

Effect of Foreclosure Status on Residential Selling Price: Comment

Thomas M. Carroll*
Terrence M. Clauretie**
Helen R. Neill***

Abstract. In this comment we examine the conclusion by Forgey, Rutherford and VanBuskirk (1994) “that the foreclosed properties sold at a 23% discount,” using a sample of nearly 2,000 residential property sales from the Las Vegas, Nevada area. We found that when not controlling for location with a set of dummy variables for zip codes, HUD foreclosed properties sold for between 12.18% and 13.96% below a random sample of properties not within one block of foreclosed properties. When controlling for location, using a set of thirty-one dummy variables for zip codes, the foreclosure discount fell to between 8.45% and 9.72%. When controlling for the common characteristics between foreclosed properties and their neighbors, we found foreclosure discounts are very small (between .17% and 2.58%) and no longer statistically significant. We conclude that foreclosure does not provide an opportunity for arbitrage profits, and this study does reinforce the findings of other studies that conclude real estate markets operate efficiently.

Introduction

In a recent article in this *Journal*, Forgey, Rutherford and VanBuskirk (1994) (hereafter FRV) analyzed the sales price of foreclosed properties in Arlington, Texas. They found “that foreclosed properties sold at a 23% discount.” This finding flies in the face of both the efficient market hypothesis and previous evidence that residential markets are relatively efficient (see Gau, 1984, 1985; Johnson and Kaserman, 1983; Shilling, 1990).

To see why this is so, consider that the seller in a foreclosed property is usually a financial institution or government agency, such as the Department of Housing and Urban Development (HUD). There is no reason to suspect, a priori, that the selling price of a property will be affected by who the seller may be. A 23% discount should afford ample opportunities to make excess profits, even after transaction costs, by trading in such properties. In the extreme, imagine buying a property at a 23% discount because it is foreclosed and then selling it without any repairs whatsoever the next day under nonforeclosed status. Are we to believe there will be a 23% appreciation in this short period?

We suggest that foreclosed properties sell at a discount because of the condition of the property itself, the condition of the neighborhood within which the property lies, or both.¹ If this is the case, then the findings of FRV result from a failure to adequately

*Department of Economics, University of Nevada-Las Vegas, Las Vegas, Nevada 89154.

**Department of Finance, University of Nevada-Las Vegas.

***Department of Economics, University of Nevada-Las Vegas.

Date Revised—December 1994; Accepted—May 1995.

control for the condition of the property or the condition of the neighborhood. We submit that the variables used to control for neighborhood quality in the FRV study are inadequate for this purpose for several reasons and suggest other methods. Below we offer reasons for the inadequacy of FRV's control variables.

In this paper we test for market efficiency by first replicating FVR's model with data from Las Vegas, Nevada, and then, using this same data, control for the effects of the neighborhood conditions with different control variables. We replicate FRV's findings when we do not control for the effects of neighborhood quality. However, contrary to their results, we find that the discount for foreclosed properties is much less when neighborhood effects are adequately controlled. The actual discount because of foreclosure status is well within the range of transactions costs, suggesting that the market for residential properties is efficient.

In this next section we discuss an alternative control technique to that of FRV. In the following section we discuss the data and empirical results and follow with a concluding section.

Controlling for Neighborhood Quality

Any model designed to isolate the effect of an independent variable on a dependent variable must adequately control for the effects of other independent variables.² We believe FRV fail to do so in their model of foreclosure effects on selling prices. They use three control variables: *ZIP*, *MAPSCO* and *AREA*. *ZIP* is the numerical value of one of approximately fourteen zip codes within Arlington, Texas. *MAPSCO* represents numerous small area numbers designed for mapping purposes, and *AREA* represents about thirteen area numbers assigned by the local multiple listing service (MLS). None of these location variables were designed to capture neighborhood effects, although the smaller segments may begin to capture neighborhood effects by chance. Further, FRV enter these variables as cardinal variables and not as nominal variables. Since zip codes are merely numbers assigned arbitrarily to postal service districts, treating zip as a cardinal number is a serious flaw in their analysis. Two of the three variables contain insignificant signs and for all three their coefficients cannot be interpreted. In our model we treat the thirty-two zip codes in the Las Vegas area as qualitative distinctions by using a set of thirty-one dummy variables. We show that the magnitude of the "foreclosure effect" is sensitive to the correct treatment of zip codes as nominal variables.

To control for the quality of neighborhood one should select a control group of non-foreclosed properties from as close a proximity to the foreclosed properties as possible. Such a selection avoids the use of arbitrary boundaries represented by areas such as zip codes or real estate map coordinates. In our model, we select two control groups: non-foreclosed properties within one block of each foreclosed property, and a random sample of properties not located near foreclosure properties. We test the robustness of the FRV model by first treating foreclosures and neighborhood sales as distinct groups, that is, with nonoverlapping indicator variables. We then recode the neighborhood indicators to include foreclosed properties, since these properties are obviously in their own neighborhoods. This recoding allows us to distinguish between location and foreclosure effects.

Data

Our sample consists of 1,974 residential properties sold in the Las Vegas Valley between 1990 and 1993. Of these properties, 385 were sold by HUD and 19 were sold by private financial institutions after being foreclosed. We consider both HUD- and bank-foreclosed properties separately to further investigate any difference in selling price between these types. To control for neighborhood effects we included properties within one block of each foreclosed property; our sample contains 931 (47% of the sample) in close proximity to HUD-foreclosed properties and 38 (2%) situated nearby bank-foreclosed properties. Such close proximity of the non-foreclosed properties to the foreclosed houses will control for neighborhood effect much better than arbitrary boundaries such as zip codes. The remainder of the sample consists of 602 properties not within a block of foreclosed properties. All data were selected from *Microscan*, a database containing 319,451 properties including 196,000 single-family residence sales in Clark County, Nevada. A summary of the data is presented in Exhibit 1.

The average price of houses in our sample was \$88,268 (\$65,640 in 1982–1984 dollars) and the average age was 15.74 years. The average house also had 1.87 baths and 3.08 bedrooms. The typical house had one story and the overwhelming majority had an attached garage. The average prevailing mortgage interest rate was 8.89%³ and a typical house had 1,485 square feet. Thirteen percent of the residences had swimming pools and 11.55% were built on slab foundations.

Exhibit 1
Descriptive Statistics

	Mean	Median	Std Dev.	Maximum	Minimum	Observs.
<i>PRICE</i>	\$88,268	\$81,000	\$45,941	\$750,000	\$10,074	1974
<i>RPRICE</i>	\$65,640	\$60,030	\$34,158	\$552,690	\$7,175	1974
<i>AGE</i>	15.74	13	12.1933	58	0	1974
<i>BATH</i>	1.87	1.75	0.4953	6	0.75	1974
<i>BEDROOM</i>	3.08	3	0.7481	8	1	1974
<i>FIREPLACE</i>	0.57	1.00	0.60	4	0	1974
<i>FLOORS</i>	1.24	1	0.4281	2	1	1974
<i>GARAGE</i>	0.98	1	0.6862	3	0	1974
<i>MRATE</i>	8.89	8.94	1.0724	10.48	6.8	1974
<i>SQFT</i>	1485.50	1350	532.875	8235	603	1974
<i>POOL</i>	13.12%	0	33.77%	1	0	1974
<i>SLAB</i>	11.55%	0	31.97%	1	0	1974
<i>BANK</i>	.96%	0	9.77%	1	0	1974
<i>BANKNAB</i>	1.93%	0	13.74%	1	0	1974
<i>HUD</i>	19.45%	0	39.59%	1	0	1974
<i>HUDNAB</i>	47.16%	0	49.93%	1	0	1974
<i>BANKNAB2</i>	2.89%	0	16.75%	1	0	1974
<i>HUDNAB2</i>	66.67%	1	47.15%	1	0	1974

The Model and Empirical Results

In an attempt to replicate the findings of FRV we fit the following linear and log-linear models:

$$RPRICE = f(AGE, BATH, BR, FP, FL, GAR, MR, SQFT, POOL, SLAB, ZIP, BANK, BANKNAB, HUD, HUDNAB) \quad (1)$$

$$\ln(RPRICE) = g(AGE, BATH, BR, FP, FL, GAR, MR, \ln(SQFT), POOL, SLAB, ZIP, BANK, BANKNAB, HUD, HUDNAB), \quad (2)$$

where:

- SP* = sales price of residential property, deflated by consumer price index for housing (1982–84=100);⁴
- AGE* = age of the property in years;
- BATH* = number of bathrooms;
- BR* = number of bedrooms;
- FP* = number of fireplaces;
- FL* = number of floors (stories) in the house;
- GAR* = 1 if house has garage, 0 otherwise;
- MR* = prevailing mortgage interest rate the month the property was sold;
- SQFT* = total square foot in building (measured as natural logarithm in log-linear regression⁵);
- POOL* = 1 if house has pool, 0 otherwise;
- SLAB* = 1 if foundation is slab, 0 otherwise;
- ZIP* = a set of thirty-one dummy variables for different zip codes in the Las Vegas Valley;
- BANK* = 1 if foreclosure sale by commercial bank;
- BANKNAB* = 1 if property is within one block⁶ of bank-foreclosed property;
- HUD* = 1 if foreclosure sale by U.S. Department of Housing and Urban Development;
- HUDNAB* = neighbor of HUD foreclosure.

Equations (1) and (2) are modified to account for the fact that *BANK*-foreclosed properties (*BANK*=1) and *HUD*-foreclosed properties (*HUD*=1) are located in their own neighborhood. As equations (1) and (2) are written, one cannot tell whether the coefficient on *BANK* reflects bank foreclosures or the effect of hidden neighborhood characteristics. Similarly, one cannot determine whether the coefficient on *HUD* reflects the effect of *HUD* foreclosures or the effects of hidden neighborhood characteristics. Clearly neighborhood indicator variables should be recoded to include both foreclosed properties and non-foreclosed properties in the same neighborhood.

We mitigate this problem by replacing the neighborhood dummy variables in equations (1) and (2) with new variables coded as follows:

$BANKNAB2 = BANKNAB + BANK$.⁷ That is, *BANKNAB* is coded as 1 for both bank foreclosures and neighboring non-foreclosed properties. This allows us to interpret the coefficient on *BANK* as measuring the effect of foreclosure independent of neighborhood effects.

$HUDNAB2 = HUDNAB + HUD$. As explained above, this means that the coefficient on HUD measures the foreclosure effect independent of neighborhood characteristics. Our revised equations become:

$$RPRICE = f(AGE, BATH, BR, FP, FL, GAR, MR, SQFT, POOL, SLAB, ZIP, BANK, BANKNAB2, HUD, HUDNAB2); \quad (1a)$$

$$\ln(RPRICE) = f(AGE, BATH, BR, FP, FL, GAR, MR, \ln(SQFT), POOL, SLAB, ZIP, BANK, BANKNAB2, HUD, HUDNAB2). \quad (2a)$$

Exhibit 2 presents the linear and log-linear regressions for the sample of 1,974 houses, wherein the indicator variables for foreclosed properties (HUD and $BANK$) are treated as mutually exclusive relative to the indicator variables for close neighbors of foreclosed properties ($HUDNAB$ and $BANKNAB$).

The results are also presented with and without the inclusion of a set of thirty-one dummy variables for zip codes in the Las Vegas Valley. The individual coefficients and t -statistics for the coefficients for the zip code indicators are not reported for brevity and the collective significance of those indicators is evaluated with the use of an F -statistic.⁸

In the linear equation, all coefficients are statistically significant at the .01 level except for the indicator variables for slab construction, bank-foreclosed properties, and neighbors of bank-foreclosed properties. Adding the zip code indicators yields a statistically significant improvement in the adjusted R^2 and generally reduces the t -statistics for the

Exhibit 2 Regressions: Foreclosures and Neighborhood Properties Separate

Variable	Linear Regressions				Log-Linear Regressions			
	Dependent Variable: $RPRICE$				Dependent Variable: $\ln(RPRICE)$			
	Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.
Intercept	6873.81	1.40	-3344.87	-.17	6.6517	24.12	7.1711	17.58
AGE	-547.96	-11.75	-393.45	-7.13	-.0097	-13.05	-.0068	-7.89
BATH	4462.17	2.98	4167.86	2.78	.0812	3.41	.0605	2.59
BR	-6250.33	-8.41	-5007.35	-6.65	-.0572	-4.63	-.0258	-2.08
FP	6183.21	6.39	5565.08	5.76	.1272	8.12	.1065	6.95
FL	-5325.46	-4.26	-5309.21	-4.30	-.0537	-2.69	-.0568	-2.92
GAR	3634.53	5.26	3322.45	4.92	.0346	3.10	.0248	2.32
MR	2582.03	6.11	2672.10	6.53	.0417	6.19	.0422	6.58
SQFT*	39.75	28.43	36.89	26.17	.5745	13.83	.4788	11.59
POOL	9608.02	6.66	9415.94	6.67	.1526	6.66	.1338	6.07
SLAB	-675.72	-.47	-739.73	-.53	-.0212	-.92	-.0281	-1.28
BANK	-2256.43	-.48	-1750.56	-.38	-.0306	.41	.0596	.83
BANKNAB	-5734.41	-1.70	-7421.88	-2.19	-.0873	-1.62	-.0867	-1.63
HUD	-9026.51	-6.47	-6280.68	-4.56	-.1299	-5.85	-.0883	-4.09
HUDNAB	-7325.57	-6.40	-4590.51	-4.01	-.1228	-6.74	-.0867	-4.84
ZIP			$F = 6.44$				$F = 8.81$	
R^2		.6588		.6908		.5325		.5905
Adjusted R^2		.6564		.6836		.5292		.5810
F-statistic		270.20		95.74		159.40		61.79
Observations		1974		1974		1974		1974

*SQFT entered as natural logarithm in log-linear equation

housing characteristics. The exception in the linear equation is that the addition of zip code indicators renders the indicator for bank neighbors (but not bank-foreclosed properties) statistically significant at the .05 level. Even with zip code indicators, the indicators for *HUD* foreclosures and *HUD* neighbors are both statistically significant. Furthermore, the results are similar to those found by FRV. It must be mentioned, however, that the results so far do not indicate whether *HUD* properties sell at significantly less than other properties because of foreclosure per se or because of their location.

In the log-linear equation we find much the same results. All of the variables are statistically significant except for the indicators for slab construction, bank foreclosures, and neighbors of bank foreclosures. *HUD* foreclosures appear to sell for 12.18% less, and *HUD* neighbors sell for 11.56% less than houses not located near foreclosed properties. Introducing the set of zip code dummies reduces the discount on *HUD* properties to 8.45% and the markdown on *HUD* neighbors to 8.3%. Again, it is unclear from Exhibit 2 whether *HUD* properties sell for less because of their foreclosure status or because of their location. The significance of the indicator variables for *HUD* foreclosures might be due to foreclosure per se or they may proxy the location of those properties.

The results of Exhibit 3 clear up some of the confusion in Exhibit 2. In Exhibit 3 the indicator variable *HUDNAB*, which is coded 1 for *HUD* neighbors and coded 0 for *HUD* properties, is replaced by *HUDNAB2*, which equals the sum of *HUD* and *HUDNAB*. That is, *HUDNAB2* is coded as 1 for properties that are either *HUD* foreclosed or in close proximity to *HUD*-foreclosed properties. Similarly, *BANKNAB2* is coded as 1 for both bank foreclosures and their neighbors. This substitution allows the interpretation of the coefficient on *HUD* (and *BANK*) as reflecting the effect of foreclosure alone and not the effect of the location of the foreclosed property. In all cases the coefficient of the *HUDNAB2* is statistically significant, while the coefficient on *HUD* is statistically insignificant. Once we have controlled for the location of *HUD* properties, foreclosure, by itself, has no significant impact on property values. In the linear model, *HUD* foreclosures sell at an average discount of only \$1,700, compared to the significant discount of \$7,300 for *HUD* neighborhood properties (including *HUD* foreclosures). Introduction of the dummy variables for zip codes⁹ reduces the neighborhood effect of \$4,600. In the log-linear model without zip code dummies, houses in *HUD* neighborhoods sell for 11.58% less than houses with the same characteristics, while *HUD* foreclosures sell for an additional 8.38% discount. Introducing the zip code dummies reduces the discount on properties in *HUD* neighborhoods to 8.31%, with *HUD* foreclosures selling for a mere .17% less than other neighborhood properties.

This finding reinforces previous studies that found the housing market to be reasonably efficient. That *HUD* foreclosures and their neighbors both sell for significantly less than non-foreclosed properties has to do with their hidden characteristics—most likely the characteristics of their shared neighborhoods. This implies that buying a *HUD* foreclosed property would require expensive repairs and/or restoration before the house could be resold at a higher price. Purchase and resale of *HUD* foreclosures would return about one-sixth of 1% of their potential value. This is hardly a lucrative arbitrage prospect.

Conclusion

Forgey et al. (1994) purport to show that “foreclosure sales” result in a substantial (23%) discount on the residential selling price. To do so they investigated a large sample

Exhibit 3
Regressions: Foreclosed and Neighborhood Properties Combined

Variable	Linear Regressions				Log-Linear Regressions			
	Dependent Variable: <i>RPRICE</i>				Dependent Variable: $\ln(RPRICE)$			
	Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.
Intercept	6978.53	1.42	-3297.49	-.17	6.6540	24.13	7.1717	17.58
<i>AGE</i>	-548.52	-11.77	-393.84	-7.14	-.0097	-13.06	-.0068	-7.90
<i>BATH</i>	4439.44	2.96	4153.83	2.77	.0808	3.39	.0602	2.58
<i>BR</i>	-6253.35	-8.42	-5008.92	-6.65	-.0573	-4.63	-.0258	-2.08
<i>FP</i>	6176.71	6.38	5562.18	5.76	.1271	8.11	.1064	6.95
<i>FL</i>	-5323.03	-4.26	-5309.94	-4.30	-.0536	-2.68	-.0568	-2.92
<i>GAR</i>	3632.94	5.25	3321.83	4.92	.0345	3.10	.0248	2.32
<i>MR</i>	2578.68	6.11	2670.04	6.53	.0416	6.18	.0421	6.57
<i>SQFT*</i>	39.76	28.44	36.90	26.18	.5744	13.83	.4788	11.59
<i>POOL</i>	9610.68	6.66	9417.36	6.67	.1527	6.66	.1338	6.07
<i>SLAB</i>	-680.79	-.47	-742.47	-.53	-.0212	-.93	-.0282	-1.28
<i>BANK</i>	3486.03	.62	5675.16	1.05	.1181	1.31	.1464	1.72
<i>BANKNAB2</i>	-5744.24	-1.70	-7426.37	-2.19	-.0071	-.36	-.0868	-1.63
<i>HUD</i>	-1699.35	-1.39	-1693.02	-1.44	-.0875	-1.62	-.0017	-.09
<i>HUDNAB2</i>	-7343.82	-6.41	-4591.49	-4.01	-.1231	-6.75	-.0868	-4.84
<i>ZIP</i>				<i>F</i> = 6.44				<i>F</i> = 8.81
<i>R</i> ²		.6588		.6908		.5326		.5905
Adjusted <i>R</i> ²		.6564		.6836		.5292		.5810
<i>F</i> -statistic		270.23		95.74		159.42		61.79
Observations		1974		1974		1974		1974

* *SQFT* entered as natural logarithm in log-linear regression

of properties sold through the multiple listing service in Arlington, Texas. It is unclear from their article whether these “foreclosure sales” represent hurried sales by irrational sellers with positive equity attempting to beat an impending foreclosure date, or sales by *HUD*, banks, or other owners of property whose owners actually defaulted. If foreclosure, per se, results in such a large discount, we wonder why such properties were not purchased long before they made the multiple listing service. If real estate professionals handling these foreclosure sales do not recognize such profitable arbitrage prospects, who would?

We believe that FRV’s findings are suspect because they inappropriately control for the location characteristics of foreclosed (or soon-to-be-foreclosed) properties. Treating zip codes and numerical map designations as cardinal, continuous numbers fail to adequately control for the location characteristics of foreclosed properties. We have shown that treating zip codes as a set of nominal (dummy) variables results in a substantial reduction in the apparent discount for foreclosed properties. A dummy variable designating houses in close proximity of foreclosed properties is consistently significant in both the linear and log-linear models, with and without zip codes included in the regression. When the variable is recoded to include foreclosed properties themselves (which are obviously in their own neighborhood), we find that the dummy variable for *HUD* foreclosures becomes insignificant, as the dummy variable for *BANK* foreclosures always is. We must conclude that foreclosed properties sell for lower prices because of hidden defects (leading to

negative equity when discovered) or due to neighborhood characteristics. Arbitrage possibilities appear to be unattainable in the efficient market sense.

Notes

¹There are two reasons why foreclosure would occur: (1) the owner has negative equity—the market price is less than the loan balance, or (2) the owner has positive equity but places the property on the market too late to achieve the optimal price. The second case presents arbitrage profits, the first does not.

²FRV conclude their paper with an irrelevant discussion of multicollinearity, when their main problem is misspecification due to omitted variables and treating nominal variables as cardinal ones. One cannot control for location by calculating a slope coefficient for zip code, as if places with higher zip code values had “more location”.

³Our data does not report the actual interest rate paid by the buyer, so we used the monthly average rate for new mortgages, obtained from *Citibase Macroeconomic Data Base*.

⁴Deflating the sales price removes the effect of housing inflation (measured by the housing CPI) from the price appreciation effect of time.

⁵Using the logarithm for square feet allows the coefficient to be interpreted as the elasticity of price with respect to building size.

⁶Neighbors of foreclosed properties were identified by sorting properties by address. All property sales within one block of foreclosed properties were labeled as neighbors. Properties further than a block from any foreclosed property were categorized as “market” properties.

⁷If *BANK* is coded as 1, or if *BANKNAB* is coded as 1, then *BANKNAB2* is coded as 1. *BANKNAB2* is coded as zero only if both *BANKNAB* and *BANK* are coded as 0.

⁸The formula for the *F*-statistic is:

$$F = \frac{(SSE_r - SSE_u) / (df_r - df_u)}{SSE_u / df_u},$$

where *SSE_r* is the restricted sum of the squared errors, *SSR_u* is the unrestricted sum of squared errors, and *df* is the degrees of freedom in each equation, respectively.

⁹*HUD* foreclosures are heavily concentrated among the thirty-one zip codes. Two-thirds of the *HUD* foreclosures are located in six zip codes, and three-fourths are located in ten zip codes. There are six zip codes that contain no *HUD* foreclosures.

References

- Forgey, F. A., R. C. Rutherford and M. L. VanBuskirk, Effect of Foreclosure Status on Residential Selling Price, *Journal of Real Estate Research*, 1994, 9:3, 313–18.
- Gau, G., Public Information and Abnormal Returns in Real Estate Investment, *AREUEA Journal*, 1985, 13:1, 15–31.
- , Weak Form Test of Efficiency in Real Estate Investment Markets, *Financial Review*, 1984, 19:4, 301–20.
- Johnson, R. and D. Kaserman, Housing Market Capitalization of Energy-Saving Durable Goods Investment, *Economic Inquiry*, 1983, 21:3, 374–86.
- Shilling, J. D., J. D. Benjamin and C. F. Sirmans, Estimating Net Realizable Value for Distressed Real Estate, *Journal of Real Estate Research*, 1990, 5:1, 129–40.

The authors wish to thank Cindy Durtschi for assistance in collecting the data set. The authors also wish to acknowledge financial support for this research from the First Interstate Bank Institute for Business Leadership in Las Vegas, Nevada. Conclusions reached are those of the authors and not those of First Interstate Bank.