Vol. 37, No. 2

ECONOMIC REVIEW

2001 Quarter 2

FEDERAL RESERVE BANK
OF CLEVELAND

http://clevelandfed.org/research/review/

ECONOMIC REVIEW

2001 Quarter 2
Vol. 37, No. 2

Economic Review is published quarterly by the Research Department of the Federal Reserve Bank of Cleveland. To receive copies or to be placed on the mailing list, e-mail your request to: *4d.subscriptions@clev.frb.org* or fax it to 216/579-3050.

Economic Review is available electronically through the Cleveland Fed's site on the World Wide Web: *http://www.clev.frb.org.*

We welcome your comments, questions, and suggestions. E-mail us at *editor@clev.frb.org.*

ISSN 0013-0281

Unitary Thrifts: A Performance Analysis

by James B. Thomson

James B. Thomson is a vice president and economist at the Federal Reserve Bank of Cleveland.

Introduction

Title IV of the Financial Services Modernization Act of 1999 closed an important loophole in U.S. banking law: the unitary thrift holding company exemption from laws that prevented the mingling of banking and commerce.[1] The loophole had allowed nondepository institutions, including nonfinancial corporations, to own unitary thrift holding companies (UTHCs). Only depository institutions could own multiple bank and thrift holding companies or unitary bank holding companies. The period over which the exemption was in force provides us with the opportunity to examine the implications of commingling banking and commerce.

The focus of this paper is a comparison of the performance of thrifts in UTHCs with that of other thrifts. Because little stock market data on UTHCs is available, this study constructs performance measures using balance sheet and income statement data from Thrift Financial Reports.[2]

Overall, we find important differences between UTHC subsidiaries and other thrifts. UTHC thrifts tend to outperform other thrifts during the sample period studied. Moreover, UTHC thrifts appear to have more diversified revenues streams, loan and asset portfolios, and funding sources than do non-UTHC thrifts. Furthermore, differences do not suggest UTHC thrifts pose a greater risk to the federal financial safety net; UTHC thrifts produce similar returns on book equity as non-UTHC thrifts but with higher levels of capitalization. Therefore, we find no evidence consistent with a need to close the UTHC loophole, which could reduce the contestability of banking markets. Overall, these results suggest that some limited commingling of banking and commerce might not pose undue risks to the federal financial safety net. But because thrifts have limited commercial lending powers, these results must be interpreted cautiously when evaluating less restrictive environments.

I. History of the Unitary Thrift Loophole

Holding companies grew in importance in the 1940s and early 1950s as banks used this organizational form to circumvent regulatory restrictions on branching and other activities. Most thrifts could not adopt the holding company form. Subsidiary depository institutions of the holding company had to be stock chartered, but most savings institutions—and all federally chartered thrifts—were

1 Public Law No. 106-102.

2 Most UTHCs are wholly owned subsidiaries of larger organizations, so stock market data do not exist for these firms.

organized as mutuals. Nonetheless, thrifts that were stock-chartered used the holding company form.

Concerned about the growth of multibank holding companies, Congress passed the Bank Holding Company Act of 1956 to regulate them and extended the regulation to unitary bank holding companies in 1970. The regulations restricted the ownership of holding companies, their interstate expansion, and the activities permitted to nonbank subsidiaries. The Reigle-Neal Act of 1994 and the Financial Services Modernization Act of 1999 reduced or eliminated these restrictions, except the restrictions on ownership.

Congressional attempts to reign in the growth of thrift holding companies began with the Spence Act of 1959. This act placed a moratorium on the formation of multiple thrift holding companies and the acquisition of additional thrifts by existing holding companies. The act did not extend to multithrift holding companies that already existed, and it did not prohibit the formation of UTHCs or regulate their behavior. The 1967 Savings and Loan Holding Company Amendments removed the moratorium and subjected multithrift holding companies to regulation. Despite the objection of the Federal Home Loan Bank Board, the federal regulator of thrifts and thrift holding companies, Congress exempted UTHCs from regulation.[3]

Three factors likely played into the decision to exempt UTHCs from regulation. First, at the time thrift holding company regulations were enacted, unitary bank holding companies were also exempt from holding company regulation. Second, thrifts had very restricted lending powers, and allowing nondepository firms to own unitary thrift holding companies would not give rise to the concerns about conflict of interest and concentration of power that the commingling of commercial banking and commerce would.[4] Finally, rising interest rates in the second half of the 1960s threatened the solvency of a number of thrift institutions. Instead of acting to resolve the statutorily induced duration mismatch between thrift assets and liabilities, policymakers took a series of actions aimed at providing thrifts with access to inexpensive deposits—including extending Regulation Q ceilings to savings deposits in thrifts. In other words, policymakers limited the ability of small savers to earn positive real rates of return on their savings in order to assure the funding of the housing finance industry.[5] In this environment, it is not surprising that policymakers would continue to allow nondepository financial firms and nonfinancial firms (henceforth, nonbanks) to own UTHCs. The UTHC exemption would be viewed as a method of providing additional sources of equity capital for the thrift industry as it provided a limited form of entry into banking by nondepository institutions.

II. Empirical Strategy and Sample

This paper investigates whether thrifts in UTHCs perform differently than other types of thrifts, and if so, why. It is possible that the UTHC form is a more efficient arrangement because it gives the firm greater flexibility in arranging links between commercial and banking activities. On the other hand, the links could give the thrifts an advantage only at the expense of taxpayers, by extending the federal financial safety net subsidy to the nondepository owners.[6]

We compare the earnings, leverage, and portfolio risk of UTHC thrifts and non-UTHC thrifts.[7] We also compare the performance of thrifts owned by depository and nondepository institutions. Ideally, we would directly test for differences in risk-adjusted return on equity and the likelihood of failure. However, using balance sheet and income data from Thrift Financial Reports precludes us from constructing direct measures of risk variables, and the sample period does not contain sufficient numbers of thrift closings to fit a hazard function to the data. Hence, our tests measure dimensions of risk and return using proxy variables that are drawn (in large part) from the bank performance and bank failure and early warning literature.[8] The list of variables used and their definitions appears in table 1.

The sample consists of all SAIF-insured thrift institutions that filed at least one complete OTS Thrift Financial Report between March 31, 1993, and December 31, 1999, and also were not in conservatorship. The sample period begins after the Thrift Financial Report was substantially revised as a result of the Federal Deposit Insurance Corporation Improvement Act of 1991 and ends before the UTHC exemption was closed by the Financial Services Modernization Act of 1999. Thrifts are divided into three groups, those in UTHCs, those in

3 See Office of Thrift Supervision (1998a, 1997).

4 See Office of Thrift Supervision (1998a) and Boyd, Chang, and Smith (1998).

5 See Kane (1970, 1977, and 1985, chapter 4).

6 The empirical strategy is similar to that used by Valnek (1999) to compare the performance of stock retail banks and mutual building societies in the United Kingdom.

7 For a discussion of alternative ways to gauge performance using accounting data, see Sinkey (1998, chapter 3).

8 See Cole (1972), Gillis et al. (1980), and Thomson (1991, 1992).

TABLE 1

Variable Definitions

CAPA	Equity capital/total assets
LAGCAPA	CAPA lagged one quarter
CLASSASM	Classified assets/earning assets
FEESHR	Fee income/operating income
FIXEDA	Fixed assets/total assets
FHLBADVA	Federal Home Loan Bank advances/ total assets
INSDEPA	Insured deposits/total assets
INSLNA	Loans to insiders/total assets
INTSHR	Interest income/total income
LGAP (Liquidity gap)	Total cash, deposits due from depository institutions, and investment securities minus borrowings/total assets
LIABHERF (Liability Herfindahl index)	$10000 \cdot$[(deposits/total liabilities)^2 + (borrowings/total liabilities)^2 + (other liabilities/total liabilities)^2]
LNHERF (Loan Herfindahl index)	$10000 \cdot$ [(credit card receivables/total loans)^2 + (other consumer loans/total loans)^2 + (home equity loans/total loans)^2 + (commercial loans/total loans)^2 + (1–4 family mortgage loans/ total loans)^2 + (other mortgage loans/total loans)^2 + (all other loans/total loans)^2]
MPSA	Mortgage pool securities/total assets
MDERA	Mortgage derivative securities/total assets
NIM	Interest income minus interest expense/ earning assets
NTCHRGOF	Net charge-offs/earning assets
OVRHD	Operating expense/total assets
QTL	Mortgage-related assets/total assets
REGLIQR (Regulatory liquidity)	Liquid assets meeting criteria set forth by the Office of Thrift Supervision/total liabilities
ROA	Net income after taxes and extraordinary items/total assets
ROE	Net income after taxes and extraordinary items/common equity
SBLA	Total small business loans/total assets
SIZE	Natural log of total assets
SUBDEBTA	Subordinated debt/total assets
TIER1CAP	Tier 1 capital/total assets
TRUSTA	Trust assets/total assets

Intercept and Slope Dummy Variables

DUMUT	1 if thrift is in a UTHC, 0 otherwise
DUMUD	1 if thrift is in a UTHC owned by a depository institution, 0 otherwise
DUMUN	1 if thrift is in a UTHC owned by a nondepository institution, 0 otherwise
XXXXUT	XXXX • DUMUT
XXXXUD	XXXX • DUMUD
XXXXUN	XXXX • DUMUN

multithrift holding companies, and independent thrifts. Thrifts in the UTHC sample were further partitioned into two groups according to affiliation, using the Office of Thrift Supervision's holding company file. The first consists of UTHCs owned or controlled by depository institutions (henceforth, banks), and the second consists of thrifts in UTHCs that are affiliated with nonbanks.[9]

III. Results

Differences in Return on Equity

The simplest method for evaluating performance is to examine differences in return on equity (ROE) across the different types of thrifts in the sample. As seen in table 2, the ROE of thrifts owned by UTHCs does not differ significantly from that of other thrifts during the full sample period. But when assessing whether the UTHC loophole represented a dangerous mixing of banking and commerce, differences across UTHCs by ownership

9 The thrift holding company database can be found on the Office of Thrift Supervision Web site, <www.ots.treas.gov/applications.html>.

TABLE 2

Return on Equity

	Non-UTHC Thrifts	UTHC Thrifts	Nonbank UTHCs	Bank UTHCs
Mean	0.0181	0.0200	0.0251	0.0195
Std. error	0.0013	0.0016	0.0020	0.0017
P-value	0.3566		0.0375	

SOURCES: Office of Thrift Supervision, Thrift Financial Reports; and author's calculations.

TABLE 3

Components of ROE

ROA

	Non-UTHC Thrifts	UTHC Thrifts	Nonbank UTHCs	Bank UTHCs
Mean	0.0018	0.0019	0.0028	0.0018
Std. error	0.0000	0.0001	0.0008	0.0001
P-value	0.3680		0.2561	

CAPA

	Non-UTHC Thrifts	UTHC Thrifts	Nonbank UTHCs	Bank UTHCs
Mean	0.1037	1.1110	0.1621	0.1061
Std. error	0.0004	0.0010	0.0079	0.0007
P-value	<0.0001		<0.0001	

SOURCES: Office of Thrift Supervision, Thrift Financial Reports; and author's calculations.

type are probably more relevant. After all, ownership of a UTHC expands the set of activities a nondepository institution can engage in, while UTHC ownership by banks and thrifts represents a more modest expansion of their existing franchise. Thrifts in nonbank-UTHCs outperformed thrifts in bank-UTHCs during the sample period, and the difference is significant.

Unfortunately, simple comparisons of ROE can be misleading. Given the positive relationship between risk and required return, a higher ROE may simply reflect higher risk. Even after controlling for risk, comparing the ROE of the different subsamples does not provide sufficient information to understand the underlying factors driving the performance of earnings measures such as ROE. For example, the higher ROE for UTHCs owned by nondepository firms may reflect a conflict of interest arising from the commingling of banking and commerce (albeit on a limited basis), a superior set of real options held by the nondepository UTHCs, or simply a higher degree of leverage employed by these firms.

Differences in Return on Assets and the Ratio of Capital to Assets

The first step in sorting out the different possible influences on ROE is to decompose it into its main components: return on assets (ROA) and the amount of capital per dollar of assets (CAPA).

$$(1) \qquad ROE = \frac{ROA}{CAPA}$$

The optimum degree of risk borne by a firm is a function of the risks emanating from its asset portfolio and the degree of leverage employed. Equation (1) illustrates how ROE is driven by these two simultaneous sets of decisions by the thrift's management, with the investment/operating decision represented by ROA and the financing decision embodied on CAPA.

Overall, the first stage of ROE decomposition reported in table 3 suggests that thrifts in UTHCs, particularly those in UTHCs owned by nonbanks, perform better because they have higher ROA, not higher leverage. While we can't reject the hypothesis that higher risk produced the higher ROE, these univariate results are not consistent with increased leverage driving the results. The first column of table 3 shows that the difference in ROA between UTHC thrifts and non-UTHC thrifts is not statistically significant. UTHC thrifts have higher capital-to-asset ratios than non-UTHC thrifts, and the difference is significant during the sample period.

Insignificant differences in ROE and the significantly higher level of CAPA for the UTHC thrifts relative to other thrifts suggest that performance-related differences between UTHC thrifts and non-UTHC thrifts are due to differences in ROA. That is, differences in performance across these two samples are driven by differences in operating/investment decisions and not by leverage. A comparison of tables 2 and 3 shows that in all three samples, higher ROE is accompanied by lower leverage (higher CAPA).

Similar results hold for the two UTHC subsamples. Thrifts in nonbank UTHCs have higher ROA, although the difference is not statistically significant, and they have significantly higher CAPA. Taken together, these findings suggest that important differences in performance exist across the UTHC subsamples, and these differences are related to operating and investment decisions. Decomposing ROE further should tell us more about how ROA affects ROE.

Differences in the Factors That Affect Return on Assets

The second stage of the ROE decomposition examines the underlying controllable factors that drive ROA—business mix, income production, asset quality, expense control, and tax management—by specifying and estimating the following equation:

$$
\begin{aligned}
(2)\ \mathrm{ROA}_{it} = {} & \beta_0 + \beta_1\mathrm{NIM}_{it} + \beta_2\mathrm{NTCHRGOF}_{it} \\
& + \beta_3\mathrm{OVRHD}_{it} + \beta_4\mathrm{SIZE}_{it} \\
& + \beta_5\mathrm{LAGCAPA}_{it} + \beta_6\mathrm{FEESHR}_{it} \\
& + \beta_7\mathrm{LGAP}_{it} + \beta_8\mathrm{QTL}_{it} + \beta_9\mathrm{INSDEPA}_{it} \\
& + \beta_{10}\mathrm{DUMUT}_{it} + \beta_{11}\mathrm{NIMUT}_{it} \\
& + \beta_{12}\mathrm{NTCHRGOFUT}_{it} \\
& + \beta_{13}\mathrm{OVRHDUT}_{it} + \beta_{14}\mathrm{SIZEUT}_{it} \\
& + \beta_{15}\mathrm{LAGCAPAUT}_{it} + \beta_{16}\mathrm{FEESHRUT}_{it} \\
& + \beta_{17}\mathrm{LGAPUT}_{it} + \beta_{18}\mathrm{QTLUT}_{it} \\
& + \beta_{19}\mathrm{INSDEPAUT}_{it} + \varepsilon_{it}
\end{aligned}
$$

Table 1 explains how these variables were constructed. To proxy for business mix, we included variables for the proportion of a thrift's assets that are mortgage related (QTL), the proportion of its funding that comes from insured deposits (INSDEPA), the proportion of its income that comes from fee-based services (FEESHR), and the size of its liquidity gap (LGAP). Asset quality is captured by the proportion of net charge-offs to interest-earning assets (NTCHRGOF), although this variable may also be related to tax management or expense preference behavior—where the reporting of losses is delayed until profits are higher than average to reduce taxes (see Greenwalt and Sinkey [1988]). Income production is proxied for primarily by the spread between interest income and interest expense scaled by average assets (NIM), although income from fee-based services (FEESHR) may also be related. The ratio of operating expenses to assets (OVRHD) is included to capture expense control. We included a measure of the relative size of the thrift's overall balance sheet (SIZE) because economies of scale may play a role in providing a number of financial products and services. The previous quarter's capital-to-asset ratio (LAGCAPA) is included because the investment/operating decision and the financing decision of depository institutions are not independent. To test the hypothesis that earnings-related measures of performance don't differ across the groups of thrifts, we include intercept and slope dummy variables for the thrift type. For thrifts in the UTHC sample, the effect of each independent variable is measured by combining its coefficient and the coefficient on the corresponding dummy variable.

Equation (2) is estimated using ordinary least squares. The results are presented in table 4 and appear to reject the hypothesis that no performance differences exist, as the coefficients on all the dummy variables except the proportion of mortgage-related assets (QTLUT) are significant at the 1 percent level. A closer inspection of the results, however, is needed to ascertain whether performance-related differences stem from extensions of the federal safety net or an increased set of options available to UTHCs. To do this, we examined differences between UTHC and non-UTHC thrift ROA as captured by the proxies for business mix, asset quality, and the financing decision.

The regression results show business mix is important in determining the earnings performance of thrifts; all of the variables that proxy for this factor are significant. Three of the four corresponding dummy variables are significant and of opposite sign, indicating that differences in business mix between thrifts in UTHCs and other thrifts drive differences in earnings performance. In fact, the dummy variables for fee income and liquidity gap are larger in absolute value, meaning these variables affect ROA in opposite ways and to different degrees depending on the type of thrift.

Non-UTHC thrifts appear to have lower ROA if they rely more heavily on fee income for revenues (FEESHR is negative and significant). One should be careful, however, in interpreting these results in terms of performance. If fee income and interest income are sufficiently uncorrelated, thrifts with higher levels of fee income may have less variable revenues and higher risk-adjusted returns. However, the coefficient on FEESHRUT and the combined effect of FEESHR for UTHC thrifts is significantly positive. The different relationship between FEESHR and earnings for UTHC thrifts may trace to scale economies in the production of fee-based lines of business—the mean of SIZE for UTHC thrifts is significantly larger than for non-UTHC thrifts during the sample period—or to economies of scope and cross-selling opportunities with other businesses conducted in the holding company or by its parent firm. If we accept the argument that fee income is not highly correlated with interest income—that is, fee income reduces the variability of revenues—then the fact that fee-based services generate better performance for UTHC thrifts than for other thrifts is not consistent with the hypothesis that performance differences trace to increased risk and safety-net subsidies.

The positive and significant sign on LGAP suggests that non-UTHC thrifts benefit from the flexibility option associated with liquidity (as measured by the difference between liquid assets and short-term nondeposit liabilities), and they perform better the higher it is. For UTHC thrifts, liquidity

T A B L E 4

OLS Estimation of Equation 2

Dependent Variable: ROA

	Coefficient	t-Statistic	Prob > t
Intercept	–0.0014	–2.26	0.024
NIM	0.6308	83.58	<0.0001
NTCHRGOF	0.0895	13.00	<0.0001
OVRHD	–1.0328	–126.68	<0.0001
SIZE	0.0001	2.90	0.0037
LAGCAPA	0.0127	23.69	<0.0001
FEESHR	–0.0002	–1.95	0.0517
LGAP	0.0023	8.38	<0.0001
QTL	0.0013	3.62	0.0003
INSDEPA	–0.0012	–8.40	<0.0001
DUMUT	–0.0131	–9.59	<0.0001
NIMUT	0.4028	13.70	<0.0001
NTCHRGOFUT	–0.2689	–31.15	<0.0001
OVRHDUT	1.0474	101.56	<0.0001
SIZEUT	0.0004	5.70	<0.0001
LAGCAPAUT	–0.0177	–14.06	<0.0001
FEESHRUT	0.0026	4.68	<0.0001
LGAPUT	–0.0050	–11.06	<0.0001
QTLUT	0.0007	1.06	0.2892
INSDEPAUT	0.0007	2.70	0.007
R-Square	0.6265		
Root MSE	0.0063		
Dependent mean	0.0014		
Coeff Var	438.5867		
F value	2,259.58		
Prob > F	<0.0001		
Number of observations	25,614		

SOURCES: Office of Thrift Supervision, Thrift Financial Reports; and author's calculations.

has the opposite effect on ROA; the coefficient on LGAPUT and the net effect of LGAP on ROA for UTHC thrifts is significantly negative. There are two possible explanations for why this is so. On one hand, the parent holding company may serve as a source of strength and liquidity for its thrift subsidiary, so UTHC thrifts do not need as much on-balance-sheet liquidity to conduct their operations. That is, holding companies may provide their thrift subsidiaries with access to other funding sources which, at the margin, may be less expensive than raising additional retail deposits.[10] In addition, UTHC thrifts are larger and more likely to face deposit constraints than non-UTHC thrifts. On the other hand, UTHC thrifts may find it desirable to take on more liquidity risk to increase the value of their federal deposit guarantees. Obviously, the first explanation would be consistent with the greater-options hypothesis, while the latter would be consistent with the safety-net-subsidy hypothesis. But differences in leverage and the composition of nondeposit funding between the two samples of thrifts do not support the safety-net-subsidy hypothesis. After all, UTHC thrifts are significantly less leveraged than non-UTHC thrifts, and a large part of the funding difference arises because UTHC thrifts rely more on Federal Home Loan Bank advances.[11]

It is somewhat curious that for non-UTHC thrifts, the proportion of assets funded by insured deposits is significantly negatively related to ROA. Moreover, while the dummy variable for this factor is positive, it is smaller in absolute value than INSDEPA, and the relationship between INSDEPA and ROA for percentage of funding remains negative and significant for UTHC thrifts. Care needs to be taken in interpreting these results since insured deposits are a stable funding source and are likely to reduce profit variability. UTHC thrift earnings appear to be less sensitive to changes in the insured deposit base than other thrifts (the coefficient on INSDEPAUT is positive). This may be due to funding advantages associated with holding company affiliation, which reduces the marginal cost of funding additional assets. In other words, differences in earnings performance based on INSDEPA are not likely to trace to increased safety-net subsidies but to greater availability of deposit substitutes (enhanced funding options) associated with using the holding company organizational form.

ROA is positively related to the concentration of thrift assets in mortgage-related loans and securities (QTL). This positive relationship may indicate thrifts have specialized expertise in mortgage-related assets and a competitive advantage in housing-finance markets. On the other hand, the Competitive Equality Banking Act established a qualified-thrift-lender requirement that required a minimum level of investment in qualified assets (primarily housing-finance-related assets and, after 1996, small business and agricultural loans). Given the qualified-thrift-lender requirement, the positive and significant coefficient on QTL may be proxying for regulatory taxes—that is, thrifts with high QTL would be less subject to regulatory interference. The coefficient on QTLUT is not significant, however, and this sheds some doubt on that interpretation. If a thrift in a UTHC fails to meet the QTL test, the UTHC is subject to more stringent and invasive bank holding company regulation. In the case of nonbank-UTHC

10 Most of the difference in the liquidity gap between samples derives from differences in funding, as liquid assets made up similar portions of both group's assets—20 percent for non-UTHC thrifts and 18 percent for UTHC thrifts.

11 Over the sample period, thrifts in the UTHC sample financed 12 percent of their assets with Federal Home Loan Bank advances on average, while advances were 4 percent of non-UTHC thrifts' assets.

thrifts, penalties for failing the qualified-thrift-lender test could include forced divestiture of the thrift subsidiary.

The positive and significant effect of the proportion of a thrift's net charge-offs to interest-earning assets suggests that non-UTHC thrifts perform better the more they engage in expense preference behavior—the timing of loss recognition to smooth income for tax purposes. Note that the opposite is true for UTHC thrifts (the coefficient on the dummy variable is negative and significant, and the overall effect for UTHC thrifts is significantly negative), which is consistent with low net charge-offs as an indicator of asset quality. The net charge-off variable might operate differently across the subsamples because holding companies are able to utilize leverage at the parent-company level to arbitrage taxes. Thus, thrifts in holding companies have less of an incentive to engage in expense preference behavior. Unfortunately, because non-UTHC thrifts also engage in this behavior, it is impossible to interpret differences in the coefficients on NTCHRGOF across the two groups of thrifts as an indicator of differences in asset quality.

The coefficient on LAGCAPA is significantly positive—which is consistent with the findings in table 3 that firms with high ROE also had high capital-to-asset ratios. The negative coefficient on LAGCAPAUT indicates that UTHC-thrift earnings are less sensitive to the level of capital than non-UTHC thrifts. Two factors likely drive these results. First, UTHC thrifts have higher levels of capital than independent thrifts. Second, to the extent that the parent holding company serves as a source of strength to its thrift subsidiary, we would expect the thrift-level financing decision to have less of an impact on performance.

Three other differences emerge from the ROA decomposition. First, UTHC-thrift earnings performance is significantly more responsive to changes in the net interest margin than non-UTHC thrifts. The coefficient on the proxy for efficiency, OVRHD, has a large negative effect on ROA. However, this earnings factor has no effect on earnings performance for UTHC thrifts (the coefficient on OVRHDUT is of the same magnitude and of opposite sign as the coefficient on OVRHD). Finally, larger thrifts have higher earnings, and this effect is significantly greater for UTHC thrifts (the coefficients on SIZE and SIZEUT are both positive and significant).

Differences between Bank- and Nonbank-UTHC Thrifts

Overall, the results of the ROA decomposition over the full sample of thrifts are consistent with the hypothesis that performance-related differences exist between UTHC thrifts and non-UTHC thrifts. However, the nature of the differences across these two samples does not provide sufficient information to decide which of the two possible causes is responsible. Given that the UTHC exemption may be more valuable, or at least the bundle of real options associated with owning a UTHC thrift is likely different, for nonbank owners of UTHCs than for depository institutions, we perform the ROA decomposition again, but over the UTHC sample only. We re-estimate equation (2), including intercept and dummy variables for type of UTHC ownership to explore differences between thrifts owned by nonbank UTHCs and those owned by bank UTHCs. The results of this regression are presented in table 5.

Business Mix

Because we are exploring differences based on ownership type, we focus on the coefficients of the nonbank ownership dummy variables. We first consider whether different choices of business mix are responsible for differences in performance. The results of this regression suggest that nonbank UTHC thrifts use a different business mix than bank UTHC thrifts. Interest income appears to be a less important determinant of ROA for thrifts owned by nonbank firms (NIMUN is negative and significant, but of smaller magnitude than NIM). Fee income seems more important (the coefficients on FEESHRUN and SIZEUN are positive and significant, and the overall relationship between the share of fee income and ROA for nonbank-UTHC thrifts is positive). These results are consistent with nonbank UTHCs holding a different set of real options than depository institutions. In other words, as the logit regression analysis that follows will confirm, it is the options other than lending powers afforded by depository institution charters—such as access to the payments system—that have the most value to the nonbank acquirers of UTHCs.

Several other important differences emerge between nonbank UTHCs and bank UTHCs. First, asset quality matters more to earnings performance for nonbank-UTHC thrifts than for bank-UTHC thrifts (the signs on NTCHRGOF and NTCHRGOFUN are significantly negative). Second, the coefficients on OVRHD and OVRHDUN are significant and of opposite sign. The overall impact of efficiency for nonbank-UTHC thrifts is positive and significant. In other words, the lack of a significant efficiency effect on earnings performance in the first estimation of equation (2) for UTHC thrifts was due to conflicting effects within the UTHC sample. Given that OVRHD is constructed as the ratio of operating expense to total assets, the differences in the effect of this proxy on

TABLE 5

OLS Re-estimation of Equation 2

Dependent Variable: ROA

	Coefficient	t-Statistic	Prob > t
Intercept	0.0010	0.67	0.5028
NIM	1.2667	34.60	<0.0001
NTCHRGOF	–0.1240	–19.57	<0.0001
OVRHD	–0.3590	–23.63	<0.0001
SIZE	0.0001	1.52	0.1298
LAGCAPA	–0.0105	–6.36	<0.0001
FEESHR	–0.0038	–5.74	<0.0001
LGAP	–0.0036	–8.35	<0.0001
QTL	–0.0014	–1.78	0.0755
INSDEPA	–0.0004	–1.51	0.1301
DUMUN	–0.0872	–19.69	<0.0001
NIMUN	–0.2580	–3.08	0.0021
NTCHRGOFUN	–1.1784	–8.08	<0.0001
OVRHDUN	0.5198	28.96	<0.0001
SIZEUN	0.0040	15.42	<0.0001
LAGCAPAUN	0.0105	3.62	0.0003
FEESHRUN	0.0347	17.40	<0.0001
LGAPUN	0.0097	5.76	<0.0001
QTLUN	0.0012	0.72	0.4729
INSDEPAUN	–0.0062	–2.93	0.0034
R-Square	0.5007		
Root MSE	0.0069		
Dependent mean	0.0020		
Coeff Var	350.51		
F value	375.37		
Prob > F	<.0001		
Number of observations	7,132		

SOURCES: Office of Thrift Supervision, Thrift Financial Reports; and author's calculations.

earnings across the nonbank- and bank-UTHC samples may reflect differences in business mix and strategy. If nonbank-UTHC thrifts place greater emphasis on fee-based products and services, higher levels of OVRHD may be picking up increased activity in these areas, and increases in operating expenses per dollar of assets would be positively related to earnings performance. Hence, the significantly positive sign on OVRHDUN is consistent with the hypothesis that nonbank UTHCs hold a different set of real options than bank UTHCs and non-UTHC thrifts.

Bank-UTHC thrifts exhibit a significantly negative relationship between ROA and the previous quarter's capital-to-assets ratio (LAGCAPA) and a negative but not significant relationship between ROA and the share of insured deposits (INSDEPA). For nonbank-UTHC thrifts, ROA is significantly negatively related to the share of insured deposits (the coefficient on INSDEPAUN is negative and significant). However, the net effect of the capital-to-assets ratio on earnings is not significantly different from zero for nonbank-UTHC thrifts (the coefficient on LAGCAPAUN is positive and significant and of the same magnitude as the coefficient on LAGCAPA). Finally, liquidity has opposite effects on ROA in the two groups. It is negative for bank-UTHC thrifts and positive for nonbank-UTHC thrifts (the coefficients on LGAP and LGAPUN are significantly negative and positive, respectively). The impact of differences in the financing decision on performance as proxied for by ROA for the bank-UTHC thrifts and the nonbank-UTHC thrifts is not consistent with the hypothesis that this commingling of banking and commerce increases the loss exposure of the taxpayer to the federal financial safety net.

Finally, the fact that the concentration of mortgage-backed assets (QTL) is negative and marginally significant and QTLUN is positive but not significant is not consistent with our earlier interpretation of QTL as a regulatory variable. Bank-UTHCs would already be subject to the more stringent bank holding company regulation, and thus penalties associated with the violation of the qualified-thrift-lender test would have less impact on bank UTHCs than nonbank ones.

Level of Risk

The second part of examining differences in performance is to look at the level of risk of the institutions according to their organizational structure. A lack of market data and a relatively short time series for the accounting data make a direct examination of risk problematic. Consequently, we pursue an alternative strategy of examining differences in a number of risk proxies constructed from thrift balance sheet data. The approach is to devise an empirical model that explains organizational type using proxy variables for different risk characteristics and business strategies constructed from thrift balance sheet and income statement data. Equations (3) and (4) specify the model.

To control for thrift-level structural effects that may be related to holding company affiliation, such as scale of operation and geographic presence, we include, regressors for business volume (SIZE) and proportion of fixed-to-total assets (FIXEDA). (See table 1 for a description of the variables.) Because measures of capital adequacy and liquidity have been shown to be related to the probability that a bank will be closed, we include regressors for the ratio of capital to assets (TIER1CAP) and regulatory liquidity (REGLIQR). These are also included to capture regulatory restrictions on leverage and liquidity.

To capture the diversification of loan portfolio and funding sources, the Herfindahl indicies LNHERF and LIABHERF are included. Higher levels of these variables suggest higher levels of balance sheet risk. Diversification of revenue streams is proxied for by the proportion of interest income to total income (INTSHR)—we assume that interest income and noninterest income are not highly correlated. Higher levels of INTSHR would suggest higher variability of revenues. Subordinated debt has been held up by some as a potential source of market discipline for depository institutions, so we included the proportion of it to total assets (SUBDEBTA). The ratio of Federal Home Loan Bank advances to total assets (FHLBADVA) is included to proxy for funding strategy because this type of funding represents a subsidized alternative to deposits for funding assets, albeit to the extent the thrift has sufficient eligible collateral in the form of mortgage assets. Two regressors are included to proxy for asset quality, the ratio of insider loans to total loans (INSLNA) and the ratio of classified assets to total earning assets (CLASSASM). Both loans to insiders as a percent of assets and measures of problem assets have been shown to be positively related to bank closings. To capture thrifts' use of alternative business lines, we include two proxies. The proportion of small business loans to total assets (SBLA) represents thrifts' new small business and agricultural lending powers, and the proportion of trust assets to total assets (TRUSTA) represents fee-based activities. Finally, because mortgage derivative securities represent potential hedges against risks arising from mortgage lending, the thrift's proportion of these securities (MDERA) is included, and because mortgage pool securities represent an alternative to direct mortgage holdings—an asset that is typically more liquid but riskier than a traditional home mortgage loan—the proportion of these (MPSA) is included as well.

Equation (3) seeks to explain differences between UTHC and non-UTHC thrifts, and (4) examines differences between bank-UTHC and nonbank-UTHC thrifts. Using the logistic regression procedure in SAS, we estimate equations (3) and (4) over the full sample and the UTHC sample, respectively. The results appear in table 6.

$$
\begin{aligned}
(3)\ \mathrm{DUMUT}_{it} = {} & \phi_0 + \phi_1\mathrm{SIZE}_{it} + \phi_2\mathrm{TIER1CAP}_{it} \\
& + \phi_3\mathrm{LNHERF}_{it} + \phi_4\mathrm{LIABHERF}_{it} \\
& + \phi_5\mathrm{FIXEDA}_{it} + \phi_6\mathrm{SBLA}_{it} \\
& + \phi_7\mathrm{INTSHR}_{it} + \phi_8\mathrm{SUBDEBTA}_{it} \\
& + \phi_9\mathrm{FHLBADVA}_{it} + \phi_{10}\mathrm{INSLNA}_{it} \\
& + \phi_{11}\mathrm{REGLIQR}_{it} + \phi_{12}\mathrm{TRUSTA}_{it} \\
& + \phi_{13}\mathrm{MDERA}_{it} + \phi_{14}\mathrm{MPSA}_{it} \\
& + \phi_{15}\mathrm{CLASSASM}_{it} + \mu_{it}
\end{aligned}
$$

$$
\begin{aligned}
(4)\ \mathrm{DUMUD}_{it} = {} & \phi_0 + \phi_1\mathrm{SIZE}_{it} + \phi_2\mathrm{TIER1CAP}_{it} \\
& + \phi_3\mathrm{LNHERF}_{it} + \phi_4\mathrm{LIABHERF}_{it} \\
& + \phi_5\mathrm{FIXEDA}_{it} + \phi_6\mathrm{SBLA}_{it} \\
& + \phi_7\mathrm{INTSHR}_{it} + \phi_8\mathrm{SUBDEBTA}_{it} \\
& + \phi_9\mathrm{FHLBADVA}_{it} + \phi_{10}\mathrm{INSLNA}_{it} \\
& + \phi_{11}\mathrm{REGLIQR}_{it} + \phi_{12}\mathrm{TRUSTA}_{it} \\
& + \phi_{13}\mathrm{MDERA}_{it} + \phi_{14}\mathrm{MPSA}_{it} \\
& + \phi_{15}\mathrm{CLASSASM}_{it} + \mu_{it}
\end{aligned}
$$

The results from equation 3 show important differences between UTHC thrifts and others in the proxies for balance sheet structure and risk. Twelve of the 15 regressors and the intercept term are significant, a result that is not consistent with the hypothesis that no differences in performance exist between the two types of thrifts. UTHC thrifts are significantly larger and hold significantly more tier 1 capital than thrifts in the non-UTHC sample. In addition, the negative and significant coefficient on LNHERF is consistent with UTHC thrifts having a more diversified loan portfolio. None of these results is consistent with the hypothesis that UTHCs increase taxpayer risk.

It is also interesting to note that the coefficient on LIABHERF is negative and significant, which is consistent with UTHC thrifts having more diversified funding sources. However, to the extent that the lower-liability Herfindahl results from a higher dependence on Federal Home Loan Bank advances —as indicated by the positive and significant coefficient on FHLBADVA—UTHC thrifts do not necessarily rely less on funding sources subsidized by implicit U.S. government guarantees.

UTHC thrifts appear to have a higher ratio of fixed assets to total assets, which may indicate that they maintain larger branching networks (FIXEDA is positive and significant). Given that thrifts have more liberal branching powers than banks, both before and after the Reigle-Neal Act of 1994, this result is not surprising. Moreover, the positive and significant coefficient on SBLA is also consistent with this explanation. That is, it is commonly held that an office presence in the community is needed to make the relationship-based small business loan. Therefore, we would expect thrifts with higher investments in fixed assets, presumably branches, to also have higher levels of small business loans.

Thrifts in UTHCs have significantly lower ratios of qualifying liquid assets[12] to liabilities than non-UTHC thrifts. In other words, the negative and significant coefficient on REGLIQR is consistent with the higher levels of liquidity risk undertaken by UTHCs. However, some caution should be used in interpreting this result. First, to the extent that UTHC thrifts have larger branch networks than

12 Liquid assets for purposes of regulatory liquidity must conform to the eligibility criteria as expressed in OTS Regulations 566.1(g).

TABLE 6

Logit Regression Results

	Equation 4: Full Sample Dependent Variable: DUMUT			Equation 5: UTHC Sample Dependent Variable: DUMUD		
	Parameter estimate	Chi-Square	Prob > Chi Square	Parameter estimate	Chi-Square	Prob > Chi Square
Intercept	–5.3877	218.4102	<0.0001	–0.3226	0.1428	0.7055
SIZE	0.6370	1457.7063	<0.0001	0.0327	0.5675	0.4513
TIER1CAP	10.0303	686.6470	<0.0001	–1.1386	2.1779	0.1400
LNHERF	–0.0002	407.9386	<0.0001	–0.0001	36.1343	<0.0001
LIABHERF	–0.0003	183.2916	<0.0001	0.0001	2.6080	0.1063
FIXEDA	7.7109	26.0477	<0.0001	7.7869	2.8527	0.0912
SBLA	1.1632	5.3536	0.0207	1.2726	0.6252	0.4291
INTSHR	–0.0384	0.6350	0.4255	2.5124	40.8492	<0.0001
SUBDEBTA	7.7712	2.1190	0.1455	–19.6167	15.8421	<0.0001
FHLBADVA	1.7655	21.3132	<0.0001	1.7099	6.1526	0.0131
INSLNA	–6.8246	5.3927	0.0202	46.0876	9.6365	0.0019
REGLIQR	–0.0035	6.7873	0.0092	–0.0005	0.0223	0.8812
TRUSTA	0.1395	28.4154	<0.0001	–0.0134	0.5456	0.4601
MDERA	1.1045	19.5965	<0.0001	–0.3963	0.7169	0.3972
MPSA	–0.7827	24.7668	<0.0001	2.3233	23.6342	<0.0001
CLASSASM	–0.1634	0.1335	0.7149	–2.9886	11.5626	0.0007
"–2 LOG L"		21,931.422			3,777.77	
AIC		21,963.422			3,753.194	
Likelihood ratio		8,376.3025,	DF: 15		84.5255,	DF: 15
Prob > Chi Square		<0.0001			<0.0001	
Number of observations		25,655			7,122	

SOURCES: Office of Thrift Supervision, Thrift Financial Reports; and author's calculations.

non-UTHC thrifts, UTHC thrifts would need a secondary reserve less. Second, because holding companies are a source of liquidity for their subsidiary thrifts, we would also expect thrifts in holding companies to need liquid assets as a secondary reserve less.

The positive and significant sign on the proportion of trust assets (TRUSTA) indicates that UTHC thrifts are more active in services that generate fee income. Note, proxies for other fee-related business lines were omitted from this regression because of the high degree of colinearity between these variables and TRUSTA—that is, thrifts that engaged in one fee-based activity tend to be engaged in the others. This result is consistent with FEESHR being more important for UTHC thrifts as a determinant of ROA in the equation (2) regression. Moreover, increased reliance on nonlending business is likely to reduce the variability of revenues and the risk of loss to depositors and deposit insurance funds. This result does not support the hypothesis of no performance-related differences.

Three other regression coefficients suggest that UTHC thrifts may hold less risk than their non-UTHC counterparts. First, UTHC thrifts hold more mortgage derivative contracts (MDERA is positive and significant). To the extent that these are used to hedge risks, such as prepayment risks, from the mortgage portfolio, UTHC thrifts would hold less risk than their less-hedged counterparts. Second, non-UTHC thrifts have higher holdings of mortgage pool securities than UTHC thrifts (MPSA is negative and significant). Finally, UTHC thrifts appear less likely to have risk exposure to insiders (INSLNA is negative and significant). Thomson (1991, 1992) interprets the ratio of inside loans to total loans as an indicator of fraud and shows that INSLNA is related to the probability of bank failure. Hence, the INSLNA result could indicate potential problems in asset quality for non-UTHC thrifts.

Finally, it is important to note that the data could not discriminate between UTHC and non-UTHC thrifts in three areas: the share of revenues represented by interest income (INTSHR), the reliance on subordinated debt as a funding source (SUBDEBTA), and the proportion of assets classified by regulators as problems (CLASSASM).

Overall, the results of the logit regression over the full sample suggest that some important differences exist between UTHC thrifts and other thrifts. However, none of the significant differences between the two samples suggests that UTHC thrifts engage in riskier activities or pose a greater expected loss to the financial safety net than non-UTHC thrifts. In fact, just the opposite appears to be true. UTHC thrifts are more diversified in terms of their lending portfolios and sources of revenue. In addition, UTHC thrifts are less leveraged and appear to hedge more of their nondiversifiable risks than do non-UTHC thrifts.

As noted above, the UTHC loophole would be of most value to nonbank owners of UTHC thrifts. Therefore, before we draw definitive conclusions, we need to examine differences between UTHC thrifts by ownership type. To do this, we estimate equation (4) over the UTHC sample. Interestingly, eight of the fifteen regression coefficients and the intercept term are not significantly different from zero—as opposed to three for equation (3). Moreover, all three of the insignificant regressors from equation (3) are significant explanatory variables in equation (4).

The regression results suggest that bank-UTHC thrifts hold slightly more diversified loan portfolios than nonbank-UTHC thrifts (LNHERF is negative and significant). But bank-UTHC thrifts also hold higher levels of mortgage pool securities (MPSA is positive and significant), and this likely offsets increased loan portfolio diversification and results in a higher concentration of their mortgage-related risks. In addition, nonbank-UTHC thrifts rely less on interest income than bank-UTHC thrifts (INTSHR is positive and significant). Overall, these results suggest that the asset portfolios of nonbank-UTHC thrifts are less exposed to mortgage-related risks and that their revenue streams are more diversified than bank-UTHC thrifts. These results are not inconsistent with the hypothesis that performance differences trace to different real options held by nonbank UTHCs.

Two important differences between our two samples of UTHC thrifts emerge from the funding side of their balance sheets. First, bank-owned UTHCs rely more heavily on subsidized Federal Home Loan Bank advances (FHLBADVA is positive and significant). Second, nonbank-UTHCs appear to be more influenced by market-based forms of discipline than bank-owned UTHCs (SUBDEBTA is negative and significant). To the extent that subordinated debt serves as a source of market discipline, the positive and significant coefficient (albeit at the 90 percent level) on SUBDEBTA is not consistent with the hypothesis of no performance-related differences.[13] However, it is possible that this variable is picking up a preference by banks to issue subordinated debt at the holding-company level instead of the level of the thrift subsidiary.

Finally, two conflicting effects are found for asset quality. As noted before, previous research finds INSLNA to be positively related to the probability of failure. Hence, the positive and significant coefficient on INSLNA for bank-UTHC thrifts indicates they have potential asset quality problems. On the other hand, the negative and significant coefficient on CLASSASM suggests nonbank-UTHC thrifts have higher levels of substandard assets on their books.

In general, the results from equations (3) and (4) do not appear to be consistent with the hypothesis that the UTHC loophole increases the risk to taxpayers by extending the safety-net subsidy to nondepository firms. We find no evidence from balance-sheet proxies that the asset and liability decisions of UTHC thrifts, particularly nonbank-UTHC thrifts, pose higher risks. In fact, we tend to see that nonbank-UTHC thrifts rely less on mortgage-related assets and more on nonsubsidized sources of funds than their counterparts.

IV. Policy Conclusions

Debates over the merits of commingling banking and commerce have focused on the potential efficiency gains associated with enhanced bundles of real options afforded by cross-industry mergers. These efficiency gains are weighed against potential efficiency losses due to possible conflicts of interest, extension of the financial safety net (and the attendant moral hazard incentives) to nonfinancial activities, and increased concentration of economic power.

The UTHC exemption represents the commingling of banking and commerce in the United States. The degree of commingling is limited because the commercial lending powers of thrift institutions are restricted and because nonbanks must satisfy the qualified-thrift-lender requirement to own UTHCs. The Financial Services Modernization Act of 1999 effectively eliminated this exemption, preventing the formation of additional nonbank-owned UTHCs.

This paper examines the merits of the UTHC exemption as a limited form of commingled banking and commerce. It tests for differences in performance of non-UTHC thrifts and two categories of UTHC thrifts. While performance differences are found,

13 See Federal Reserve System (1999).

none suggests nonbank ownership of thrift holding companies poses a risk to the federal financial safety net, at least no greater a risk than a bank-owned UTHC or an independent thrift poses. Furthermore, we find evidence of performance differences that suggest nonbank-UTHC thrifts hold a different set of real options than unaffiliated thrifts and hence have the potential for gains in economic efficiency. Therefore, our results do not provide an economic justification for the Financial Services Modernization Act's elimination of the UTHC exemption. Given the limited scope of powers afforded to nonbank owners of UTHCs, these results do not extend to proposals for more general forms of universal banking.

References

Boyd, John H., Chun Chang, and Bruce D. Smith. "Moral Hazard under Commercial Banking and Universal Banking," *Journal of Money, Credit, and Banking*, vol. 30, no. 3 (August, part 2, 1998), pp. 426–68.

Cole, David W. "Return on Equity Model for Banks," *The Banker's Magazine*, vol. 55 (Summer 1972), pp. 40–7.

Federal Reserve System. "Using Subordinated Debt as an Instrument of Market Discipline," Board of Governors of the Federal Reserve System, staff study 172, 1999.

Gillis, Harvey N., Rufus W. Lumry, III, and Thomas J. Oswald. "A New Approach to Analyzing Bank Performance," *The Banker's Magazine*, vol. 163 (March-April 1980), pp. 67–73.

Greenawalt, Mary B., and Joseph F. Sinkey, Jr. "Bank Loan-Loss Provisions and the Income-Smoothing Hypothesis: An Empirical Analysis, 1976–1984," *Journal of Financial Services Research,* vol. 1, no. 4 (December 1988), pp. 301–18.

Kane, Edward J. "Short-Changing the Small Saver: Federal Government Discrimination against the Small Saver during the Vietnam War," *Journal of Money, Credit, and Banking*, vol. 2, no. 4 (November 1970), pp. 513–22.

———. "Good Intentions and Unintended Evil: The Case against Selective Credit Allocation," *Journal of Money, Credit, and Banking*, vol. 9, no. 1 (February, part 1, 1977), pp. 55–69.

———. *The Gathering Crisis in Federal Deposit Insurance*, Cambridge, Mass.: MIT Press, 1985.

Office of the Thrift Supervision. "Holding Companies in the Thrift Industry," background paper, <www.ots.treas.gov/docs/48031.html>, April 1977.

———. "Holding Companies in the Thrift Industry," supplement to the 1977 background paper, <www.ots.treas.gov/docs/48032.pdf>, January 1998a.

———. "Historical Framework for Regulation of Activities of Unitary Savings and Loan Holding Companies," studies and papers, <www.ots.treas.gov/docs/48035.html>, 1998b.

S**inkey, Joseph F., Jr.** *Commercial Bank Financial Management*, 5th edition, New York, N.Y.: Prentice Hall, Inc., 1998.

Thomson, James B. "Predicting Bank Failures in the 1980s," Federal Reserve Bank of Cleveland, *Economic Review*, vol. 27, no. 1 (Quarter 1, 1991), pp. 9–20.

———. "Modeling the Bank Regulator's Closure Option: A Two-Step Logit Regression Approach," *Journal of Financial Services Research,* vol. 6, no 1 (May 1992), pp. 5–23.

Valnek, Tomas. "The Comparative Performance of Mutual Building Societies and Stock Retail Banks," *Journal of Banking and Finance,* vol. 23, no. 6 (June 1999), pp. 925–38.

The Sources and Nature of Long-Term Memory in Aggregate Output

by Joseph G. Haubrich and Andrew W. Lo

Joseph G. Haubrich is an economic advisor at the Federal Reserve Bank of Cleveland and Andrew W. Lo is the Harris & Harris Group Professor at the MIT Sloan School of Management. This research was partially supported by the MIT Laboratory for Financial Engineering, the National Science Foundation (SES–8520054, SES–8821583, SBR–9709976), the Rodney L. White Fellowship at the Wharton School (Lo), the John M. Olin Fellowship at the NBER (Lo), and the University of Pennsylvania Research Foundation (Haubrich). The authors thank Don Andrews, Pierre Perron, Fallaw Sowell, and seminar participants at Columbia University, the NBER Summer Institute, the Penn Macro Lunch Group, and the University of Rochester for helpful comments, and Arzu Ilhan for research assistance.

Introduction

Questions about the persistence of economic shocks currently occupy an important place in economics. Much of the recent controversy has centered on "unit roots," determining whether aggregate time series are better approximated by fluctuations around a deterministic trend or by a random walk plus a stationary component. The mixed empirical results reflect the general difficulties in measuring low-frequency components. Persistence, however, has richer and more relevant facets than the asymptotic behavior at the heart of the unit root debate. In particular, fractionally differenced stochastic processes parsimoniously capture an important type of long-range dependence midway between the quick decay of an ARMA process and the infinite persistence of a random walk. Fractional differencing allows something of a return to the classical NBER business cycle program exemplified by Wesley Claire Mitchell, who urged examinations of stylized facts at all frequencies.

Though useful in areas such as international finance (Diebold, Husted, and Rush [1991] and Baillie, Bollerslev, and Mikkelsen [1993]), fractionally differenced processes have had less success in macroeconomics. For GDP at least, it is hard to estimate the appropriate fractional parameter with any precision. One promising technique (Geweke and Porter-Hudak [1983] and Diebold and Rudebusch [1989]) also has serious small-sample bias, which limits its usefulness (Agiakloglou, Newbold, and Wohar [1993]). Although both Deibold and Rudebusch (1989) and Sowell (1992) find point estimates that suggest long-term dependence, they cannot reject either extreme of finite-order ARMA or a random walk. The estimation problems raise again the question posed by Christiano and Eichenbaum (1990) about unit roots: do we know and do we care? In this paper we provide an affirmative answer to both questions.

Do we know? Applying the modified rescaled range (R/S) statistic confronts the data with a test at once both more precise and more robust than previous estimation techniques. The R/S statistic has shown its versatility and usefulness in a variety of different contexts (Lo [1991] and Haubrich [1993]). Operationally, we can determine what we know about our tests for long-range dependence by using Monte Carlo simulations of their size and power. Not surprisingly, with typical macroeconomic sample sizes we cannot distinguish between fractional exponents of 1.000 and 0.999, but we can distinguish between exponents of 0 and 0.333.

Do we care? Persistence matters directly for making predictions and forecasts. It matters in a more

The Sources and Nature of Long-Term Memory in Aggregate Output

by Joseph G. Haubrich and Andrew W. Lo

Joseph G. Haubrich is an economic advisor at the Federal Reserve Bank of Cleveland and Andrew W. Lo is the Harris & Harris Group Professor at the MIT Sloan School of Management. This research was partially supported by the MIT Laboratory for Financial Engineering, the National Science Foundation (SES–8520054, SES–8821583, SBR–9709976), the Rodney L. White Fellowship at the Wharton School (Lo), the John M. Olin Fellowship at the NBER (Lo), and the University of Pennsylvania Research Foundation (Haubrich). The authors thank Don Andrews, Pierre Perron, Fallaw Sowell, and seminar participants at Columbia University, the NBER Summer Institute, the Penn Macro Lunch Group, and the University of Rochester for helpful comments, and Arzu Ilhan for research assistance.

Introduction

Questions about the persistence of economic shocks currently occupy an important place in economics. Much of the recent controversy has centered on "unit roots," determining whether aggregate time series are better approximated by fluctuations around a deterministic trend or by a random walk plus a stationary component. The mixed empirical results reflect the general difficulties in measuring low-frequency components. Persistence, however, has richer and more relevant facets than the asymptotic behavior at the heart of the unit root debate. In particular, fractionally differenced stochastic processes parsimoniously capture an important type of long-range dependence midway between the quick decay of an ARMA process and the infinite persistence of a random walk. Fractional differencing allows something of a return to the classical NBER business cycle program exemplified by Wesley Claire Mitchell, who urged examinations of stylized facts at all frequencies.

Though useful in areas such as international finance (Diebold, Husted, and Rush [1991] and Baillie, Bollerslev, and Mikkelsen [1993]), fractionally differenced processes have had less success in macroeconomics. For GDP at least, it is hard to estimate the appropriate fractional parameter with any precision. One promising technique (Geweke and Porter-Hudak [1983] and Diebold and Rudebusch [1989]) also has serious small-sample bias, which limits its usefulness (Agiakloglou, Newbold, and Wohar [1993]). Although both Deibold and Rudebusch (1989) and Sowell (1992) find point estimates that suggest long-term dependence, they cannot reject either extreme of finite-order ARMA or a random walk. The estimation problems raise again the question posed by Christiano and Eichenbaum (1990) about unit roots: do we know and do we care? In this paper we provide an affirmative answer to both questions.

Do we know? Applying the modified rescaled range (R/S) statistic confronts the data with a test at once both more precise and more robust than previous estimation techniques. The R/S statistic has shown its versatility and usefulness in a variety of different contexts (Lo [1991] and Haubrich [1993]). Operationally, we can determine what we know about our tests for long-range dependence by using Monte Carlo simulations of their size and power. Not surprisingly, with typical macroeconomic sample sizes we cannot distinguish between fractional exponents of 1.000 and 0.999, but we can distinguish between exponents of 0 and 0.333.

Do we care? Persistence matters directly for making predictions and forecasts. It matters in a more subtle way when making econometric inferences. This is especially true for fractionally differenced processes, which though stationary, are not "strong-mixing" and so in a well-defined probabilistic sense behave very differently than more standard ARMA processes. Furthermore, because the optimal filter depends on the characteristics of the underlying process (Christiano and Fitzgerald [1999] and Baxter and King [1999]), long-term persistence (sometimes called long memory) matters even for estimates of higher frequency objects such as business cycle properties. While for some purposes persistence is less important than how agents decompose shocks into permanent and temporary components (Quah [1990]), the results still depend on the persistence of the time series decomposed and on the persistence of the temporary component. In addition, the univariate approach, in contrast to Quah, has the advantage that it does not assume agents observe more than the econometrician.

Finally, beyond the value of purely statistical explorations, a compelling reason to search for fractional differencing comes from economic theory. We show how fractional differencing arises as a natural consequence of aggregation in real business cycle models. This holds out the promise of more specific guidance in looking for long-range dependence, and conversely, adds another criterion to judge and calibrate macroeconomic models.

This paper examines the stochastic properties of aggregate output from the standpoint of fractionally integrated models. We introduce this type of process in section I, reviewing its main properties, advantages, and weaknesses. Section II develops a simple macroeconomic model that exhibits long-range dependence. Section III employs the modified rescaled range statistic to search for long-range dependence in the data. We conclude in section IV.

I. Review of Fractional Techniques in Statistics

Macroeconomic time series look like neither a random walk nor white noise, suggesting that some compromise or hybrid between white noise and its integral may be useful. Such a concept has been given content through the development of the fractional calculus, that is, differentiation and integration to noninteger orders.[1] The fractional integral of order between 0 and 1 may be viewed as a filter that smooths white noise to a lesser degree than the ordinary integral; it yields a series that is rougher than a random walk but smoother than white noise. Granger and Joyeux (1980) and Hosking (1981) develop the time-series implications of fractional differencing in discrete time. For expositional purposes we review the more relevant properties in this section.

Perhaps the most intuitive exposition of fractionally differenced time series is via their infinite-order autoregressive and moving-average representations. Let X_t satisfy:

$$(1)\quad (1-L)^d X_t = \varepsilon_t ,$$

where ε_t is white noise, d is the degree of differencing, and L denotes the lag operator. If $d = 0$, then X_t is white noise, whereas X_t is a random walk if $d = 1$. However, as Granger and Joyeux (1980) and Hosking (1981) have shown, d need not be an integer. Using the binomial theorem, the AR representation of X_t becomes:

$$(2)\quad A(L)\, X_t = \sum_{k=0}^{\infty} A_k$$

$$(3)\quad L^k X_t = \sum_{k=0}^{\infty} A_k X_{t-k} = \varepsilon_t ,$$

where $A_k \equiv (-1)^k \binom{d}{k}$. The AR coefficients are often re-expressed more directly in terms of the gamma function:

$$(4)\quad A^k \equiv (-1)^k \binom{d}{k} = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)} .$$

By manipulating equation (1) mechanically, X_t may also be viewed as an infinite-order MA process since:

$$(5)\quad X_t = (1-L)^{-d}\varepsilon_t = B(L)\,\varepsilon_t$$

$$B_k = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)} .$$

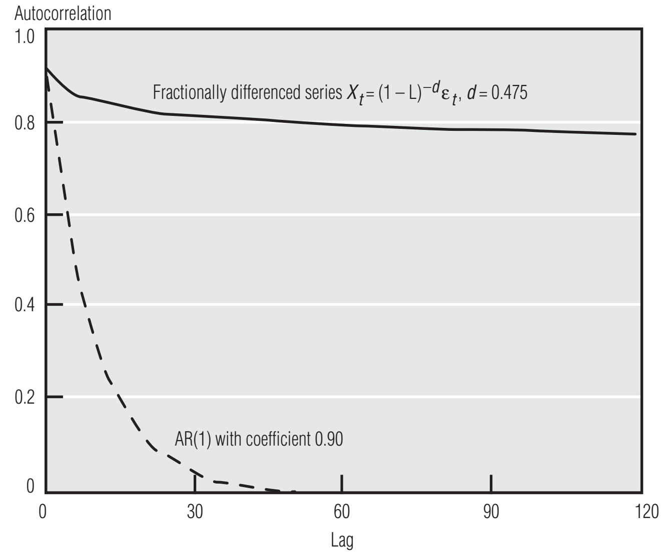
The particular time-series properties of X_t depend intimately on the value of the differencing parameter d. For example, Granger and Joyeux (1980) and Hosking (1981) show that when d is less than $\frac{1}{2}$, X_t is stationary; when d is greater than $-\frac{1}{2}$, X_t is invertible. Although the specification in equation (1) is a fractional integral of pure white noise, the extension to fractional ARIMA models is clear.

The AR and MA representations of fractionally differenced time series have many applications and illustrate the central properties of fractional processes, particularly long-range dependence. The MA coefficients, B_k, give the effect of a shock k periods ahead and indicate the extent to which current levels of the process depend on past values.

1 The idea of fractional differentiation is an old one, dating back to an oblique reference by Leibniz in 1695, but the subject lay dormant until the nineteenth century when Abel, Liouville, and Riemann developed it more fully. Extensive applications have only arisen in this century; see, for example, Oldham and Spanier (1974), who also present an extensive historical discussion. Kolmogorov (1940) was apparently the first to notice its applications in probability and statistics.

F I G U R E 1

Autocorrelation Functions
p_j for $(1 - L)^{-.475}\ \varepsilon_t$



SOURCE: Authors' calculations.

How fast this dependence decays furnishes valuable information about the process. Using Stirling's approximation, we have:

$$(6)\ B_k \approx \frac{k^{d-1}}{\Gamma(d)}$$

for large k. Comparing this with the decay of an AR(1) process highlights a central feature of fractional processes: they decay hyperbolically, at rate k^{d-1}, rather than at the exponential rate of ρ^k for an AR(1). For example, compare in figure 1 the autocorrelation function of the fractionally differenced series $(1 - L)^{-0.475}X_t = \varepsilon_t$ with the AR(1) $X_t = 0.9X_{t-1} + \varepsilon_t$. Although they both have first-order autocorrelations of 0.90, the AR(1)'s autocorrelation function decays much more rapidly.

These representations also show how standard econometric methods can fail to detect fractional processes, necessitating the methods of section III. Although a high order ARMA process can mimic the hyperbolic decay of a fractionally differenced series in finite samples, the large number of parameters required would give the estimation a poor rating from the usual Akaike or Schwartz criteria. An explicitly fractional process, however, captures that pattern with a single parameter, d. Granger and Joyeux (1980) and Geweke and Porter-Hudak (1983) provide empirical support for this by showing that fractional models often outpredict fitted ARMA models.

For large k, the value of B_k measures the response of X_{t+k} to an innovation at time t, a natural metric for persistence. From equation 6, it is immediate that for $0 < d < 1$, $lim_{k\to\infty} B_k = 0$, and asymptotically there is no persistence in a fractionally differenced series, even though the autocorrelations die out very slowly.[2]

This holds true not only for $d < \frac{1}{2}$ (the stationary case), but also for $\frac{1}{2} < d < 1$ (the nonstationary case).

From these calculations, it is apparent that the long-run dependence of fractional processes relates to the slow decay of the autocorrelations, not to any permanent effect. This distinction is important; an IMA(1,1) can have small but positive persistence, but the coefficients will never mimic the slow decay of a fractional process.

The spectrum, or spectral density (denoted $f(\omega)$) of a fractionally differenced process reflects these properties. It exhibits a peak at 0 (unlike the flat spectrum of an ARMA process), but one not as sharp as the random walk's. Given $X_t = (1 - L)^{-d}\varepsilon_t$, the series is clearly the output of a linear system with a white-noise input, so that the spectrum of X_t is:

$$(7)\ f(\omega) = \frac{1}{|1 - z|^{2d}}\frac{\sigma^2}{2\pi} \quad \text{where}$$

$$z \equiv e^{-i\omega}, \qquad \sigma^2 \equiv E\left[\varepsilon_t^2\right].$$

The identity $|1 - z|^2 = 2[1 - \cos(\omega)]$ implies that for small ω we have:

$$(8)\ f(\omega) = c\,\omega^{-2d}, \qquad c \equiv \frac{\sigma^2}{2\pi}.$$

This encompasses the two extremes of a white noise (or a finite ARMA) process and a random walk. For white noise, $d = 0$, and $f(\omega) = c$, while for a random walk, $d = 1$, and the spectrum is inversely proportional to ω^2. A class of processes of current interest in the physics literature, called $1/f$ noise, matches fractionally integrated noise with $d = \frac{1}{2}$.

■ **2** There has been some confusion in the literature on this point. Geweke and Porter-Hudak (1983) argue that $\lim_{k\to\infty} B_k > 0$, which, in their terminology, is expressed as $C(1) > 0$. They correctly point out that Granger and Joyeux (1980) have made an error, but then incorrectly claim that $C(1) = 1/\Gamma(d)$. If our equation 6 is correct, then it is apparent that $C(1) = 0$ (which agrees with Granger [1980] and Hosking [1981]). Therefore, the focus of the conflict lies in the approximation of the ratio $\Gamma(k + d)/\Gamma(k + 1)$ for large k. We have used Stirling's approximation. However, a more elegant derivation follows from the functional analytic definition of the gamma function as the solution to the following recursive relation (see, for example, Iyanaga and Kawada [1980, section 179.A]):

$$\Gamma(x + 1) = x\,\Gamma(x)$$

and the conditions:

$$\Gamma(1) = 1 \qquad \lim_{n\to\infty} \frac{\Gamma(x + n)}{n^x\Gamma(n)} = 1.$$

II. A Simple Macroeconomic Model with Long-Term Dependence

Economic insight requires more than a consensus on the Wold representation of GDP; it demands a falsifiable model based on the tastes and technology of the actual economy. As Wesley Claire Mitchell (1927, p. 230) wrote, "We stand to learn more about economic oscillations at large and about business cycles in particular, if we approach the problem of trends as theorists, than if we confine ourselves to strictly empirical work."

Thus, before testing for long-run dependence, we develop a simple model where aggregate output exhibits long-run dependence. This model presents one reason that macroeconomic data might show the particular stochastic structure for which we test. It also shows that models can restrict the fractional differencing properties of time series, so that our test holds promise for distinguishing between competing theories. Furthermore, the maximizing model presented below connects long-range dependence to central economic concepts of productivity, aggregation, and the limits of the representative agent paradigm.

A Simple Real Model

One plausible mechanism for generating long-run dependence in output, which we will mention here and not pursue, is that production shocks themselves follow a fractionally integrated process. This explanation for persistence follows that used by Kydland and Prescott (1982). In general, such an approach begs the question, but in the present case evidence from geophysical and meteorological records suggests that many economically important shocks have long-run correlation properties. Mandelbrot and Wallis (1969), for instance, find long-run dependence in rainfall, riverflows, earthquakes, and weather (measured by tree rings and sediment deposits).[3]

A more satisfactory model explains the time-series properties of data by producing them despite white-noise shocks. This section develops such a model with long-run dependence, using a linear quadratic version of the real business cycle model of Long and Plosser (1983) and aggregation results due to Granger (1980).[4] In our multisector model, the output of each industry (or island) will follow an AR(1) process. Aggregate output with N sectors will not follow an AR(1) but rather an ARMA(N, N–1). This makes dynamics with even a moderate number of sectors unmanageable. Under fairly general conditions, however, a simple fractional process will closely approximate the true ARMA specification.

Consider a model economy with many goods and a representative agent who chooses a production and consumption plan. The infinitely lived agent inhabits a linear quadratic version of the real business cycle model. The agent has quadratic utility, with a lifetime utility function of $U = \sum \beta^t u(C_t)$, where C_t is an $N \times 1$ vector denoting period-t consumption of each of the N goods in our economy. Each period's utility function, $u(C_t)$, is given by

$$(9) \quad u(C_t) = C_t' \iota - \tfrac{1}{2} C_t' B C_t,$$

where ι is an $N \times 1$ vector of ones. In anticipation of the aggregation considered later, we assume B to be diagonal so that $C_t' BC_t = \sum b_{ii} C_{it}^2$. The agents face a resource constraint: total output Y_t may be either consumed or saved, thus:

$$(10) \quad C_t + S_t \iota = Y_t,$$

where the i,j-th entry S_{ijt} of the $N \times N$ matrix S_t denotes the quantity of good j invested in process i at time t, and it is assumed that any good Y_{jt} may be consumed or invested. Output is determined by a random linear technology:

$$(11) \quad Y_t = AS_t + \varepsilon_t,$$

where A is the matrix of input–output coefficients a_{ij}, and ε_t is a (vector) random production shock, whose value is realized at the beginning of period $t+1$. To focus on long-range dependence we restrict A's form. Thus, each sector uses only its own output as input, yielding a diagonal A matrix and allowing us to simplify notation by defining $a_i \equiv a_{ii}$. This might occur, for example, with a number of distinct islands producing different goods. To further simplify the problem, all commodities are perishable and capital depreciates at a rate of 100 percent.

In this case, the dynamic programming problem of solving for optimal consumption and investment policies reduces to the familiar optimal stochastic linear regulator problem (see Sargent [1987], section 1.8, for an excellent exposition). Given the simple diagonal form of the A matrix, which corresponds to an assumption that each sector uses only its own output as input, the problem simplifies even further.

3 For a related mechanism creating fractional intergration by aggregating shocks of differing duration, see Parke (1999). Abadir and Talmain (undated) use aggregation over heterogeneous firms in a setting of monopolistic competition.

4 Dupor (1999) is skeptical of the ability of multisector models to match aggregate time series data but does not consider long-range dependence.

The chosen quantities of consumption and investment/savings have the following closed-form solutions:

$$(12)\ S_{it} = \frac{b_i}{b_i - 2\beta P_i a_i^2} Y_{it} + \frac{\beta q_i a_i - 1}{b_i - 2\beta P_i a_i^2}$$

$$(13)\ C_{it} = \frac{2\beta P_i a_i^2}{2\beta P_i a_i^2 - b_i} Y_{it} + \frac{\beta q_i a_i - 1}{2\beta P_i a_i^2 - b_i},$$

where:

$$(14)\ P_i \equiv b_i \left[\frac{a_i - \sqrt{(1+4\beta)a_i^2 - 4}}{4\beta a_i}\right],$$

and q_i are fixed constants given by the matrix Riccati equation that results from the recursive definition of the value function.

The simple form of the optimal consumption and investment decision rules comes from the quadratic preferences and the linear production function. Two qualitative features bear emphasizing. First, higher output today will increase both current consumption and current investment, thus increasing future output. Even with 100 percent depreciation, no durable commodities, and independent and identically distributed (i.i.d.) production shocks, the time-to-build feature of investment induces serial correlation. Second, the optimal choices do not depend on the uncertainty present. This certainty-equivalence feature is an artifact of the linear-quadratic combination.

The time series of output can now be calculated from the production function, equation (11), and the decision rule, equation (12). Quantity dynamics then come from the difference equation:

$$(15)\ Y_{it+1} = \frac{a_i b_i}{b_i - 2\beta P_i a_i^2} Y_{it} + K_i + \varepsilon_{it+1}$$

or

$$(16)\ Y_{it+1} = \alpha_i Y_{it} + K_i + \varepsilon_{it+1},$$

where α_i is a function of the utility parameters and of a_i, the input-output coefficient of the industry, and K_i is some fixed constant. The key qualitative property of quantity dynamics summarized by equation (16) is that output, Y_{it}, follows an AR(1) process. Higher output today implies higher output in the future. That effect dies off at a rate that depends on the parameter α_i, which in turn depends on the underlying preferences and technology.

The simple output dynamics for a single industry or island neither mimics business cycles nor exhibits long-run dependence. However, aggregate output, the sum across all sectors, will show such dependence, which we demonstrate here by applying the aggregation results of Granger (1980,1988).

It is well-known that the sum of two series, X_t and Y_t, each AR(1) with independent error, is an ARMA(2,1) process. Simple induction then implies that the sum of *N* independent AR(1) processes with distinct parameters has an ARMA(*N*, *N*–1) representation. With over six million registered businesses in America, the dynamics can be incredibly rich, and the number of parameters unmanageably huge. The common response to this problem is to pretend that many different firms (or islands) have the same AR(1) representation for output, which reduces the dimension of aggregate ARMA process. This "cancelling of roots" requires identical autoregressive parameters. An alternative approach, due to Granger, reduces the scope of the problem by showing that the ARMA process approximates a fractionally integrated process, and thus summarizes the many ARMA parameters in a parsimonious manner. Though we consider the case of independent sectors, dependence is easily handled.

Consider the case of *N* sectors, with the productivity shock for each serially uncorrelated and independent across sectors. Furthermore, let the sectors differ according to the productivity coefficient, a_i. This implies differences in α_i, the autoregressive parameter for sector *i*'s output, Y_{it}. One of our key results is that under some distributional assumptions on the α_i's, aggregate output, Y_t^a, follows a fractionally integrated process, where:

$$(17)\ Y_t^a \equiv \sum_{i=1}^{N} Y_{it}.$$

To show this, we approach the problem from the frequency domain and apply spectral methods, which often simplify problems of aggregation.[5] Let $f(\omega)$ denote the spectrum (spectral density function) of a random variable, and let $z = e^{-i\omega}$. From the definition of the spectrum as the Fourier transform of the autocovariance function, the spectrum of Y_{it} is:

$$(18)\ f_i(\omega) = \frac{1}{|1-\alpha_i z|^2}\ \frac{\sigma_i^2}{2\pi}.$$

Similarly, independence implies that the spectrum of Y_t^a is

$$(19)\ f_i(\omega) = \sum_{i=1}^{N} f_i(\omega).$$

5 See Theil (1954).

The α_i's measure an industry's average output for given input. This attribute of the production function can be thought of as a drawing from nature, as can the variance of the productivity shocks, ε_{it}, for each sector. Thus, it makes sense to think of the a_i's as independently drawn from a distribution $G(a)$ and the α_i's as drawn from $F(\alpha)$. Provided that the ε_{it} shocks are independent of the distribution of α_i's, the spectral density of the sum can be written as:

$$(20)\ f_i(\omega) = \frac{N}{2\pi} E[\sigma^2] \bullet \int \frac{1}{|1-\alpha_i z|^2} dF(\alpha).$$

If the distribution $F(\alpha)$ is discrete, so that it takes on $m(< N)$ values, Y_t^a will be an ARMA $(m,m-1)$ process. A more general distribution leads to a process no finite ARMA model can represent. To further specify the process, take a particular distribution for F, in this case a variant of the beta distribution.[6] In particular, let α^2 have a beta distribution $\beta(p,q)$, which yields the following density function for α:

$$(21)\ dF(\alpha) = \begin{cases} \frac{2}{\beta(p,q)} \alpha^{2p-1}(1-\alpha^2)^{q-1} d\alpha, & 0 \le \alpha \le 1; \\ 0, & \text{otherwise.} \end{cases}$$

with $p,q > 0$.[7] Obtaining the Wold representation of the resulting process requires a little more work. First note that:

$$(22)\ 1/|1-\alpha z|^2 = \frac{1}{2(1-\alpha^2)} \left[\frac{1+\alpha z}{1-\alpha z} + \frac{1+\alpha \bar{z}}{1-\alpha \bar{z}}\right],$$

where $\bar{z}$ denotes the complex conjugate of z, and the terms in parentheses can be further expanded by long division. Substituting this expansion and the beta distribution, equation (21), into the expression for the spectrum and simplifying (using the relation $z + \bar{z} = 2\cos(\omega)$) yields:

$$(23)\ f(\omega) = \int_0^1 [2 + 2\sum_{k=1}^{\infty} \alpha^k \cos(k\,\omega)].$$

Then the coefficient of $\cos(k\omega)$ is

$$(24)\ \int_0^1 \frac{2\alpha^k}{\beta(p,q)} \alpha^{2p-1}(a-\alpha^2)^{q-1} d\alpha.$$

Since the spectral density is the Fourier transform of the autocovariance function, equation (24) is the k-th autocovariance of Y_t^a. Furthermore, because the integral defines a beta function, equation (24) simplifies to $\beta(p+k/2, q-1)/\beta(p,q)$. Dividing by the variance gives the autocorrelation coefficients, which reduce to

$$(25)\ \rho(k) = \frac{\Gamma(p+q-1)}{\Gamma(p)} \frac{\Gamma(p+\frac{k}{2})}{\Gamma(p+\frac{k}{2}+q+1)},$$

which, again using the result from Stirling's approximation, $\Gamma(a+k)/\Gamma(b+k) \approx k^{a-b}$, is proportional (for large lags) to k^{1-q}. Thus aggregate output Y_t^a follows a fractionally integrated process of order $d = 1 - \frac{q}{2}$. Furthermore, as an approximation for long lags, this does not necessarily rule out interesting correlations at higher frequencies, such as those of the business cycle. Similarly, co-movements can arise, as the fractionally integrated income process may induce fractional integration in other observed time series. Two additional points are worth emphasizing. First, the beta distribution need not be over (0,1) to obtain these results, only over $(a,1)$. Second, it is indeed possible to vary the a_i's so that α_i has a beta distribution.

In principle, all parameters of the model may be estimated, from the distribution of production function parameters to the variance of output shocks. Empirical estimates of production function parameters (such as those in Jorgenson, Gollop, and Fraumeni [1987]) reveal a large dispersion, suggesting the plausibility and significance of the simple model presented in this section.

Although the original motivation of our real business cycle model was to illustrate how long-range dependence could arise naturally in an economic system, our results have broader implications for general macroeconomic modeling. They show that moving to a multiple-sector real business cycle model introduces not unmanageable complexity, but qualitatively new behavior that can be quite manageable. Our findings also show that calibrations aimed at matching only a few first and second moments can similarly hide major differences between models and the data, missing long-range dependence properties. While widening the theoretical horizons of the paradigm, they therefore also widen the potential testing of such theories.

6 Granger (1980) conjectures that this particular distribution is not essential.

7 For a discussion of the variety of shapes the beta distribution takes as p and q vary, see Johnson and Kotz (1970).

III. R/S Analysis of Real Output

The results of section II show that simple aggregation may be one source of long-range dependence in the business cycle. In this section we employ a method for detecting long memory and apply it to real GDP. The technique is based on a simple generalization of a statistic first proposed by the English hydrologist Harold Edwin Hurst (1951), which has subsequently been refined by Mandelbrot (1972, 1975) and others.[8]

Our generalization of Mandelbrot's statistic (called the "rescaled range" or "range over standard deviation" or R/S) enables us to distinguish between short- and long-run dependence, in a sense to be made precise below.

We define our notions of short and long memory and present the test statistic below. Then we present the empirical results for real GDP; we find long-range dependence in log-linearly detrended output, but considerably less dependence in the growth rates. To interpret these results, we perform several Monte Carlo experiments under two null and two alternative hypotheses and report these results.

The Rescaled Range Statistic

We test for fractional differencing using Lo's modification of the modified rescaled range (R/S) statistic. In particular, we define short-range dependence as Rosenblatt's (1956) concept of "strong-mixing," a measure of the decline in statistical dependence of two events separated by successively longer spans of time. Heuristically, a time series is strong-mixing if the maximal dependence between any two events becomes trivial as more time elapses between them. By controlling the rate at which the dependence between future events and those of the distant past declines, it is possible to extend the usual laws of large numbers and central limit theorems to dependent sequences of random variables. Such mixing conditions have been used extensively by White (1982), White and Domowitz (1984), and Phillips (1987) for example, to relax the assumptions that ensure consistency and asymptotic normality of various econometric estimators. We adopt this notion of short-range dependence as part of our null hypothesis. As Phillips (1987) observes, these conditions are satisfied by a great many stochastic processes, including all Gaussian finite-order stationary ARMA models. Moreover, the inclusion of a moment condition also allows for heterogeneously distributed sequences (such as those exhibiting heteroscedasticity), an especially important extension in view of the nonstationarities of real GDP.

Fractionally differenced models, however, possess autocorrelation functions that decay at much slower rates than those of weakly dependent processes and violate the conditions of strong mixing. More formally, let X_t denote the first difference of log-GDP; we assume that:

(26) $X_t = \mu + \varepsilon_t$,

where μ is an arbitrary but fixed parameter. For the null hypothesis H, assume that the sequence of disturbances, $\{\varepsilon_t\}$, satisfies the following conditions:

(A1) $E[\varepsilon_t] = 0$ *for all t.*

(A2) $\sup_t E[|\varepsilon_t|^\beta] < \infty$ *for some* $\beta > 2$.

(A3) $\sigma^2 = \lim_{n\to\infty} E\left[\frac{1}{n}\left(\sum_{j=1}^{n}\varepsilon_j\right)^2\right]$ *exists and* $\sigma^2 > 0$.

(A4) $\{\varepsilon_t\}$ *is strong-mixing with mixing coefficients* α_k *that satisfy:*[9]

$$\sum_{k=1}^{\infty} \alpha_k^{1-\frac{2}{\beta}} < \infty .$$

Condition (A1) is standard. Conditions (A2)–(A4) allow dependence and heteroskedasticity, but prevent them from being too dominant. Thus, short-range dependent processes such as finite-order ARMA models are included in this null hypothesis, as are models with conditional heteroskedasticity. Unlike the statistic used by Mandelbrot, the modified R/S statistic is robust to short-range dependence. A more detailed discussion of these conditions may be found in Phillips (1987) and Lo (1991).

To construct the modified R/S statistic, consider a sample $X_1, X_2, \ldots, X_n$ and let $\bar{X}_n$ denote the sample mean $\frac{1}{n}\sum_j X_j$. Then the modified R/S statistic, which we shall call Q_n, is given by:

$$(27)\quad Q_n \equiv \frac{1}{\hat{\sigma}_n(q)}\left[Max_{1\le k\le n}\sum_{j=1}^{k}(X_j - \bar{X}_n) - \mathrm{Min}_{1\le k\le n}\sum_{j=1}^{k}(X_j - \bar{X}_n)\right],$$

8 See Mandelbrot and Taqqu (1979) for further references.

9 Let $\{\varepsilon_t(\omega)\}$ be a stochastic process on the probability space $(\Omega, \mathcal{F}, P)$ and define:

$$\alpha(\mathcal{A},\mathcal{B}) \equiv \sup_{\{A\in\mathcal{A}, B\in\mathcal{B}\}} |P(A\cap B) - P(A)P(B)| \qquad \mathcal{A}\subset\mathcal{F}, \mathcal{B}\subset\mathcal{F}$$

The quantity $\alpha(\mathcal{A},\mathcal{B})$ is a measure of the dependence between the two σ-fields $\mathcal{A}$ and $\mathcal{B}$ in $\mathcal{F}$. Denote by $\mathcal{B}_s^t$ the Borel σ-field generated by $\{\varepsilon_s(\omega),\ldots,\varepsilon_t(\omega)\}$, i.e., $\mathcal{B}_s^t \equiv \sigma(\varepsilon_s(\omega),\ldots,\varepsilon_t(\omega)) \subset \mathcal{F}$. Define the coefficients α_k as:

$$\alpha_k \equiv \sup_j \alpha(\mathcal{B}_{-\infty}^{j}, \mathcal{B}_{j+k}^{\infty})$$

Then $\{\varepsilon_t(\omega)\}$ is said to be strong-mixing if $\lim_{k\to\infty} \alpha_k = 0$. For further details, see Rosenblatt (1956), White (1984), and the papers in Eberlein and Taqqu (1986).

TABLE 1(a)

Fractiles of the Distribution $F_V(v)$

$P(V < v)$	.005	.025	.050	.100	.200	.300	.400	.500
v	0.721	0.809	0.861	0.927	1.018	1.090	1.157	1.223
$P(V < v)$	.543	.600	.700	.800	.900	.950	.975	.995
v	$\sqrt{\frac{\pi}{2}}$	1.294	1.374	1.473	1.620	1.474	1.862	2.098

SOURCE: Authors' calculations.

TABLE 1(b)

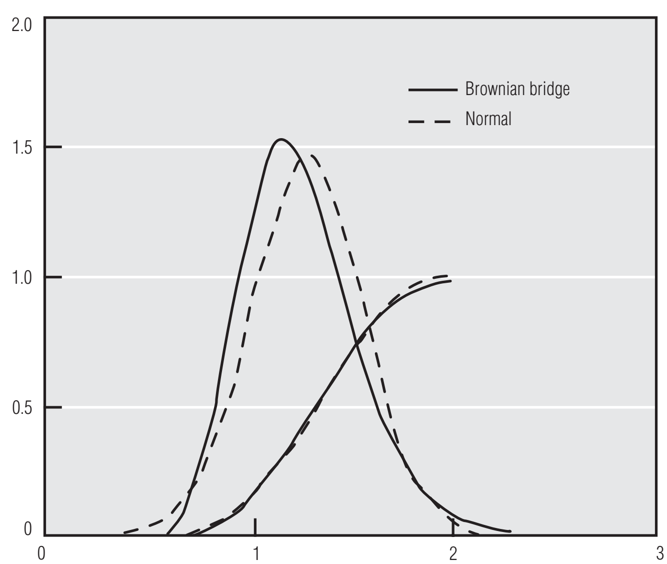
Symmetric Confidence Intervals about the Mean

$P\left(\sqrt{\frac{\pi}{2}} - \gamma < V < \sqrt{\frac{\pi}{2}} + \gamma\right)$	γ
.001	0.748
.050	0.519
.100	0.432
.500	0.185

SOURCE: Authors' calculations.

FIGURE 2

Distribution and Density Function of the Range V of a Brownian Bridge



SOURCE: Authors' calculations.

where

$$(28)\quad \hat{\sigma}_n^2(q) \equiv \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2 + \frac{2}{n}\sum_{j=1}^{q}\omega_j(q)\left\{\sum_{i=j+1}^{n}(X_i - \bar{X}_n)(X_{i-j} - \bar{X}_n)\right\}$$

$$(29)\quad = \hat{\sigma}_x^2 + 2\sum_{j=1}^{q}\omega_j(q)\,\hat{\gamma}_j$$

$$\omega_j(q) \equiv 1 - \frac{j}{q+1} \qquad q < n$$

and $\hat{\sigma}_x^2$ and $\hat{\gamma}_j$ are the usual sample variance and autocovariance estimators of X. Q_n is the range of partial sums of deviations of X_j from its mean, normalized by an estimator of the partial sum's standard deviation divided by n. The estimator $\hat{\sigma}_n(q)$ involves not only sums of squared deviations of X_j, but also its weighted autocovariances up to lag q; the weights $\omega_j(q)$ are those suggested by Newey and West (1987), and always yield a positive estimator $\hat{\sigma}_n^2(q)$.[10]

Intuitively, the numerator in equation (27) measures the memory in the process via the partial sums. White noise does not stay long above the mean: positive values are soon offset by negative values. A random walk will stay above or below 0 for a long time, and the partial sums (positive or negative) will grow quickly, making the range large. Fractional processes fall in between. Mandelbrot (1972) refers to their behavior as the "Joseph effect"—seven fat and seven lean years. The denominator normalizes not only by the variance but also by a weighted average of the autocovariances. This innovation over Hurst's (1951) R/S statistic provides the robustness to short-range dependence.

The choice of the truncation lag, q, is a delicate matter. Although q must increase with (albeit at a slower rate than) the sample size, Monte Carlo evidence suggests that when q becomes large relative to the number of observations, asymptotic approximations may fail dramatically.[11] However, q cannot be too small or the effects of higher-order autocorrelations may not be captured. The choice of q is clearly an empirical issue and must therefore be chosen with some consideration of the data at hand.

The partial sums of white noise constitute a random walk, so $Q_n(q)$ grows without bound as n increases. A further normalization makes the statistic easier to work with and interpret:

$$(30)\quad V_n(q) \equiv \frac{Q_n(q)}{\sqrt{n}}$$

10 $\hat{\sigma}_n^2(q)$ is also an estimator of the special density function of X_t at frequency zero, using a Bartlett window.

11 See, for example, Lo and MacKinlay (1989).

TABLE 2

R/S Analysis of Real GDP

Series	$\tilde{V}_n$	$V_n(1)$	$V_n(2)$	$V_n(3)$	$V_n(4)$	$V_n(5)$	$V_n(6)$	$V_n(7)$	$V_n(8)$
US Log First-Difference	1.092	0.973	0.933	0.934	0.971	1.032	1.082	1.115	1.139
Percentage Bias of $\tilde{V}_n$		(6.0)	(8.2)	(8.1)	(6.1)	(2.9)	(0.5)	(–1.0)	(–2.1)
S Log Detrended	2.374	1.741	1.479	1.337	1.252	1.198	1.160	1.134	1.116
Percentage Bias of $\tilde{V}_n$		(16.8)	(26.7)	(33.2)	(37.7)	(40.8)	(43.1)	(44.7)	(45.8)
CAN Log First-Difference	1.254	1.116	1.042	1.018	1.024	1.045	1.073	1.096	1.132
Percentage Bias of $\tilde{V}_n$		(6.0)	(9.7)	(11.0)	(10.7)	(9.5)	(8.1)	(7.0)	(5.2)
CAN Log Detrended	3.410	2.458	2.048	1.813	1.660	1.552	1.472	1.410	1.360
Percentage Bias of $\tilde{V}_n$		(17.8)	(29.1)	(37.2)	(43.3)	(48.2)	(52.2)	(55.5)	(58.3)
GER Log First-Difference	1.357	1.185	1.159	1.158	1.176	1.203	1.235	1.278	1.325
Percentage Bias of $\tilde{V}_n$		(7.0)	(8.2)	(8.3)	(7.4)	(6.2)	(4.8)	(3.1)	(1.2)
GER Log Detrended	4.241	3.052	2.539	2.242	2.044	1.903	1.796	1.712	1.643
Percentage Bias of $\tilde{V}_n$		(17.9)	(29.2)	(37.5)	(44.0)	(49.3)	(53.7)	(57.4)	(60.6)
UK Log First-Difference	1.051	0.907	0.851	0.853	0.887	0.920	0.961	0.993	1.031
Percentage Bias of $\tilde{V}_n$		(7.7)	(11.1)	(11.0)	(8.9)	(6.9)	(4.6)	(2.9)	(1.0)
UK Log Detrended	4.637	3.327	2.760	2.431	2.213	2.055	1.933	1.837	1.757
Percentage Bias of $\tilde{V}_n$		(18.1)	(29.6)	(38.1)	(44.8)	(50.2)	(54.9)	(58.9)	(62.5)
GDP Log First-Difference	1.391	1.201	1.107	1.066	1.057	1.069	1.087	1.109	1.132
Percentage Bias of $\tilde{V}_n$		(7.6)	(12.1)	(14.3)	(14.7)	(14.1)	(13.1)	(12.0)	(10.9)
GDP Log Detrended	5.612	3.999	3.290	2.873	2.592	2.387	2.229	2.103	2.000
Percentage Bias of $\tilde{V}_n$		(18.5)	(30.6)	(39.8)	(47.1)	(53.3)	(58.7)	(63.4)	(67.5)

NOTE: US, CAN, GER, and UK refer to the annual Maddison series for those countries 1870–1994. GDP refers to the quarterly U.S. series 1947:QI–1999:QI. The classical rescaled range $\tilde{V}_n$ and the modified rescaled range $V_n(q)$ are reported. Under a null hypothesis of short-range dependence, the limiting distribution of $V_n(q)$ is the range of a Brownian bridge, which has a mean of $\sqrt{\pi/2}$. Fractiles are given in table 1; the 95 percent confidence interval with equal probabilities in both tails is [0.809, 1.862]. Entries in the %-Bias rows are computed as $[\tilde{V}_n / V_n(q)^{1/2} - 1] \bullet 100$ and are estimates of the bias of the classical R/S statistic in the presence of short-term dependence.
SOURCE: Authors' calculations.

The limiting distribution of $V_n(q)$ is derived by Lo (1991), and its most commonly-used values are reported in tables 1(a) and 1(b). Table 1(a) reports the fractiles of the limiting distribution while table 1(b) reports the symmetric confidence intervals about the mean. The moments of the limiting distribution are also easily computed using f_V, the density of the random variable V to which $V_n(q)$ converges in distribution; it is straightforward to show that $E[V] = \sqrt{\frac{\pi}{2}}$, and $E[V^2] = \frac{\pi^2}{6}$; thus the mean and standard deviation of V are approximately 1.25 and 0.27, respectively. The distribution and density functions are plotted in figure 2. Observe that the distribution is positively skewed, and most of its mass falls between $\frac{3}{4}$ and 2.

Empirical Results for Real Output

We apply our test to several time series of real output: quarterly U.S. postwar real GDP from 1947:QI to 1999:QI, and the annual Maddison (1995) OECD series for the United States, Canada, Germany, and the United Kingdom from 1870 to 1994. These results are reported in table 2. Entries in the first numerical column are estimates of the classical rescaled range, $\tilde{V}_n$, which is not robust to short-range dependence. The next eight columns are estimates of the modified rescaled range $V_n(q)$ for values of q from 1 to 8. Recall that q is the truncation lag of the estimator of the spectral density at frequency zero. Reported in parentheses below the entries for $V_n(q)$ is an estimate of the percentage bias of the statistic $\tilde{V}_n$, which is computed as $100 \cdot [(\tilde{V}_n/V_n(q)) - 1]$.

The first row of numerical entries in table 2 indicate that the null hypothesis of short-range dependence for the first-difference of log-GDP cannot be rejected for any value of q. The classical rescaled range statistic also supports the null hypothesis, as do the results for the Maddison series. On the other hand, when we log-linearly detrend real GDP, the results differ considerably. Looking at the results for the annual data in table 2 shows that short-range dependence may be rejected for log-linearly detrended output using the classical statistic for the United States and with q values from 1 to 2 for Canada, 1 to 5 for Germany, and 1 to 6 for the United Kingdom. For quarterly U.S. data, short-term dependence is rejected for all q up to 8. That the rejections are weaker for larger q is not surprising since additional noise arises from estimating higher-order autocorrelations.

The values reported in table 2 are qualitatively consistent with other empirical investigations of fractional processes in GNP, such as Diebold and Rudebusch (1989) and Sowell (1992). For first-differences, the R/S statistic falls below the mean, suggesting a negative fractional exponent, or in level terms, an exponent between 0 and 1. Furthermore, though earlier papers produce point estimates, they do not lead to a rejection of the hypothesis of short-term dependence because of imprecise estimates. For example, the 2 standard deviation error bounds for two point estimates of Diebold and Rudebusch (1989), $d = 0.9$ and 0.52, are [0.42, 1.38] and [–0.06, 1.10], respectively.

Taken together, these results confirm the unit root findings of Campbell and Mankiw (1987), Nelson and Plosser (1982), Perron and Phillips (1987), and Stock and Watson (1986). That there are more significant autocorrelations in log-linearly detrended GDP is precisely the spurious periodicity suggested by Nelson and Kang (1981). Moreover, the trend plus stationary noise model of GDP is not contained in our null hypothesis; hence our failure to reject the null hypothesis is also consistent with the unit root model.[12] To see this, observe that if log-GDP y_t were trend stationary, that is,

$$(31) \quad y_t = \alpha + \beta t + \eta_t$$

where η_t is stationary white noise, then its first-difference, X_t, is simply $X_t = \beta + \varepsilon_t$, where $\varepsilon_t \equiv \eta_t - \eta_{t-1}$. But this innovations process violates our assumption (A3) and is therefore not contained in our null hypothesis.

Sowell (1992) has used estimates of d to argue that the trend-stationary model is correct. Following the lead of Nelson and Plosser (1982), Sowell checks if the d parameter for the first-differenced series is close to 0 as the unit root specification suggests, or close to –1 as the trend-stationary specification suggests. His estimate of d is in the general range of –0.6 to 0.2, providing some evidence that the trend-stationary interpretation is correct. Even in his case though, the standard errors tend to be rather large, on the order of 0.3. Although our procedure yields no point estimate of d, our results do seem to rule out the trend-stationary case.

To conclude that the data support the null hypothesis because our statistic fails to reject it is, of course, premature since the size and power of our test in finite samples is yet to be determined. We perform illustrative Monte Carlo experiments and report the results in the next section.

The Size and Power of the Test

To evaluate the size and power of our test in finite samples, we perform several illustrative Monte Carlo experiments for a sample size of 208 observations, corresponding to the number of quarterly observations of real GDP growth from 1947:QII to 1999:QI.[13] We simulate two null hypotheses: independently and identically distributed increments, and increments that follow an ARMA(2,2) process. Under the i.i.d. null hypothesis, we fix the mean and standard deviation of our random deviates to match the sample mean and standard deviation of our quarterly data set: 8.221×10^{-3} and 1.0477×10^{-2}, respectively. To choose parameter values for the ARMA(2,2) simulation, we estimate the model:

$$(32) \quad (1 - \phi_1 L - \phi_2 L^2) y_t = \mu + (1 + \theta_1 L + \theta_2 L^2)\varepsilon_t \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

using nonlinear least squares. The parameter estimates are (with standard errors in parentheses):

$$\hat{\phi}_1 = \underset{(0.1678)}{1.3423}, \quad \hat{\theta}_1 = \underset{(0.1839)}{1.0554}$$

$$\hat{\phi}_2 = \underset{(0.1198)}{-0.7065}, \quad \hat{\theta}_2 = \underset{(0.1377)}{-0.5200}$$

$$\hat{\mu} = \underset{(0.0008)}{0.0082}, \quad \hat{\sigma}_\varepsilon = 0.0097$$

12 Of course, this may be the result of low power against stationary but near-integrated processes, and it must be addressed by Monte Carlo experiments.

13 Simulations were performed on a DEC Alphaserver 2100 4/275 using a Gauss random number generator; each experiment comprised 10,000 replications.

TABLE 3

Finite Sample Distribution of the Modified R/S Statistic under IID and ARMA (2,2) Null Hypotheses for the First-Difference of Real Log GNP

q	Min.	Max.	Mean	S.D.	Size 1%–Test	Size 5%–Test	Size 10%–Test
IID Null							
0	0.527	2.468	1.175	0.266	0.002	0.030	0.061
1.5	0.525	2.457	1.171	0.253	0.015	0.069	0.121
1	0.548	2.342	1.177	0.258	0.001	0.027	0.052
2	0.548	2.251	1.180	0.251	0.001	0.024	0.054
3	0.555	2.203	1.183	0.245	0.000	0.021	0.052
4	0.572	2.156	1.187	0.240	0.000	0.020	0.050
5	0.592	2.098	1.190	0.234	0.000	0.018	0.046
6	0.622	2.058	1.193	0.228	0.000	0.015	0.044
7	0.637	2.031	1.197	0.223	0.000	0.012	0.041
8	0.657	1.981	1.200	0.218	0.000	0.010	0.038
ARMA (2,2) Null							
0	0.654	2.864	1.411	0.314	0.025	0.152	0.245
6.8	0.610	2.200	1.177	0.229	0.009	0.041	0.084
1	0.570	2.473	1.227	0.269	0.004	0.039	0.083
2	0.533	2.269	1.134	0.246	0.001	0.014	0.035
3	0.517	2.190	1.094	0.233	0.000	0.007	0.021
4	0.522	2.139	1.086	0.226	0.000	0.006	0.016
5	0.543	2.101	1.099	0.223	0.000	0.005	0.017
6	0.562	2.035	1.123	0.221	0.000	0.006	0.020
7	0.587	2.011	1.149	0.220	0.000	0.007	0.024
8	0.620	1.995	1.171	0.218	0.000	0.008	0.029

NOTE: The Monte Carlo experiments under the two null hypotheses are independent and consist of 10,000 replications each, for a sample size $n = 208$. Parameters of the i.i.d. simulations were chosen to match the sample mean and variance of quarterly real GNP growth rates from 1947:QII to 1999:QI; parameters of the ARMA (2,2) were chosen to match point estimates of an ARMA (2,2) model fitted to the same data set. Entries in the column labelled "q" indicate the number of lags used to compute the R/S statistic; a lag of 0 corresponds to Mandelbrot's classical rescaled range, and a noninteger lag value corresponds to the average (across replications) lag value used according to Andrews's (1991) optimal lag formula. Standard errors for the empirical size may be computed using the usual normal approximation; they are 9.95×10^{-4}, 2.18×10^{-3}, and 3.00×10^{-3} for the 1, 5, and 10 percent tests, respectively.
SOURCE: Authors' calculations.

Table 3 reports the results of both null simulations. It is apparent from the i.i.d. null panel of table 3 that when serial correlation is not a problem, the classical and modified rescaled range statistics perform similarly. The 5 percent test using the classical statistic rejects 3 percent of the time: the modified R/S with $q = 4$ rejects 2 percent of the time. As the number of lags increases to 8, the test becomes more conservative. Under the ARMA(2,2) null hypothesis, however, it is apparent that modifying the R/S by the spectral density estimator $\hat{\sigma}_n^2(q)$ is critical; the size of a 5 percent test based on the classical R/S is 15.2 percent, whereas the corresponding size using the modified R/S statistic with $q = 1$ is 3.9 percent. As before, the test becomes more conservative when q is increased.

Table 3 also reports the size of tests using the modified rescaled range when the lag length q is chosen optimally using Andrews's (1991) procedure. This data-dependent procedure entails computing the first-order autocorrelation coefficient $\hat{\rho}(1)$ and then setting the lag length to be the integer-value of $\bar{M}_n$, where:[14]

$$(33)\quad \bar{M}_n \equiv \left(\frac{3\hat{\alpha}n}{2}\right)^{1/3} \qquad \hat{\alpha} \equiv \frac{4\hat{\rho}^2}{(1-\hat{\rho}^2)^2}\,.$$

Under the i.i.d. null hypothesis, Andrews's formula yields a 5 percent test with empirical size 6.9 percent; under the ARMA(2,2) alternative, the corresponding size is 4.1 percent. Although significantly different from the nominal value, the empirical size of tests based on Andrews's formula may not be economically important. In addition to its optimality properties, the procedure has the advantage of eliminating a dimension of arbitrariness in performing the test.

14 In addition, Andrews's procedure requires weighting the autocovariances by $1 - \frac{j}{\bar{M}_n}$ $(j = 1,\ldots,[\bar{M}_n])$ in contrast to Newey and West's (1987) $1 - \frac{j}{q+1}$ $(j = 1,\ldots,q)$, where q is an integer and $\bar{M}_n$ need not be.

TABLE 4

Power of the Modified R/S Statistics under a Gaussian Fractionally Differenced Alternative with Differencing Parameters d = 1/3, – 1/3

q	Min.	Max.	Mean	S.D.	Power 1%–Test	Power 5%–Test	Power 10%–Test
d = 1/3							
0	0.888	5.296	2.551	0.673	0.722	0.842	0.890
6.1	0.665	2.569	1.577	0.305	0.039	0.193	0.310
1	0.825	4.112	2.149	0.528	0.511	0.680	0.758
2	0.752	3.497	1.936	0.452	0.355	0.543	0.638
3	0.712	3.126	1.799	0.403	0.244	0.427	0.535
4	0.687	2.877	1.701	0.367	0.156	0.339	0.446
5	0.675	2.616	1.630	0.344	0.097	0.268	0.379
6	0.669	2.469	1.571	0.321	0.051	0.203	0.3171
7	0.666	2.350	1.523	0.302	0.020	0.148	0.256
8	0.663	2.294	1.481	0.281	0.007	0.095	0.196
d = –1/3							
0	0.339	1.009	0.583	0.095	0.917	0.979	0.993
3.9	0.467	1.598	0.814	0.132	0.257	0.525	0.671
1	0.398	1.136	0.673	0.108	0.698	0.890	0.944
2	0.443	1.282	0.741	0.117	0.474	0.736	0.848
4	0.518	1.468	0.844	0.129	0.167	0.430	0.594
5	0.550	1.499	0.884	0.132	0.091	0.312	0.467
6	0.576	1.573	0.922	0.136	0.046	0.2134	0.358
7	0.613	1.633	0.957	0.139	0.021	0.138	0.263
8	0.596	1.578	0.989	0.143	0.011	0.092	0.190

NOTE: The Monte Carlo experiments under the two alternative hypotheses are independent and consist of 10,000 replications each, for sample size n = 208. Parameters of the simulations were chosen to match the sample mean and variance of quarterly real GDP growth rates from 1947:QII to 1999:QI. Entries in the column labeled "q" indicate the number of lags used to compute the R/S statistic; a lag of 0 corresponds to Mandelbrot's classical range, and a noninteger lag value corresponds to the average (across replications) lag value used according to Andrews's (1991) optimal lag formula.
SOURCE: Authors' calculations.

Table 4 reports power simulations under two fractionally differenced alternatives: $(1 - L)^d \varepsilon_t = \eta_t$, where d =1/3, –1/3. Hosking (1981) has shown that the autocovariance function $\gamma_\varepsilon(k)$ of ε_t is given by:

$$(34)\quad \gamma_\varepsilon(k) = \frac{\Gamma(1-2d)\,\Gamma(d+k)}{\Gamma(d)\,\Gamma(1-d)\,\Gamma(1-d+k)}\,\hat{\sigma}^2_\eta$$

$$d \in \left(-\frac{1}{2}, \frac{1}{2}\right).$$

Realizations of fractionally differenced time series (of length 208) are simulated by premultiplying vectors of independent standard normal random variates by the Cholesky factorization of the (208 x 208) covariance matrix whose entries are given by equation (34). To calibrate the simulations, σ^2_η is chosen to yield unit variance ε_t's, the $\{\varepsilon_t\}$ series is then multiplied by the sample standard deviation of real GDP growth from 1947:QII to 1999:QI, and to this series is added the sample mean of real GDP growth over the same sample period. The resulting time series is used to compute the power of the rescaled range; table 4 reports the results.

For small values of q, tests based on the modified rescaled range have reasonable power against both fractionally differenced alternatives. For example, using one lag, the 5 percent test has 68 percent power against the d = 1/3 alternative, and 89 percent power against the d = –1/3 alternative. As the lag length is increased, the test's power declines.

Note that tests based on the classical rescaled range are significantly more powerful than those using the modified R/S statistic. This, however, is of little value when distinguishing between long-range versus short-range dependence since the test using the classical statistic also has power against some stationary finite-order ARMA processes. Finally, note that tests using Andrews's truncation lag formula have reasonable power against the d = –1/3 alternative but are considerably weaker against the more relevant d = 1/3 alternative.

The simulation evidence in tables 3 and 4 suggests that our empirical results do indeed support the short-range dependence of GDP with a unit root. Our failure to reject the null hypothesis does not seem to be explicable by a lack of power against long-memory alternatives. Of course, our simulations were illustrative and by no means exhaustive; additional Monte Carlo experiments must be performed before a full assessment of the test's size and power is complete. Nevertheless, our modest simulations indicate that there is little empirical evidence in favor of long-term memory in GDP growth rates.

IV. Conclusions

This paper has suggested a new approach to the stochastic structure of aggregate output. Traditional dissatisfaction with the conventional methods—from observations about the typical spectral shape of economic time series, to the discovery of cycles at all periods—calls for such a reformulation. Indeed, recent controversy over deterministic versus stochastic trends and the persistence of shocks underscores the difficulties even modern methods have of identifying the long-run properties of the data.

Fractionally integrated random processes provide one explicit approach to the problem of long-range dependence; naming and characterizing this aspect is the first step in studying the problem scientifically. Controlling for its presence improves our ability to isolate business cycles from trends and to assess the propriety of that decomposition. To the extent that it explains output, long-range dependence deserves study in its own right. Furthermore, Singleton (1988) has pointed out that dynamic macroeconomic models often inextricably link predictions about business cycles, trends, and seasonal effects. So, too, is long-range dependence linked: a fractionally integrated process arises quite naturally in a dynamic linear model via aggregation. This model not only predicts the existence of fractional noise, but also suggests the character of its parameters. This class of models leads to testable restrictions on the nature of long-range dependence in aggregate data and holds the promise of policy evaluation.

Advocating a new class of stochastic processes would be a fruitless task if its members were intractable. In fact, manipulating such processes causes few problems. We constructed an optimizing linear dynamic model that exhibits fractionally integrated noise and provided an explicit test for such long-range dependence. Modifying a statistic of Hurst and Mandelbrot gives us a statistic robust to short-range dependence, and this modified R/S statistic possesses a well-defined limiting distribution, which we have tabulated. Illustrative computer simulations indicate that this test has power against at least two specific alternative hypotheses of long memory.

Two main conclusions arise from the empirical work and Monte Carlo experiments. First, the evidence does not support long-range dependence in GDP—the greater power of the modified R/S test may explain why our results contradict earlier work that purported to find long-range dependence. Rejections of the short-range dependence null hypothesis occur only with detrended data, and this is consistent with the well-known problem of spurious periodicities induced by log-linear detrending. Second, since a trend-stationary model is not contained in our null hypothesis, our failure to reject may also be viewed as supporting the first-difference stationary model of GDP, with the additional implication that the resulting stationary process is weakly dependent at most. This supports and extends the conclusion of Adelman (1956) that, at least within the confines of the available data, there is little evidence of long-range dependence in the business cycle. Nevertheless, Haubrich (1993) finds indirect evidence for long-range dependence using aggregate consumption series, and hence the empirical relevance of long memory for economic phenomena remains an open question that deserves further investigation.

References

Abadir, Karim, and Bariel Talmain. "Aggregation, Persistence, and Volatility in a Macromodel," University of York (undated).

Adelman, Irma. "Long Cycles: Fact or Artifact?" *American Economic Review*, vol. 55, no. 3 (June 1965), pp. 444–63.

Andrews, Donald W. K. "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, vol. 59, no. 3 (May 1991), pp. 817–58.

Agiakloglou, Christos, Paul Newbold, and Mark Wohar. "Bias in an Estimator of the Fractional Difference Parameter," *Journal of Time Series Analysis,* vol. 14, no. 3 (1993), pp. 235–46.

Baillie, Richard T., Tim Bollerslev, and Hans Ole Mikkelsen. "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, vol. 74, no. 1 (September 1996), pp. 3–30.

Baxter, Marianne, and Robert G. King. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, vol 81, no. 4 (November 1999), pp. 575–93.

Campbell, John Y., and N. Gregory Mankiw. "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, vol. 102, no. 4 (November 1987), pp. 857–80.

Christiano, Lawrence J., and Martin Eichenbaum. "Unit Roots in Real GNP: Do We Know, and Do We Care?" *Carnegie-Rochester Conference Series on Public Policy*, vol. 32, (Spring 1990), pp. 7–61.

Christiano, Lawrence J., and Terry J. Fitzgerald. "The Band Pass Filter," National Bureau of Economic Research, Working Paper No. 7257, July 1999.

Cochrane, John H. "How Big is the Random Walk in GNP?" *Journal of Political Economy*, vol. 96, no. 5 (October 1988), pp. 893–920.

Diebold, Francis X., Steven Husted, and Mark Rush. "Real Exchange Rates under the Gold Standard," *Journal of Political Economy,* vol. 99, no. 6 (December 1991), pp. 1252–71.

Diebold, Francis X., and Glenn D. Rudebusch. "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics,* vol. 24, no. 2 (September 1989), pp. 189–209.

Dupor, Bill. "Aggregation and Irrelevance in Multi-sector models," *Journal of Monetary Economics,* vol. 43, no. 2 (April 1999), pp. 391–409.

Eberlein, Ernst, and Murad Taqqu. *Dependence in Probability and Statistics*. Progress in Probability and Statistics 11. Boston: Birkhauser, 1986.

Fisher, Irving. "Our Unstable Dollar and the So-Called Business Cycle," *Journal of the American Statistical Association*, vol. 20, new series no. 149 (March 1925), pp. 179–202.

Geweke, John, and Susan Porter-Hudak. "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, vol. 4, no. 4 (1983), pp. 221–38.

Granger, Clive W. J. "The Typical Spectral Shape of an Economic Variable," *Econometrica*, vol. 34, no. 1 (January 1966), pp. 150–61.

———. "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, vol. 14, no. 2 (October 1980), pp. 227–38.

———. "Aggregation of Time Series Variables—A Survey," Federal Reserve Bank of Minneapolis Institute for Empirical Macroeconomics, Discussion paper 1, 1988.

Granger, Clive W. J. and Roselyne Joyeux. "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, vol. 1, no. 1 (1980), pp. 15–29.

Haubrich, Joseph G. "Consumption and Fractional Differencing: Old and New Anomalies," *Review of Economics and Statistics*, vol. 75, no. 4 (November 1993), pp. 767–72.

Herrndorf, Norbert. "A Functional Central Limit Theorem for Strongly Mixing Sequences of Random Variables," *Zeitschrift für Wahrscheinlichkeitstheorie verwandte Gebiete*, vol. 69, no. 4 (1985), pp. 541–50.

Hosking, J.R.M. "Fractional Differencing," *Biometrika*, vol. 68, no. 1 (April 1981), pp. 165–76.

Hurst, Harold E. "Long-Term Storage Capacity of Reservoirs," *Transactions of the American Society of Civil Engineers*, vol. 116, paper no. 2447 (1951), pp. 770–99.

Iyanaga, Shokichi, and Yukiyosi Kawada (editors). *Encyclopedic Dictionary of Mathematics,* by the Mathematical Society of Japan. Cambridge: MIT Press, 1977.

Jorgenson, Dale W., Frank M. Gollop, and Barbara M. Fraumeni. *Productivity and U.S. Economic Growth.* Harvard Economic Studies, vol. 159. Cambridge: Harvard University Press, 1987.

Kolmogorov, Andrei N. "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum," *Comptes Rendus (Doklady) de l'Académie des Sciences de l' URSS*, vol. 26, no. 2 (January 1940), pp. 115–18.

Kuznets, Simon. *Economic Growth and Structure.* New York: Norton, 1965

Kydland, Finn E., and Edward C. Prescott. "Time to Build and Aggregate Fluctuations," *Econometrica,* vol. 50, no. 6 (November 1982), pp. 1345–70.

Lo, Andrew W. "Long-Term Memory in Stock Market Prices," *Econometrica*, vol. 59, no. 5 (September 1991), pp. 1279–313.

Lo, Andrew W., and A. Craig MacKinlay. "The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation," *Journal of Econometrics*, vol. 40, no. 2 (February 1989), pp. 203–38.

Long, John B., Jr., and Charles I. Plosser. "Real Business Cycles," *Journal of Political Economy,* vol. 91, no. 1 (February 1983), pp. 39–69.

Lucas, Robert E., Jr. *Models of Business Cycles.* New York: Basil Blackwell, 1987.

Maddison, Angus. *Monitoring the World Economy 1820–1992.* New York: Organisation for Economic Co-operation and Development, 1995.

Mandelbrot, Benoit. "Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis," *Annals of Economic and Social Measurement,* vol. 1, no. 3 (July 1972), pp. 259–90.

———. "Limit Theorems on the Self-Normalized Range for Weakly and Strongly Dependent Processes," *Zeitschrift für Wahrscheinlichkeitstheorie verwandte Gebiete*, vol. 31, no. 2 (1975), pp. 271–85.

Mandelbrot, Benoit, and Murad Taqqu. "Robust R/S Analysis of Long-Run Serial Correlation," *Bulletin of the International Statistical Institute,* vol. 48, book 2 (1979), pp. 59–104.

Mandelbrot, Benoit, and James Wallis. "Noah, Joseph, and Operational Hydrology," *Water Resources Research*, vol. 4, no. 5 (October 1968), pp. 909–18.

———. "Some Long-Run Properties of Geophysical Records," *Water Resources Research*, vol. 5, no. 2 (April 1969), pp. 321–40.

Mitchell, Wesley Claire. *Business Cycles: The Problem and Its Setting,* NBER Studies in Business Cycles No. 1, New York: National Bureau of Economic Research, 1927.

Nelson, Charles R., and Heejoon Kang. "Spurious Periodicity in Inappropriately Detrended Time Series," *Econometrica*, vol. 49, no. 3 (May 1981), pp. 741–51.

Nelson, Charles R., and Charles I. Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, vol. 10, no. 2 (September 1982), pp. 139–62.

Newey, Whitney K., and Kenneth D. West. "A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, vol. 55, no. 3 (May 1987), pp. 703–08.

Oldham, Keith B., and Jerome Spanier. *The Fractional Calculus.* New York: Academic Press, 1974.

Parke, William R. "What is Fractional Integration?" *Review of Economics and Statistics*, vol. 81, no. 4 (November 1999), pp. 632–38.

Perron, Pierre, and Peter C. B. Phillips. "Does GNP Have a Unit Root? A Reevaluation," *Economics Letters*, vol. 23, no. 2 (1987), pp. 139–45.

Phillips, Peter C. B. "Time Series Regression with a Unit Root," *Econometrica*, vol. 55, no. 2 (March 1987), pp. 277–301.

Rosenblatt, Murray. “A Central Limit Theorem and a Strong Mixing Condition,” *Proceedings of the National Academy of Sciences,* vol. 42, (1956), pp. 43–7.

Sargent, Thomas J. *Dynamic Macroeconomic Theory.* Cambridge: Harvard University Press, 1987.

Singleton, Kenneth J. “Econometric Issues in the Analysis of Equilibrium Business Cycle Models,” *Journal of Monetary Economics*, vol. 21, no. 2/3 (March/May 1988), pp. 361–86.

Slutzky, Eugen. “The Summation of Random Causes as the Source of Cyclic Processes,” *Econometrica,* vol. 5, no. 2 (April 1937), pp. 105–46.

Sowell, Fallaw. “Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models,” *Journal of Econometrics,* vol. 53, no 1/3 (July/September 1992), pp. 165–188.

———. “Modeling Long-Run Behavior with the Fractional ARIMA Model,” *Journal of Monetary Economics*, vol. 29, no. 2 (April 1992), pp. 277–302.

———. “The Fractional Unit Root Distribution,” *Econometrica*, vol. 58, no. 2 (March 1990), pp. 495–506.

Stock, James, and Mark Watson. “Does GNP Have a Unit Root?” *Economics Letters,* vol. 22, no. 2/3 (1986), pp. 147–51.

White, Halbert. “A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity,” *Econometrica*, vol. 48, no. 4 (May 1980), pp. 817–38.

———. *Asymptotic Theory for Econometricians.* Orlando, Fl.: Academic, 1984.

White, Halbert, and Ian Domowitz. “Nonlinear Regression with Dependent Observations,” *Econometrica*, vol. 52, no.1 (January 1984), pp. 143–62.

Yajima, Yoshihiro. “On Estimation of Long-Memory Time Series Models,” *Australian Journal of Statistics,* vol. 27, no. 3 (1985), pp. 303–20.

———. “On Estimation of a Regression Model with Long-Memory Stationary Errors,” *Annals of Statistics*, vol. 16, no. 2 (June 1988), pp. 791–807.