted to those developments. Liquidity-constrained consumers already may have changed their desired consumption, but actual consumption may have to await increases in actual income and the increased cash flow and ability to borrow that comes with it.

The liquidity constraint impinging on an individual may be relatively short-lived or very long-lived. Borrowing may be constrained to less-than-optimal levels for households with expected (average, discounted) lifetime earnings that are above their actual, current earnings for periods extending into years. Given the typical upward tilt in the age-earnings profile, most young households would be expected to be substantial net debtors for many years. Kotlikoff (1988) shows that in practice, however, there is very little net borrowing by the young, whose consumption tracks earnings very closely at least up to age 45. And it may be that not only the young are liquidity constrained. Wilcox (1989) cites several studies that suggest that a substantial portion of the elderly may be liquidity constrained.[7]

Liquidity constraints resulting from the combination of lending criteria based on current income and the usual upward slope of the age-earnings profile then may lead to a very important and relatively constant share of households whose spending is constrained. For an individual, this type of liquidity constraint may become less binding as actual earnings approach potential earnings and as financial assets are accumulated.[8] The number of households subject to this form of liquidity constraint may change as time passes, probably slowly and in tandem with the ratio of young to total households.

In addition to the households subject to this age-related constraint, a changing fraction of households is likely to experience varying degrees of liquidity constraint in response to variations in unemployment and nominal interest rates.[9] One reason consumers may become liquidity constrained is that lenders widely follow a practice of restricting consumer borrowing so as to keep payment-to-income ratios below some ceiling level. A recent American Bankers Association textbook on consumer lending suggests that a borrower's capacity to repay a loan can be measured by the payment-to-income ratio.[10] This means that applications for credit are likely to be disapproved if the ratio of total loan payments to income breaches a ceiling, for example, of 40 percent.[11] Note that, in practice, this policy refers to *current* payment-to-*current* income. One consequence of this policy is that lenders generally refuse to extend credit to the currently unemployed.[12]

Using a payment-to-income rule means that movements in current income relative to permanent income can lead to consumers being liquidity constrained. This practice suggests a capital market imperfection that may explain why current income and cash flow affect consumption when the permanent income theory suggests they should not.

These credit practices also suggest that the extent to which consumption is liquidity constrained will vary with nominal interest rates. Lending policies that predetermine a payment-to-income ratio ceiling reduce the *real* amount of credit that would be made available to a borrower as the *nominal* interest rate rises.

Consider a $10,000, 48-month fully-amortizing loan. Suppose that the real interest rate is six percent.[13] If the expected inflation rate built into interest rates were zero percent, the resulting six percent loan would entail monthly payments of $235.[14] Since the actual and expected inflation rate is zero for this period, the actual payments and their real, or inflation-adjusted, values are both $235 per month for 48 months. The lower horizontal line in Figure 1 shows that payment, which is level in dollar, or nominal, terms and in real terms.

Suppose now that the actual inflation rate and the expected inflation rate incorporated into nominal interest rates rises from zero to five percent and that the real interest rate remains at six percent. The resulting eleven percent market interest rate means that the same loan now carries a $258 payment, an increase of 10 percent. This higher, nominal amount is shown as the upper horizontal line in Figure 1. This repayment pattern has the same six percent real return over the life of the contract. The real burden of those higher nominal payments is shown as the diagonal line in Figure 1. Note that relative to the constant real payments in the zero-percent inflation case, the level dollar payments in the five-percent inflation case, in real terms, are higher early on and lower later in the life of the loan.

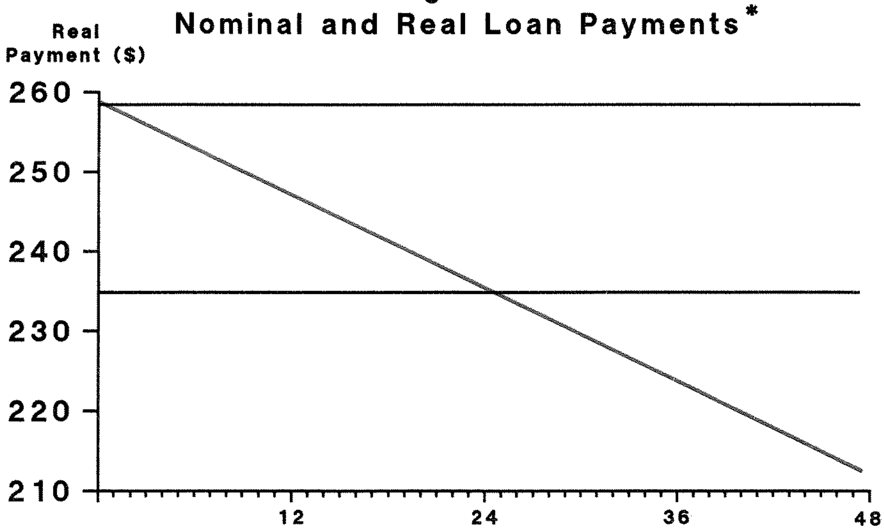
Level nominal payments during a period of inflation imply falling payments in real terms over time. An increase in expected inflation leads to an immediate one-time increase to higher nominal payments. The problem is that the onset of inflation does not have the same effect on household incomes. Incomes generally rise gradually as the level of prices rises. They do not jump by the same 10 percent in the first month that payments do. Thus, an increase in nominal interest rates stemming from an increase in the inflation premium would raise the initial real burden of this loan by virtually 10 percent. Suppose that a lender followed a payment-to-income rule and that the borrower would be permitted in either inflation scenario to borrow an amount that would imply a $235 payment. In the zero percent inflation case, a loan of $10,000 would be granted; in the five percent inflation case, a loan of only $9,087 would be granted.

This happens even though households and financial institutions both may think that they are adjusting for inflation. By basing their decisions on the ratio of two (nominal or real) flows, which often leads to an inflation-adjusted magnitude, they may be attempting to make a "real" decision. They are not. The reason is straightforward: the dollar payment per dollar of credit extended rises with the nominal interest rate. The only type of loan repayment schedule currently available provides for level dollar repayments. As time passes and inflation raises the level of wages and prices, those level payments constitute falling real payments.[15] Since later nominal repayments will be less in real terms, earlier ones must be greater to preserve the same average real payment and real rate of interest.

Over time, lenders might be expected to make lending policy parameters "realistic" to maintain optimal real borrowing limits. In practice, adjustments in lending policy take place so slowly that the aggregate amount of consumption that is liquidity constrained is likely to rise with the level of inflation. To the extent that nominal interest rates respond to (expected) inflation, payment-to-income ceilings would have to rise and fall with inflation to avoid tightening of liquidity constraints. It does appear that, on average, consumer credit parameters may have become somewhat looser in higher inflation periods. It does not appear, however, that they became tighter as inflation fell over the past ten years. In any event, lending parameters seem to be adjusted slowly enough, if ever, that as nominal interest rates move in response to inflation, more households become subject to these interest-rate-related restrictions.

The extension of loan maturities may reflect an attempt to overcome the high, initial, real payments brought on by inflation-related increases in nominal interest rates. The evidence does seem to be that loan maturities have consistently lengthened over the past four decades, but that seems to have gone on apart from the rise and fall of inflation. Regardless, attempting to solve the real-payment-tilt problem with longer maturities is indirect and inefficient since longer loans have lower, but still level, dollar payments.

Figure 1
Nominal and Real Loan Payments*



*Based on $10,000, 48-month, fully-amortizing loan with a 6% real interest rate.

II. Micro Evidence

One potential way to measure the extent of aggregate liquidity constraint is to look at data on loan applications disapproved.[16] However, the results of such a study will almost certainly understate the degree of liquidity constraint since borrowing constraints related to actual income may not be imposed only by lenders but also (self-imposed) by households. To the extent they recognize that lenders impose payment-to-income restrictions, households are likely to adjust current borrowing behavior as a hedge against temporary declines in actual income in the future. This may explain why, for example, so many households purchase lines of credit through credit card fees. Such lines may provide access to the credit that lenders otherwise would refuse to extend under circumstances that would lead a borrower to seek it. By tempering their debt accumulation and by maintaining lines of credit, households effectively have credit-access insurance.

Likewise, households may not bother applying for credit when they (accurately) forecast that they do not fall under lenders' payment-to-income ceilings. Most households are likely to believe (accurately) that their prospects for obtaining additional credit are dim when they are unemployed, for example, and therefore may not even apply. Finally, even though the constrained may end up obtaining credit, the amount they borrow will be less than if they were not constrained. In the mortgage market, for example, it is common for potential borrowers to seek prior estimates from lenders directly or indirectly of the maximum mortgage they would qualify for and then to adjust home purchases accordingly.

As an alternative approach to assessing the extent to which consumers are liquidity constrained, a number of empirical studies have investigated the consumption behavior of individual households. The important question for our purposes is whether liquidity-constrained consumers are numerous enough or receive a large enough share of aggregate income that aggregate consumption patterns are importantly affected. Below, I briefly review the results of studies of the spending patterns of individuals with that in mind.

Hall and Mishkin (1982) used data from the early 1970s on income and consumption of food by individual families. They conclude that 80 percent of these actual consumption expenditures appear to move in the manner prescribed by the liquidity-*unconstrained,* permanent income hypothesis. Their results further suggest that much of the deviation of actual from unconstrained consumption is due to the inability to borrow to overcome temporary income shortfalls. Hall and Mishkin conclude that "food consumption behaves as if constraints on borrowing were relatively unimportant."

Hayashi (1985) analyzed consumption behavior of two groups of consumers in the early 1960s: those who had high savings and those who did not. The hypothesis was that consumers who had accumulated wealth were unlikely to be constrained in their consumption behavior since they could self-finance more consumption by saving less, or even by dissaving. In contrast, those who did not have high savings were more likely to find their consumption restricted if liquidity constraints did, in fact, exist.

When the empirical model that best tracked the consumption of the high savers was used to predict the consumption of the low savers, Hayashi found that low savers tended to spend *less* (and save more) than the model forecasted. The rationale given for that result is that liquidity constraints prevented the group with low accumulated savings from consuming as much as they otherwise would have chosen; their inability to borrow precluded spending. The group whose consumption seemed most constrained was young households. That result is consistent with the typical upward tilt in age-earnings and in age-wealth profiles.

Mariger (1986) also concluded "that liquidity constraints are quite prevalent." Again employing the early 1960s data set, his estimates suggest that about twenty percent of families were liquidity constrained. These families accounted for about one-sixth of aggregate consumption.

Zeldes (1988) uses food consumption data collected from the late 1960s through the early 1980s to assess whether binding liquidity constraints have been widespread. Like Hayashi, he splits the individual-family data set in two. The specific criterion is whether the family has a non-negligible wealth-to-income ratio. He then estimates whether either group seems to exhibit consumption behavior that is consistent with the presence of borrowing constraints. He finds that, as Fitzgerald and Hemingway first conjectured, the rich are different.[17] Their consumption displayed no indication of being constrained by an inability to borrow, whereas that of the group with the low wealth-to-income ratio did.

These studies each find that a minority of households has been influenced by either binding current, or potentially-binding future, liquidity constraints. Taken together, these studies seem to make a compelling case for the practical importance of liquidity constraints. First, the fraction of households deemed to be constrained was a substantial minority. Finding that about 20 percent of consumers had been constrained implies serious deviation from the unconstrained model of aggregate consumption. Since the unconstrained consumers are able optimally to smooth their consumption, it may be that the remaining consumers account for a very large fraction of total consumption variability.[18]

Second, of all the items in the household budget, expenditure for food would seem to be one of the *least* likely to be liquidity constrained. Food consumption is usually considered very income *in*elastic; it is the last purchase to be sacrificed when income falls. Evidence that the food expenditures of twenty percent of all households are constrained suggests that expenditures on the remaining categories of expenditure may be vastly more affected.[19]

III. Macro Evidence

Given the evidence from studies of the spending patterns of individual households, this section investigates whether liquidity constraints have important effects on aggregate consumption behavior. According to the liquidity-unconstrained version of the permanent income hypothesis of consumption, no lagged values of any variable should help predict the growth rate of consumption. Table 1 provides evidence to the contrary. It presents the results of regressing consumption on the first four lags of (personal disposable) income. Rows 1, 2, and 3 present the results of using lagged income to predict total consumption expenditures (C), the sum of consumer expenditures on nondurables and on services (CNS), and consumer expenditures on durables (CD), respectively. All variables in Table 1 are expressed as real, per capita, seasonally-adjusted percentage changes at annual rates. Current growth rates are based on current relative to prior-quarter levels.

The first two lagged income coefficients tend to be sizeable and significant. Lags three and four tend to be smaller and negative. Not surprisingly, the reaction of durable goods expenditures to lagged income is considerably different from that of nondurables and services. For nondurables and services, it is reasonable to take both the size and the timing of the consumption services that flow from them to be the same as expenditures on them. When it comes to durables, such an assumption is patently unreasonable.[20] A $20,000 expenditure this quarter for a new automobile is almost entirely investment and little consumption. The continuing flow of consumption services from past durable goods purchases, therefore, suggests that a positive response of durable goods expenditures to income is likely to be followed by negative ones. That is what Row 3 shows. After large and significant positive responses to the first two lags of income, large, negative responses appear.

The F-statistics in each row test whether the lagged values of income significantly help predict consumption. Since each of the calculated F-statistics exceeds the .05 significance level (critical value of 2.37), I conclude that each of the measures of consumption is predicted by lagged, actual-income movements.[21] This predictability of consumption growth argues against the simplest version of the permanent income theory as a sufficient explanation for consumption.

Table 1

Lagged Income Predicts Consumption

Quarterly Data, 1949Q3–1988Q3
(t-statistics in parentheses)

Dependent Variable	Constant	Coefficient on Income Lagged 1	2	3	4	F	R^2	D.W.
1. Total Consumption Expenditures	1.51 (4.37)	0.200 (3.30)	0.147 (2.50)	−0.054 (−0.92)	−0.023 (−0.39)	4.46	.105	2.11
2. Expenditures on Nondurables and Services	1.49 (6.14)	0.132 (3.09)	0.042 (1.01)	−0.004 (−0.09)	0.024 (0.59)	2.87	.070	1.88
3. Expenditures on Durables	1.51 (0.85)	0.672 (2.16)	0.939 (3.13)	−0.411 (−1.37)	−0.352 (−1.16)	4.29	.102	2.33

The results in Table 1, however, do not point to the reasons the permanent income hypothesis might be violated. The results presented in Table 2 suggest that lagged income and lagged consumption appear to affect consumption to the extent that each serves as a proxy for current income. Table 2 shows the results of arbitrarily splitting the sample period into decades and testing whether lagged values of consumption and income help predict consumption and income. (Reported F-statistics above 2.65 in Table 2 indicate statistical confidence above the .95 level.) When current income is predicted by its own lags, as in the 1950s and, to a lesser extent, the 1960s, lagged income serves as an effective proxy for current income. Those are also the periods when consumption is predicted by lagged income.[22] When income is not predicted by past income, as in the 1970s and 1980s, lagged income does not serve as an effective proxy for current income. In those periods, consumption is not predicted by past income.

Similar findings pertain to lagged consumption. When actual income is predicted by, and therefore effectively proxied by, lagged consumption, as in the 1980s, consumption is significantly related to its own lags. When lagged consumption does not serve as an effective proxy for current income, lagged consumption does not significantly predict current consumption. The exception to this pattern is that, in the 1970s, when income is not predicted by lags of consumption, lagged consumption still helps predict consumption.[23] These results indicate that it is not lagged income or lagged consumption *per se* that affects consumption. Instead, what they suggest is that consumption reacts to current income to an extent greater than is warranted by the permanent income hypothesis.

Table 2

Income and Consumption Predictability over Various Sample Periods

	Dependent Variable	Regressors (lags of)	F-Statistic 1950s	1960s	1970s	1980s
1.	Consumption	Income	4.50	4.63	1.39	1.85
2.	Income	Income	3.21	2.18	0.78	0.71
3.	Consumption	Consumption	1.17	0.46	4.06	2.34
4.	Income	Consumption	1.56	1.06	1.74	3.31

Testing for Liquidity Constraints

The finding that consumption does not correspond to the predictions of the simplest version of the permanent income hypothesis does not necessarily invalidate it. As noted earlier, consumption growth also may be affected by changes in the real interest rate and be completely in accord with the permanent income hypothesis. The apparent violation reported above may reflect that effect. However, as I argue, an alternative explanation for the results in Tables 1 and 2 is that consumption is importantly affected by liquidity constraints associated with payment-to-income ceilings on borrowing.

This hypothesis implies that changes in the *nominal* interest rate (i), as distinguished from changes in the real interest rate, and changes in the unemployment rate (U), which proxy for changes in current income, affect the growth rate of aggregate consumption expenditures on nondurables and services, apart from their effects on permanent income.[24] Suppose that the propensity to consume out of permanent income apart from the liquidity constraints associated with i and U is k^*, that the actual propensity to consume is k, and that the two are related by:[25]

$$k = k^* e^{(\alpha U + \beta i)} \tag{6}$$

On the assumption that α and β are negative, equation (6) embodies the hypothesis that, as either the unemployment rate or the nominal interest rate rises, consumption is reduced relative to the level implied by permanent income. The reason is that as either rises, liquidity constraints will bind on more households and more consumption per household. Taking logarithms of equation (6), then substituting into equation (1), and allowing for the real interest rate effects discussed earlier produces:

$$\Delta log(C) = \gamma + \delta r^e_{-1} + \alpha \Delta U + \beta \Delta i + \mu \tag{7}$$

Table 3 shows the results of estimating equation (7) and some variations of it. To ensure consistent estimates and valid statistical inference of the effect of each of these right-hand-side variables on consumption growth, an instrumental variables estimation technique was used.[26] This approach addresses two problems. First, the current changes in unemployment and interest rates (and presumably in almost all other macroeconomic variables) are very likely to be correlated with current revisions to permanent income and are, therefore, correlated with the equation's error term.

Second, even if households are continuously obeying the permanent income hypothesis, quarterly-average measures of consumption growth will be correlated with the previous period's consumption growth, which is determined by the previous period's new information. This apparent violation of the permanent income hypothesis emanates from time-averaging of data, not from a violation of the theory. The induced autocorrelation implies that, when time-averaged data are employed (as they are here), the consumption equation error term will be correlated with one-period-lagged variables, such as the one-period-lagged expected real interest rate that equation (7) suggests is relevant. Thus, one-period-lagged variables are not valid instruments.

The variables used as instruments are a constant term and lags two through five of the first differences of the unemployment rate, the expected inflation rate, the auto loan rate, the Treasury bill rate, and the real, after-tax Treasury bill rate.[27] Expected real interest rates were derived by subtracting the one-year expected inflation rate from the nominal interest rate.[28]

The top and bottom halves of Table 3 use as the consumption measure the sum of expenditures on consumer nondurables and services (CNS) and expenditures on consumer durables (CD), respectively. Row 1 presents results from estimating equation (5), the modified (unconstrained) permanent income hypothesis, using the real, after-tax Treasury bill interest rate as the measure of the lending rate that consumers face.[29] As can be seen from these results, the real interest rate does not have a significant effect on consumption. Row 2 adds the two presumed determinants of the degree of economy-wide liquidity constraint: the unemployment rate and an interest rate at which consumers can borrow. The borrowing rate that households face is taken to be the nominal, before-tax, interest rate on auto loans. Row 3 includes both the borrowing and the lending interest rates households face. The lending rate is taken to be the Treasury bill rate. The borrowing and lending rates that consumers face did not always move closely over this period, in part due to

Table 3

Effects of Expected Real Interest Rates, Unemployment Rates, and Nominal Interest Rates on Consumption

1949Q3–1988Q3

Estimation Method: Instrumental Variables

(t-statistics in parentheses)

		Coefficient on							
			First Difference of Current:						
Dependent Variable	Constant	Level of Lagged Expected Real After-Tax Lending Rate	Unemployment Rate	Nominal Borrowing Interest Rate	Nominal Lending Interest Rate	R^2	S.E.E.	D.W.
1. Expenditures on Nondurables and Services	1.96 (9.92)	−0.099 (−0.81)	—	—	—	0.010	2.25	1.71
2. Expenditures on Nondurables and Services	2.05 (10.08)	−0.120 (−0.87)	−1.78 (−2.70)	−2.12 (−2.58)	—	0.115	2.14	1.99
3. Expenditures on Nondurables and Services	2.03 (9.32)	−0.089 (−0.55)	−1.73 (−2.58)	−2.10 (−2.54)	0.130 (0.36)	0.119	2.14	2.00
4. Expenditures on Nondurables and Services	1.96 (11.48)	—	−1.99 (−3.31)	−1.81 (−2.48)	—	0.132	2.11	2.00
5. Expenditures on Durables	2.76 (1.87)	0.786 (0.86)	—	—	—	−0.002	16.76	2.18
6. Expenditures on Durables	3.51 (2.47)	0.478 (0.50)	−10.00 (−2.17)	−15.00 (−2.61)	—	0.212	14.95	2.40
7. Expenditures on Durables	3.70 (2.41)	0.262 (0.23)	−10.32 (−2.18)	−15.20 (−2.61)	−0.904 (−0.36)	0.206	15.11	2.38
8. Expenditures on Durables	3.38 (3.19)	—	−9.14 (−2.13)	−16.24 (−3.12)	—	0.196	15.06	2.37

regulations, in part due to maturity differences. The estimates show that the borrowing rate clearly is the dominant financial force.[30] That the entire interest-rate impact takes place through the nominal borrowing rate is just what a payment-to-income constraint leads us to expect.

Consistently, the estimated effect of (lagged, expected) real interest rates on nondurables and services expenditure is negative, small, and statistically indistinguishable from zero. This coincides with most earlier research and holds regardless whether liquidity constraint proxies are included. By contrast, the liquidity constraint proxies are large and significant.[31]

Rows 5–8 use expenditures on durables as the dependent variable. These results are qualitatively similar to those for nondurables and services. The unemployment rate and interest rate coefficients are large and clearly statistically significant in each row. Again, the borrowing rate dominates the lending rate. The real-rate effects are positive and much larger than for nondurables and services, but never approach statistical significance.

The estimates in Table 3 also embody some equalities that support the payment-to-income-constraint hypothesis. Since the coefficients for durables are about five times as large as those for nondurable and services spending and since the level of the latter category is about five times as large as that of the former, the estimated reduction in spending due to a rise in either the unemployment rate or the nominal interest rate is about the same for both spending categories.

Second, equality of the interest-rate- and unemployment-rate-spending elasticities cannot be rejected. Tests of that hypothesis for expenditures on nondurables and services and for expenditures on durables generate t-statistics of -0.8 and -1.8, respectively, each of which is below the critical value for a .05 significance test.[32] These results are consistent with the hypothesis concerning payment-to-income constraints since equal percentage changes in the interest rate and in the unemployment rate should have (approximately) the same effect on the payment-to-income ratio and therefore on the degree of liquidity constraint.

Third, when Rows 4 and 8 of the table are re-estimated using the two individual components of the nominal interest rate, the expected real rate and the expected inflation rate, the estimated coefficients on the components are close to that on their sum. Statistical tests of this equality within each spending category do not call for rejection; the respective t-statistics were -0.6 and 1.7, considerably below the critical value for confidence at the .95 level. Thus, it seems that it is not the expected inflation rate nor the real rate component of the nominal interest rate, but the nominal interest rate *in toto*, that affects consumer spending. This finding is consistent with the hypothesis advanced here that the nominal interest rate affects payments and through this channel, affects spending.

Equation (6) can be rewritten as

$$k/k^* = e^{(\alpha U + \beta i)} = e^{\alpha U}e^{\beta i} = k_U\, k_i \tag{8}$$

The heavy line in Figure 2 plots the constraint for nondur-

Figure 2
Constraints on Consumption of Nondurables and Services

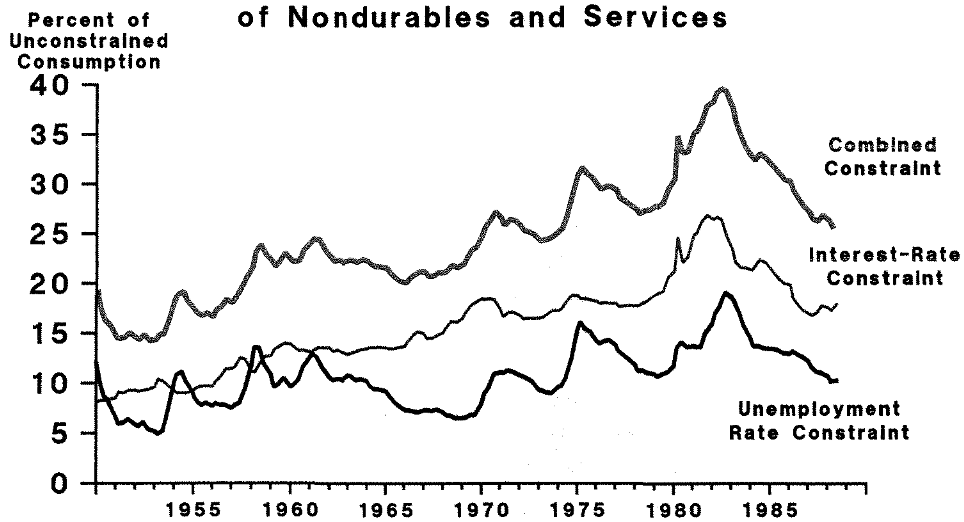
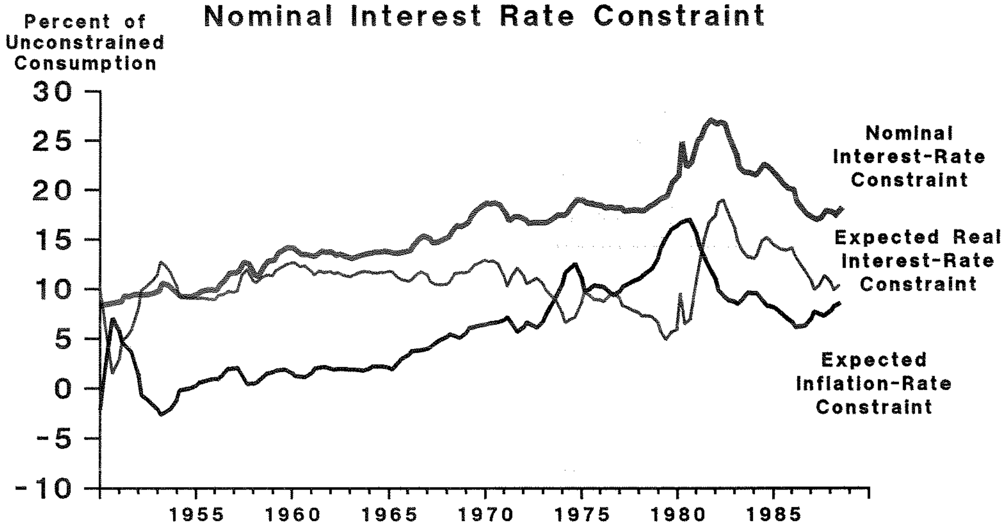


Figure 3
Decomposition of the Nominal Interest Rate Constraint



ables and services spending implied by the estimates in Row 4 of Table 3. This constraint shows the joint effects on the propensity to consume out of permanent income over time due to changes in unemployment and interest rates. In the 1970s, this constraint rose and then declined briskly after 1980. Apparently, the effects of rising nominal interest rates have offset those of falling unemployment rates over the last few years, stalling the decline in the constraint.

The separate effects of unemployment and interest rates are also plotted in Figure 2. The estimates suggest that the constraint related to the nominal interest rate generally has had a much larger effect on spending than that related to the unemployment rate. The reason for this may be that many more households are affected by payment changes than by income changes.

Equation (9) separates k_i into its components, the effects of changes in expected real rates, k_r, and of changes in expected inflation, k_p.

$$k_i = e^{\beta i} = e^{\beta (r+p)} = e^{\beta r} e^{\beta p} = k_r k_p \tag{9}$$

All three are plotted in Figure 3. Figure 3 reflects the dominant role of expected inflation in interest rate movements and consequently in changes in the constraint. The secular rise of expected inflation until the 1980s is estimated alone to have increased the constraint by about 15 percentage points. Since the early 1980s, lower expected inflation and lower real rates have combined to reduce liquidity constraints operating through nominal interest rates, thereby freeing consumers to spend more.

IV. Interpretations and Implications

The evidence presented here points to large and reliable responses of aggregate consumption expenditures to changes in unemployment and in nominal interest rates. This is consistent with the view that large numbers of households find themselves liquidity constrained. I argue that this constraint emanates from households being constrained in their ability to borrow as a result of lenders' payment-to-income restrictions.

These borrowing restrictions become more binding as nominal interest rates rise. Given that the major factor driving nominal rates has been expected inflation, households become increasingly liquidity constrained as expected inflation rate rises. This prevents households from carrying out their utility-maximizing consumption. This is an example of a large, real cost of expected inflation, one that may help account for households' generally inexplicable antipathy to *expected* inflation.

The presence of, and even the possibility of future, liquidity constraints means that aggregate spending will behave very differently than is predicted by models that ignore such constraints. The impacts of fiscal and monetary policies, short-run or long-run, will depend upon the amount and nature of liquidity constraints faced by various households. Since the degree of liquidity constraint is likely to differ systematically by group, assessment of these policies should allow for differential effects on the young and the old, the rich and the poor, the homeowner and the renter. Without allowance for these distributional considerations, the workings of the economy and the effects of policies are likely to be misunderstood.

In the absence of liquidity constraints, changes in fiscal policy that the public perceives to be temporary would not be expected to affect consumption much, especially that of the young. In the presence of liquidity constraints, however, the change in cash flows can greatly alter households' ability to achieve desired consumption, especially that of the young. A temporary income tax reduction, for example, might change permanent income very little, but might raise consumption virtually dollar-for-dollar, as consumers found themselves less bound by liquidity constraints.[33] The widespread presence of liquidity constraints also affects the effectiveness of monetary policy, since nominal interest rate changes have large effects on expenditures. In fact, in contrast to much recent theorizing, anticipated monetary policy may have larger effects than unanticipated policy, since it is the anticipated part of inflation that primarily affects nominal interest rates.

The substantial response of aggregate consumption to nominal-interest- and expected-inflation rates suggests that a large number of households are genuinely bound by the cash-flow constraint associated with those factors. This paper has not addressed the reasons that such constraints persist. To the extent they are associated with institutional and psychological factors that can be overcome, the financial services industry can satisfy a heretofore-constrained, enormous household demand for credit. The easiest way to tap this unmet demand is to introduce financial instruments the payments of which are geared to the upward tilt in household income associated with long-run aggregate productivity increases, long-run individual real income increases, and, especially, increases in the average level of prices.

NOTES

1. The variabilities referred to are the standard errors of the estimate generated by separately regressing the ratio of each spending component (in real terms) to detrended real GNP on a linear trend with the 1947Q2–1988Q3 quarterly data. Detrended real GNP consists of the exponentiated fitted values obtained by regressing the log of real GNP on a constant, a linear trend, and the square of the linear trend.

2. For example, a nominal interest rate variable is significant when added to the FRBSF econometric model's current equation for consumer expenditures on nondurables and services. The durables expenditure equation already has a nominal rate variable included, in order to handle credit rationing effects. For a typical example of significant estimated nominal interest rate effects on aggregate consumption, see Blinder and Deaton (1985).

3. Consumers and households are referred to interchangeably, as are liquidity or borrowing or financing constraints. I do not attempt to ascertain why lenders have chosen their lending criteria or why they so seldom change them.

4. Of course, the optimal forecast may be that actual income will change. During a period of perceived-to-be-temporary unemployment, the optimal forecast will be that income will rise on average over time. The optimal forecast of what average income will be over a fixed period will be revised as new information arrives, but the direction and size of those revisions will not be predictable. Optimal forecasts have the property that no one, including the forecaster, can forecast how the forecast will change.

5. Liquidity constraints may affect not only households. Econometric estimates of business investment spending long have found substantial effects of various financial variables which can be interpreted as indicating whether firms are likely to face liquidity constraints. Recent microeconomic evidence concurs with the earlier aggregate estimates that support the significant role played by these constraints on business borrowing. Given the size of the economic units involved and the relatively more collaterizable nature of the assets being financed by businesses, it is easy to imagine that individual households also might be subject to financing constraints.

6. Household expenditures are financed with some combination of current income, changes in gross household debt, and changes in gross household savings (assets). To illustrate the typical financing pattern, three regressions were performed. In each, the change in personal consumption expenditures was regressed on the change in one of the methods of financing. The data were monthly, current-dollar, seasonally adjusted at an annual rate, and covered the period 1975:03–1988:10. Each regression contained one regressor but no constant term. Income was calculated as the sum of the consumption and net saving measures. Consumption was taken to be total personal consumption expenditures. The flow of net saving was taken to be personal saving. The flow of debt was taken to be the net (extensions minus repayments) change in consumer installment credit.

The results show that an additional dollar of consumption typically is financed with an estimated \$0.62 of additional income, \$0.13 of reduced (gross) saving, and \$0.25 of additional debt. Although aggregate household assets exceed household liabilities, the distribution of financial assets is very skewed, with most households owning very few. One consequence is that, per dollar change in consumer spending, the change in the flow of credit is about twice as large as that of (gross) saving. Thus, changing credit flows are an integral part of changes in consumption. These estimates also hint that the borrowing, rather than the lending, rate that households face may be more relevant for household spending.

7. Applying lending criteria based on current income to recipients of social security income, which is tied to the CPI by statute, seems especially perplexing.

8. Credit-granting processes that do not use age as a determining factor then may inadvertently discriminate against the young.

9. Tobin and Dolde (1971) discuss some types of liquidity constraints and, through simulations, assess their effects on consumption. Walsh (1986) models the fraction of consumers whose consumption is liquidity constrained due to an inability borrow against future income. The fraction is affected by the level of wealth and also fluctuates with actual aggregate income. He concludes that it is inappropriate in such circumstances to treat aggregate consumption as being the outcome of fixed shares of constrained and unconstrained households.

10. See Beares (1987). One example in Beares (1987) shows how the average maturity of a consumer's loans can be lengthened in order to reduce the payment-to-income ratio and thereby enable the lender to extend credit to a loan applicant that it would otherwise turn down. This does not mean that a payment-to-income ceiling is the sole criterion. Various consumer lending textbooks refer to various criteria, e.g., the six "c's" of credit. The literature and discussions with consumer lending officials do suggest that lending policies set ceilings on payment-to-income ratios, and rely much less on debt-to-income ratios. In spite of that, it is apparently common in this industry and its literature to refer to debt-to-income ratios when meaning payment-to-income ratios.

11. Such rules may have developed from either households' or financial institutions' fear of default. Households' recognition that they may be subject to future constraints may also influence their current behavior.

12. The unemployed referred to here are those who would not generally be deemed to be guaranteed re-employment. We mean to exclude the seasonally unemployed and those on definitely temporary layoff, for example, but it is not apparent that much is made of this distinction in the credit process.

13. Suppose, for the sake of this example, that income taxes are irrelevant.

14. In this example all dollar amounts have been rounded to the nearest dollar. Calculations were based on unrounded amounts.

15. It is commonly, and mistakenly, thought that adjustable rate loans cure this problem. They do not. Such loans allow payments to vary with the interest rate, and therefore indirectly with the inflation rate. In a period of steady inflation, they do not, however, imply constant real payments over the repayment period. Similarly, graduated payment mortgages provide for payments that are lower initially and rise for a short period. The rate of increase of those payments is not tied to the inflation rate. After the initial period, payments typically are constant in nominal terms for the remainder of the term of the mortgage, and therefore likewise fall in real terms if there is inflation.

16. The American Bankers Association's annual *Retail Bank Credit Report* for the years 1979, 1980, and 1981 (only) does contain a table headed "Reasons why borrowers may fail to qualify for financing during (that year)." The candidate reasons were inadequate income, insufficient equity, inadequate income management, and other. For large banks, the percentage judged to have "inadequate income" rose over those years from 31 to 43 to 47 percent.

17. F. Scott Fitzgerald: "You know, Ernest, the rich are different from us."

Ernest Hemingway: "Yes, they have money." (attributed)

18. Optimal smoothing is not the same as total smoothing. Since permanent income varies, optimal consumption will too.

19. Hall and Mishkin state that their results for food imply nothing about the behavior of other expenditures.

20. It is, however, reasonable to question how durable much of what is classified in the national income accounts as services and nondurables is.

21. Table 1 shows the ability of lagged income to predict consumption. That contrasts with the results presented by Hall (1978), whose sample period ended with 1977Q1 data. When my sample period is terminated at 1977Q1, I get results much like his (F-statistic = 2.19). Re-estimating with a sample that is just one year longer, however, implies rejection of the unpredictability hypothesis. Both longer and shorter samples generally lead to rejection of the unpredictability hypothesis. Variables other than income also may lead to that result.

22. From now on, unless otherwise noted, consumption refers to the sum of nondurables and services, expenditures on which correspond most closely to the flow of consumption services. Total consumption service flow would also include the flow of services from the outstanding stock of consumer durables. Since our objective here is to explain expenditures, we consider expenditures on only these two consumption categories.

23. This hints that there are factors other than changes in current income that are associated with the violation of the permanent income hypothesis.

24. The unemployment rate is used here to proxy for aggregate income. Flavin (1985) concludes that the unemployment rate is superior empirically to a measure of the deviation of actual income from its permanent value in accounting for consumption.

25. Even if there were no interest- and unemployment-rate-related liquidity constraints, those arising from the age-earnings profile might exist. In what follows, I take k^* to be equal to one.

26. For a more complete discussion of these issues, see Hall (1988) and Campbell and Mankiw (1987).

27. Specifically, I used the investment yield to maturity on a three-month Treasury bill. The finance rate on 48 month

automobile loans comes from DRI. Until 1984, I used the average marginal personal income tax rate on interest income from Peek and Wilcox (1987). The rates for 1984, 1985, and 1986 have been calculated in the same manner. The rates for 1987 and 1988 are assumed to be 27 and 25 percent, respectively. For each quarter I used the corresponding annual rate. The tax rate for the upcoming quarter is applied to the interest rate for the current quarter.

28. The expected inflation data are taken from the Livingston survey, which records expectations each June and December. I used those values for the second and fourth quarter observations, respectively. First and third quarter observations were obtained by interpolating between second and fourth quarter observations.

29. Using lags of the explanatory variables required shortening the sample slightly. Starting the sample after the Korean War, omitting 1975 (due to the one-time income tax rebate), and omitting 1980 (due to the imposition and removal of credit controls) each seemed to make little difference to the estimates.

30. To the extent that credit rationing effects stemming from adverse selection become more severe as open-market nominal interest rates rose, I would have expected the Treasury bill yield better to have explained consumption. The reason is that if banks do not raise loan rates commensurately with open market rates, the open market rate would likely better capture the combined constraint effects of bank rules (as already proxied by the borrowing rate) and of credit rationing (as captured by the difference between the lending and borrowing rates). In fact, these estimates provide little evidence to support that effect.

31. An alternative functional form for k is one which is linear, as opposed to the linear-in-logarithms form in (6). Specifying $k = k^* + \alpha U + \beta i$ necessitates using a nonlinear instrumental variables estimation technique. Doing so produces qualitatively similar results to those presented in Table 3. Estimates of the linear form were also obtained while including the squares of the unemployment rate and of the interest rate. This allowed the data to suggest whether liquidity constraints bind more, or less, than proportionately as rates rise. Significantly negative coefficients estimated for the squared unemployment and interest rate terms would imply that an accelerating-constraint hypothesis fit the data better. The squared terms turned out to be negative, but insignificant. Thus no strong conclusion about the appropriate functional form for k emerged.

32. Since these consumption equations do not deliver constant elasticities, the test is for equality of the elasticities at the respective sample means of the unemployment and interest rates.

33. Even a permanent, anticipated, balanced-budget shift in tax policy could affect aggregate spending. One way to do so would be to have some parameters of the tax code be age-specific, perhaps by making average income tax rates rise with the age of the taxpayer. To the extent that young-household cash flows were enhanced and balanced by reductions for older households, the average amount of constraint would be loosened.

REFERENCES

Beares, Paul. *Consumer Lending*. American Bankers Association, Washington, D.C., 1987.

Blinder, Alan S. and Angus Deaton. "The Time Series Consumption Function Revisited," *Brookings Papers on Economic Activity*. Washington, D.C., 1985:2.

Campbell, John and N. Gregory Mankiw. "Permanent Income, Current Income, and Consumption," *National Bureau of Economic Research Working Paper* No. 2436. November 1987.

Flavin, Marjorie. "Excess Sensitivity of Consumption to Current Income: Liquidity Constraints or Myopia?," *Canadian Journal of Economics*, 18, February 1985.

Hall, Robert E. "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 86, April 1988.

_________ and Frederick S. Mishkin. "The Sensitivity of Consumption to Transitory Income Estimates from Panel Data on Households," *Econometrica*, 50, March 1982.

Hayashi, Fumio. "The Effect of Liquidity Constraints on Consumption: A Cross-Sectional Analysis," *Quarterly Journal of Economics*, 100, February 1985.

Kotlikoff, Laurence J. "Intergenerational Transfers and Savings," *Journal of Economic Perspectives*, 2, Spring 1988.

Mariger, Randall P. *Consumption Behavior and the Effects of Government Fiscal Policies*. Cambridge, Massachusetts: Harvard University Press, 1986.

Peek, Joe and James A. Wilcox. "Monetary Policy Regime Changes and the Reduced Form for Interest Rates," *Journal of Money, Credit, and Banking*, 19, August 1987.

Tobin, James and Walter Dolde. "Wealth, Liquidity and Consumption, Consumer Spending and Monetary Policy: The Linkages." *Federal Reserve Bank of Boston Conference Series*, No. 5, June 1971.

Walsh, Carl E. "Borrowing Restrictions and Wealth Constraints: Implications for Aggregate Consumption," *Working Paper*, Federal Reserve Bank of San Francisco, September 1986.

Wilcox, David W. "Social Security Benefits, Consumption, Expenditure, and the Life Cycle Hypothesis," *Journal of Political Economy*, 97, April 1989.

Zeldes, Stephen P. "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, 97, April 1989.