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:

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

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:

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

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).

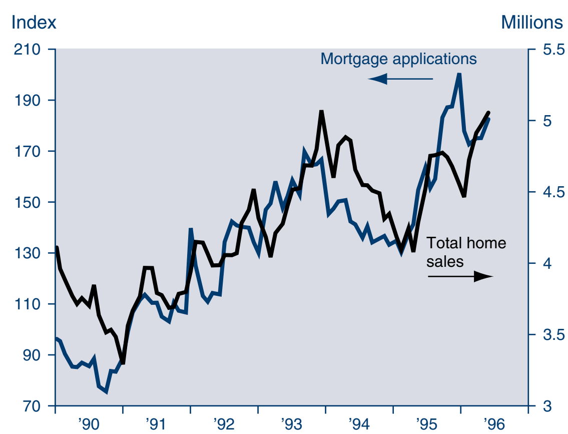
Because significance test results are sometimes sensitive to the choice of lag length, three approaches to picking lag lengths are tested. However, since the empirical results are unaffected by the choice of lag length, the tables report F statistics on regressions using lags selected with the Akaike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

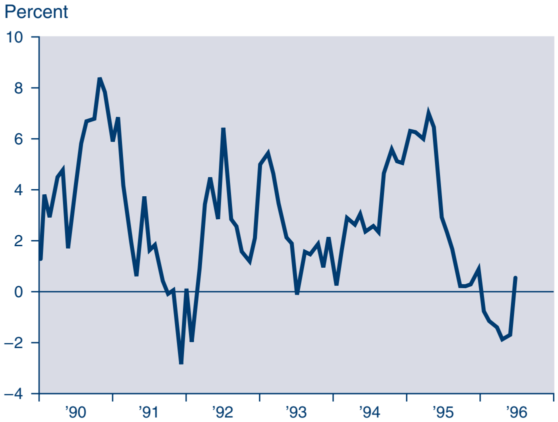
Figure 1

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 $1 \le k \ge 8$, 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** (–3.89)	.431** (12.95)	–3.650**
AFFORD	–4.053** (–5.30)	1.142** (7.21)	–3.801*
RMORT	1.509** (17.57)	–.022 (–.73)	–2.707

*(**) 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:

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

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)
$\sum_{i=1}^{2}\Delta RMORT_{t-i}$			–.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 $\Delta RMORT$.

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.

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.

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)	
$\Sigma^2_{i=1}\Delta MAPP_{t-1}$						.153+
						(2.71)
$\Delta AFFORD_{t-1}$		–.358+		–.253+	–.226	–.243
		(–1.91)		(–1.73)	(–1.53)	(–1.63)
$\Sigma^2_{i=1}\Delta RMORT_{t-i}$			–.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 $\Delta RMORT$.

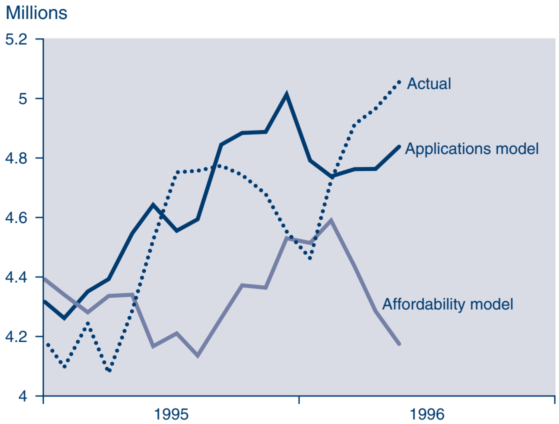
Because the real mortgage rate has some negative observations, its log is not always defined. For this reason, the level of *RMORT* is used, which is analogous to testing for a semirate elasticity of home sales. Qualitative results are similar using levels and first differences of levels of all the variables.

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.

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.

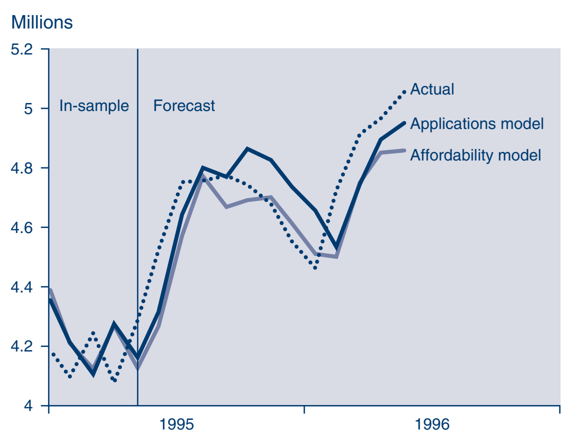
Multivariate forecasts. To shed more light on the practical use of the MBA index as an indicator, ex post forecasts are constructed based on three multivariate models and are plotted in two separate charts. These forecasts use actual data and apply coefficients estimated from these models using an in-sample 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.