Journal article by Richard K. Green; Real Estate Economics, Vol. 25, 1997

Follow the Leader: How Changes in Residential and Non-residential Investment Predict Changes in GDP

by Richard K. Green
Graduate School of Business
University of Wisconsin-Madison
Madison, WI 53706
(608) 262-5227 (P)
(608) 265-2738 (F)
rgreen@bus.wisc.edu

March 7, 1996

Abstract

This paper examines the effect of different kinds of investments on the business cycle. Specifically, it examines whether residential and non-residential investment Granger cause GDP, and whether GDP Granger causes each of these types of investments. Under a wide variety of specifications, residential investment causes, but is not caused by GDP, while non-residential investment does not cause, but is caused by GDP. Thus housing leads and other types of investment lag the business cycle. The results also suggest that policies designed to funnel capital away from housing into plant and equipment could produce severe short-run dislocations.

Follow the Leader: How Changes in Residential and Non-residential Investment Predict Changes in GDP

by Richard K. Green
Graduate School of Business
University of Wisconsin -- Madison

March 7, 1996

Introduction

Economists have generally maintained that the United States has invested too much in housing (see Mills (1987) and Hendershott (1989) for several citations). Specifically, they have noted how tax distortions have caused capital to flow out of plant and equipment investment and into housing. Some have gone so far as to speculate that the success of some of the Pacific Rim economies has been a function of the ability of governments to keep capital away from housing and forcing it instead into plant and equipment.[1]

If it is true that the United States has in fact overinvested in housing, we should find that such overinvestment has had a deleterious effect on Gross Domestic Product (GDP); in particular, we would expect to see housing investment to have a smaller impact on GDP than other types of investment. Mills (1987) tested this proposition indirectly by measuring the marginal product of housing capital relative to non-housing capital. His well-known finding was that returns to housing capital are little more than half the returns to non-housing capital. This, of course, suggests a large misallocation of capital toward housing, and a concomitant reduction in GDP growth.

But long-term growth is only one major goal of macroeconomic policy--stability is another.

Thus it might be of some interest to determine the impact of the two types of investment contemplated in Mills--residential and non-residential--on the peaks and troughs of the business cycle. A natural series of test procedures is available for this purpose: tests of Granger Causality.

The organization of the paper is as follows: it will briefly describe Sims-Granger tests and provide a rationale for the specific tests of the relationship between residential investment, non-residential investment, and GDP. It will then turn to the data set used to estimate the model and discuss the procedures for determining the types of relationships we may test using Granger causality and other time series techniques. The paper will next present model results and the implications of those results. It will finally discuss the implications of the findings..

Granger Tests

The concept of Granger causality is a simple one: if lagged values of a stationary variable X1 can improve our ability to predict another stationary variable X2 after controlling for lagged values of X2, X1 is said to "Granger Cause" X2. An excellent discussion of this technique's importance and implications is put forth in Sargent (1979). Granger causality has been used in the past to test the relationship between total investment (i.e., the sum of residential and non-residential investment) and GDP (see Gupta (1988) and Acemoglu (1993)). The technique also has its detractors; I will attempt to deal with their criticisms in my discussion of the results' implications.

The test of Granger causality I use in this paper is straightforward. We fit the equation

$$X2(t)=\sum_{i=1}^{J} a_j X2(t-j)+\sum_{i=1}^{J} b_j X1(t-j)+\alpha+\beta T+u(t) \tag{1}$$

and test the hypothesis:

$$HO\colon b_j=0, j=1,.....,J \qquad (2)$$

We may test the above hypothesis using a straightforward F-test.

At least three tests have been developed for Granger causality (Granger (1969), Sims (1972), and Geweke, Meese and Dent (1982)). A widely-cited article by Guilkey and Salemi (1982) uses experimental data to demonstrate that Granger's original test--the test described above -- outperforms the others in terms of drawing correct inferences about whether to reject or accept the hypothesis of Granger Causality. The Guilkey and Salemi work was also important because it presented experimental evidence showing that the Granger test works quite well at detecting both the presence and direction of Granger causality.

In our context, we will test whether net residential investment Granger causes GDP, and whether net non-residential investment Granger causes GDP. We will also test whether GDP Granger causes net residential and net non-residential investment. An appendix to the Guilkey and Salemi paper shows that if there is little feedback from GDP to residential and non-residential investment, we may be confident about the inferences we draw about whether residential or non-residential investment "causes" GDP.

The number of lags chosen in estimating (1) will have an impact on the decision to reject or accept the hypothesis given in (2). As a result, we take two approaches to specifying the number of lags. First, we simply follow the recommendation arising from the experimental results in Guilkey and Salami, and specify six lags. We then reject or accept the hypothesis given in (2) at the 99 percent confidence level.[2] Second, we "allow the data to speak," and use F-tests to determine the optimal number of lags for the models at the 99 percent level of confidence. Once again, we accept or reject the hypothesis given in (2) at the 99 percent confidence level.

Finally, it is important to check the structural stability of the parameters in (1) across time. Although we may, for instance, on average find causality from X1 to X2 over a long period, we may also find that causality disappears within subperiods.

A well known method for testing coefficient stability is the cusum of squares test put forth in Brown, Durbin, and Evans (1975). The cusum of squares procedure as described in Greene (1992) follows. Let

$$e_t = y_t - x_t' b_{t-1} \tag{3}$$

where b_{t-1} is the least squares coefficient vector using all observations up to but not including $[y_t, x_t]$. We create a set of scaled residuals

$$w_r = \frac{e_r}{[1 + x_r'(X_{r-1}' X_{r-1}) x_r]^{1/2}} \tag{4}$$

for r = k+1 to r = T-(k+1), where k is the number of regressors and T is the total number of

observations. Then we create the series of Cusum of squares test statistics:

$$S_t = \frac{\sum_{t=K+1}^{t} w_r^2}{\sum_{t=K+1}^{T} w_r^2} \tag{5}$$

The expected value of S_t is approximately (t-k)/(T-K), and Greene claims that for T-K greater than 100, 95 percent confidence limits can be obtained using (t-k)/(T-K) ± .11. Should the test statistic S_t cross the confidence limit, we have evidence of coefficient instability, and therefore the presence of a new set of regression relationships.

Data and Regression Strategy

The data from which I estimate the model originate with three quarterly time series from the years 1959-1992 from Citibase: real GDP, real private domestic non-residential investment, and real domestic residential investment. All series are in 1987 dollars and all series are seasonally adjusted.

Figure 1 makes clear that none of these three series is time-stationary. This presents us with a problem as to how to properly specify a model for testing the presence or absence of Granger Causality. Hamilton (1994) suggests several strategies for dealing with this issue.

One could estimate the vector-autoregression model in (1) in levels, and perform standard t-tests and F-tests on the coefficients. Unfortunately, while individual coefficients may be tested using a standard normal distribution, an F test of a joint null hypothesis that more than one coefficient is equal to zero has a nonstandard limiting distribution.

As an alternative, one could estimate (1) in first differences. However, if the data are actually stationary (not really an issue here), or if the dynamic relationship between a combination of the data is stationary (a potential concern), using first differences to test for Granger causality will produce a misspecified model. That is, if residential investment and GDP or non-residential investment and GDP are *cointegrated*, we will not wish to use a model in first differences.

Hamilton suggests that yet a third approach is to determine the nature of the non-stationarity and develop systems of equations which are in fact stationary for performing the hypothesis tests. In the context of Granger causality, this means estimating an error-corrections model. Once again, however, such a strategy could lead to misspecification.

I will therefore by-and-large follow Hamilton's final suggestion, which is essentially to try more than one of the above specifications, and to see if they produce consistent results. Should they do so, misspecification ceases to be a major issue. Specifically, I perform Granger Tests on detrended data (these may be looked at as short-run tests) and on the data themselves (long-run tests); I also estimate an error correction model. For the model that have data in levels, I test for Granger Causality at the 99 percent level, rather than the 95 percent level, because the test statistic will not have an F-distribution. Experimental data from Ohanian (1988) show that testing at the 95 percent level causes spurious rejection 20 percent of the time. Raising the threshold for rejection therefore seems appropriate.[3]

We are also able to immediately reject one specification: an estimate of the model in first differences. To see why, we now turn to tests of co-integration.

Cointegration tests

We seek to find whether GDP and residential investment are cointegrated and whether GDP and non-residential investment are cointegrated. To do so, we perform the following regressions:

$$Investment_{jt} = \alpha + \beta_j * GDP_t + u_{jt} \tag{6}$$

where j =1,2 represents non-residential and residential investment, respectively, and the u_{tj} represent corresponding error terms for the regressions. We perform a Dickey-Fuller unit-root test on the u_{jt}. If for either of the investment series we reject the null hypothesis that the error term has unit root, we may say that that series is cointegrated with GDP and thus rule out the first-difference specification. We also may then be confident that the error-correction model will be appropriate.

The unit root test procedure is shown in Hamilton (p. 585), and will be described briefly here. The residual series is regressed against itself lagged once and against a series of lagged first differences sufficient to make auto-correlation disappear. We then test the hypothesis of whether the coefficient on the lagged term is equal to one. Dickey-Fuller (1979) showed that the distribution of the test statistic for this hypothesis is not normally distributed, and they developed a table of appropriate critical values for just this test. The critical value for a model that includes a constant and that has 119 observations is -2.88. The test statistic for the model for non-residential investment and GDP is -3.94; test statistic for residential investment is -3.52. Thus we may say that the both types of investment are cointegrated with GDP, and therefore reject using the model in first differences, and use instead the error correction model.

Granger Results: Short Run Tests

The first series of tests used lag lengths of six. Once again, the experimental data in Guilkey and Salami suggest that this is a reasonable place to begin. To recapitulate, the tests investigate (1) whether non-residential investment Granger causes GDP, (2) whether GDP Granger causes non-residential investment, (3) whether residential investment Granger causes GDP, and (4) whether GDP Granger causes residential investment.

Results of these tests are presented in Table 1, and are at least in part surprising. First, we cannot reject at the ninety-nine percent level of confidence the null hypothesis that non-residential investment does not cause or predict GDP. Indeed, we cannot even reject this hypothesis at the 90 percent level. This contrasts with the results in Acemoglu (1993), who finds that total investment does Granger cause GDP.

We can reject the null hypothesis, however, that GDP does not cause non-residential investment. In their experiments, Guilkey and Salemi show that in very small samples, strong causation from X2 to X1, as that notation is used in equation (1), can lead one to incorrectly accept the null hypothesis that X1 does not cause X2. Our finding that GDP causes non-residential investment might therefore give us pause about not rejecting the null hypothesis that non-residential investment does not cause GDP. In our case, however, such worries are likely misplaced. First, note that our inability to reject the null that non-residential investment does not cause GDP is not a particularly close call; in fact, we would have to climb down to the 75 percent level of confidence before accepting the null. Needless to say, using such a standard would expose us to a much larger probability of Type 1 error.

Moreover, the "small" sample size Guilkey and Salemi refer to is 75; at a sample size of 200, causation from X2 to X1 seems to have no impact on inferences about causation from X1 to X2. Our sample size is 119, something of a middle ground. Finally, a check of partial correlations on the

Table 1
Tests of Granger Causality
1959-1992
Lag Length = 6

	short run F	long run F	error cor F	num. d.f.	den. d.f.	99% c.v.
Non-residential investment causes GDP	1.37	1.64	1.17	6	109	2.95
GDP causes Non-residential investment	2.54	3.21	.82	6	109	2.95
Residential investment causes GDP	4.71	5.02	3.46	6	109	2.95
GDP causes Residential investment	0.80	2.14	.74	6	109	2.95

F tests are tests of HO: $b_j=0$, $j=1,\dots J$, for the equation $X2(t)=\sum_{j=1}^{J} a_j X2(t-j)+\sum_{j=1}^{J} b_j X1(t-j) + \alpha+\beta T +u(t)$ for short run, long run, and error corrections models, respectively.

X2 is the variable being "caused." X1 is the variable doing the "causing". T is a time trend. Error correction model is in first differences; long run model puts $\beta=0$.

lagged GDP terms in the regression testing whether GDP causes non-residential investment suggests that while GDP's influence is significant, it is not particularly strong. Added together, and after controlling for lagged non-residential investment, the lagged values of GDP explain roughly 11 percent of the variance in non-residential investment. On the other hand, Guilkey and Salami's definition of strong causation has lagged values of X1 explaining 40 percent of the variation in X2.

In contrast to the results for non-residential investment, we can reject the null hypothesis that residential investment does not cause GDP. An F-statistic of 4.71 easily clears the 99 percent critical value of 2.95. At the same time, we cannot reject the null hypothesis that GDP does not cause residential investment. This second finding is important: the fact that residential investment appears to be exogenous with respect to GDP means that the parameters suggesting that residential investment Granger causes GDP are not "mongrel" parameters (Sargent (1979)). This implies that residential investment, rather than some omitted variable acting through residential investment, is a true "cause" of GDP.

These results are striking, but we will wait until we have completed the remainder of our tests before discussing implications.

Having tested for Granger causality at six lags, we now turn to F-tests of differing lag specifications. The results of these tests are reported in Table 2. We find that at the 99 percent level of confidence, we may specify three of our models as having two lags: the models testing whether non-residential investment causes GDP, whether residential investment causes GDP, and whether GDP causes residential investment. The F-tests show that we should use a lag length of three for testing whether GDP causes non-residential investment.

Our test results, presented in Table 3, are substantially the same as before: we cannot reject

Table 2 - Panel 1
Tests of Lag Lengths

	Non-res. inv. causes GDP		GDP causes non-res. inv.		Res. inv. causes GDP		GDP causes non-res. inv.	
Lags tested	F	d.f.	F	d.f.	F	d.f.	F	d.f.
6 vs 5	0.04	2, 109	0.76	2, 109	0.02	2, 109	1.87	2, 109
5 vs 4	1.22	2, 111	0.79	2, 111	1.26	2, 111	0.75	2, 111
4 vs 3	1.57	2, 113	0.22	2, 113	0.94	2, 113	0.53	2, 113
3 vs 2	0.79	2, 115	4.70	2, 115	1.27	2, 115	1.68	2, 115
2 vs 1	4.80	2, 117	5.42	2, 117	14.9	2, 117	21.6	2, 117

Note: in all instances, the 99% critical value is between 4.78 and 4.82.

F tests of whether extra lag contributes to model's explanatory power using short run model described in Table 1.

Table 2 - Panel 2
Tests of Lag Lengths

	Non-res. inv. causes GDP		GDP causes non-res. inv.		Res. inv. causes GDP		GDP causes non-res. inv.	
Lags tested	F	d.f.	F	d.f.	F	d.f.	F	d.f.
6 vs 5	0.15	2, 114	0.84	2, 114	0.04	2, 114	2.05	2, 114
5 vs 4	0.96	2, 117	0.91	2, 117	1.08	2, 117	0.38	2, 117
4 vs 3	1.88	2, 120	0.61	2, 120	0.40	2, 120	0.81	2, 120
3 vs 2	0.58	2, 123	4.83	2, 123	0.35	2, 123	2.75	2, 123
2 vs 1	4.40	2, 125	28.1	2, 125	16.7	2, 125	23.8	2, 125

Note: in all instances, the 99% critical value is between 4.78 and 4.82.

Table 3
Tests of Granger Causality
1959-1992
Lag Length based on test in Table 2

	Short Run F	Long Run F	E.C. F	num. d.f. (lags)	den. d.f.	99% c.v.
Non-res. investment causes GDP	3.03	3.67	1.98	2	117	4.78
GDP causes Non-res. investment	4.99	7.92	4.48	3	115	3.94
Res. investment causes GDP	13.7	15.2	6.74	2	117	4.78
GDP causes Res. investment	0.84	2.78	0.85	2	117	4.78

F tests are tests of HO: $b_j=0$, $j=1,\ldots J$, for the equation $X2(t)=\sum_{j=1}^{J} a_j X2(t-j)+\sum_{j=1}^{J} b_j X1(t-j) + \alpha+\beta T + u(t)$ for short run, long run, and error corrections models, respectively.

X2 is the variable being "caused." X1 is the variable doing the "causing". T is a time trend. Error correction model is in first differences; long run model puts $\beta=0$.

the nulls that non-residential investment does not cause GDP and that GDP does not cause residential investment. We also can reject the nulls that residential investment causes GDP and that GDP causes non-residential investment.

Note, however, that while we cannot reject the hypothesis that non-residential investment does not cause GDP at the 99 percent confidence level, we would reject the hypothesis at the 90 percent level. This would suggest that we should not be too hasty about asserting that non-residential investment does not cause GDP. Now turn to table 4, which reports the regression coefficients and their standard errors from the shorter lag length models. For the model testing whether non-residential investment causes GDP, we find that the coefficients on the lagged non-residential investment variables sum to a negative number. Such a result would suggest that not rejecting the null is perhaps not so unreasonable after all.

Cusum Tests and Tests of Granger Causality from Two Regimes

Now turn to figure 1, which depicts the results of the Cusum of square tests depicted in equations (3)-(5). These results are based upon the two-lag equation testing whether housing investment Granger causes GDP. Note that in the first quarter of 1980, the test statistic S_t crosses the upper 95-percent confidence bound about the expected value of S_t under the null hypothesis that the coefficients remain stable over time. This suggests that a "regime shift" occurs in the early 1980's and casts some suspicion on the results reported above. We therefore run Granger causality tests on data from two sub-periods from our data set: 1959:1 to 1980:1, and 1980:2 to 1992:4.

We again test the models using six lags and using the number of lags suggested by a series of F tests. These results are reported in tables 5 through 7. Despite the fact that the cusum of squares

Table 4
Granger Causality Regression Results
Using "Short" Lag Lengths
1959-1992

Explanatory Variables	Dependent Variable			
	GDP	GDP	Non-res. Investment	Residential Investment
GDP_{-1}	0.93 (9.3)	1.17 (10.0)	0.73 (3.0)	-0.46 (0.8)
GDP_{-2}	0.02 (0.0)	-0.16 (1.5)	-0.37 (1.1)	0.31 (0.6)
GDP_{-3}			-0.16 (0.6)	
Non-res. $inv._{-1}$		0.02 (0.5)	1.09 (10.0)	
Non-res. $inv._{-2}$		-0.05 (1.2)	0.00 (0.0)	
Non-res. $inv._{-3}$			-0.23 (2.3)	
Residential $inv._{-1}$	0.09 (4.5)			1.43 (14.3)
Residential $inv._{-2}$	-0.07 (3.5)			-0.54 (5.4)
Residential $inv._{-3}$				
Adjusted R^2	.95	.94	.95	.91

Note: T-statistics are in parenthesis.

Figure 1
Cusum Test of Coefficient Stability

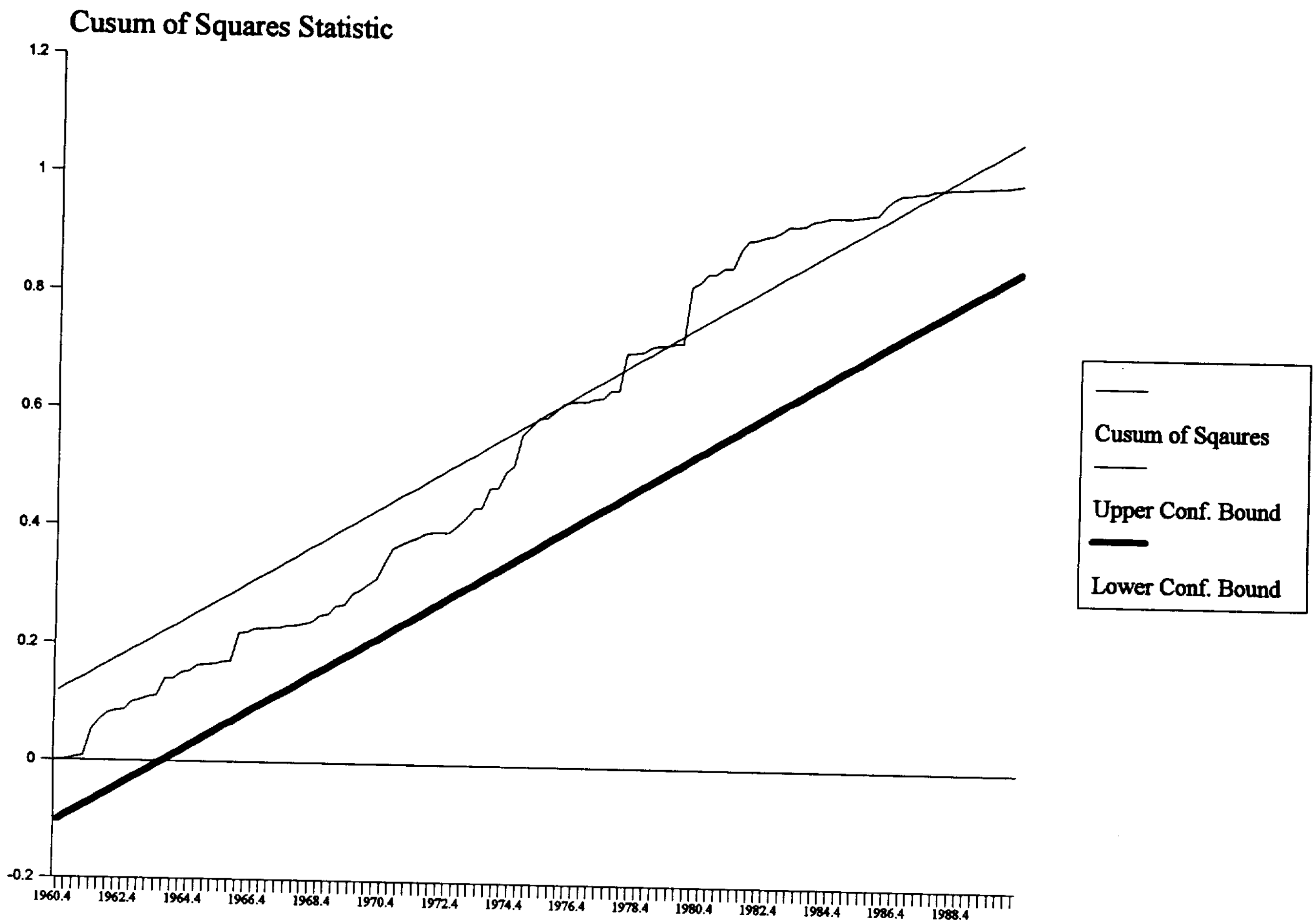

Table 5
Tests of Short Run Granger Causality
Two Subperiods
1959-1980:1 1980:2-1992:4
Lag Length = 6

Panel 1
1959-1980:1

	F	num. d.f.	den. d.f.	99% c.v.
Non-res. investment causes GDP	1.26	6	67	3.10
GDP causes Non-res. investment	2.78	6	67	3.10
Res. investment causes GDP	2.14	6	67	3.10
GDP causes Res. investment	0.23	6	67	3.10

Panel 2
1980:2-1992:4

	F	num. d.f.	den. d.f.	99% c.v.
Non-res. investment causes GDP	0.22	6	23	3.72
GDP causes Non-res. investment	2.70	6	23	3.72
Res. investment causes GDP	3.25	6	23	3.72
GDP causes Res. investment	1.13	6	23	3.72

F tests are tests of HO: $b_j=0$, $j=1,\ldots J$, for the equation $X2(t)=\sum_{j=1}^{J} a_j X2(t-j)+\sum_{j=1}^{J} b_j X1(t-j) + \alpha+\beta T +u(t)$.

Table 6
Tests of Lag Lengths (Short Run Model)
1959:1-1980:1 and 1980:2-1992:4

Panel 1
1959:1-1980:1

	Non-res. inv. causes GDP		GDP causes non-res. inv.		Res. inv. causes GDP		GDP causes non-res. inv.	
Lags tested	F	d.f.	F	d.f.	F	d.f.	F	d.f.
6 vs 5	0.12	2, 67	1.01	2, 67	0.12	2, 67	1.36	2, 67
5 vs 4	0.76	2, 69	1.57	2, 69	0.85	2, 69	0.03	2, 69
4 vs 3	1.48	2, 71	0.28	2, 71	0.95	2, 71	0.55	2, 71
3 vs 2	0.74	2, 73	3.16	2, 73	1.32	2, 73	2.63	2, 73
2 vs 1	2.35	2, 75	2.50	2, 75	7.35	2, 75	12.4	2, 75

Note: in all instances, the 95% critical value is between 3.11 and 3.14.

Table 6 (cont.)
Tests of Lag Lengths

Panel 2
1980:2-1992:4

	Non-res. inv. causes GDP		GDP causes non-res. inv.		Res. inv. causes GDP		GDP causes non-res. inv.	
Lags tested	F	d.f.	F	d.f.	F	d.f.	F	d.f.
6 vs 5	0.01	2, 23	1.04	2, 23	0.93	2, 23	0.01	2, 23
5 vs 4	0.40	2, 25	1.47	2, 25	0.61	2, 25	1.10	2, 25
4 vs 3	1.05	2, 27	2.27	2, 27	2.91	2, 27	1.56	2, 27
3 vs 2	0.90	2, 29	5.12	2, 29	0.12	2, 29	0.87	2, 29
2 vs 1	5.76	2, 31	5.42	2, 31	2.08	2, 31	11.0	2, 31

Table 7
Tests of Granger Causality (Short Run Model)
Two Subperiods
1959-1980:1 1980:2-1992:4
Lag Length based on test in Table 6

Panel 1
1959-1980:1

	F	num. d.f. (lags)	den. d.f.	99% c.v.
Non-res. investment causes GDP	3.58	2	75	4.93
GDP causes Non-res. investment	3.74	3	73	4.09
Res. investment causes GDP	6.35	2	75	4.93
GDP causes Res. investment	0.73	2	75	4.93

Panel 2
1980:2-1992:4

	F	num. d.f. (lags)	den. d.f.	99% c.v.
Non-res. investment causes GDP	.54	2	31	5.39
GDP causes Non-res. investment	6.64	3	29	4.51
Res. investment causes GDP	8.36	2	31	5.39
GDP causes Res. investment	0.84	2	31	5.39

F tests are tests of HO: $b_j=0$, $j=1,...J$, for the equation $X2(t)=\sum_{j=1}^{J} a_j X2(t-j)+\sum_{j=1}^{J} b_j X1(t-j) + \alpha+\beta T +u(t)$.

test suggests a regime shift in 1980, the results are substantially similar to those generated from the entire time series. Apparently coefficients changed, meaning that we would not want to use data from the 1960's to predict GDP in the 1990's. But the causality relationships remained by and large the same.

There is one notable exception. For the period 1959:1-1980:1, we cannot quite reject the null hypothesis that residential investment does not Granger cause GDP at the 99 percent level of confidence, although this is just barely the case. On the other hand, we do reject the null for the 1959:1-1980:1 period in the two lag model, and also reject the nulls in the six lag model and two lag model for the 1980:1 period to 1992:4 period.

Granger Results: Long Run Tests

I will not spend much time discussing the long-run model, because its results are so similar to those presented in detail for the short-run model (see Tables 1 through 3). This gives us our first bit of reassurance that the surprising results above were not simply the product of mis-specification.

Error Correction Model

When two series are cointegrated, we can test for Granger Causality using Granger and Engle's (1987) error-correction model. The form of the model is

$$\Delta X1 = \alpha + \sum_{i=1}^{k} B_{i1} \Delta X1_{t-i} + \sum_{i=1}^{k} B_{i2} \Delta X2_{t-i} + B_z Z_{t-1} + \eta_t \tag{7}$$

where Z is a vector of error terms from the equation (6). We then use standard F tests on the β_2 coefficients. As before, the limiting distribution of the test statistic of the joint hypothesis is not normal, so we test the hypothesis at the 99 percent level. The model in (7) allows us to model long-run equilibrium relationships (e.g., the relationship between investment and GDP) while allowing for periods of disequilibria (the "errors" that get corrected).

Once again, our results are virtually identical (see Tables 1 and 3): the exception is that GDP no longer seems to Granger Cause non-residential investment.

Conclusion

The results portrayed in this paper are striking: across a wide range of specifications, residential investment appears to Granger cause GDP and is exogenous of GDP, while non-residential investment appears not to Granger cause GDP. Thus residential investment seems to lead the nation into and out of recession, while non-residential investment does not. The question remains as to why this should be true.

A natural way of answering this question could be to criticize the very notion of Granger causality.[4] Many have criticized using Granger causality as a method for drawing inferences about the relationships among time-series. These criticisms have tended to fall into three categories. First, some have noted that the technique is essentially ad hoc, and devoid of an underlying economic theory -- we might summarize this view as "correlation does not necessarily imply causation."

Others suggest that unless the relationship between the dependent variable and explanatory variable is truly dynamic, the Granger causality model is by its very nature a misspecification of the relationship between the variables.

Finally, some have criticized the unsatisfactory statistical properties of the test procedures for Granger causality.

While there is something to be said for the philosophical argument that near-constant presence of X2 before X1 does not establish a causal relationship, we should note that Granger's definition of causality is formally similar to a number of proposed precise definitions of causality (see Sims (1977)). In the current context, even if we look at the results generated only as stylized facts, they must be considered reasonably compelling. Yet let us admit the possibility of a different interpretation: perhaps residential investment, like stock prices and interest rates, is a good predictor of GDP because it is a series that reflects forward looking behavior: presumably households will not increase their expenditures on housing unless they expect to prosper in the future. Building a house is a natural mechanism for doing this. Thus the series can do a good job of *predicting* GDP without necessarily *causing* GDP.

With respect to the issue of dynamics, a long tradition of theoretical models (Tobin (1965) and its successors) have made explicit the potentially dynamic nature of the relationship between investment and GDP.

As for the issue of the statistical properties of Granger causality tests, I deal with the issue directly by giving explicit rationales for a wide variety of models and test specifications and by performing stability tests of the model coefficients. Having done all of this, I found remarkably consistent outcomes.

The question remains, then, why models of Granger causality suggest residential investment, (in contrast to non-residential investment) "causes" GDP.

First we should note that the fact that residential investment Granger causes GDP does not necessarily imply that residential investment is more "productive" than non-residential investment. The impact of a particular investment on long-term growth is better inferred from cross sectional data. Indeed, the fact that non-residential investment does not Granger cause GDP requires us to sever the link between Granger causality and long-term economic performance. Otherwise, we would have to accept the implausible argument that non-residential investment has no impact on productivity.

Nevertheless, the strength of residential investment as a predictor of the business cycle is striking. One potential reason for this has already been suggested: perhaps residential investment is just a predictor, rather than a causer. But perhaps the problem can also be resolved when we consider potential "exogenous forces" that lead to the economically exogenous movements in residential investment: the income tax treatment of residential investment and the regulatory treatment of housing finance institutions. For example, when the tax law gives residential investment special treatment through such things as accelerated depreciation and generous treatment of passive losses and capital gains, it stimulates the residential investment market by attracting capital to it. One could cite myriad specific cases from between 1981 and 1986 of residential real estate projects that were wonderful tax shelters, but had little, if any, economic justification. These projects would earn the low returns found in Mills (and thus would place a drag on long-term productivity); they would also provide the people who built them with high paying jobs, and have a reasonably large multiplier effect over a period of several years. But the bill would at some point come due: the overbuilding financed largely through favorable tax policy could not be sustained, residential construction markets could collapse, and the positive effects of residential construction on the economy's overall economic health could unwind. Although the collapse of the multi-family residential construction market in the late 1980's was surely in part a function of the changes in tax policy embodied in the Tax Reform Act of 1986, it was also at least in part a function of severe overbuilding in many regions of the country.

Consequently, residential investment could well lead GDP while being exogenous to GDP. We may not infer from this that high levels of spending on residential investment every year would inevitably lead to higher permanent levels of GDP: it seems apparent that high levels of such investment cannot be sustained.

As for explaining the poor predictive power of non-residential investment, we perhaps should consider the importance of inventories to the business cycle. Inventories rise going into recession, and start being drained as the economy recovers. Only when these inventories are drained do companies need to begin adding to their capital stocks to produce more goods. Thus it is entirely possible that GDP will predict while not being predicted by non-residential investment.

In any event, an important policy implication remains. Clearly, any change in tax laws that would cause a severe and immediate reduction in residential investment could produce severe dislocations in the short run. Thus policies with the goal of reducing "overinvestment" in housing should be pursued with caution.

Acknowledgments

I wish to thank Jim Shilling, Larry Ozanne, and three anonymous referees for useful comments. Any errors that remain are my own.

References Cited

Acemoglu, D. 1993. Learning About Others' Actions and the Investment Accelerator. *Economic Journal,* pp 318-328.

Box, G. and G. Jenkins 1976. *Time Series Analysis.* Holden Day: Oakland.

Brown, Durbin, and Evans 1975. Techniques for Testing the Constancy of Regression Relationships over Time. *Journal of the Royal Statistical Society, Series B*, pp 149-162.

Geweke, J., R. Meese, and W. Dent 1982. Comparing Alternative Tests of Causality in Temporal Systems: Analytical Results and Experimental Evidence," *Review of Economics and Statistics*

Granger, C. 1969. Investigating Causal Relationships by Econometric Models and Cross-spectral Methods," *Econometrica* 37:3 pp 424-438.

Green, R., S.Malpezzi, and K. Vandell 1994. Urban Regulation and the Price of Housing in Korea. *Journal of Housing Economics* 3:4 pp. 330-356.

Greene, W. 1992. *Limdep Version 6.0 User's Manual and Reference Guide*. Econometric Software, Inc: Bellport, NY.

Guilkey, D. and M. Salemi 1982. Small Sample Properties of Three Tests for Granger-Causal Ordering in a Bivariate Stochastic System. *Review of Economics and Statistics* 64:4, pp. 668-708.

Gupta, S. 1988. Profits, Investment, and Causality: An Examination of Alternative Paradigms. *Southern Economic Journal*, 55:1, pp 9-20.

Hamilton, J.D. 1994. *Time Series Analysis*. Princeton:Princeton.

Mills, E. 1987. Has the United States Overinvested in Housing? *AREUEA Journal*, 15:1, pp. 601-16.

Mills, E. 1989. Social Returns to Housing and Other Fixed Capital. *AREUEA Journal*, 17:2, pp. 197-217.

Ohanian, L. 1988. The Spurious Effects of Unit Roots on Autoregressions: A Monte Carlo Study. *Journal of Econometrics* 39:pp. 251-66

Sargent, T. 1979. *Macroeconomics*. San Diego:Academic Press.

Sims, C. 1972. Money, Income and Causality. *American Economic Review* 62:4, pp. 540-552.

Sims, C. 1977. Exogeneity and Causal Ordering in Macroeconomic Models, in C. Sims, ed. *New Methods in Business Cycle Research:Proceedings from a Conference*, Minneapolis: Federal Reserve Bank of Minneapolis. pp.45-109.

Tobin, J. 1965. Money and Economic Growth. *Econometrica*, 33:4, pp. 671-684.

Toda, H. and P. Phillips 1993.Vector Autoregressions and Causality, *Econometrica* , 61:6, pp. 1367-1394.

United States Department of Commerce (1989) Revisions to National Income and Products Accounts. *Survey of Current Business* 73:8, pp. 9-51.

Endnotes

1. See Green, Malpezzi and Vandell (1994) for citations.

2. The reason for testing at the 99 percent level of confidence will be given in the discussion below on the nature of time series data.

3. Although the authors did not characterize them as such, results in Toda and Phillips (1993) also suggest that while the test statistic will not be characterized as an F-distribution, it will be very close to an F-distribution. Thus testing at the 99 percent level should be sufficiently prudent in our attempt to avoid Type II error.

4. A nice summary of such critiques may be found in Gupta (1988).