Are Home Values Affected by Sinkhole Proximity? Results of a Hedonic Price Model

E. Spencer Fleury
Department of Geography
University of South Florida

Introduction

With its limestone bedrock, warm climate and high precipitation levels, Florida provides near-ideal conditions for sinkhole development. Additional contributing factors in certain areas include high rates of urbanization and overpumping of groundwater to meet agricultural demand. While sinkholes can be found throughout Florida, distribution is not even across the state, with the highest concentrations occurring in the west-central region, to the north and east of Tampa Bay (USGS, 1999).

Though they lack the high profile and sheer destructive force of hurricanes, floods, and other natural hazards, sinkholes have on occasion generated significant damage to buildings, roads, and other human-built structures, and should be considered natural hazards in their own right. In sinkhole-prone areas where market insurance against sinkhole damage is available, economic theory suggests that homes located there should be valued somewhat lower than homes located in areas where sinkholes are rare or nonexistent, in recognition of both the risk faced by the homeowner in a sinkhole-prone area, and the cost of insuring one's property against that risk. Working with sinkhole and Census data from the Tampa Bay, Florida region in 1990, this paper uses a hedonic price model to look for a statistically significant relationship between the presence of sinkholes (and, in a separate set of regressions, the density of sinkholes) in a neighborhood and the value of homes in that neighborhood. The model did not find evidence of either type relationship.

Background

The decision to use 1990 data for this analysis has its roots in a policy decision made in 1991. In 1969, the state of Florida put into place a reinsurance facility that could cover the risk of property loss to sinkholes. At that point, Florida homeowners in sinkhole-prone areas had two options: they could either purchase sinkhole insurance, or they could gamble that their property would not be damaged by sinkholes. Either way, the risk of living in a sinkhole-prone environment was borne entirely by the homeowner. However, very few people purchased this optional available coverage (Maroney, et al, 2005), and in 1991 the Florida Statutes were amended to automatically include sinkhole coverage in every homeowner's insurance policy, at no specified additional cost (Eastman, et al, 1995). Even though a 1993 study found that the problem of sinkhole losses was largely confined to the Tampa Bay area, the amended statutes made no distinction between homeowner policies issued for sinkhole-prone areas, and those issued for parts of the state where sinkholes were all but unknown (Maroney, et al, 2005).

Though it was probably not the intent of the Florida Legislature to do so, by passing this piece of legislation lawmakers actually enacted a mechanism to encourage people to engage in a higher-risk behavior—purchasing a home in a sinkhole-prone area—while the full costs of those behaviors are distributed among people who choose not to engage in that same higher-risk behavior. Because homebuyers will be forced to pay for sinkhole insurance (a cost which is undoubtedly built into the price of each policy), they have no incentive to minimize their own risk by moving to an area where sinkholes are less likely to damage their property. The distribution of sinkhole risk was therefore altered and is now shared by those homeowners who face little to no risk of sinkhole damage, but are still obligated to pay into the insurance pool. Thus, the cost of living in sinkhole-prone areas is artificially lowered, which means that homebuyers are theoretically more likely to relocate there than they otherwise may have been.

Methodology and data

In real estate economics, hedonic regression models are often used when a researcher wishes to control for the value of amenities such as square footage, number of bedrooms, and location, among others. For this reason, hedonic models are frequently used to examine questions related to the impact of various hazards on home prices (see Nourse, 1967; Palm, 1981; Brookshire, et al, 1985; Tobin and Montz, 1994; Kiel and McClain, 1995; Dale, et al, 1999; among others). Results obtained using hedonic regression models often contradict those of other studies using different models, which suggests that the specific nature of the hazard may be a crucial factor. A search of the literature did not turn up hedonic studies of any potential relationship between sinkholes and home prices.

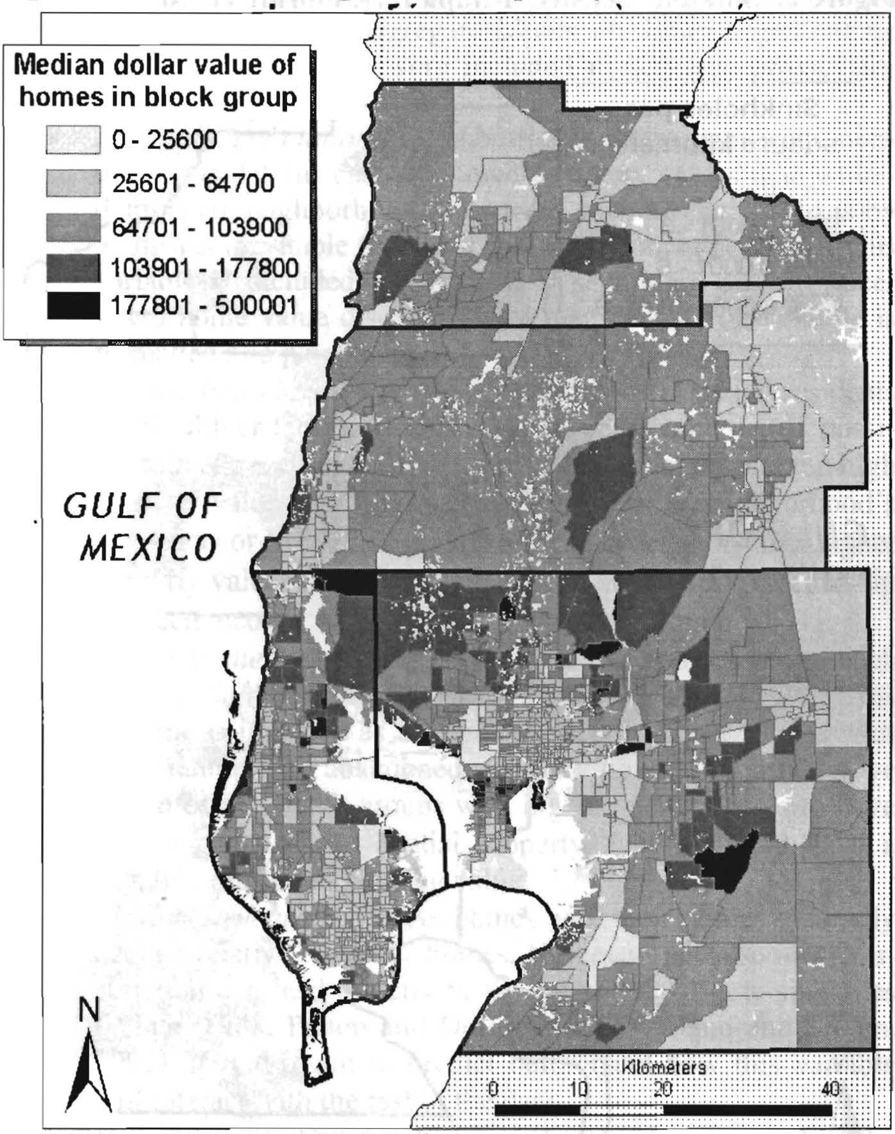
Data sources used for this analysis were the Florida Geological Survey's sinkhole database and the 1990 US Census. Block-group-level Census data (described below) for the four counties of the Tampa Bay area (Hernando, Hillsborough, Pasco and Pinellas) were used. All sinkholes reported in the four-county area between 1964 and 1990 were included in this analysis; sinkholes reported after 1990 or observations lacking a reporting year were dropped. Sinkhole locations were entered into ArcGIS, and were linked to the block group in which they were located. Ordinary least squares (OLS) and Probit regressions were then run in order to characterize any potential relationship between sinkhole location and home values. These regressions were run once using regionwide data; the OLS regressions were also run once for each county.

The OLS and Probit regressions actually examine different, yet still related, questions. The OLS regressions investigate the relationship of home values to sinkhole *density* within each block group; the Probit regression instead focuses on the mere *presence* of sinkholes in a block group, with no adjustment for either the number of sinkholes or the geographic size of the block group. The median home value variable is included in both the OLS and Probit regressions, though as the dependent variable in the former and as an explanatory variable in the latter. Figure 1 illustrates the distribution of median home values across the region in 1990.

In addition to median home value, the following variables are included in the model:

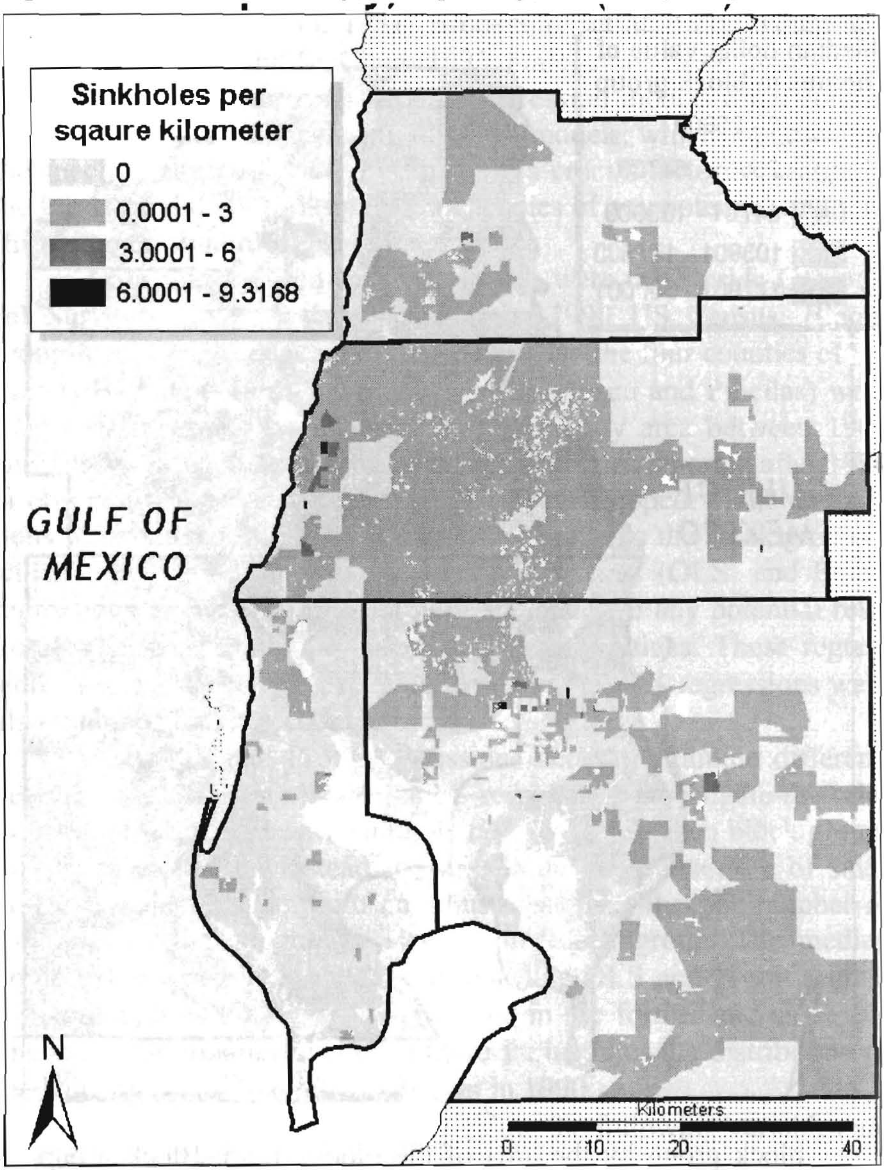
- *Sinkhole density*: The value for this variable is derived from dividing the number of sinkholes reported in each

Figure 1. Median housing value, Tampa Bay, Florida (1990).



block group by the area of the block group. Block group sizes vary widely across the region; this variable was created as a method of mitigating this problem. It is also the

Figure 2. Sinkhole density, Tampa Bay, Florida (1990).



key explanatory variable in the OLS model. Sinkhole density across the Tampa Bay region is shown in figure 2.

- *Number of bedrooms*: Because median square footage data were not available, aggregate counts of homes organized by number of bedrooms were included to serve as proxies for home size.
- *Nonwhite population*: Neighborhoods with significant non-white populations can have lower property values than white-dominated neighborhoods; however, such a relationship is by no means inevitable (Palmore, 1966; Boston, et al, 1972). This variable is included as a means of separating out any race-based home value disparities that may occur. The variable is formatted as a percentage.
- *Median household income*: Neighborhoods in which residents are wealthier tend to have higher property values. It is possible to make a causal argument in either direction (are the high property values the result of the wealth of the neighborhood's residents, or are wealthier residents attracted by the higher property values); however, the exact nature of the relationship between income and home prices is not relevant here.
- *Vacancy rates*: This variable is formatted as a percent of each block group's housing stock that was vacant in 1990. This variable is included as a means to identify block groups with large numbers of abandoned or empty houses. Intuitively, we would expect block groups with higher vacancy rates to have less demand for residential property, which should have a negative impact on housing value.
- *Median structural age*: As homes age, their values generally decline relative to newer homes. However, because new construction can lead directly to the formation of new sinkholes (White, 1988; Patton and DeHan, 1998; Soriano and Simon, 2002), it is difficult to predict beforehand how this variable will interact with the rest of the data.
- *Sinks_present*: This is a binary indicator variable used as the dependent variable in the Probit regression. It is not included in the OLS regression. Its value is 1 for block groups where a sinkhole had been reported prior to 1991, and is 0 for block

groups with no reported sinkholes.

- *Total housing units:* This is included as an explanatory variable in the Probit regression. It is included as a means of controlling for the greater likelihood of sinkhole reporting in block groups with higher populations, as well as the possibility of new sinkhole generation brought on by higher levels of new construction in growing areas.

This analysis makes use of a level of aggregation that some readers might find troubling. Specifically, both sinkhole occurrence and median home values are measured at the Census block group level; some might ask why the actual sale prices of individual homes were not plotted and mapped in relation to the nearest reported sinkhole. While this almost certainly would have been the preferred method of proceeding, data limitations forced this approach. For one thing, home sale data in Hillsborough County is no longer available for years prior to the late 1990s; for another, the FGS sinkhole database has always depended on voluntary reporting of sinkholes, and therefore suffers from a certain lack of comprehensiveness (the 1990s in particular were lean years for the database, as funding for database maintenance dried up for much of that decade). These two factors led to the development of the methodology used here, one that is not as precise as the ideal method but can still tell us something about the relationship between sinkhole density and home values.

Results

Regionwide, both the OLS (table 1) and the Probit regressions (table 2) generated statistically significant results; however, in no case were any of the variables of interest significant. While median home value in a given block group shows a statistical relationship to every other explanatory variable in the model, there does not seem to be a connection between median home values and sinkhole densities. This suggests that the discounting predicted by economic theory did not occur here. (Some of the other explanatory variables—in particular, the vacancy rate—did not generate the kinds of results we might have expected before running the regressions. And while these results merit examination, they go beyond the scope of this paper, and thus

Table 1. Regionwide OLS estimates.
Dependent variable: median home value.

Variable	Coefficient	t-statistic
const	34459.4	5.9670
Sinkhole density	-1004.23	-0.9534
Median household income	2.69503	28.8575*
Vacancy rate	521.565	5.0075*
Nonwhite population (pct. of total)	-114.312	-2.9157*
Homes with one or two bedrooms (pct. of total)	-275.316	-4.5196*
Homes with three or four bedrooms (pct. of total)	-535.285	-8.8066*
Homes with five or more bedrooms (pct. of total)	2139.38	5.9000*
Median age of housing units	-234.38	-3.4488*

n = 1483
Adjusted R^2 = 0.566116

**significant at p=0.05*

Table 2. Probit estimates using binary dependent variable *sink_present*.

Variable	Coefficient	t-statistic
const	-0.186665	-1.2109
Total housing units	0.000214976	3.9162*
Median home value	-1.66425e-06	-1.5712
Median age of homes	-0.0418238	-8.8272*

n = 1491
Akaike information criterion (AIC) = 1176.9
Schwarz Bayesian criterion (BIC) = 1198.13
McFadden's pseudo-R^2 = 0.113904

**significant at p=0.05*

Table 3. County-by-county OLS estimates.
Dependent variable: median home value.

	Hernando		Hillsborough	
Variable	Coeff.	t-statistic	Coeff.	t-statistic
Constant	31735.2	1.1587	55360.2	5.9406
Sinkhole density	-3099.19	-1.1186	741.351	0.4497
Median household income	2.29319	5.4061*	2.4366	18.1359*
Vacancy rate	345.571	1.4545	186.415	0.8851
Nonwhite population (pct. of total)	36.8033	0.1801	-175.021	-2.8967*
Homes with one or two bedrooms (pct. of total)	-213.233	-0.8374	-373.12	-3.2903*
Homes with three or four bedrooms (pct. of total)	-7.96901	-0.0298	-720.688	-7.5710*
Homes with five or more bedrooms (pct. of total)	1217.25	0.7610	1817.31	3.1781*
Median age of housing units	-1036.7	-3.2087*	-25.8778	-0.2473
	n = 79		n = 696	
	Adj. R^2:	0.509	Adj. R^2:	0.523
	Pasco		**Pinellas**	
Variable	Coeff.	t-statistic	Coeff.	t-statistic
Constant	46384.1	2.6189	4241.16	0.5373
Sinkhole density	-977.498	-0.8833	-2155.6	-0.6799
Median household income	1.1523	3.2194*	3.7057	20.9963*
Vacancy rate	517.546	2.5159*	820.553	5.7169*
Nonwhite population (pct. of total)	-281.261	-2.0141*	8.95306	0.1567
Homes with one or two bedrooms (pct. of total)	-69.4777	-0.4504	-218.262	-2.6124*
Homes with three or four bedrooms (pct. of total)	213.011	1.1446	-636.588	-6.3491*
Homes with five or more bedrooms (pct of total)	423.644	0.3693	1925.19	3.9087*
Median age of housing units	-1035.5	-4.7487*	-363.374	-3.3427*
	n = 148		n = 568	
	Adj. R^2:	0.705	Adj. R^2:	0.704

**significant at p=0.05*

will not be discussed here.)

A Probit regression, using the binary variable *sinks_present* as its dependent variable, was run in order to generate a second set of results to compare with the OLS results. The Probit regression offered no evidence contradicting the results of the OLS regression. Here, the median structure age and total housing units variables are both statistically significant and show the expected signs: positive for total housing units, negative for median structural age. These results lend some support for the hypotheses that sinkhole reporting is tied to population (sinkholes are more likely to be reported in areas where there are more people to find them). Median home value—our variable of interest in this regression—is not statistically significant at p=0.05, thus confirming the results of the OLS regression.

County-by-county regressions (table 3) were included here to account for the localized nature of housing markets; it seemed possible that any statistically significant relationship between sinkhole density and home values could potentially be obscured by examining the results only at the wider regional level. However, the results of the county-by-county regressions mirrored those of the regionwide analysis. The only explanatory variable that was significant for each countywide regression was median household income, which displayed a positive correlation with the dependent variable in each case. Conversely, the only dependent variable to show no statistical significance in any of the countywide regressions was sinkhole density.

Based on these results, it seems likely that there is no significant relationship between home values and sinkhole location or density at anything greater than an extremely localized scale.

Possible explanations, and potential directions for future research

The most obvious possible explanation for the results described here is that homeowners and homebuyers may not generally consider sinkholes to be a significant threat to their property. This possibility presents an obvious and straightforward avenue for additional research, which could be addressed via surveys and focus groups of homeowners and potential homebuyers in sinkhole-prone areas.

It is also possible that homebuyers may not generally be aware of the locations of sinkholes. The most accurate information on sinkhole locations is often proprietary information held by insurance companies, who have historically been reluctant to share it. The existence of publicly-available information—like the FGS sinkhole database, for example—may not be widely known among the general public.

Of course, regardless of its relatively low public profile, the FGS sinkhole database is still not a complete list of sinkholes within the state. The database relies on information provided by individuals who find a sinkhole and report it. In order for this to happen, a person with information on the location of a sinkhole must know where to report it, or to report it at all. Additionally, database maintenance often depends on the provision of adequate funding by the state government, which is volatile from year to year. The inherent shortcomings of this database may have resulted in an inaccurate picture of sinkhole location and density across the region.

Finally, some sinkholes may be used as "water features" in new residential developments, as a means of adding value to nearby properties. Most homebuyers are unlikely to distinguish between a man-made lake or pond, or a previously-existing sinkhole that has been intentionally converted to that purpose. Because of that, and because these water features are often seen as desirable amenities among homebuyers, it is at least conceivable that any negative impacts of sinkhole proximity on home prices in other areas (particularly, in areas without newer, high-end developments making use of water features) have been obscured.

This paper has demonstrated a lack of statistical evidence pointing to any relationship between home values and the presence of sinkholes in the Tampa Bay area. These results could be due to homebuyer preferences, accessibility of relevant information, or a lack of available data for analysis. Any future research into the question of how sinkholes influence real estate markets should attempt to shed some light onto the underlying cause of the results presented here.

References

Boston, John; Rigsby, Leo C.; Zald, Mayer N. 1972. The impact of race on housing markets: A critical review. Social Problems, vol. 19, no. 3., pp. 382-393.

Brookshire, David S.; Thayer, Mark A.; Tschirhart, John A.; Schulze, William D., 1985. A test of the expected utility model: evidence from earthquake risks. The Journal of Political Economy: vol. 93, no. 2, pp. 369-389.

Dale, Larry; Murdoch, James C.; Thayer, Mark A.; Waddell, Paul A, 1999. Do property values rebound from environmental stigmas? Evidence from Dallas. Land Economics, vol. 75, no. 2, pp. 311-326.

Eastman, Kevin L; Butler, Ann M.; Lilly, Claude C. 1995. The effect of mandating sinkhole coverage in Florida homeowners insurance policies. CPCU Journal, September 1995, pp. 165-176.

Kiel, Katherine; McClain, Katherine, 1995. House prices during siting decision stages: The case of an incinerator from rumor through operation. Journal of Environmental Economics and Management: Vol. 28, pp. 241-255.

Maroney, Patrick F.; Cole, Cassandra R.; Corbett, Richard B.; Dumm, Randy E.; Eastman, Kevin L.; McCullough, Kathleen A. 2005. Preliminary Report: Insurance Study of Sinkholes. Tallahassee, FL.

Nourse, Hugh O., 1967. The effect of air pollution on home values. Land Economics: vol. 43, no. 2, pp. 181-189.

Palm, Risa I. 1981. Public response to earthquake hazard information. Annals of the Association of American Geographers, v. 71, no .3, pp. 389-399.

Palmore, Erdman. 1966. Integration and property values in Washington, DC. Phylon, v. 27, no. 1, pp. 15-19.

Patton, D.; DeHan, B., 1998. Water Issues: Global, National, State, Ecosystem. In Fernald, Edward A.; Purdum, Elizabeth D., eds., Water Resources Atlas of Florida. Florida State University.

Soriano, M.A.; Simon, J.J., 2002. Subsidence rates and urban damages in alluvial dolines of the Central Ebro basin (NE Spain). Environmental Geology, vol. 42, no. 5, pp. 476-484.

Tobin, Graham A.; Montz, Burrell E., 1994. The flood hazard and dynamics of the urban residential land market. Water Resources Bulletin, v. 30, no. 4, pp. 673-685.

United States Census Bureau, 1990. Summary Tape File 3.

United States Geologic Survey, 1999. Sinkholes of West-Central Florida.

White, William B., 1988. Geomorphology and Hydrology of Karst Terrains. Oxford University Press. Oxford, UK.