Economic Review

FEDERAL RESERVE BANK OF DALLAS
FOURTH QUARTER 1996

Neighborhood School Characteristics: What Signals Quality to Homebuyers?

Kathy J. Hayes and Lori L. Taylor

Trade Deficits: Causes and Consequences

David M. Gould and Roy J. Ruffin

Can Mortgage Applications Help Predict Home Sales?

John V. Duca

Economic Review
Federal Reserve Bank of Dallas

The *Economic Review* is published by the Federal Reserve Bank of Dallas. The views expressed are those of the authors and do not necessarily reflect the positions of the Federal Reserve Bank of Dallas or the Federal Reserve System.

Subscriptions are available free of charge. Please send requests for single-copy and multiple-copy subscriptions, back issues, and address changes to the Public Affairs Department, Federal Reserve Bank of Dallas, P.O. Box 655906, Dallas, TX 75265-5906, (214) 922-5257.

Contents

Popular wisdom and economic research suggest that the quality of the neighborhood school should be an important determinant of housing values. Many researchers have found that housing values are higher where school spending or student test scores are higher. However, few economists consider these characteristics good indicators of school quality. Meanwhile, no one has examined whether the economists' notion of school quality—the school's marginal effect on students—is a school characteristic that matters to homebuyers.

Using a model of new home purchases and historical data on homes in the Dallas Independent School District (DISD), Kathy Hayes and Lori Taylor demonstrate that property values do reflect the characteristics of the neighborhood school. They present evidence that property values reflect student test scores but not school expenditures. Interestingly, they also find that the relationship between test scores and property values arises from an underlying relationship between property values and the marginal effects of schools. Thus, their analysis suggests that homebuyers and economists share the same definition of school quality.

According to conventional wisdom, trade balances reflect a country's competitive strength—the lower the trade deficit, the stronger the country's industries and the higher its rate of economic growth. In this article, David Gould and Roy Ruffin review the history of the conventional wisdom and empirically examine whether large overall trade deficits or bilateral trade imbalances are associated with lower rates of economic growth. They find that, once the fundamental determinants of growth have been accounted for, trade imbalances have little effect on rates of economic growth.

In this article, John Duca finds that the Mortgage Bankers Association (MBA) index of home mortgage applications can help forecast home sales. Alone, the index is a good, albeit imperfect, predictor of total home sales. But when included along with housing affordability and real, after-tax mortgage rate data, the index does not add extra information if one disregards differences in data release lags.

The index is available roughly three to four weeks ahead of the two alternative indicators. Taking into account its greater timeliness, it provides some extra information on home sales beyond that in the two other indicators considered. Given this qualification, the MBA index can help predict overall home sales. In addition, the long-run equilibrium relationships suggest that its usefulness may increase in the future. Nevertheless, the index should be used cautiously. It is still relatively new, and evidence suggests it may be misleading under some circumstances.

Neighborhood School Characteristics: What Signals Quality to Homebuyers?

Kathy J. Hayes
Research Associate
Federal Reserve Bank of Dallas
and Professor of Economics
Southern Methodist University

Lori L. Taylor
Senior Economist and Policy Advisor
Federal Reserve Bank of Dallas

Analysis suggests that homebuyers and economists share the same definition of school quality.

Most people are familiar with the adage that real estate values are determined by three basic characteristics—location, location, location. Economists consider this cliché only a modest exaggeration because research suggests that locational characteristics can explain much of the variation in residential property values. Not surprisingly, home prices tend to be lower in communities with high property taxes and higher in communities with low crime rates. Home prices fall as the commute to the central business district increases and rise as the amount of air pollution decreases. Locations near a city park command a premium, while locations near the city dump sell at a discount.

Popular wisdom and economic research suggest that the quality of the neighborhood school should also be an important locational characteristic. Many researchers have found that property values are higher where school spending is higher (for example, Oates 1969; Sonstelie and Portney 1980; and Bradbury, Case, and Mayer 1995). Other researchers have found a positive relationship between housing values and the test performance of students at the corresponding school (for example, Jud and Watts 1981, Rosen and Fullerton 1977, and Walden 1990). However, the economic literature on school quality measurement argues that the appropriate measure of school quality is the school's marginal effect on students (see Hanushek 1986), and no one has examined the relationship between marginal school effects and housing values.[1] Thus, we have an incongruity in the literature: spending and test scores seem to influence property values, but economists who study schools would not generally consider these characteristics measures of school quality. Meanwhile, the literature has been silent on whether the economists' notion of school quality is a locational characteristic that matters to homebuyers.

In this article, we attempt to identify the influence of neighborhood schools on the value of residential homes. Using a hedonic model of home purchases and historical data on homes in the Dallas Independent School District (DISD), we demonstrate that school quality can be an important locational characteristic in determining housing values. We find evidence that property values in DISD reflect student test scores but not school expenditures. Interestingly, we also find that the relationship between test scores and property values arises from an underlying relationship between property values and the marginal effects of schools. Thus, our analysis suggests that homebuyers and economists share the same definition of school quality.

A simple model of housing values

A house is a collection of desirable characteristics such as shelter, comfort, and location. Therefore, the price that buyers are willing to pay for a house should be related to the prices they are willing to pay for its component characteristics. By treating a house as the sum of its parts, a hedonic housing model generates estimates of the consumer's willingness to pay for each component characteristic.

Our hedonic model of housing prices in a single labor market is adapted from Rosen (1974). In this simplified model, consumers attempt to maximize their own happiness, taking the housing stock as given. Consumers derive satisfaction from consuming all sorts of housing characteristics ($Z = z_1, z_2,....z_n$) and a composite good (x). They earn an income (y) regardless of their chosen residence and can only consume combinations of Z and x that are affordable given that income. There are many types of consumers, and tastes for Z and x differ among those consumers according to socioeconomic characteristics (α) such as the person's age or educational attainment. In equilibrium, all consumers with identical preferences and income are able to achieve the same level of satisfaction.

After some manipulation, the individual consumer's decision-making can be described with a willingness-to-pay relationship or, more formally, a bid rent function:

$$(1) \qquad R = R(z_1, z_2....z_n; y, \alpha).$$

The value of the bid rent function represents the amount the consumer is willing to pay to rent a home with certain characteristics (Z), given the consumer's income level and socioeconomic type. Partial derivatives of the bid rent function with respect to housing characteristics represent the consumer's willingness to pay for those characteristics.

The price a potential buyer would be willing to pay for a house represents the present discounted value of the after-tax stream of bid rents.[2] If τ_R is the tax rate chosen by the jurisdiction for real estate,[3] θ represents the discounting factor, and housing is an infinitely lived asset, then the bid price of a house (P) would be

$$(2) \qquad P = \frac{R - \tau_R P}{\theta},$$

or equivalently,

$$(3) \qquad P = \frac{R(z_1, z_2....z_n; y, \alpha)}{\theta + \tau_R}.$$

The variation in incomes and socioeconomic characteristics generates a continuum of bid prices over a variety of types of homes.

In equilibrium, the sale price of any particular house equals the highest bid offered by potential consumers, regardless of their income or socioeconomic type. The hedonic price function describes this equilibrium.[4] The hedonic price function that we estimate describes the arm's length sales price as a function of the characteristics of the house and of its location.[5] The locational characteristics include neighborhood characteristics as well as local school characteristics.

The data

Data for this analysis come from three sources. Data on elementary school characteristics have been provided by DISD. Data on the characteristics of single-family homes in DISD come from the SREA Market Data Center's annual publication of residential property transactions. We restrict attention to the 288 DISD properties for which complete data are available that sold in July 1987 and were located in both the city and the county of Dallas. Data on nonschool locational characteristics come from the 1990 Census of Housing and Population.

DISD has provided data on student body characteristics, student achievement scores, and per-pupil expenditures for ninety-six elementary schools in its jurisdiction. From these data, we construct four possible indicators of school quality in 1987—current expenditures per pupil (*SPEND*), average sixth-grade achievement in mathematics on the Iowa Test of Basic Skills (*MATH687*), the marginal effect of the school on sixth-grade mathematics achievement (*SCHL687*), and the expected achievement of the student body in sixth-grade mathematics (*PEER687*). The first two of these indicators are common measures of school quality in the housing literature. The second two indicators represent a decomposition of average mathematics achievement into school effects and peer group effects (see the appendix). *SCHL687* measures the increase in student achievement in mathematics that can be attributed to the school. It corresponds to a common measure of school quality in the economics of education literature (see Hanushek and Taylor 1990, Aitkin and Longford 1986, and Boardman and Murnane 1979). *PEER687* is included as a possible indicator of school quality because research has shown that a high-achieving peer group in the school can have a positive effect on individual student performance (Summers and Wolfe 1977).

Table 1
Descriptive Statistics: A Tale of Two Cities

Variable	Northern Dallas Mean	Northern Dallas Standard deviation	Southern Dallas Mean	Southern Dallas Standard deviation
PRICE	$203,266	(204,301)	$82,502	(55,926)
SQFTLA	2,192	(1,026)	1,471	(568)
YRBUILT	58.3	(13.2)	53.5	(18.7)
POOL	.22	(.42)	.04	(.19)
FIREPL	.71	(.45)	.42	(.50)
DISTANCE	2.46	(.65)	2.11	(.86)
APARTMENTS	.18	(.20)	.26	(.23)
PRIVSCHL	.39	(.21)	.10	(.08)
NEIGHBORS	–1.47	(1.34)	1.59	(1.62)
MEDIAN INCOME	$52,819	(26,841)	$27,256	(7,735)
COLLEGE	.72	(.15)	.40	(.20)
BLUE-COLLAR	.11	(.09)	.31	(.13)
UNDER 12	.12	(.03)	.18	(.05)
OVER 65	.19	(.06)	.11	(.04)
HISPANIC	.10	(.12)	.32	(.25)
BLACK	.03	(.05)	.27	(.29)
SPEND	$2,498	(381)	$2,068	(232)
MATH687	76.97	(5.27)	69.56	(4.26)
SCHL687	29.55	(4.30)	26.86	(3.18)
PEER687	47.42	(3.21)	42.70	(3.07)
Number of observations	150		138	

The housing data used in this analysis include the log of the sale price of the property (*PRICE*), the year in which the home was built (*YRBUILT*), the number of square feet of living area in the structure (*SQFTLA*), and indicator variables that take on the value of one if the house has a swimming pool or a fireplace and zero otherwise (*POOL* and *FIREPL*, respectively). To capture potential nonlinearities in the relationship between the sale price and the age of the property, we also include interaction terms that take on the value of *YRBUILT* when the residence has a pool (*YR•POOL*) or fireplace (*YR•FIREPL*) and zero otherwise. We match the potential school quality indicators with housing characteristics using the SREA data on addresses and a Realtor's guide to DISD attendance zones (Positive Parents of Dallas et al. 1987).

The address data also permit us to merge in census tract characteristics from the 1990 Census of Housing and Population. The census tract data support three nonschool locational characteristics. These potential locational characteristics are the demographic characteristics of the neighborhood residents (*NEIGHBORS*),[6] the share of apartments in the neighborhood housing stock (*APARTMENTS*), and a proxy for the accessibility of private schools (the share of the elementary school population that is attending private school, denoted *PRIVSCHL*).

Finally, we used the address data to construct another nonschool locational characteristic—the linear distance to the central business district (*DISTANCE*)—and to divide the sample into two parts according to whether or not the property is located substantially north of downtown Dallas.[7]

Table 1 presents descriptive statistics for the data used in this analysis. As the table clearly indicates, there are significant differences between northern and southern Dallas.[8] On average, northern Dallas homes are more expensive, bigger, and more likely to have a pool or fireplace. Northern Dallas schools register higher on all our potential indicators of school quality. The average northern Dallas neighborhood has a smaller share of apartments in the housing stock and more access to private elementary schools than the average southern Dallas neighborhood. Meanwhile, the residents of southern Dallas neighborhoods are more likely than the residents of northern Dallas to be black or His-

Table 2
Estimates of the Hedonic Price Function

	Northern Dallas			Southern Dallas		
Variables	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
INTERCEPT	3.465**	3.123**	3.174**	3.163**	2.865**	2.867**
	(.334)	(.380)	(.391)	(.341)	(.592)	(.596)
SQFTLA	5.0E−4**	5.0E−4**	5.0E−4**	5.5E−4**	5.4E−4**	5.4E−4**
	(2.5E−5)	(2.5E−5)	(2.5E−5)	(5.7E−5)	(5.8E−5)	(5.9E−5)
YRBUILT	.007*	.006*	.007*	.008**	.008**	.008**
	(.004)	(.004)	(.004)	(.002)	(.002)	(.002)
YR•POOL	−.004	−.005	−.005	−.022**	−.023**	−.023**
	(.003)	(.003)	(.003)	(.011)	(.011)	(.011)
YR•FIREPL	−.007*	−.007*	−.007*	−.005	−.005	−.005
	(.004)	(.004)	(.004)	(.003)	(.003)	(.003)
POOL	.272	.289	.301	1.211**	1.255**	1.258**
	(.202)	(.201)	(.202)	(.571)	(.577)	(.581)
FIREPL	.448**	.433**	.441**	.431**	.419**	.420**
	(.205)	(.203)	(.204)	(.204)	(.205)	(.206)
DISTANCE	−.122**	−.146**	−.146**	−.137**	−.139**	−.138**
	(.039)	(.041)	(.041)	(.034)	(.034)	(.036)
APARTMENTS	.018	.007	.006	.074	.089	.088
	(.092)	(.092)	(.092)	(.121)	(.123)	(.124)
PRIVSCHL	.450**	.431**	.435**	1.073**	1.078**	1.075**
	(.142)	(.141)	(.141)	(.515)	(.516)	(.520)
NEIGHBORS	−.055**	−.042*	−.039*	−.042	−.041	−.041
	(.023)	(.024)	(.024)	(.029)	(.029)	(.030)
SPEND	3.3E−5	−7.8E−6	1.7E−5	−8.6E−6	−4.1E−6	−2.4E−6
	(7.0E−5)	(7.3E−5)	(8.3E−5)	(1.2E−4)	(1.2E−4)	(1.3E−4)
MATH687	—	.007*	—	—	.004	—
	—	(.004)	—	—	(.007)	—
SCHL687	—	—	.009*	—	—	.005
	—	—	(.005)	—	—	(.009)
PEER687	—	—	.004	—	—	.004
	—	—	(.007)	—	—	(.009)
Number of observations		150			138	

NOTE: Standard errors are in parentheses. The superscripts denote a coefficient that is significant at the 5-percent (**) or 10-percent (*) level.

panic, young, hold a blue-collar job, have a lower income, and to have not attended college.

The estimation and results

Because southern and northern Dallas differ so dramatically, we estimate the hedonic price function separately for the two areas using weighted least squares regression.[9] Furthermore, for comparison with the previous literature, we examine three models of the hedonic price function. In the first model, school quality is measured by per-pupil spending. In the second model, school quality is measured by both per-pupil spending and test scores. In the third model, which represents an unrestricted version of the second model, test scores are decomposed into school effects and peer group effects. We correct the standard errors from model 3 for the problem of estimated regressors (*SCHL687* and *PEER687*), using the technique suggested by Murphy and Topel (1985).[10] Table 2 presents our estimation results.

Despite the dramatic differences between northern and southern Dallas, Table 2 reveals striking similarities in the consumer's willingness to pay for housing characteristics. In both parts of the city, homebuyers pay a substantial premium for additional living space. Southern Dallas buyers tend to be slightly more sensitive to the age of the property, but homebuyers in both parts of the city have strong preferences for newer homes. Fireplaces add value to older homes, but the effect dissipates for newer homes.[11] After controlling for the age and size of the property and the presence of a fireplace, pools have a negligible effect on home prices.[12]

Northern and southern Dallas homebuyers are also similar in their willingness to pay for most nonschool locational characteristics. In both parts of the city, homebuyers are unwilling to pay for a change in the concentration of apartments (*APARTMENTS*) but are willing to pay for a shorter commute (*DISTANCE*) and greater access to private schools (*PRIVSCHL*). Furthermore, northern and southern Dallas homebuyers pay similar premiums for a shorter commute or greater access. Evaluated at the mean, a 1-percent decrease in the distance to the city center increases home prices by 0.36 percent in northern Dallas and 0.29 percent in southern Dallas, while a 1-percent increase in *PRIVSCHL* increases home prices by 0.17 percent in northern Dallas and 0.11 percent in southern Dallas.[13] Northern and southern Dallas homebuyers differ substantially in their willingness to pay for neighborhood demographics, however. Northern Dallas buyers seem willing to pay a premium for a change in resident characteristics, while southern Dallas buyers do not.

Another significant difference between northern and southern Dallas homebuyers appears in their willingness to pay for school quality. The data suggest that neither group considers school spending an indicator of school quality for which they are willing to pay. *SPEND* is insignificant across all of the model specifications for both northern and southern Dallas. However, the data indicate substantial differences in the willingness to pay for student achievement on standardized tests. As model 2 illustrates, homebuyers in northern Dallas pay a premium to live in the attendance zone of a school where students score well on standardized tests. Homebuyers in southern Dallas pay no such premium.

Given the desegregation efforts during the sample period, it is not particularly surprising that southern Dallas homebuyers were unwilling to pay a premium for the neighborhood schools.[14] Busing students away from the neighborhood school was much more common in southern Dallas than in northern Dallas (Linden 1995). Therefore, while homebuyers might have been able to rely on the attendance zone map in northern Dallas, they had less reason to expect that their choice of residence would guarantee a specific school in southern Dallas. Given the uncertainty about the stability of school attendance zones, it is more surprising that northern Dallas homebuyers were willing to pay a premium for school quality than that southern Dallas homebuyers were unwilling to pay such a premium.

One might suspect that northern Dallas homebuyers are willing to pay for school zones with good test scores because those scores indicate characteristics of the students who live in the area. If so, then the premium for test performance would arise from the attractiveness of the neighbors rather than the neighborhood school. However, as model 3 illustrates, the test score premium in northern Dallas arises from the marginal effects of the schools (*SCHL687*), not the characteristics of the student body (*PEER687*).[15] Evaluated at the mean, a 1-percent increase in *SCHL687* increases home prices by 0.26 percent. Of the characteristics that we are able to observe, only the size and age of the property and the distance from downtown have more influence than school effects on home prices in northern Dallas.

Conclusions

Using a hedonic model of property values, we examine the extent to which school quality is a locational characteristic that influences property values. We find that some homebuyers are not only cognizant of differences in school quality but also have revealed their preferences for higher quality schools by paying a premium for their home. Our analysis suggests that this premium for school quality can be among the most important determinants of housing prices.

Not all school characteristics appear to be indicators of school quality, however. We find no evidence that homebuyers are willing to pay for changes in school expenditures or student body characteristics. Instead, we find evidence that the school characteristic for which homebuyers pay a premium is the same characteristic that economists associate with school quality, namely, the marginal effect of the school on student performance.

A number of policy implications can be drawn from this research. The analysis suggests that policies that impact school effects can have a significant influence on residential property values. It also casts considerable doubt on policy analyses or policy initiatives that equate school spending with school quality. Finally, the analysis suggests that, at least as far as Dallas homebuyers are concerned, researchers are on target in trying to identify policy reforms that would increase the marginal effectiveness of schools.

Notes

We would like to thank Rebecca Bergstrasser, Stephen P. A. Brown, Thomas Fomby, Donna Ginther, Shawna Grosskopf, Joe Hirschberg, and Jim Murdoch for helpful comments and suggestions; Kelly A. George for research assistance; and the Dallas Independent School District for making its data available. Any remaining errors are our own.

[1] A few researchers, including Sonstelie and Portney (1980), have examined the relationship between property values and changes in test scores, but test score changes are generally considered a poor proxy for the marginal effects of schools.

[2] This discussion ignores the differential tax treatment of renters and owners.

[3] If assessment errors are randomly distributed, then all residences in a given government jurisdiction are taxed at the same expected rate. Restricting analysis to a single taxing jurisdiction eliminates the need to measure the potential capitalization of tax rate differentials and one can focus on estimating the hedonic price function for housing characteristics (*Z*).

[4] For a further discussion of the hedonic price function, see Bartik and Smith (1987).

[5] An arm's-length sales price can be considered an equilibrium house price for that time and location.

[6] *NEIGHBORS* is a principal components index of resident characteristics. The demographic characteristics included in the index are median income of the census tract and the shares of the population that are black, Hispanic, over 65 years of age, under 12 years of age, employed in a blue-collar occupation, and college educated. The principal components index explains 65 percent of the variation in these variables. The index is negatively correlated with median income and the population shares of elderly and college educated individuals and positively correlated with the remaining demographic characteristics.

[7] Residences north of a line along the southern border of Highland Park Independent School District were classified as being in northern Dallas. The remaining residences were classified as being in southern Dallas.

[8] The means are significantly different at the 5-percent level for all of the characteristics.

[9] The weight for northern Dallas is the reciprocal of the product of the square root of (*SQFTLA*) and the square root of (*1 – PRIVSCHL*); the weight for southern Dallas is the reciprocal of the product of the square root of (*1/YRBUILT*) and the square root of (*1 – PRIVSCHL*). Given these weights, the residuals are normally distributed and a Breusch–Pagan test can no longer detect heteroskedasticity at the 5-percent level of significance in either sample.

[10] The Murphy–Topel error correction involves using the variance–covariance matrix of the first-stage estimation to inflate the standard errors that are used in hypothesis testing in the second stage. Parameter estimates are unaffected by the correction. Specifically, one tests hypotheses using the variance–covariance matrix

$$\hat{\Sigma}_{\text{corrected}} = \hat{\Sigma}_{\text{uncorrected}} + (Z'Z)^{-1}Z'F^{*}\hat{V}(\hat{\theta})F^{*\prime}Z(Z'Z)^{-1},$$

where Z is the matrix of second-stage regressors, F^{*} is a matrix of first-stage derivatives that is weighted by the estimated coefficients on the generated regressors from the second stage, and $\hat{V}(\hat{\theta})$ is the variance–covariance matrix from the first-stage regression. Murphy and Topel demonstrate that the second term in the above equation is a positive definite matrix.

[11] It is unlikely that fireplaces, in and of themselves, have such large effects on property values. Rather, fireplaces likely proxy for other desirable home characteristics that we cannot observe in the data.

[12] Pools appear to add value in southern Dallas, but the effect may be spurious because only five southern Dallas homes in our sample have pools.

[13] These estimates come from model 3.

[14] Of course, there are other possible explanations for not finding a relationship between school quality measures and property values in southern Dallas.

[15] Omitting the potentially collinear *NEIGHBORS* from the estimation does not alter this result.

References

Aitkin, M., and N. Longford (1986), "Statistical Modeling Issues in School Effectiveness Studies," *Journal of the Royal Statistical Society*, A 149, pt. 1: 1–26.

Bartik, Timothy J., and V. Kerry Smith (1987), "Urban Amenities and Public Policy," in *Handbook of Regional and Urban Economics*, ed. Edwin S. Mills (Amsterdam: North Holland Press).

Boardman, Anthony E., and Richard J. Murnane (1979), "Using Panel Data to Improve Estimates of the Determinants of Educational Achievement," *Sociology of Education* 52 (April): 113–21.

Bradbury, Katherine L., Karl E. Case, and Christopher J. Mayer (1995), "School Quality, Local Budgets, and Property Values: A Re-Examination of Capitalization," manuscript.

Hanushek, Eric A. (1986), "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* 24 (September): 1,141–76.

——— , and Lori L. Taylor (1990), "Alternative Assessments of the Performance of Schools," *Journal of Human Resources* 25 (Spring):179–201.

Jud, G. Donald, and James M. Watts (1981), "Schools and Housing Values," *Land Economics* 57 (August): 459–70.

Linden, Glenn M. (1995), *Desegregating Schools in Dallas: Four Decades in the Federal Courts* (Dallas: Three Forks Press).

Murphy, Kevin M., and Robert H. Topel (1985), "Estimation and Inference in Two-Step Econometric Models," *Journal of Business and Economic Statistics* 3 (October): 370–79.

Oates, Wallace E. (1969), "The Effects of Property Taxes and Local Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis," *Journal of Political Economy* 77 (November/December): 957–71.

Positive Parents of Dallas, Dallas Chamber of Commerce, and Dallas Independent School District (1987), *All About DISD*.

Rosen, Harvey S., and David J. Fullerton (1977), "A Note on Local Tax Rates, Public Benefit Levels, and Property Values," *Journal of Political Economy* 85 (April): 433–40.

Rosen, Sherwin (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy* 82 (January/February): 34–55.

Sonstelie, Jon C., and Paul R. Portney (1980), "Gross Rents and Market Values: Testing the Implications of Tiebout's Hypothesis," *Journal of Urban Economics* 7 (January): 102–18.

SREA Market Center Data Inc. (1987), *North Texas Annual 1987* (Atlanta: Damar Corp.).

Summers, Anita A., and Barbara L. Wolfe (1977), "Do Schools Make a Difference?" *American Economic Review* 67 (September): 639–52.

Walden, Michael L. (1990), "Magnet Schools and the Differential Impact of School Quality on Residential Property Values," *Journal of Real Estate Research* 5 (Summer): 221–30.

Appendix

We decompose average test scores into school effects and peer group effects, following the methodology outlined in Hanushek and Taylor (1990). Thus, we hypothesize that student achievement in period T is a function of the student's complete history of school (S) and student and family (F) characteristics. However, because the relationship is recursive, we can write

$$(A.1) \quad A_{iT} = \lambda A_{iT-1} + \beta_T F_{iT} + \sum_{k=1} q_{kT} S_{ikT} + \epsilon_{iT},$$

where A_{iT} is the achievement of student i in period T, the S_{ikT} are dummy variables that equal one if the ith student attends school k in period T and equal zero otherwise, and F_{iT} represents student and family characteristics in period T. In this formulation, q_{kT} represents the value added by school k in period T and

$$(A.2) \quad \hat{A}_{iT} = \lambda A_{iT-1} + \beta_T F_{iT}$$

represents the level of student achievement that could be expected regardless of the school attended. Thus, q_{kT} is a measure of school effects, and the average $\hat{A}_{iT}$ for each school is a measure of peer group effects in that school.

Whenever student-level data are unavailable and the marginal effects of schools are independent of the student and family characteristics, equation A.1 can be estimated at the school level as

$$(A.3) \quad A_{kT} = \gamma + \tilde{\lambda} A_{kT-1} + \tilde{\beta}_T F_{kT} + \mu_{kT}.$$

In this equation, A_{kT} is average student achievement at school k in period T, F_{kT} represents average student and family characteristics at school k in period T, $\gamma + \mu_{kT} = q_{kT} + \varepsilon_{kT}$, and ε_{kT} represents the average estimation error for students at school k in period T. At this level of aggregation, $\gamma + \mu_{kT}$ is the best available proxy for school effects, and $P_{kT} = \tilde{\lambda} A_{kT-1} + \tilde{\beta}_T F_{kT}$ is the best available proxy for peer group effects. Because analysis at the school level incorporates error into the estimates of school and peer group effects, it is particularly important to treat these variables as estimated regressors in any subsequent analysis.

Table A.1
Estimating School and Peer Group Effects on Sixth-Grade Mathematics Achievement

	Parameter estimate	Standard error
INTERCEPT	26.767	6.301
MATH586	.740	.092
XCOHORT	–.083	.017
B&HISP	–.004	.002
SES	.004	.021
Number of observations	96	
R^2	.544	

DISD provided data on student body characteristics and student achievement scores for ninety-six primary schools in its jurisdiction for the years 1986 and 1987. The student body characteristics used in the analysis are the percentage of students who were black or Hispanic (*B&HISP*) and the percentage of students who were not receiving free or reduced-price lunches (the best available proxy for socioeconomic status, *SES*). The student achievement data used in the analysis are average scores on the Iowa Test of Basic Skills in mathematics. We use sixth-grade scores from 1987 (*MATH687*) and fifth-grade scores from 1986 (*MATH586*) as the measures of student achievement. The variable *XCOHORT* (the percentage increase in the number of students taking the exam) controls for changes in cohort size between 1986 and 1987.

From these data and the estimated coefficients in Table A.1, we construct measures of school and peer group effects for each of the ninety-six schools in our study. Thus, for each school, $SCHL687_k = 26.767 + \mu_{kT}$, and $PEER687_k = 0.740 \cdot MATH586_k - 0.083 \cdot XCOHORT_k - 0.004 \cdot B\&HISP_k + 0.004 \cdot SES_k$.

Trade Deficits: Causes and Consequences

David M. Gould
Senior Economist and Policy Advisor
Federal Reserve Bank of Dallas

Roy J. Ruffin
Research Associate
Federal Reserve Bank of Dallas
and
M. D. Anderson Professor of Economics
University of Houston

For the most part, trade deficits or surpluses are merely a reflection of a country's international borrowing or lending profile over time.…Neither one, by itself, is a better indicator of long-run economic growth than the other.

On September 19, 1996, the *Washington Post, Wall Street Journal*, and *New York Times* reported trade figures released by the U.S. Department of Commerce showing that the monthly U.S. trade deficit increased by $3.5 billion in July 1996. Almost unanimously, analysts quoted in the articles stated that the recent trade figures showed weakness in the U.S. economy. The news was not earth shattering, nor was the interpretation of the increasing trade deficit controversial. The conventional wisdom is that the trade balance reflects a country's competitive strength—the lower the trade deficit, the greater a country's competitive strength and the higher its economic growth.

But the conventional wisdom on trade balances stands in stark contrast to that of the economics profession in general. Standard economic thought typically regards trade deficits as the inevitable consequence of a country's preferences regarding saving and the productivity of its new capital investments. Trade deficits are not necessarily seen as a cause for concern, nor are they seen as good predictors of a country's future economic growth. For example, large trade deficits may signal higher rates of economic growth as countries import capital to expand productive capacity. However, they also may reflect a low level of savings and make countries more vulnerable to external economic shocks, such as dramatic reversals of capital inflows. Is the conventional wisdom wrong, or has the economics profession just failed to keep its theories well-grounded in fact?

Certainly, anyone can create a theory about trade deficits and speculate about how they may, or may not, be related to a nation's economic performance. The paramount question is not whether one can create a theory, but whether it is logically consistent and stands up to empirical observation.

The purpose of this article is to answer the question of whether trade deficits, bilateral as well as overall, are related to a country's economic performance. We begin by discussing the origin of popular views on trade deficits and compare these views with current economic thought on trade balances. Next, we discuss the relationship between international capital flows and trade balances and relate them to economic growth. We then empirically examine the relationship between trade deficits and long-run economic growth.

The evolution of ideas about trade balances

The mercantilists. Much of the current popular thinking on trade balances can trace its intellectual roots to a group of writers in the seventeenth and eighteenth centuries called the *mercantilists*. The mercantilists advanced the view that a country's gain from international commerce depends on having a "favorable" trade balance (favorable balance meaning that exports are greater than imports). The mercantilists were businessmen, and they looked at a country's trade balance as analogous to a firm's profit and loss statement. The greater are receipts over outlays (exports over imports), the more profitable (competitive) is the business (country). Thus, they argued that a country could benefit from protectionist policies that encouraged exports and discouraged imports. Because most international transactions during the seventeenth and eighteenth centuries were paid for with gold and silver, mercantilists were advocating a trade surplus so that the country would accumulate the precious metals and, according to their arguments, become rich.[1]

In 1752, David Hume exposed a logical inconsistency in the mercantilism doctrine through his explanation of the "specie-flow mechanism."[2] The specie-flow mechanism refers to the natural movement of money and goods under a gold standard or, indeed, any fixed exchange rate system in which the domestic money supply is inextricably linked to a reserve asset. The reserve asset need not be gold.[3]

Hume argued that an accumulation of gold from persistent trade surpluses increases the overall supply of circulating money within the country, and this would cause inflation. The increase in overall inflation also would be seen in an increase in input prices and wages. Hence, the country with the trade surplus soon would find its competitive price advantage disappearing as prices rose but the exchange rate remained constant. Automatically, through the specie-flow mechanism, the country with a trade surplus would find that its surplus shrank as its prices rose relative to other countries' prices. Any attempt to restore the trade surplus by raising tariffs or imposing other protectionist policies would simply result in another round of cost inflation, leading ultimately to a balance between exports and imports once again.

Several of the mercantilists—such as Gerard de Malynes (1601) and Sir Thomas Mun (1664)—understood the problems of maintaining a perpetual trade surplus as domestic prices rose but discounted this problem as a very long-run phenomenon and emphasized the benefits of accumulating gold as a means of exchange in a hostile and uncertain world.[4]

A few decades after Hume's original writings, economists such as Adam Smith and David Ricardo added further arguments against the mercantilistic advocacy of trade surpluses. They argued that what really matters to a country is its terms of trade—that is, the price it pays for its imports relative to the price it receives for its exports. Smith and Ricardo stood the advocacy of trade surpluses on its head when they showed that a country is better off the more imports it receives for a given number of exports and not vice versa. They argued that the mercantilistic analogy between a country's exports and a firm's sales was faulty.

Adam Smith in 1776 argued that money to an economy is different from money to an individual or firm. A business firm's objective is to maximize the difference between its imports of money and its exports of money. Money "imports" are the sale of goods and money "exports" are the purchases of labor and other inputs to production. However, for the economy as a whole, wealth consists of goods and services, not gold. Money, or gold, is useful as a medium of exchange, but it cannot be worn or eaten by a country. More money, in the medium and long run, just results in a higher level of prices. In the short run, however, Adam Smith also recognized that under the gold standard, a country's supply of gold would enable it to purchase the goods of other countries.

To some extent, therefore, the argument between the most able mercantilists and the classical economists was partly a question of emphasis—the mercantilists were concentrating on the fact that in the short run, the accumulation of money is wealth, while the classical economists were concentrating on the fact that in the long run, it is only the quantity of goods and services available that is wealth. However, the classical economists primarily were responding to the naive writings of most mercantilists, who confused the flow of money with the flow of goods in the short and long run.

National income accounting. Perhaps the great emphasis placed on national income accounting today is an important reason the naive form of mercantilism lives on in the hearts of many individuals. According to basic national income accounting, gross domestic product (GDP) is consumption (C) *plus* investment (I) *plus* government spending (G) *plus* exports (X) *minus* imports (M)—that is,

$$GDP = C + I + G + X - M.$$

This makes it appear that exports increase gross domestic product while imports reduce gross domestic product. This is erroneous because the definition of gross domestic product is just a tautology, and no conclusion about causality is possible. For example, it is equally true that the volume of goods and services available to an economy ($C+I+G$) consists of domestic output (*GDP*) *plus* imports *minus* exports—that is,

$$C + I + G = GDP + M - X.$$

Looked at in this way, a trade deficit appears to be "favorable" because we ultimately are interested in domestic spending. But this, too, is definitional. The question of whether deficits improve or hurt the economy cannot be resolved by such tautological manipulations. Theory and empirical evidence are required to evaluate whether deficits are favorable or unfavorable.

Employment and trade balances. The national income accounting view often leads many to associate trade deficits with reductions in employment. For example, some have argued that for every million dollars the United States has in its trade deficit, it costs about thirty-three American jobs, assuming that the average worker earns $30,000 a year (that is, $1,000,000/$30,000 = 33.33). So this implies that the July 1996 trade deficit of $11.7 billion cost around 390,000 jobs.[5] This calculation, however, is based on the fallacious assumption that capital inflows do not find their way into productive activity. Because a trade deficit is associated with capital inflows (to finance the deficit), the jobs lost by the deficit would be restored by the inflows of capital in expanding sectors of the economy. Gould, Ruffin, and Woodbridge (1993) correlated unemployment rates of the twenty-three OECD (Organization for Economic Cooperation and Development) countries with their import penetration ratios (the ratio of imports to GDP) and their export performance ratios (the ratio of exports to GDP) over thirty-eight years. They found that, for about half the countries, the correlation between import penetration ratios and unemployment rates (future or present) is negative (that is, higher imports are related to lower unemployment).

More importantly, however, they found that there is no instance of a significant positive or negative correlation of import penetration ratios with unemployment rates that is not the same for export performance ratios. In other words, exports and imports always had the same type of correlation with unemployment rates. Exports and imports are more related to each other than they are to other macroeconomic factors like unemployment rates because, ultimately, exports must pay for imports.

International capital movements and the balance of payments

From a public policy viewpoint, the fundamental question is: Do trade deficits reflect a malfunctioning of the economic system? If they do, perhaps limiting their size can improve a country's future standard of living. What is known, however, is that trade deficits or surpluses ultimately depend on a country's preferences regarding present and future consumption and the profitability of new capital investments. In understanding movements in the balance of trade, it helps to see their connection to movements in the balance of international capital flows. In a world of international capital mobility, trade deficits and international capital movements are the result of the same set of economic circumstances.

As first discussed by J. E. Cairnes (1874), international capital flows go through certain natural stages. The capital account balance (or the trade balance) should be seen as balancing a country's propensity to save with a country's investment opportunities and its resulting income payments, rather than as negative or positive indicators. The benefit of international capital flows and trade imbalances is that, in ordinary circumstances, they can lead to an efficient allocation of resources around the world. Net capital importers get their scarce capital more cheaply, and net capital exporters receive a higher return on their investments. In turn, capital imports finance trade deficits and trade surpluses finance capital exports.

In fact, under the right circumstances, a country can run a perpetual trade deficit or surplus. What matters for the balance of trade is how long a country has been a borrower or lender in international capital markets. How can countries maintain a perpetual trade deficit or surplus? Over time, the longer a country imports capital, the larger the interest rate payments on that capital. Eventually, a long-term debtor country will be borrowing less than its interest payments on existing debt to other countries and, in the steady-state, necessarily will have a trade surplus to pay these interest payments. A long-term creditor country will be lending less to other countries than its income receipts from other countries and will have a perpetual trade deficit. (For a fuller description of this mechanism, see the box entitled *International Capital Flows and the Balance of Trade* and the appendix.)

International Capital Flows and the Balance of Trade

The trade balance is a reflection of how long a country has been a borrower or lender in international capital markets. To see this relationship, it is helpful to examine the basic structure of a country's balance of payments. Let X = exports, M = imports, T = net gifts or unilateral transfers to foreigners, ΔB = net new borrowing from abroad, B = net indebtedness to the rest of the world, and r = the rate of interest on foreign indebtedness. A country's balance of payments must be

$$X + \Delta B = T + M + rB.$$

The left-hand side of the equation refers to receipts from foreigners; the right-hand side refers to payments to foreigners. These must always balance. If $\Delta B > 0$, a country is borrowing; if $B > 0$, a country is a net debtor. If $\Delta B < 0$, a country is lending, and if $B < 0$, a country is a net creditor. A country is considered to be a relatively short-term borrowing nation when its net indebtedness, B, is small compared with its net new borrowing, ΔB. In this case, imports will be greater than exports ($M > X$). A country is considered to be a relatively long-term borrowing nation when the interest it pays on foreign indebtedness, rB, is larger than its net new borrowing from abroad, ΔB. Here, exports are greater than imports ($X > M$). The opposite is true for a short-term or long-term creditor country.

Countries also can be in transition from a long-term creditor or debtor country to a short-term creditor or debtor country. The United States, for example, was a long-term creditor country throughout the 1970s, with trade deficits partly or wholly financed by net income payments from foreigners. However, in the 1980s, the U.S. trade deficit ballooned as both capital imports and income payments financed the deficit. The country was in transition until the net income account turned negative in 1994. Today, the United States must be regarded as a short-term debtor country. Japan, on the other hand, represents a major short-term creditor country.

Table 1 shows a snapshot of the 1994 balance of payments for several major countries. We show the net capital account, the net income account, and the trade and transfers account (the sum of net exports of goods and services and net transfers from the rest of the world). The current account (not shown) is the sum of the first and last columns; we separate the two components to illustrate the forces at work. It is very difficult to find examples of long-term creditor countries. The United Kingdom comes close, with its trade deficit and large net income from foreign investments, but the country may be entering a transition period. Austria is another example. There are more examples of long-term debtor countries, such as Canada and most of the Scandinavian countries.

Today, the United States has a relatively small obligation as far as investment income is concerned. But as we continue to be a debtor nation, the accumulated debts with the rest of the world will grow so large that the debt-service payments become larger than any amount of fresh capital borrowed by the country. If the United States continues to borrow, it will become a long-term debtor country. At this point, we will be in a perpetual balance-of-trade surplus. This must happen in order to pay the foreigners who own assets in the United States. Thus, the U.S. trade deficit in the future should completely turn around.

A key conclusion from this analysis and an examination of the relationship between capital flows and economic growth (see the appendix) is that, in the long run, there should be no link between economic growth and the trade

Table 1
The Balance of Trade and Net Capital and Income Accounts, 1994
(Millions of U.S. dollars)

Country	Trade and transfers	Capital account	Income account
Australia	$ –5,604	$ 15,860	$ –11,876
Austria	–630	–1,822	2,804
Belgium–Luxembourg	8,167	–10,452	4,853
Brazil	7,938	7,965	–9,091
Canada	3,754	8,331	–21,242
Chile	1,016	4,541	–1,773
Denmark	7,980	–5,537	–5,320
Finland	5,294	4,286	–4,226
France	19,051	–5,015	–10,962
Germany	–28,584	24,501	4,704
Japan	88,910	–86,190	40,330
Korea	–2,301	10,610	–1,554
Mexico	–17,039	12,754	–11,754
Netherlands	11,826	–6,485	1,546
Norway	5,413	–1,321	–1,769
Spain	1,496	4,449	–7,923
Sweden	6,690	6,390	–5,874
Switzerland	9,949	16,469	8,545
United Kingdom	–18,520	–24,562	16,129
United States	–140,440	120,806	–10,494

SOURCE: International Financial Statistics—capital account, line 78bjd; income account, line 78agd + line 78ahd; balance of trade and net transfers, line 78afd + line 78ajd + line 78akd.

balance. The long-run trade balance is jointly determined with the net creditor or debtor status of the country, while the long-run growth rate is determined by the growth rate of the population and technological progress. The next section is devoted to the empirical relationship between economic growth rates and trade imbalances, after controlling for other factors determining the rate of growth.

Are trade balances related to long-run economic growth?

Although the theoretical exposition above concludes that trade balances should not be related to long-run economic growth, the relevance of that theory has yet to be empirically examined. Moreover, there are other possible elements of trade balances, not discussed above, that may have implications for long-run economic growth. For example, large trade deficits imply large inflows of international capital. But international capital inflows may be subject to dramatic reversals, due to external shocks to a country's export sector and changes in foreign sentiment. In such cases, large trade deficits may be seen as an indicator of a country's vulnerability to external shocks. If large inflows of capital and trade deficits make a country more vulnerable to external economic shocks, long-run economic growth may be hampered.

While several studies have found that freer international trade (exports and imports) is an important determinant of cross-country growth rates, trade balances (the difference between exports and imports) have yet to be explored. This section examines the question of whether overall and bilateral trade balances are related to long-run rates of economic growth.

Overall trade balances. Empirically, one can imagine circumstances in which the trade balance is correlated to a nation's rate of economic growth, even though it may not cause it. Suppose, for example, that a nation is moving from a relatively closed economy to integration with the world economy—perhaps East Germany after the Berlin Wall fell in 1989. A country just opening up to world markets, like East Germany, would have a relatively high potential for future growth and would likely experience net capital inflows. But large capital inflows would be associated with large trade deficits. Consequently, there would appear to be a positive relationship between trade deficits and higher rates of economic growth. The higher rates of economic growth, however, are not caused by the larger trade deficits but by the opening up of domestic markets.

In contrast, trade deficits may be negatively related to economic growth if they reflect impediments to the market mechanism. Here again, however, the trade deficit itself is not causing lower growth but is itself determined by another factor that affects growth and the trade deficit. For example, it has been shown that the share of government consumption in GDP is negatively correlated to economic growth across countries (Barro 1991, and Levine and Renelt 1992). If a large share of government consumption tends to stimulate the demand for imports, generates a trade deficit, and reduces growth, this would show up as a negative correlation between trade deficits and economic growth, even though there is no causal relationship between the two. What is really decreasing growth is the large share of government consumption in total GDP, not the trade deficit.

Bilateral trade balances. While a country's overall trade may be balanced, a country may have bilateral deficits with many of its trading partners. Consequently, the relationship between overall trade balances and economic growth (discussed earlier) should not necessarily be the same as that between bilateral trade balances and economic growth. Nonetheless, we examine the empirical relationship between bilateral trade balances and economic growth because much popular attention has focused on this aspect of our trade account. To do the analysis, we develop a summary measure of bilateral trade balances that indicates the degree to which a country's bilateral trade flows are imbalanced (that is, bilateral exports and imports are unequal).

As is the case with overall trade imbalances, there is no theoretical reason bilateral trade imbalances should be related to economic growth. It is likely that countries that specialize in primary products will have higher bilateral imbalances than countries that specialize in manufactured goods. The reason is that a primary product producer cannot sell much to another country that produces the same primary product. On the other hand, a country that exports manufactured goods can easily sell manufactured goods to another country that exports manufactured goods because of the diversity of manufactured goods and intraindustry trade.

In fact, the correlation between our measure of bilateral imbalances (see note 12) and per capita real GDP is -0.62. This may be explained by the fact that countries with lower per capita GDPs tend to export fewer manufactured goods. Moreover, if protectionism rises with greater bilateral imbalances, bilateral imbalances

Table 2

The Role of Trade Balances in Growth

Dependent variable: average yearly real GDP per capita growth, 1960–89

	(1)	(2)	(3)	(4)	(5)
Constant	15.327	15.653	15.187	15.17	14.279
	(3.602)	(4.236)	(2.902)	(3.336)	(3.818)
ln(*Y*60)	–.837	–.933	–.801	–.863	–.943
	(–3.852)	(4.354)	(–3.546)	(–3.636)	(–4.404)
ln(*I*/*Y*)	3.251	2.970	3.048	2.963	3.149
	(8.521)	(7.634)	(7.897)	(7.486)	(7.651)
ln(*School*)	.904	.922	.850	.893	.843
	(6.651)	(6.973)	(6.185)	(6.447)	(5.607)
Exchange controls		–.005	–.004	–.005	–.005
		(–2.495)	(–1.956)	(–2.231)	(–2.414)
Share of all years in deficit			.007		
			(1.657)		
Trade deficit as a share of trade				–.004	
				(–.703)	
Bilateral imbalance as a share of trade					–.895
					(–.601)
$\bar{R}^2$	.684	.681	.664	.679	.663
RMSE	1.092	1.061	1.081	1.064	1.092
Observations	91	91	91	91	91

NOTES: *t* values are in parentheses. Real per capita growth is the least squares estimate; *Y*60 is real per capita GDP in 1960; *I*/*Y* is investment as a share of GDP, 1960–89; *School* is secondary-school enrollment rates, 1960–89; *exchange controls* is the black market premium.

SOURCES OF PRIMARY DATA: Real per capita growth and *Y*60, Summers and Heston (1991) Penn World Tables, version 5.6; *I*/*Y*, World Bank National Accounts; *School*, Barro (1991); *exchange controls*, Levine and Renelt (1992); *trade deficit as a share of total trade*, *bilateral imbalance as a share of total trade*, and *share of all years in deficit*, authors' calculations based on data from the International Monetary Fund, *Direction of Trade Statistics*.

may be negatively related to economic growth. For example, U.S. protectionism against Japanese products may rise as the U.S. bilateral trade deficit with Japan increases. Because several studies on the determinants of economic growth have found that protectionism tends to decrease long-run growth rates, there may be a negative correlation between bilateral imbalances and economic growth.[6] The next section attempts to empirically determine whether there is any relationship between overall and bilateral trade balances and economic growth when taking into consideration the underlying fundamental determinants of economic growth.

Trade balances and economic growth

The benchmark model. Before examining the role of trade balances in economic growth, we first present the results of a basic benchmark growth model. The model utilizes a formulation that is common to many of the recent cross-country empirical examinations of growth and attempts to control for the underlying determinants of long-run economic growth.[7] Equation 1 of Table 2 presents the estimation results of the benchmark model.[8] The dependent variable is the average annual real per capita GDP growth rate between 1960 and 1989,[9] and the explanatory variables are (1) the log of real GDP per capita in 1960, ln(*Y*60); (2) physical capital savings, which is the log of the share of investment in gross domestic product, ln(*I*/*Y*); and (3) a proxy for human capital savings—the log of secondary-school enrollment rates in 1960–89, ln(*School*).

The results of the benchmark model are consistent with most recent growth studies. Real GDP per capita in 1960 is negative and highly significant, suggesting income convergence conditional on human capital.[10] Physical capital savings and the proxy for human capital savings, ln(*I*/*Y*) and ln(*School*), are positive and significant at the 1-percent level, consistent with the empirical findings of Levine and Renelt (1992).

Equation 2 of Table 2 examines the role of capital controls, as proxied by black market

Table 3
Trade balances and growth

Country	Growth	Trade deficit as a share of trade	Bilateral trade imbalance as a share of trade	Share of all years in deficit
Angola	−2.7	−20.7	43.5	25.0
Chad	−2.4	28.9	49.0	89.3
Mozambique	−2.0	39.4	42.8	100.0
Madagascar	−1.8	10.2	35.5	81.8
Zambia	−1.4	−11.1	52.3	21.4
Central African Republic	−.5	−2.5	41.2	46.9
Ghana	−.5	6.3	37.2	72.7
Liberia	−.5	−15.5	39.3	4.5
Niger	−.4	4.5	37.6	86.7
Benin	−.2	55.3	39.8	100.0
Senegal	−.2	23.3	34.8	96.9
Uganda	−.2	−5.1	52.0	24.2
Guyana	0	2.8	33.2	55.2
Sierra Leone	0	11.6	47.5	68.8
Mauritania	.1	−12.1	55.8	12.5
Sudan	.1	34.4	44.9	90.9
Zaire	.2	−16.4	39.1	7.7
Somalia	.2	39.4	63.0	93.1
Bangladesh	.2	41.0	47.6	95.2
Haiti	.3	34.3	30.9	76.7
Mali	.3	36.9	47.7	96.8
Uruguay	.6	1.6	39.6	45.5
Nigeria	.7	−16.5	42.7	21.9
India	.8	18.0	31.1	93.9
Ethiopia	.8	24.0	47.8	96.3
Papua New Guinea	.8	−4.8	45.3	48.3
Average	−.3	12.2	43.4	63.7
Nepal	.9	45.4	48.1	100.0
Bolivia	1.0	−5.6	41.0	30.3
Chile	1.0	−5.8	33.0	27.3
Sri Lanka	1.1	16.8	40.0	90.9
Nicaragua	1.1	24.0	36.6	90.3
Argentina	1.1	−15.0	40.1	22.6
Malawi	1.3	17.5	44.3	96.6
El Salvador	1.3	14.1	29.7	78.8
Honduras	1.3	6.9	29.4	90.9
Guatemala	1.5	10.8	27.4	78.8
Burkina Faso	1.6	53.9	43.7	97.0
Kenya	1.6	24.7	43.1	100.0
Zimbabwe	1.6	−1.0	31.3	40.0
South Africa	1.6	−14.5	39.8	21.2
Peru	1.7	−10.7	27.3	15.2
Mauritius	1.7	6.5	60.0	66.7
Burundi	1.7	19.7	56.7	87.5
New Zealand	1.8	.7	25.3	54.5
Togo	1.8	30.8	42.3	90.9
Jamaica	1.9	20.4	34.5	100.0
Pakistan	1.9	24.1	35.9	97.0
United States	1.9	12.5	19.8	69.7
Rwanda	2.0	30.2	50.0	89.7
United Kingdom	2.2	7.3	16.8	100.0
Switzerland	2.2	4.7	22.6	93.9
Venezuela	2.2	−21.2	34.8	12.1
Average	1.6	11.4	36.7	70.8

Country	Growth	Trade deficit as a share of trade	Bilateral trade imbalance as a share of trade	Share of all years in deficit
Tanzania	2.3	27.7	36.0	69.0
Colombia	2.3	1.9	24.9	57.6
Paraguay	2.3	13.4	38.1	66.7
Philippines	2.3	15.1	26.8	93.9
Australia	2.3	−.8	33.6	42.4
Canada	2.4	−4.5	13.5	6.1
Costa Rica	2.5	11.9	30.0	100.0
Dominican Republic	2.5	23.7	39.9	54.5
Iraq	2.5	−8.9	52.0	50.0
Sweden	2.6	−2.5	19.5	54.5
Mexico	2.6	0	18.2	75.8
Ireland	2.7	−.3	21.6	75.8
Morocco	2.8	24.4	27.7	100.0
Ecuador	2.8	−9.2	35.3	39.4
Gambia	2.9	25.0	52.6	64.3
Turkey	2.9	24.1	25.3	100.0
Denmark	2.9	2.5	18.6	81.8
Netherlands	3.0	−.2	20.0	63.6
Iran	3.1	2.3	42.8	60.0
Barbados	3.1	39.7	37.3	100.0
Jordan	3.1	54.6	53.8	100.0
Suriname	3.3	2.1	39.9	58.6
Trinidad and Tobago	3.3	−8.0	49.2	46.9
Tunisia	3.3	22.8	30.2	100.0
Average	2.7	11.9	32.8	68.9
Germany, West	3.4	−7.3	15.7	0
France	3.4	3.5	17.5	93.9
Norway	3.4	−2.8	27.6	69.7
Panama	3.5	60.6	46.8	100.0
Congo	3.5	−18.3	57.1	48.5
Cameroon	3.6	2.3	36.7	54.5
Thailand	3.6	12.5	34.6	100.0
Israel	3.6	24.3	30.7	100.0
Finland	3.6	.5	19.0	69.7
Spain	3.6	21.7	25.7	97.0
Algeria	3.7	−5.4	30.0	46.7
Austria	3.9	10.9	21.1	100.0
Malaysia	4.0	−6.3	29.2	7.7
Italy	4.0	4.9	17.5	100.0
Syria	4.1	22.6	50.1	90.6
Egypt	4.3	43.7	43.2	93.9
Portugal	4.5	25.5	29.3	100.0
Greece	4.6	38.9	26.3	100.0
Brazil	4.6	−8.2	29.2	51.5
Gabon	5.2	−31.2	31.7	0
Malta	5.4	32.8	37.9	100.0
Korea, Republic of	5.6	2.3	30.8	84.4
Hong Kong	6.1	−.7	41.9	72.7
Japan	6.1	−9.8	31.6	39.4
Singapore	6.6	8.3	30.5	100.0
Average	4.3	−9.0	31.7	72.8

SOURCES OF PRIMARY DATA: Same as Table 2.

exchange rate premia, in economic growth.[11] We include a measure of capital controls in the benchmark equation to account for any negative growth effects due to the lack of capital mobility across nations. Specifically, we do not want to confuse the effects of low capital mobility with the effects of a low trade imbalance. Low capital mobility impedes the development of trade imbalances and is likely to be related to low rates of economic growth.

As model 2 shows, exchange controls decrease economic growth and the coefficient is statistically significant and economically important. Holding all else constant, the size of the coefficient suggests that a black market premium of 50 percent, for example, would decrease a country's average growth rate in the range of 0.20 to 0.25 percentage points per year.

The effects of trade balances. Can trade balances explain any variation in economic growth once capital controls and the standard determinants of growth are held constant? Before we examine this question, we first present some simple descriptive statistics on the relationship between trade balances and economic growth.

Table 3 summarizes the countries in the data set and shows their average yearly growth rate, the trade deficit as a share of total trade, the share of total years in deficit, and a measure of bilateral trade imbalances as a share of total trade. The trade deficit as a share of total trade is imports minus exports divided by total trade (imports plus exports); the share of total years in deficit is the number of years a country has had a trade deficit over the 1960–89 period divided by the number of years in the period (30); bilateral trade imbalances are measured by summing a country's bilateral trade deficits and surpluses and dividing by that country's total trade and adjusting for overall surpluses and deficits.[12] In other words, our measure of bilateral imbalances represents the percentage of a country's trade that is bilaterally imbalanced (after adjustments for total imbalances).

The countries in Table 3 are grouped according to growth rates; the slowest 25 percent of countries are in the upper left and the fastest 25 percent of countries are in the lower right. Without controlling for the important determinants of growth, there appears to be a weak positive correlation between economic growth and the percentage of years in deficit. The fastest growing countries seem to have more years in deficit than the slower growing countries, although those countries in the middle growth range are not distinguishable as having a higher or lower share of years in deficit. There is a negative correlation between bilateral imbalances and economic growth. This correlation is stronger than the previous one. The greater the bilateral imbalance, the lower the growth; there is no ambiguity in the middle growth categories. The overall trade deficit as a share of total trade also appears to be negatively related to growth, although it is not a strong relationship. A priori, it is difficult to see any strong relationship between measures of overall or bilateral trade imbalances and economic growth. However, the other factors determining growth should be taken into account before any conclusions can be properly made.

Equation 3 in Table 2 adds the share of years a country's trade account is in deficit to the benchmark model. As the results indicate, the variable is positively related to economic growth, but it is not statistically significant at the standard 5-percent level. However, taking the point estimate seriously, the size of the coefficient suggests that its economic effects are only moderate. For example, the United States, with 69.7 percent of its years in deficit, would experience an increase in its growth rate of about 0.5 percentage points per year.

The weakness of the relationship between trade balances and economic growth is shown by an alternative measure of trade deficit: trade deficit as a share of total trade, shown in equation 4 of Table 2. In this case, the coefficient is negative but is extremely small and statistically insignificant. Taking the point estimate seriously, a trade deficit that is 12 percent of total trade, which is what the United States had over the period 1960–89, would only decrease average yearly per capita real GDP growth by about 0.05 percentage points.

Equation 5 includes bilateral imbalances as a share of trade. As the results indicate, the coefficient on this variable is negative, but, with a t value less than 1, it is not statistically significant. Thus, empirical evidence is consistent with the hypothesis that bilateral trade imbalances have no particular impact on economic growth.

Conclusion

In this study, we have examined, both theoretically and empirically, the relationship between trade balances and long-run economic growth. We find that trade imbalances have little effect on rates of economic growth once we account for the fundamental determinants of economic growth.

For the most part, trade deficits or surpluses are merely a reflection of a country's international borrowing or lending profile over time. Just as companies borrow to finance investment and purchases, so do countries. A country can have a perpetual trade deficit or surplus simply because income payments from investments allow it to finance the country's desired flow of goods. Far too often, the common wisdom is that large trade deficits signal a fundamentally weak economy, when the empirical evidence suggests that there is no long-run relationship between the two. Trade deficits and surpluses are part of the efficient allocation of economic resources and international risk-sharing that is critical to the long-run health of the world economy. Neither one, by itself, is a better indicator of long-run economic growth than the other.

Notes

1 Thomas Mun (1664) pointed out that "Our yearly consumption of foreign wares to be for the value of twenty thousand pounds, and our exportations to exceed that two hundred thousand pounds, *which sum wee have therupon affirmed is brought to us in treasure to ballance the accompt*" [emphasis added]. International lending must have been relatively small in the seventeenth century.

2 In the eighteenth century, precious metals were referred to as "specie."

3 In modern times, currency boards, such as those found in Hong Kong and Argentina, use the U.S. dollar to back their currency, and many other fixed-exchange-rate regimes peg the value of their currencies to the U.S. dollar.

4 See Schumpeter (1954, 344–45 and 356–57) for an excellent discussion of mercantilistic thought.

5 For an example of this type of analysis, see Duchin and Lange (1988). They argue that eliminating the trade deficit in 1987 would have increased employment by 5.1 million jobs. This figure represented an increase of about 5 percent in total employment from a trade deficit that represented only about 3 percent of GDP.

6 See, for example, Krueger (1978); Bhagwati (1978); World Bank (1987); De Long and Summers (1991); Michaely, Papageorgiou, and Choksi (1991); Edwards (1992); Roubini and Sala-i-Martin (1992); and Gould and Ruffin (1995).

7 See, for example, Kormendi and Meguire (1985); Barro (1991); Romer (1990); Levine and Renelt (1992); Edwards (1992); Roubini and Sala-i-Martin (1992); Backus, Kehoe, and Kehoe (1992); and Mankiw, Romer, and Weil (1992). These empirical studies typically rely on a closed-economy version of the neoclassical Solow growth model. The closed-economy model would seem inappropriate in a world where capital is internationally mobile. However, an implication of the open-economy neoclassical model is that countries should experience rapid income convergence because capital can move quickly across borders and does not have to be slowly accumulated at home. But fast income convergence is not borne out by cross-country empirical evidence.

Barro, Mankiw, and Sala-i-Martin (1995) find that the transition to the long run in an open-economy neoclassical model may not be instantaneous if there are some impediments to the flow of capital across countries. Impediments to the flow of capital are likely, especially when considering the flow of human capital across nations.

8 The benchmark model utilizes a log-linear formulation for two reasons: it has a basis in Cobb–Douglas production technologies (such as Backus, Kehoe, and Kehoe 1992 and Mankiw, Romer, and Weil 1992), and this model is superior to a simple linear formulation in minimizing the mean squared error.

9 Least squares estimates are used because they are less sensitive to the end points of the growth period.

10 Although regressing average growth rates against initial income levels suggests income convergence, it does not necessarily provide statistical evidence of convergence. Quah (1990) and Friedman (1992) note that, because of regression to the mean, a negative relationship between average growth rate and initial income does not necessarily provide statistical evidence of convergence.

11 The black market exchange rate premium is the percentage by which the official exchange rate deviates from the market exchange rate and is often a good proxy for the degree to which countries attempt to control international capital flows.

12 The formula for the bilateral trade imbalance of country i is $BIM_i = \frac{\sum_{j=1}^{n} \left| X_{ij}^{*} - M_{ij} \right|}{X_i + M_i}$, where $X_{ij}^{*} = X_{ij} * \frac{M_i}{X_i}$, X_i is total exports of country i, M_i is total imports of country i, X_{ij} is exports of country i to country j, and M_{ij} is imports to country i from country j.

References

Backus, David K., Patrick J. Kehoe, and Timothy J. Kehoe (1992), "In Search of Scale Effects in Trade and Growth," *Journal of Economic Theory* 58 (December): 377–409.

Barro, Robert J. (1991), "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics* 106 (May): 407–43.

———, N. Gregory Mankiw, and Xavier Sala-i-Martin (1995), "Capital Mobility in Neoclassical Models of Growth," *American Economic Review* 85 (March): 103–15.

Bhagwati, Jagdish (1978), *Anatomy and Consequences of Exchange Control Regimes* (Cambridge, Mass.: Ballinger Publishing Co.).

Cairnes, John E. (1874), *Some Leading Principles of Political Economy Newly Expanded* (New York: Macmillan and Co.).

De Long, J. Bradford, and Lawrence H. Summers (1991), "Equipment Investment and Economic Growth," *Quarterly Journal of Economics* 106 (May): 445–502.

Dollar, David (1992), "Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976–1985," *Economic Development and Cultural Change* 40 (April): 523–44.

Duchin, Faye, and Glenn-Marie Lange (1988), "Trading Away Jobs: The Effects of the U.S. Merchandise Trade Deficit on Employment," Economic Policy Institute, Washington, D.C., October.

Edwards, Sebastian (1992), "Trade Orientation, Distortions, and Growth in Developing Countries," *Journal of Development Economics* 39 (July): 31–57.

Friedman, Milton (1992), "Do Old Fallacies Ever Die?" *Journal of Economic Literature* 30 (December): 129–32.

Gould, David M., and Roy J. Ruffin (1995), "Human Capital, Trade and Economic Growth," *Weltwirtschaftliches Archiv* 131 (3): 425–45.

———, ———, and Graeme L. Woodbridge (1993), "The Theory and Practice of Free Trade," Federal Reserve Bank of Dallas *Economic Review,* Fourth Quarter, 1–16.

Kormendi, Roger, and Philip Meguire (1985), "Macroeconomic Determinants of Growth: Cross-Country Evidence," *Journal of Monetary Economics* 16 (September): 141–63.

Krueger, Anne (1978), *Foreign Trade Regimes and Economic Development: Liberalization Attempts and Consequences* (Cambridge, Mass.: Ballinger Publishing Co.).

Levine, Ross, and David Renelt (1992), "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review* 82 (September): 942–63.

Malynes, Gerard de (1601), *A Treatise of the Canker of England's Commonwealth.*

Mankiw, N. Gregory, David Romer, and David N. Weil (1992), "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics* 107 (May): 407–37.

Michaely, M., D. Papageorgiou, and A. Choksi, eds. (1991), *Liberalizing Foreign Trade: Lessons of Experience in the Developing World*, vol. 7 (Cambridge, Mass.: Basil Blackwell).

Mun, Thomas (1664), *England's Treasure by Foreign Trade: Or, the Balance of Our Foreign Trade Is the Rule of Our Treasure.*

Phelps, Edmund S. (1966), *Golden Rules of Economic Growth* (New York: Norton).

Quah, Danny (1990), "Galton's Fallacy and Tests of the Convergence Hypothesis" (Massachusetts Institute of Technology), photocopy.

Romer, Paul M. (1990), "Human Capital and Growth: Theory and Evidence," *Carnegie–Rochester Conference Series on Public Policy* 32: 251–85.

Roubini, Nouriel, and Xavier Sala-i-Martin (1992), "Financial Repression and Economic Growth," *Journal of Development Economics* 39 (July): 5–30.

Ruffin, Roy J. (1979), "Growth and the Long-Run Theory of International Capital Movements," *American Economic Review* 69 (December): 832–42.

Schumpeter, Joseph A. (1954), *History of Economic Analysis* (New York: Oxford University Press).

Solow, Robert (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70 (February): 65–94.

Summers, Robert, and Alan Heston (1991), "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988," *Quarterly Review of Economics* 106 (May): 327–68.

World Bank (1987), *World Development Report 1987* (New York: Oxford University Press).

Appendix
A Simple Dynamic Model of Growth, the Balance of Trade, And International Capital Movements

It is not obvious that a long-term creditor country must have a trade deficit, or that a long-term debtor country must have a trade surplus. In other words, why should it be that net investment income necessarily exceeds net capital outflows for a long-term creditor country, or that net debt payments must necessarily exceed net capital inflows for a long-term debtor country? To demonstrate this claim, we consider a world consisting of two countries—home and foreign. For the sake of simplicity, both countries produce a single, identical good that can be either consumed or used as capital.[1] Moreover, to keep the notation simple, we assume both countries have the same population and that there is no depreciation of capital (capital lasts forever or is used up in consumption). To keep one country from overrunning the other, we suppose the labor forces grow at the same rate. Finally, we suppose that the single good is produced under constant returns to scale by only two factors, labor and capital.

Let k and k^* denote the *owned* capital per unit of labor in the home and foreign countries, respectively. Capital movements take place by the home country's borrowing B units of capital from the foreign country, so the capital per unit labor *located* in the home country is $k + b$, where $b = B/L$, while the capital per unit labor *located* in the foreign country is $k^* - b$. If capital is freely mobile, the equilibrium per capita stock of foreign investment, b, is determined by equating the marginal products of capital in both countries—that is,

$$f'(k + b) = g'(k^* - b) = r, \quad \text{(A.1)}$$

where f and g denote the per capita production functions in the home and foreign countries, respectively, and f' and g' denote the derivatives or the marginal products of capital.

Let s and s^* denote the constant saving rates in the home and foreign countries, and let n denote the rate of growth of the labor force in both countries. Per capita incomes are $f(k + b) - rb$ in the home country and $g(k^* - b) + rb$ in the foreign country. According to the Solow growth model (Solow 1956), the countries will be in steady-state when savings equal required investment:

$$s[f(k + b) - rb] - nk = 0, \quad \text{(A.2)}$$

and

$$s^*[g(k^* - b) + rb] + nk^* = 0. \quad \text{(A.3)}$$

Solving equations A.1–A.3 yields steady-state values of k, k^*, and b.[2] Thus, in the long run, $db/dt = 0$.

In the short run, db/dt may be nonzero. Let us look at the short-run and long-run dynamics of the balance of payments as envisioned by Cairnes (1874). Since $b = B/L$, the rate of change in the per capita stock of foreign investment is

$$db/dt = (dB/dt)/L - nb, \quad \text{(A.4)}$$

where $n = (dL/dt)/L$ and $(dB/dt)/L$ is the per capita inflow of capital to the home country from the foreign country. Equation A.3 may be used to describe the determinants of the per capita trade balance. By definition, the per capita trade surplus, $x - m$ (exports minus imports), will be per capita foreign debt service, rb, where r is the rate of interest $[r = f'(k + b)]$ – per capita capital inflows. In other words,

$$x - m = rb - (dB/dt)/L. \quad \text{(A.5)}$$

Combining equations A.4 and A.5 results in

$$x - m = (r - n)b - db/dt. \quad \text{(A.6)}$$

This is our key equation. The home country's per capita trade balance equals $(r - n)b$ minus the change in its per capita net indebtedness. In the steady-state, $db/dt = 0$, the per capita trade surplus $(x - m) = (r - n)b$. Assuming $r > n$, if the home country is a net debtor, $b > 0$, there will be a surplus. In contrast, if $b < 0$, the country will have a long-run deficit. Cairnes claimed that the net creditor's long-run trade balance would be negative, implying that $r > n$ in the long run. Remarkably, the condition that $r > n$ is the condition for dynamic economic efficiency (Phelps 1966).

In the above model, the long-run growth rate is simply equal to the population growth rate. This follows because in the steady-state, b, k, and k^* are constant; accordingly, per capita income remains constant. If we reinterpreted the model in terms of the effective labor supply and labor-augmenting technological progress, per capita income would increase by the rate of technological progress. Whatever interpretation is made, the model is then so constructed that both countries grow at exactly the same rate.

A key conclusion from this analysis is that in the long run, there should be no link between economic growth and the trade balance. The long-run trade balance is determined by the net creditor or debtor status of the country, while the long-run growth rate is determined by the growth rate of the population and technological progress.

[1] This model is based on Ruffin (1979).

[2] Ruffin (1979) demonstrates the conditions under which the above model has a unique solution.

Can Mortgage Applications Help Predict Home Sales?

John V. Duca
Research Officer
Federal Reserve Bank of Dallas

When the more timely availability of the mortgage applications index is taken into account, it adds some information about the pace of total home sales.

Good predictions of housing activity are important to both the private sector and to policymakers. Homebuilders, for example, need to gauge housing demand when considering whether to build homes before obtaining sales contracts. With respect to monetary policy, the Federal Reserve monitors data, particularly on interest-rate-sensitive and cyclically sensitive sectors like housing, to gauge the future underlying pace of aggregate demand.[1] This article assesses the usefulness of the Mortgage Bankers Association index of mortgage applications as a near-term indicator of home sales.

Aside from the ultimate uses of good housing predictions, there are at least two practical reasons for developing near-term leading indicators of housing. First, most housing data are not very timely, reflecting earlier decisions owing to lags in construction and sales, as well as to lags in the collection and release of data. Second, housing markets are sometimes difficult to predict for several reasons: a sudden rise in the interest rate may prompt people to speed up home purchases to avoid any further increases in mortgage interest rates; regulatory and institutional changes have altered the interest sensitivity of housing (for examples, see Duca forthcoming, Kahn 1989, and Mauskopf 1990); and economic growth is sometimes restrained by temporary factors that may or may not affect decisions to purchase homes.

A recent example of such difficulties occurred in 1996, when bond yields rose on news that economic growth had rebounded from the temporary effects of bad weather and government shutdowns in late 1995 and early 1996. Many analysts predicted that housing activity would fall off quickly, but levels of home sales and construction were generally stronger than expected in the spring. As one analyst put it, "By just about every available measure, growth in housing has far surpassed industry expectations and outpaced many sectors of the economy" (Pesek 1996).

There are at least three plausible explanations for this unexpected strength. First, a rebound in confidence and income may have largely offset the initial impact of higher mortgage interest rates on housing. Second, the impact of higher long-term interest rates may have been cushioned by a shift toward adjustable-rate mortgages, which have interest rates linked to lower, short-term rates. Third, the early 1996 rise in long-term rates may have induced many people to speed up their home purchases out of fear of further interest rate increases.

Each of these explanations has a somewhat different implication for housing in the second half of 1996. The first account implies that home sales will not decline too much, as does the second, provided that short-term interest rates do not rise a great deal. By contrast, the third explanation implies that home sales will fall more sharply in late 1996 or early 1997 because the strength of sales in early 1996 came at the expense of future home activity. Given the different implications of these explanations, it is useful to have good and timely near-term predictors of housing activity.

Partly to address such needs, the Mortgage Bankers Association (MBA) has, on a weekly basis since January 1990, surveyed lenders about the pace of mortgage applications for home purchases and for refinancings. Compared with home sales and housing starts data, the MBA's index provides more up-to-date information on home-buying for two reasons. First, mortgage applications typically precede home sale closings by one to two months. Second, every Thursday, the MBA releases its mortgage applications index for the prior week, whereas monthly data on housing starts (and permits) and existing home sales are released with three- and four-week lags, respectively. Given its shorter data lags, the MBA index may help analysts better forecast home sales.

In evaluating the usefulness of the MBA index, we first need to determine whether it and other housing indicators provide information about future changes in home sales. In addition to this index, two alternative indicators are considered: a housing affordability index and a real, after-tax mortgage interest rate. After establishing that each indicator leads home sales, I test whether mortgage applications add information about future home sales beyond what the affordability index and mortgage interest rates signal. The final part of this article summarizes the findings by providing an overall assessment of the MBA index.

Do mortgage applications lead home sales?

This section presents the basic empirical approach used to assess whether mortgage applications lead home sales. After I describe the data used, I run unit root tests and test lead–lag relationships.

Basic specification. To test whether a variable Y is a leading indicator of a variable X, the following type of regression, called a Granger causality, or lead–lag, test, is run:

$$(1) \qquad X_t = \text{constant} + \sum_i \delta_{xi} X_{t-i} + \sum_j \delta_{yj} Y_{t-j},$$

where the δ_{xi} and δ_{yj} are estimated coefficients. If the lags of Y are jointly significant according to an F test, then Y is a leading indicator of X. If, however, X and Y have a unit root and are cointegrated (have a common trend), then one needs to test whether the lagged error-correction term and/or the lags of changes in Y (ΔY) are jointly significant in the following regression:

$$(2) \qquad \Delta X_t = \text{constant} + \gamma EC_{t-1} + \sum_i \delta_{xi} \Delta X_{t-i} + \sum_j \delta_{yj} \Delta Y_{t-j},$$

where EC is an error-correction term that captures the long-run relationship between contemporaneous values of X and Y. After describing the indicators and home sales variables, I show that these variables have unit roots (implying that first differences need to be used) and that two of the indicators are cointegrated with home sales (implying that equation 2 should be used to test for lead–lag relationships for those variables).

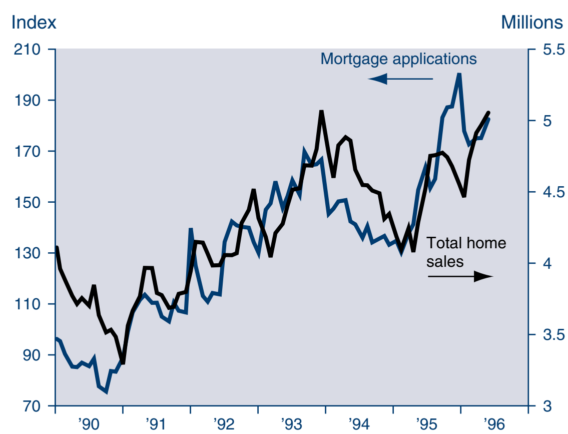
Because significance test results are sometimes sensitive to the choice of lag length, three approaches to picking lag lengths are tested. However, since the empirical results are unaffected by the choice of lag length, the tables report F statistics on regressions using lags selected with the Akaike criterion.[2]

Data and variables. Four data series are used in this study: total home sales, mortgage applications, real mortgage interest rates, and housing affordability.

Total home sales. Total home sales (THS) are measured by the sum of existing home sales (with data from the National Association of Realtors) and new home sales (with data from

Figure 1

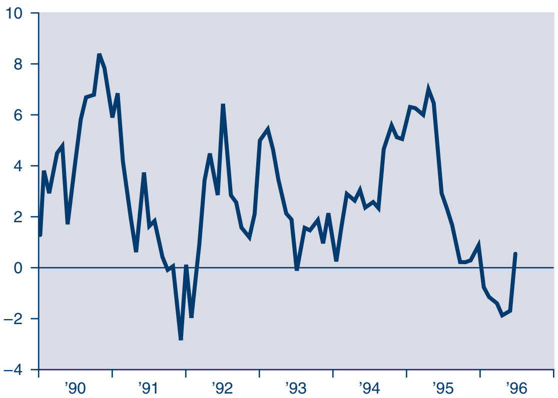
Total Home Sales and Mortgage Applications Index Trend Together



SOURCES: Mortgage Bankers Association; National Association of Realtors; U.S. Census Bureau.

Figure 2

Real, After-Tax Mortgage Rates Have Varied Much in the 1990s



SOURCES: Author's calculations; Federal Home Loan Mortgage Corp.; National Association of Realtors.

the U.S. Census Bureau). This sum is used instead of existing homes sales because the mortgage applications index and the housing affordability index do not distinguish between new and existing home sales.[3]

Mortgage applications. Mortgage applications (*MAPP*) are measured by the monthly average of the weekly MBA index of mortgage applications for home purchases, where weekly data are converted into monthly averages on a business week basis and the weekly data are seasonally adjusted using factors estimated by Federal Reserve Board staff.[4] As shown in Figure 1, the MBA index began moving slightly ahead of total home sales during the 1995 second-half surge in home sales, much as it did before the mid-1993 jump in total sales.

Real mortgage interest rates. The real, after-tax mortgage rate (*RMORT*) equals

$$(3)\qquad RMORT = [(1 - t) \times \textit{mortgage rate}] - \textit{housing inflation}$$
$$= [(1 - .28) \times \textit{mortgage rate}] - \textit{housing inflation},$$

where *t* is the marginal income tax rate (assumed to be 0.28 for most homeowners), *mortgage rate* is the conventional thirty-year fixed rate (contract commitment rate data from the Federal Home Loan Mortgage Corp.), and *housing inflation* is measured as the twelve-month percentage change in the median price of existing homes.[5] The tax adjustment reflects the tax deductibility of mortgage interest, and subtracting housing inflation attempts to adjust mortgage costs for a measure of expected housing price gains. Figure 2 shows that this real interest rate measure has varied much in the 1990s.

Housing affordability. The final indicator tested is the composite, housing affordability index (*AFFORD*) from the National Association of Realtors (NAR). This index is the ratio of median family income to the income needed to qualify for a typical mortgage, expressed as a percentage (that is, a reading of 100 means the ratio is 1:1). The qualifying income is based on a thirty-year mortgage on a median priced home for which the homeowner provides a 20-percent down payment, pays a mortgage interest rate equal to the initial rate averaged across fixed-rate and adjustable-rate mortgages, and has a monthly mortgage payment equal to 28 percent of monthly income. As income rises relative to mortgage payments, the index rises, a reflection that the median family is better able to afford a "typical" home. Mirroring the recovery of home sales since the 1990–91 recession, this index has trended up since 1990 (*Figure 3*).

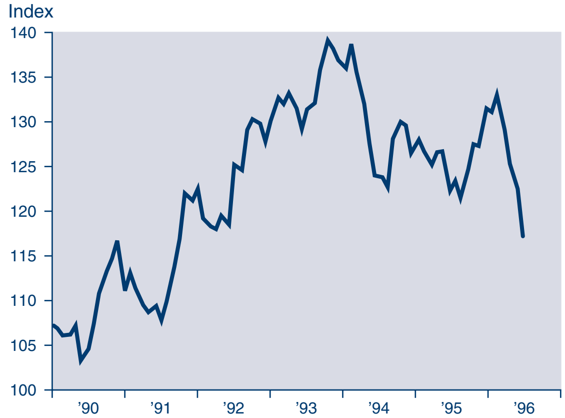
Unit roots and stationarity. Before running Granger causality tests, augmented Dickey–Fuller tests are run to see whether the levels or first differences of the indicators and home sales variables are stationary. Specifically, if one cannot reject the hypothesis that the coefficient on the term ρ on the lagged value of the variable *Y* equals 1 in the following regression, then *Y* is nonstationary:

$$(4)\qquad \Delta Y_t = \text{constant} + (\rho - 1)Y_{t-1} + \lambda_1 \Delta Y_{t-1} \ldots + \lambda_i \Delta Y_{t-i},$$

where Δ is the first difference of a variable, and the Greek letters denote parameters that are estimated. To test for unit roots allowing for a

Figure 3

Housing Affordability Has Risen Since 1990



SOURCES: National Association of Realtors.

Table 1
Augmented Dickey–Fuller Test Results

	τ statistics without trend	τ statistics with trend	Lag order (*k*)
Log levels			
THS	–1.12	–3.07	1
MAPP	–1.04	–2.53	1
AFFORD	–1.91	–1.51	1
Levels			
THS	–1.10	–3.06	1
MAPP	–.93	–2.62	1
AFFORD	–1.90	–1.52	1
RMORT	1.39	–1.25	8
First differences of logs			
THS	–5.88**	–5.88**	1
MAPP	–6.93**	–6.88**	1
AFFORD	–4.45**	–4.58**	1
First differences of levels			
THS	–5.97**	–5.98**	1
MAPP	–6.99**	–6.95**	1
AFFORD	–3.80**	–4.36**	8
RMORT	–12.08**	–12.26**	1

*(**) denotes significance at the 5- (10-) percent level. Level data: January 1990 to May 1996.

NOTES: The lag length *k* is determined by the Schwartz information criterion for $1 \leq k \geq 8$, which yields the same lags as the Akaike criterion.

Because the level of the real mortgage rate has some negative observations, the log of this variable is not continuously defined. For this reason, the level of *RMORT* is used in cointegration and causality tests involving the log of total home sales. Qualitative results are similar using levels and first differences of levels of all the variables.

SOURCES: *THS* = total home sales, existing (National Association of Realtors) + new (U.S. Census Bureau); *MAPP* = index of mortgage applications (Mortgage Bankers Association); *AFFORD* = home affordability index (National Association of Realtors).

Table 2
Dynamic OLS Cointegration Tests

$X = \text{constant} + \beta_y Y + \sum_{i=-8}^{8} \gamma_{xy} \Delta X_{t-i} + \sum_{i=-8}^{8} \delta_{xy} \Delta Y_{t-i}$

Dependent variable: total home sales (in logs)

Variables	Constant	β_y	Dickey–Fuller τ statistics
MAPP	–.635**	.431**	–3.650**
	(–3.89)	(12.95)	
AFFORD	–4.053**	1.142**	–3.801*
	(–5.30)	(7.21)	
RMORT	1.509**	–.022	–2.707
	(17.57)	(–.73)	

*(**) denotes significance at the 5- (10-) percent level.

NOTES: Raw monthly data: January 1990 to May 1996. The error-correction terms used are estimated by Stock and Watson's dynamic OLS, with leads and lags equal to eight. Cointegration tests assess whether nonstationary variables are significantly related to one another over the long run. The cointegrating vectors indicate the long-run equilibrium relationships between the variables.

Because the level of the real mortgage rate has some negative observations, the log of this variable is not continuously defined. For this reason, the level of *RMORT* is used in regressions of the log of total home sales, which is analogous to testing for a long-run semirate elasticity of home sales. Qualitative results are similar using levels and first differences of levels of all the variables.

time trend, I add a linear time trend term to equation 4 and test the joint hypothesis that the time trend equals zero and the term ρ equals 1. If the test statistic (the τ statistics in Table 1) for the joint hypothesis is significant, then the variable is stationary according to critical values specified in Dickey and Fuller (1979). The τ statistics in Table 1 indicate that the logs and levels of each of these variables are nonstationary, but the first differences of the logs and levels of these variables are stationary.

Because the indicator and sales variables have unit roots, we should check whether these variables are cointegrated (that is, are significantly related over the long run). Following the dynamic ordinary least squares (dynamic OLS) approach of Stock and Watson (1993), tests are run to see whether each indicator is cointegrated with home sales and whether various combinations of indicators are cointegrated as well, using lag and lead lengths of eight.[6] For variables X and Y that have unit roots, these tests involve running the following type of regression:

(5) $$X = \alpha_x + \beta_y Y + \sum_{i=-8}^{8} \gamma_{xi} \Delta X_{t-i} + \varepsilon_{xt}.$$

If the constant (α_x) and β_y term are significant and if augmented Dickey–Fuller tests confirm that the cointegrating residuals are stationary, then X and Y are cointegrated. Since home sales should rise with either mortgage applications or affordability, the β_y coefficients are expected to have positive signs for these indicators. In contrast, sales should be negatively related to real mortgage rates, implying a negative β_y coefficient on *RMORT*.

Results (*Table 2*) indicate that homes sales are cointegrated with *MAPP* and *AFFORD*, with the anticipated positive signs on the β_y coefficients. In contrast to these indicators, *RMORT* is not cointegrated with total home sales.

Testing whether housing indicators lead home sales. Causality test results are in Table 3, where housing indicators are evaluated individually in bivariate lead–lag tests based on running equation 2 for tests involving *AFFORD* and *MAPP*.[7] Because *RMORT* and *THS* are not cointegrated, causality tests involving *RMORT* are based on equation 1.

There are six important patterns of findings. First, each home sales indicator contains statistically significant information about future movements in home sales, as indicated by the significant coefficients on EC_x for *MAPP* and *AFFORD*, the joint significance of the EC_x and δ_{xy} terms for *MAPP* and *AFFORD*, and the joint significance of the δ_{xy} coefficients for *RMORT*. Second, the significance of the applications and affordability indexes stems from highly significant lagged error-correction terms, rather than from $t - 1$ changes in these variables. This finding implies that the levels of—rather than the changes in—both of these indexes are most informative. Third, the implied lead times, denoted by *k* in Table 3, are plausible: two months for real mortgage rates and one month for both affordability and applications. Fourth, there is some evidence of bidirectional causality in that home sales have statistically significant information about future changes in each of the indicators. In contrast to results for causality from the MBA and NAR indexes to home sales, the causality running from home sales to these indicators stems from the significance of lagged changes, rather than the significance of the lagged error-correction term. This finding is consistent with results in other tests (not shown) in which nonstationary logs and levels of *MAPP*, *AFFORD*, and *RMORT* lead home sales, but home sales do not lead the three indicators. Fifth, the evidence of causality from home sales to housing indicators is weaker than in the opposite direction, as reflected by the *F* statistics for each combination of variables used. The sixth interesting pattern is that evidence of home sales leading housing indicators is weakest for the MBA index, as reflected not only in the smaller *F* statistics on the lagged change in home sales, but also in the joint *in*significance of the lagged sales change and error-correction term.

What could account for bidirectional causality between total home sales and the housing indicators? One very plausible explanation for housing affordability and real mortgage rates is that home sales have a lagged effect on future housing prices, which, in turn, affects the affordability of housing and the housing price appreciation term used in constructing the real mortgage interest rate. This account is consistent with the negative sign on the $t - 1$ change in home sales in causality tests of *AFFORD* and *RMORT* (coefficients are not shown in the tables to conserve space). For example, a sustained run-up in home sales will eventually cause a pickup in home price inflation, which, in turn, reduces affordability and the real mortgage interest rate.

With respect to the mortgage applications and housing affordability, reverse causality could conceivably arise from sudden shifts in the timing of home purchases. Normally, home sales and mortgage applications have swings lasting several months, giving rise to positive correlations between current and past values of each series. Consider, then, what happens if many families suddenly hasten their planned home purchases at the expense of future purchases. This month's surge in applications will, via a positive autocorrelation in applications, lead a Granger model of applications to predict more strength next month. However, the negative "payback" effect on next month's sales and applications from the temporary speed-up will induce the model to estimate that this month's jump in sales will cause a decline in applications next month. If such a surge reflects people's reaction to a sudden change in affordability, then a Granger model of affordability will also estimate a negative future response to current home sales growth for analogous reasons. This account is consistent with the negative estimated effects of the $t - 1$ lag of home sales growth on the percentage changes in mortgage applications and housing affordability (coefficient estimates are not shown in the tables to conserve space).

Overall, the bivariate tests in Table 3 are mixed in terms of whether the mortgage applications index is better than the housing affordability index as an indicator of total home sales. On the one hand, the affordability index is more statistically significant than the

Table 3
Bivariate Causality (Lead–Lag) Tests

Specifications for tests involving MAPP *and* AFFORD*:*

$$\Delta X = \text{constant} + EC_x[X - \alpha_{xy} - \beta_{xy} Y]_{t-1} + \textstyle\sum_{i=1}^{k} \gamma_{xy} \Delta X_{t-i} + \sum_{i=1}^{k} \delta_{xy} \Delta Y_{t-i}$$

Specifications for tests involving RMORT*:*

$$\Delta X = \text{constant} + \textstyle\sum_{i=1}^{k} \gamma_{xy} \Delta X_{t-i} + \sum_{i=1}^{k} \delta_{xy} \Delta Y_{t-i}$$

Direction of timing	$EC_x = 0$	$\sum_{i=1}^{k}\delta_{xy} = 0$	$EC_x = 0$ and $\sum_{i=1}^{k}\delta_{xy} = 0$	***k***
MAPP –> THS	45.71**	.54	29.96**	1
THS –> MAPP	.87	4.43*	2.22	1
AFFORD –> THS	74.94**	3.67*	66.74**	1
THS –> AFFORD	.21	5.62*	3.52*	1
RMORT –> THS	n.a.	14.24**	n.a.	2
THS –> RMORT	n.a.	5.10**	n.a.	1

*(**,+) denotes significance at the 5- (1-, 10-) percent level.

n.a. denotes not applicable, as *RMORT* is not cointegrated with *THS*.

NOTES: The raw data used span January 1990 to June 1996, implying a sample of September 1990 to June 1996. All variables are in logs, except *RMORT*. The error-correction terms used are estimated by Stock and Watson's dynamic OLS, with leads and lags equal to eight. *F* statistics for the Granger causality tests are reported along with their *p* values in parentheses.

Because the level of the real mortgage rate has some negative observations, the log of this variable is not continuously defined. For this reason, the level of *RMORT* is used in regressions of the log of total home sales, which is analogous to testing for a long-run semirate elasticity of home sales. Qualitative results are similar using levels and first differences of levels of all the variables.

Table 4

Models of the Percentage Change in Total Home Sale That Overlook the More Timely Availability of Mortgage Applications

	Bivariate models		Multivariate models		
Variables	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
constant	−.005+	.003	−.003+	.002	.003
	(−1.78)	(1.05)	(−1.83)	(1.05)	(1.21)
$ECMAPP_{t-1}$	−.328**		−.269**		
	(−6.76)		(−8.48)		
$ECAFFORD_{t-1}$		−.206**		−.186**	−.201*
		(−8.66)		(−4.83)	(−5.78)
ΔTHS_{t-1}	.128	.186**	−.024	−.020	.030
	(1.30)	(10.54)	(−.26)	(−.28)	(.54)
$\Delta MAPP_{t-1}$	−.044		.003	.068	
	(−.73)		(.07)	(1.53)	
$\Delta AFFORD_{t-1}$		−.358+	−.235	−.273+	−.253+
		(−1.91)	(−1.62)	(−1.92)	(−1.73)
$\Sigma_{i=1}^{2} \Delta RMORT_{t-i}$			−.010**	−.015**	−.014**
			(13.22)	(43.45)	(17.22)
$\bar{R}^2$	.238	.156	.334	.346	.339
SSE	.0560	.0621	.0465	.0457	.0469
$Q(19)$	21.37	17.34	12.34	11.32	12.04

* (**,+) denotes significance at the 5- (1-, 10-) percent level.

NOTES: Bivariate sample: March 1990 to May 1996. Multivariate sample: April 1990 to May 1996. All variables are in logs, except *RMORT*. The error−correction terms, based on the cointegrating vectors reported in Table 2, are estimated by Stock and Watson's dynamic OLS, with leads and lags equal to eight. *T* statistics in parentheses for individual variables and *F* statistics in parentheses for the lags of Δ*RMORT*.

Because the real mortgage rate has some negative observations, its log is not always defined. For this reason, the level of *RMORT* is used, which is analogous to testing for a semirate elasticity of home sales. Qualitative results are similar using levels and first differences of levels of all the variables.

other indicators in causality tests running from housing indicators to home sales. On the other hand, there is more statistically significant evidence of causality running from home sales to affordability than from home sales to mortgage applications.

Do mortgage applications contain information not reflected in alternative indicators?

To determine whether mortgage applications contain information about home sales not reflected in housing affordability and real, after-tax mortgage rate data, several groups of regressions are run with the percentage change (Δlog) of monthly home sales as the dependent variable. Although percentage changes of most monthly series tend to be very noisy and to have lower model fits than models of quarterly data, percentage changes are used, given the nonstationarity of the variables over the short sample period.[8] Monthly rather than quarterly data are used because this article focuses on assessing the short-term information the mortgage applications index may contain—especially since monthly MBA data are available three to four weeks ahead of most other housing data.

The first set of regressions evaluates the three indicators in full-sample regressions that assume the indicators are available at the same time. The second set of runs is similar, except that the greater timeliness of the mortgage applications index is taken into account. In the third set of runs, two multivariate models are evaluated in ex post forecasts. Based on the forecasts, this section concludes with a discussion of possible conditions under which the applications index may give a biased signal of home sales.

In-sample results assuming the same timing of data. The first set of models (*Table 4*) assumes that data on *RMORT*, *AFFORD*, and *MAPP* are available at the same time. The first two models correspond to the bivariate causality models in Table 3 used to assess whether the MBA index or affordability index lead home sales. Model 1 includes an error-correction term based on the cointegrating vector for home sales and mortgage applications (*ECMAPP*), along with lags of first differences of sales and mortgage applications, where lag lengths are based on the Akaike information criterion. The second model incorporates an error-correction term based on the cointegrating vector for home sales and affordability (*ECAFFORD*), along with lags of first differences of sales and affordability. The remaining three models are multivariate models. The third model corresponds to model 1, except that it includes the $t-1$ lag of the log first difference in affordability along with the $t-1$ and $t-2$ lags of the change in real mortgage rates, where lag lengths are also based on the Akaike information criterion.[9] Similarly, the fourth model corresponds to model 2, except that it includes the $t-1$ lag of the log first difference in mortgage applications along with the $t-1$ and $t-2$ lags of the change in real mortgage interest rates. The fifth model is similar to model 4, except that it completely excludes the mortgage applications index.[10]

Table 4 shows several noteworthy findings.

First, the error-correction coefficient on $ECMAPP_{t-1}$ in model 1 is larger in size and more significant than the error-correction term ($ECAFFORD_{t-1}$) in the corresponding bivariate model (*model 2*) that includes mortgage applications rather than housing affordability. Second, the more significant error-correction term in model 1 likely accounts for the much better fit ($\bar{R}^2$) of model 1 versus model 2 because the one-month lag of the change in applications is statistically insignificant in model 1, whereas the one-month lag of the change in affordability is marginally significant in model 2. Third, a comparison of the $\bar{R}^2$s and the sum of squared errors (SSE) across the multivariate models (3, 4, and 5) reveals that the applications index adds no substantial extra information about total home sales in the presence of lagged changes in real mortgage rates. Overall, the in-sample results imply that while the mortgage applications index adds information about future total home sales in bivariate models, it adds no marginal information in the presence of lagged changes in real mortgage interest rates, assuming that all variables are available at the same time.

Table 5

Models of the Percentage Change in Total Home Sales That Reflect the More Timely Availability of Mortgage Applications

	Bivariate models		Multivariate models			
Variables	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
constant	–.005^{+}	.003	–.004	.003	.004^{+}	.001
	(–1.78)	(1.05)	(–1.57)	(1.21)	(–1.73)	(.70)
$ECMAPP_{t-1}$	–.328**		–.229**		–2.61**	
	(–6.76)		(–5.21)		(–5.33)	
$ECAFFORD_{t-1}$		–.206**		–.201*		–.163**
		(–8.66)		(–5.78)		(–3.51)
ΔTHS_{t-1}	.128	.186**	.054	.030	.045	.021
	(1.30)	(10.54)	(.68)	(.54)	(.63)	(.30)
$\Delta MAPP_t$	.138**		.114**		.112**	
	(3.78)		(3.11)		(3.37)	
$\sum_{i=1}^{2}\Delta MAPP_{t-1}$						.153^{+}
						(2.71)
$\Delta AFFORD_{t-1}$		–.358^{+}		–.253^{+}	–.226	–.243
		(–1.91)		(–1.73)	(–1.53)	(–1.63)
$\sum_{i=1}^{2}\Delta RMORT_{t-i}$			–.010**	–.014**	–.009**	–.014**
			(5.11)	(17.22)	(5.72)	(11.73)
$\bar{R}^2$	.311	.156	.378	.339	.386	.363
SSE	.0507	.0621	.0441	.0469	.0429	.0438
Q(19)	15.16	17.34	14.26	12.04	14.26	13.26

* (**,$^{+}$) denotes significance at the 5- (1-, 10-) percent level.

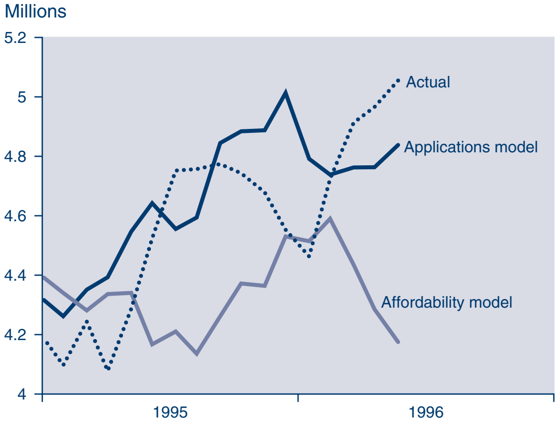
NOTES: Bivariate sample: March 1990 to May 1996. Multivariate sample: April 1990 to May 1996. All variables are in logs, except *RMORT*. The error–correction terms, based on the cointegrating vectors reported in Table 2, are estimated by Stock and Watson's dynamic OLS, with leads and lags equal to eight. *T* statistics in parentheses for individual variables and *F* statistics in parentheses for the lags of $\Delta RMORT$.

Because the real mortgage rate has some negative observations, its log is not always defined. For this reason, the level of *RMORT* is used, which is analogous to testing for a semirate elasticity of home sales. Qualitative results are similar using levels and first differences of levels of all the variables.

Accounting for the greater timeliness of mortgage applications data. The regressions in Table 4 overlook the fact that mortgage applications data are available roughly three weeks before the other indicators. Specifically, the MBA index comes out with less than a one-week lag, whereas existing and new home sales data are released with a three- to four-week lag, as are data needed to construct the real mortgage rate and home affordability measures. For example, by the first Thursday of November 1996, complete MBA index data through October 1996 would be available and could be used to predict October 1996 housing sales data that will be released in early December. In contrast, data on home sales would be available only through September 1996. Thus, if one were to predict home sales for October 1996 at the beginning of November, one would only be able to use data on home sales, real mortgage rates, and home affordability through September and MBA index data through October.

Some models in Table 5 incorporate this timing advantage by replacing the $t - 1$ lag of $\Delta MAPP$ in several models in Table 4 with the contemporaneous change. These models can be used to predict the previous month's sales at the end of the first week of the current month, three to four weeks ahead of the data release. Two key results arise. First, unlike its $t - 1$ lag, the month t change in mortgage applications is always statistically significant. Second, in contrast to Table 4, the multivariate models with error-correction terms based on applications outperform corresponding models using error-correction terms based on affordability (model 3 versus model 4, and model 5 versus model 6). Thus, when the greater timeliness of the MBA applications index is taken into account, it does add statistically significant, albeit economically modest, information on total home sales in the

Figure 4
Actual and Equilibrium Total Home Sales



SOURCES: Author's calculations; Mortgage Bankers Association; National Association of Realtors; U.S. Census Bureau.

presence of lagged real mortgage rate changes. Nevertheless, the high degree of noisiness in the growth rate of total home sales makes it a difficult series to precisely predict, as evidenced by the low $\bar{R}^2$s. Models of the level of home sales activity have better fits but are plagued by the difficulty of making statistical inferences from models using nonstationary variables. One way around this problem is to use growth rate predictions to construct implied levels forecasts, as illustrated in the next section.

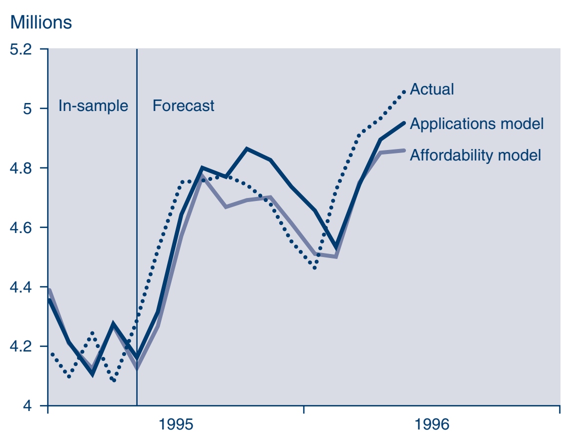
Multivariate forecasts. To shed more light on the practical use of the MBA index as an indicator, ex post forecasts are constructed based on three multivariate models and are plotted in two separate charts. These forecasts use actual data and apply coefficients estimated from these models using an in-sample period of February 1990 to May 1995. The first model is the multivariate model 4 from Table 5, which omits information from the MBA index and uses lagged changes in housing affordability, home sales, and real mortgage interest rates. The second is model 3 from Table 4, which uses the MBA index to define the error-correction term, along with lagged changes in housing affordability, home sales, mortgage applications, and real mortgage interest rates. The last specification is model 5 from Table 5, which is identical to model 3 from Table 4, except that the contemporaneous first difference of mortgage applications replaces the $t - 1$ lag to reflect the more timely release of the MBA index.

The sums of squared forecast errors are roughly equal for the first two models (0.01139 for model 4 from Table 5 and 0.01136 for model 3 from Table 4), whereas the SSE from model 5 in Table 5 is nearly 20 percent lower (0.00937). Thus, one can conclude from these ex post forecasts that the advantage of using the MBA index stems from its greater timeliness. Nevertheless, further analysis indicates that the long-run relationship between mortgage applications and home sales has held up better out-of-sample than that between affordability and home sales. This finding is shown in Figure 4, which plots sales along with the equilibrium levels implied by the error-correction terms from model 4 and 5 in Table 5, the latter of which is common to model 3 in Table 4. Clearly, the mortgage applications index yields equilibrium levels that more closely oscillate with actual home sales, suggesting that its usefulness, relative to that of the affordability index, may increase in the future.

To shed more light on these ex post forecasts, Figure 5 plots the actual level of total home sales along with the levels implied by the forecasts of models 4 and 5 from Table 5. Although the models are regressions of the percentage change in sales, implied levels forecasts are perhaps more relevant because the noisiness of percentage changes makes the levels data more indicative of the overall tone of housing activity. In this chart, the implied level for month t equals the actual level of home sales in month $t - 1$ multiplied by the sum of 1 and the forecasted percentage change in sales for month t.

Two patterns are apparent in Figure 5. First, the applications model (*model 5*) better tracks the rise in home sales during the fall of 1995 and the spring of 1996. Second, this model does worse in the winter of 1995–96, when it overpredicts sales activity in a period when unusually bad weather or government shutdowns could have distorted the normal pattern of mortgage applications and closings.

Figure 5
Forecasts of Total Home Sales



SOURCES: Author's calculations; Mortgage Bankers Association; National Association of Realtors; U.S. Census Bureau.

Why did the MBA index overstate home sales in the winter of 1995–96? Because virtually every indicator can sometimes distort the type of economic activity it is intended to track, it is important to understand how and why an indicator may provide a biased picture. From this perspective, it may be helpful to examine potential reasons mortgage applications overstated the pace of home sales last winter, especially given the index's short history. Such possible explanations may help us identify future episodes when the index could give a biased signal of housing.

The overpredictions of sales last winter from the applications models could reflect several factors. In particular, the combination of delays from bad weather and sharp changes in interest rates may have caused mortgage commitments to expire and led some people to reapply for mortgages later. Thus, some applications made in late 1995 may have never shown up in home sales, while some of February's strength in applications may have reflected reapplications from expired loan commitments. Considering the index's brief history, it is not feasible to rigorously assess the impact of weather (see Goodman 1987 and Cammarota 1988 regarding the estimation of weather effects). Nevertheless, the poorer performance of the applications model last winter, coupled with its better performance in the fall of 1995 and the spring of 1996, suggests that the MBA index may be less reliable during periods of severe weather.

An alternative and perhaps more plausible explanation is that the federal government shutdowns during the winter of 1995–96 limited the availability of FHA- and VA-insured mortgages and caused some households to shift toward conventional mortgages. Since the index tracks conventional mortgage applications, this past strength in the index may have reflected an increase in conventional market share stemming from government shutdowns, rather than a rise in total housing sales activity. Correspondingly, the FHA and VA share of all mortgage originations fell in the first quarter of 1996 to a level (13.2 percent) that was 1-percentage point below its year-earlier level.[11] However, because these originations data are not seasonally adjusted and include mortgage refinancings, and because data for all of 1996 are not yet available, this evidence is suggestive rather than conclusive. Nevertheless, the large forecast errors from the mortgage applications models last winter give us some insight as to what conditions *could* cause the index to give a distorted picture of home sales activity.

Conclusion

Results show that, by itself, the MBA index of mortgage applications for home purchases is a good, albeit imperfect, predictor of total home sales that clearly outperforms a housing affordability index. In addition, the long-run equilibrium relationships suggest that the usefulness of the MBA index may increase in the future. However, when included with housing affordability and real, after-tax mortgage interest rate data, the index adds no extra information when its greater timeliness is ignored. This last result is not surprising, given that the housing literature has established that home-buying and, thus, mortgage applications, are primarily driven by income, mortgage interest rates, and housing appreciation, all of which are reflected in the other two housing indicators. However, when the more timely availability of the mortgage applications index is taken into account, it adds some information about the pace of total home sales. With this critical qualification in mind, the MBA's index of mortgage applications for home purchases can help forecast total home sales in the near term. For example, when this article was written, the index pointed to a slight decline in total home sales in the summer only of 1996 from the high and unsustainable level of May 1996. Market analysts, however, generally had predicted a more sizable decline in home sales than had actually occurred.

Nevertheless, even after accounting for its short lead time, the MBA applications index should be used cautiously. The index has a relatively brief history, and some evidence suggests that its performance may falter in periods of severe weather or when home sales are affected by unusual shifts in the conventional share of mortgage originations.

Notes

I would like to thank my referees, D'Ann Petersen and Stephen Prowse, for providing detailed comments; Ken Emery for extending suggestions; and Michelle Burchfiel and Jean Zhang for providing research assistance. Any remaining errors are my own.

1 By aiming for a moderate, stable pace of aggregate demand growth, the Federal Reserve seeks to create a low-inflation, stable environment that is conducive to promoting its main goal of sustainable economic growth.

2 Three approaches to setting lag lengths were tried. First, lag lengths on both ΔX and ΔY were arbitrarily set at two, four, six, and eight months. Second, lag lengths were picked according to Akaike's (1969) FPE criterion, which limits the lags to lengths that balance the information gained from including more lags relative to the number of lags that are included. Third, lag lengths

were chosen based on the Schwartz criterion, which, relative to the Akaike criterion, puts slightly more weight on the number of observations and slightly less weight on the number of regressors. Nevertheless, the Akaike and Schwartz criteria picked the same lag lengths, and the qualitative results were unaffected by using these information criteria instead of the alternative arbitrary lag lengths.

[3] Qualitative results were similar using the National Association of Realtors' definition of total single-family sales, which equals existing home sales plus 97 percent of single-family housing starts. Single-family starts exceed new home sales because some of the starts are for homes that are planned as rentals and because some of the starts are for homes that are ordered by landowners and are not technically sold.

[4] The techniques used by the Federal Reserve staff prevent calendar anomalies (holidays and year-end dates) from biasing the estimated seasonal factors, in contrast to the less involved approach used by the MBA.

[5] One drawback of using the thirty-year fixed mortgage rate to define *RMORT* is that shifts between adjustable- and fixed-rate mortgages could cushion the impact of changes in fixed mortgage rates on housing. A housing affordability index, which is described elsewhere in this article, avoids this potential problem by using the average rate on adjustable- and fixed-rate mortgages to measure housing affordability. The problem may be limited, however, because in estimating housing construction since 1960, Duca (forthcoming) found little difference in results between defining a real mortgage interest rate based on a fixed mortgage rate and one based on an average of adjustable and fixed rates.

[6] Cointegration results were qualitatively similar using the approach of Johansen and Juselius (1990) to estimating cointegrating vectors.

[7] The computer programs used were adapted from those employed by Emery and Chang (1996).

[8] The preferred models, which use seasonally adjusted data, have corrected $\bar{R}^2$s around 0.35. Goodman obtains higher $\bar{R}^2$s for separate models of the percentage changes in new (around 0.50) and existing (around 0.74) home sales (see Goodman 1987, columns B and C in appendix tables 2 and 3, pages 655 and 656). However, Goodman uses data that are not seasonally adjusted because his study focuses on estimating weather effects. Also, the fits are boosted relative to those in my study because Goodman includes eleven highly significant monthly dummy variables to control for seasonal variation.

[9] As with the causality test results, the Akaike and Schwartz information criteria implied the same lag lengths in every case.

[10] The cointegrating vectors do not include real mortgage rates because the vector estimated for home sales, housing affordability, and real mortgage rates yielded counterintuitive signs. In addition, the cointegrating vectors do not combine information from mortgage applications and affordability for two reasons. First, such a vector had a negative, counterintuitive sign on affordability. Second, in second-stage models of home sales growth, models using the "combined" error-correction term had worse fits than models using the bivariate error-correction terms.

[11] The combined VA and FHA share of mortgage originations averaged 15 percent over 1994 and 1995, ranging between 11 and 22 percent during this period.

References

Akaike, H. (1969), "Fitting Autoregressive Models for Prediction," *Annals of the Institute of Statistical Mathematics* 21 (2): 243–47.

Cammarota, Mark T. (1988), "The Impact of Unseasonable Weather on Housing Starts," *Economic Activity Working Paper* no. 86 (Board of Governors of the Federal Reserve, May).

Dickey, David A., and Wayne A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* 74 (June): 427–31.

Duca, John V. (forthcoming), "Deposit Deregulation and the Sensitivity of Housing," *Journal of Housing Economics*.

Emery, Kenneth M., and Chih-Ping Chang (1996), "Do Wages Help Predict Inflation?" Federal Reserve Bank of Dallas *Economic Review*, First Quarter, 2–9.

Goodman, John L. (1987), "Housing and the Weather," *AREUEA Journal* 15 (Spring): 638–63.

Johansen, Soren, and Katarina Juselius (1990), "Maximum Likelihood Estimation and Inference on Cointegration—with Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics* 52 (May): 169–210.

Kahn, George A. (1989), "The Changing Interest Sensitivity of the U.S. Economy," Federal Reserve Bank of Kansas City *Economic Review*, November, 13–34.

Mauskopf, Eileen (1990), "The Transmission Channels of Monetary Policy: How Have They Changed?" *Federal Reserve Bulletin*, December, 985–1008.

Pesek, William, Jr. (1996), "U.S. Housing Sector Remains Strong Despite Rising Rates," *Dow Jones Capital Markets Report*, Dow Jones News Service, July 9.

Stock, James, and Mark Watson (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica* 61 (July): 783–820.