THE EFFECT OF TRANSPORTATION, LOCATION, AND AFFORDABILITY RELATED SUSTAINABILITY FEATURES ON MORTGAGE DEFAULT PREDICTION AND RISK IN MULTIFAMILY RENTAL HOUSING

Gary Pivo, University of Arizona
May 29, 2013[1]

Introduction

This study examines the relationship between a certain set of sustainability features and default risk in multifamily housing. The sustainability features studied affect both social and environmental concerns including affordability, walkability, auto dependence, exposure to pollution, and proximity to protected open space. All the features studied, except housing affordability, are a function of property location. Various physical and operational characteristics of buildings, such as energy and water efficiency, which are also important to sustainability, could not be studied due to data limitations.

The results show that the sustainability features studied here may be used to improve the prediction of mortgage default. The results also show that the sustainability features reduce default risk in multifamily rental properties. Properties in less auto-dependent residential locations, for example, where 30.0% or more of the workers living in the area commute to work by subway or elevated train, were 58.0% less likely to default, compared to other locations, all else being equal. Similarly significant findings are reported for all the sustainability features examined in the study.

The results also show that in the pool of loans examined, two key original loan terms – the loan-to-value (LTV) ratio and the debt service coverage ratio – have not in the past fully reflected the degree to which the sustainability features studied here mitigate default risk. This is suggested by a logistic regression model of default risk, based on more than 37,000 loans. The sustainability features that were studied lowered the predicted probability of default, even after controlling for the original LTV and debt service coverage ratios and other variables that could affect the ratios, like neighborhood income and regional location. If lenders thought, when they originated the loans, that the presence of these sustainability

© 2012 Gary Pivo. All rights reserved.

[1] Dr. Pivo is Professor of Urban Planning and Professor of Renewable Natural Resources at the University of Arizona, Tucson, Arizona. This paper is incidental to, and leverages the findings of, a consulting engagement performed by Hoyt Advisory Services (HAS) for Fannie Mae. The author wishes to thank Fannie Mae and HAS for their support and comments on earlier drafts. He also is grateful to Dr. Jeffrey Fisher and Dr. Grant Thrall for their help with data. As always, the author is solely responsible for the content, which does not necessarily reflect the opinions of Fannie Mae or HAS.

features would lower default risk, they could have accounted for it as part of the routine underwriting process by approving larger loans or lowering interest rates for those properties with the sustainability features, all else being equal (Grovenstein et al. 2005). That would have caused the ex-post default risk for the more sustainable properties to end up being similar to the risk for the conventional ones. If that had been done, then adding the sustainability features to a default prediction model, which already included the original LTV and debt service coverage ratios, would not have improved the ability of the model to predict default because the information carried by the sustainability regressors would have already been included in the LTV and debt service coverage ratios. That is to say, the effect of the sustainability variables would have already been fully endogenous to the loan origination process (Archer et al. 2002). However, the fact that the additional sustainability variables did improve the model when they were added indicates that they were not fully accounted for in the key loan ratios at loan origination. This is not to say that lenders completely ignored them. Indeed, some of them, such as being in a transit-oriented location, are commonly recognized by lenders as locational advantages.[2] But the evidence reported here suggests that loan terms for properties with certain sustainability attributes could have been made more attractive to borrowers, up to the program limits,[3] without making the loans on the properties riskier than loans on more conventional buildings. Put another way, this paper shows that properties with certain sustainability features are a better risk than previously thought and that in the past those features have not been given sufficient credit in the loan origination process.

Mortgage Default and Multifamily Housing

Mortgage loan defaults are a risk for multifamily real estate investors. In a study of 495 multifamily mortgages securitized by the Resolution Trust Corporation and the Federal Deposit Insurance Corporation and originating between 1989 and 1995, Archer et al. (2002) found default rates[4] by origination year that ranged from 3.7% to 50.0%. In another study of more than 7,000 multifamily loans acquired by Fannie Mae and Freddie Mac between 1983 and 1995, Goldberg and Capone (1998) found default rates[5] that ranged from 0.0% to 3.0% per year. Recently, in the fourth quarter of 2011, the default rate for multifamily mortgages held by the nation's depository institutions was 3.7%.[6] In the present study, the 90-day default rate in the study sample was 0.86%.

Previous studies show that cash flow and property value are major risk factors for multifamily loan default. Default risk increases if declining cash flow prevents loan repayment or if falling property value

---

[2] Locational advantages, such as access to mass transit and convenient shopping, are standard considerations for loan underwriters. For properties with locational advantages, underwriters would typically be comfortable with more aggressive income growth and occupancy projections, which generally result in more confidence in the borrower's ability to refinance a larger amount at the end of the loan term, which in turn leads to acceptance of a lower initial debt service coverage ratio and/or higher loan to value ratio.

[3] Lending programs are normally limited to a certain maximum loan to value ratio and minimum debt service coverage ratio and these constrain any adjustments possible for properties with sustainability features, although banking officials could relax those constraints for the most sustainable properties if there was sufficient evidence that it would not cause excessive risk.

[4] Default was defined in that study as payments ever being late by 90 days or more.

[5] Default included foreclosure, third-party sale, note sale, and short sale events.

[6] http://observer.com/2011/03/banks-commercial-mortgage-default-rates-fellnow-what/

produces negative net equity (Titman and Torous 1989, Kau et al. 1990, Vandell 1984, Vandell 1992, Vandell et al. 1993, Goldberg and Capone 1998, Goldberg and Capone 2002, Archer et al. 2002). In these studies, cash flow and equity are commonly measured in terms of debt service coverage ratio (DSCR), or the ratio of income to required loan payments, and the LTV, or the ratio of loan amount to property value. A lower DSCR and a higher LTV, both at origination and over the life of the loan, are linked to greater default risk.

Sustainable Buildings and Sustainable Investing

Traditionally, the "bottom line" of greatest interest to real estate investors is financial performance. However, many investors understand that real estate also affects other economic, social, and environmental outcomes of interest to various stakeholders (Pivo and McNamara 2005, Roberts 2009). When investments are evaluated in terms of their economic, social, and environmental outcomes, the process is increasingly referred to as "triple bottom line" investing (Hubbard 2009, Slaper and Hall 2011, Sridhar 2012). Investing that seeks good outcomes across all three bottom lines is referred to as "sustainable," "ethical," or "socially responsible" investing (Krosinsky and Robins 2008). By one estimate, real estate is one of the fastest growing types of sustainable and responsible investing (Social Investment Forum 2011).

According to Lutzkendorf and Lorenz (2007), a "sustainable building is…a building that contributes – through its characteristics and attributes – to sustainable development." The most widely recognized definition of sustainable development is "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission on Environment and Development, 1987). There are many theories about what is needed to achieve sustainable development, but they generally imply that sustainable buildings should reduce the material input required per unit of economic production (Hinterberger et al. 1997), prevent society from exceeding the carrying capacity of the man-made and natural systems needed to support people and other livings things (Rees 1992), and improve our wealth and well-being without reducing the stock of economic, social, and natural capital available to future generations (Pearce and Atkinson 1993).

The application of sustainable and responsible investment principles to property investing and development is increasingly referred to as Responsible Property Investing (RPI) (Pivo and McNamara 2005). RPI can include literally dozens of different property development and management practices that seek to address ecological integrity, community development, and human well-being in the course of profitable real estate investing. Examples of RPI practices include improving building energy efficiency, using green cleaning supplies, providing building staff with fair benefits and wages, and engaging with tenants to promote the use of public transit and car-pooling options.

Researchers have attempted to identify the attributes of buildings that should contribute to sustainable development. For example, using an interactive group process that included an international panel of more than 50 experts and stakeholders, Pivo (2007a) concluded that "the panel would emphasize the creation of less automobile-dependent and more energy-efficient cities where worker well-being and urban revitalization are priorities." Similarly, after consultations, literature review, and focus groups, Ellison and Sayce (2007) recommended defining the sustainability of buildings in terms of energy and water use efficiency, pollution and climate control, waste production, adaptability, accessibility, occupier responsibility, and contextual fit.

Features that determine how buildings affect sustainability can be placed into two broad categories. The first includes features that relate to the building itself or its operation, such as geographic orientation, design, intensity, technology, occupant behavior, property management, and affordability. The second category includes features that describe how the building is located relative to critical natural hazards and resources, pollution hotspots, public transit, and desired destinations including jobs, shopping, schools, and parks. Both groups of features influence the social and environmental performance of our built environment. However, as noted above, this paper emphasizes features from the second group because of those data were available for the study.

Several building rating tools exist in the marketplace to help practitioners score the sustainability of properties. Examples include the Building Research Establishment Environmental Assessment Method (BREEAM) in the UK and Europe, the US Green Building Council's Leadership in Energy and Environmental Design (LEED) and the Whole Building Design Guide in the USA, and Green Star in Australia. They differ in their rating criteria, but the most common dimensions include energy efficiency, landscape ecology and habitat conservation, indoor environmental quality, water conservation, and pollution controls (Reed et al. 2009). As this list suggests, these rating tools focus more on the environmental dimension of sustainability than on a triple bottom line approach that seeks better economic, social, and environmental outcomes. Within these dimensions, specific indicators address both building and location considerations. For example, LEED addresses energy issues by awarding points for both the efficiency of its heating, ventilation, and cooling system, and its accessibility by public transit.

The Connection between Sustainability and Mortgage Default

A growing number of econometric studies show that buildings with sustainability features outperform more conventional properties in terms of cash flows and values. Higher values appear to result from both stronger cash flows and lower capitalization rates, suggesting that sustainable properties are favored in both the space (i.e., rental) markets and the capital markets (Pivo and Fisher, 2010). Several different sustainability features have been linked to these effects. They include LEED certification, ENERGY STAR labeling, historic preservation, design excellence, proximity to open space, good transit service, walkability, revitalizing urban locations, and reduced exposure to pollution and natural hazards (Vandell and Lane 1989, McGreal et al. 2006, Simons and Saginor 2006, Harrison et al. 2001, Miller et al. 2008, Eichholtz et al. 2009, Pivo and Fisher 2010, Pivo and Fisher 2011).

Perhaps this should be unsurprising. After all, sustainability features promote qualities long associated with higher valuation in real estate including health, safety, operating efficiency, occupant amenities, and accessibility.[7]

If more sustainable buildings produce better cash flows and values, then they should also exhibit lower default risk because, as noted above, default risk is inversely related to cash flow and value. However, as was also noted above, adding information on sustainability features to the loan origination process would only be helpful if their impact on cash flow and value was not already fully accounted for in the loan origination process. If the financial benefits of sustainability features were fully reflected in the cash flow and value projections made when loans were originated, then rational lenders may have increased the size of the loans or reduced the loan interest rates for the more sustainable properties (Grovenstein et al. 2005). The higher loan amounts or lower interest rates for the more sustainable properties would in turn have increased their LTV or lowered their DSCR ratios, resulting in a probability of default similar to that found in more conventional properties. If, however, the financial benefits of sustainability had not been fully accounted for when the loans were originated, but instead future cash flows and values for more sustainable properties were expected to be closer to those of conventional properties than they actually turned out to be, then as any unrecognized financial benefits from the sustainability features materialized post-origination, the average post-origination DSCR would turn out to be higher and the average post-origination LTV would turn out to be lower for the more sustainable properties than for the conventional ones. This would produce a lower default rate for the sustainable properties compared to the conventional ones, because there would be a lower risk of negative cash flow or a lower risk of negative equity. Even if some of the sustainability features were previously recognized as locational advantages by lenders, as long as their full effect on cash flows, values, and default risk were not fully reflected in the original loan terms, the more sustainable properties would have a lower post-origination default rate. In a modeling context, if the financial benefits of sustainability were not completely reflected in the original DSCR and LTV, then sustainability features could significantly reduce the risk of default when holding original DSCR and LTV constant. Put another way, if sustainability features help buffer properties from losing cash flow and value in the face of energy shocks, recessions, competition, or other events that can push cash flows and values into the default "danger zone" (Bradley et al. 2001), then knowing that a property is more or less sustainable should help lenders mitigate default risk.

This leads to the hypotheses tested for this study.

> *Hypothesis 1*: *If certain transportation-, location-, and housing-related sustainability features are added to a model of default risk, the accuracy of the model will be improved*.

[7] This emerging understanding that sustainability promotes value is creating common ground between those interested in achieving good financial outcomes from real estate investing (i.e., single bottom line investors) and those focused on also achieving good social and environmental outcomes (i.e., triple bottom line investors) (Pivo 2007b).

This hypothesis is based on the expectation that most, if not all, of the impacts that sustainability features have on future cash flows and values were not been fully anticipated when multifamily loans were originated. Even though sustainability features related to locational advantages may have been considered by underwriters, their positive effect on cash flow or values was not fully accounted for in the original loan terms.

> *Hypothesis 2*: *Sustainability features will be associated with a lower risk of default, ceteris paribus.*

This should be the case because prior research has shown that sustainability features tend to be associated with increased cash flows and property values, which are in turn related to lower default risk.

METHODS

Logistic Regression Model

A logistic regression model was used to test the hypotheses. The logistic regression model has been used in several prior studies to estimate the effects of explanatory variables on the probability of mortgage default (Vandell et al. 1993, Goldberg and Capone, 1998, Goldberg and Capone 2002, Archer et al. 2002, Ruaterkus et al., 2010).

Logistic regression is a statistical method for predicting the value of a bivariate dependent variable (Menard 1995). A bivariate variable is one with two possible values, such as pass/fail, or in the present study, default/no default. The value of the dependent variable predicted by a logistic regression model is the probability that a case will fall into the higher of the two categories of the dependent variable, which normally indicates the event occurred, given the values for the case on the independent variables. In other words, it is the probability that an event will occur under various conditions characterized by the independent variables. The predicted value of the dependent variable is based on observed relationships between it and the independent variable or variables used in the study.

Logistic regression models are typically used to determine whether the classification of cases into one of the two categories of the dependent variable can be predicted better from information on the independent variables than if the cases were randomly classified into the categories. "Goodness of Fit" refers to how well all the explanatory variables in a logistic regression model, taken together, predict the dependent variable. So, in terms of logistic regression, Hypothesis 1 is that a logistic regression model for default will have a better goodness of fit if sustainability features are included along with other more conventional explanatory variables in the regression equation.

The estimated parameters of a logistic regression equation can be interpreted as the change in the dependent variable that can be expected from a one-unit change in the dependent variable, holding other dependent variables constant. So again, looking at Hypothesis 2 in terms of logistic regression analysis, the expectation is that sustainability features will have estimated parameters that show they reduce the probability of default by a significant amount.

The most common alternative to the logistic regression model in mortgage default research is the proportional hazard model. Hazard models explain the time that passes before some event occurs in terms of covariates associated with that quantity of time. They have been used to estimate the probability that a mortgage with certain characteristics will default in a given period if there has been no default up until that period (Vandell et al. 1993, Ciochetti et al. 2002, Teo 2004). Other methods for predicting default also have been explored including neural networks (Episcopos et al. 1998) and a maximum entropy approach (Stokes and Gloy 2007). A common view of the hazard model is that it is less sensitive to bias from database censuring than logistic regression. Censoring occurs when cases are removed from the database prior to observation (e.g., when a loan is paid off or foreclosed and sold prior to observation) or when the event of interest happens after observation occurs (e.g., when a loan defaults after the study observation date). However, as pointed out by Archer et al. (2002), bias is only an issue in logistic regression when the explanatory variables have a different effect on the censored and uncensored cases. In the present study, there is no reason to expect that sustainability features affected the odds of default differently in censored and uncensored cases. Hazard models also require a time series dataset that reports the occurrence of defaults over time and such a dataset was unavailable at the start of the present study. One effort to predict mortgage pre-payment using both approaches found that the logistic regression model made better predictions (Pericili et al. 1996), while in another study on insolvency among insurers, the two models produced equally accurate predictions (Lee and Urrutia 1996). So, while it would be interesting to repeat this study using a hazard model, there is no *a priori* reason to assume that the logistic regression method used here produced results that are inferior to those that would have come from another method.

To build the logistic regression model for this study, data from Fannie Mae was combined with data from other sources. The Fannie Mae data included information on all the loans in its multifamily portfolio at the end of Q32011. In the study, each loan was treated as a separate case or observation. For each case, data were available on the loan age, type, terms, and lender, on various financial, physical, and locational attributes of the property, and on the number of days the loan was delinquent, if any. More details on these variables and those from other sources are discussed further below.

Following Archer et al. (2002), cases in the Fannie Mae database with extreme values on certain variables were excluded from the study in order to filter out possible measurement error. The extreme value filters ensured that all loans used in the study had an original note interest rate greater than the 10-year constant maturity risk-free rate at their origination date, an original LTV ratio of 100% or less, an original debt service coverage ratio greater than 0.9% and less than 5.0%, and an original note interest rate greater than 3.0% and less than 15.0%. After these filters were applied, there were 37,385 loans in the sample out of 42,474 loans originally provided for the study by Fannie Mae (including affordable, student, and senior citizen properties). The sample included mortgages with fixed and adjustable rates and with a wide variety of seasoning, originating anywhere from September, 1971 through September, 2011.

The variables evaluated in the models are described in the next section. Table 1 gives their definitions and summary statistics.

Variables in the Study

Default

DEFAULT was a binary variable indicating whether or not a loan was in default as of Q32011. Loans were classified as in default for this study if they were delinquent on their payments by 90 days or more as of Q32011. This definition matches that used by Archer et al. (2002) who pointed out that such a broad definition is useful because other resolutions in addition to foreclosure can be used to resolve defaults and they all involve delinquency-related costs to the lender.

Sustainability Variables

Seven sustainability variables could be analyzed for this study. As discussed above, sustainability is a multi-dimensional construct with multiple distinct but related dimensions treated as a single theoretical concept. Unfortunately, the variables used in this study do not capture all the dimensions normally included in building sustainability. Data on some key sustainability dimensions, such as energy and water efficiency, were unavailable. They should be studied in the future because they probably affect default, due to their influence cash flows and values.

For some of the sustainability variables, a dummy that indicated whether their value fell above or below a cut-point produced a better result in the model than a continuous variable. This is common when a change in the dependent variable associated with a 1-unit change in the independent variable is nonlinear and there is a suspected or assumed threshold effect (Williams et al. 2006). Observed data can be used to find cut-points that best differentiate between high- and low-risk groups. In this study, where nonlinearity was suspected, possible cut-points were examined by categorizing the relevant continuous sustainability variable into quantiles (e.g., quartiles or deciles) and comparing the default rates for each group. Default rates significantly higher for all groups above or below a certain quantile indicated a discontinuity and suggested a likely cut-point where a threshold effect may occur. The cut-point was then used to create a new dummy variable for testing in the model. In several cases this procedure was used iteratively to find the "optimal" cut-point which produced a dichotomous variable that was most useful in predicting default.

In a practical setting, cut-points can be more useful than continuous indicators because they allow a simple risk classification into "high" and "low" and communicate clearly the threshold above (or below) which default risk will be consistently above (or below) average.

Less Auto Dependent Commuting

Three variables were used to capture the nature of the journey to work by residents living in the census tract where each property was located. These variables address the air, water, and wildlife issues linked to commuting. They also address related social issues including traffic accidents, physical activity, and social interaction. COMMUTE TIME was the average commute time in minutes for workers 16 years of age and older living in the tract who worked outside the home. Though commute time can be affected by congestion levels and mode choice, longer commute times generally indicate that people in the tract travel greater distances to work. SUBWAY30 indicates whether a property was in a census tract where at least 30% of the residents take a subway or elevated train to work. PCTWALK indicates the percentage of workers in the census tract who walk to work. The data for these variables came from the 2000 US Census. Newer data, from the American Community Survey, are now being released but were unavailable for this study. It is unlikely, however, that changes in commuting patterns over the past decade were sufficient to significantly alter the study results.

Walkability

RETAIL16 captured the walkability of the area where an apartment building was located. Walkability can be thought of as the degree to which an area within walking distance of a property encourages walking trips from the property to other destinations. It has been linked to a variety of social and environmental benefits (Pivo and Fisher 2011). RETAIL16 was a dummy variable that indicated whether the property was in a census block group with at least 16 retail establishments in 2011. The data were obtained from Claritas, which estimated the number of establishments in 2011 based on the most recent US Economic Census (completed in 2007). As already noted, walking is related to several social and environmental benefits (Federal Highway Administration 2012).

Affordability

AFFORDABLE indicates whether or not the loan is part of the Fannie Mae Targeted Affordable Segment, which focuses on financing properties that are under a regulatory agreement that provides long-term affordability, such as properties with rent subsidies or income restrictions. Family well-being (which is an important element of social capital in sustainability terms) can be in jeopardy if too much of the household budget is required for housing, leaving too little left for food, health care, childcare, transportation, clothing, and other essential household needs (Bratt 2002).

Freeway

FREEWAY1000FT is a dummy indicating whether the property is located within 1,000 feet of the edge of a freeway corridor. There is growing evidence that living close to a freeway produces greater risk for autism, cancers, and respiratory disease, particularly as a result of exposure to diesel fumes (Gauderman et al. 2007, Volk et al. 2011, Office of Health Hazard Assessment 2012, Cakmak et al. 2012). For this study, geospatial data on highway locations was obtained from the US Bureau of Transportation Statistics, 2011 National Transportation Atlas Database and used to compute the distance from each multifamily building to the edge of the nearest freeway.

Protected Areas

PROTECTED1MILE indicates whether a property is located within one mile of a Protected Area as indicated by the US Protected Area Database. The database describes public lands at all government levels that are held for conservation purposes and voluntarily provided privately protected areas. Protected open space helps sustain resource based industries, recreation, wildlife, watersheds, and other valuable ecosystem services such as the absorption of greenhouse gases and the mitigation of urban heat islands. Access to parks and recreational resources has also been linked to less childhood obesity (Wolch et al. 2011).

Control Variables

In this study, the expectation was that if certain sustainability features were related to default risk, it is because they affect cash flow and valuation to a degree not accounted for in the DSCR or LTV ratios at loan origination. However, it could also be true that the sustainability features are correlated with other factors that affect financial outcomes, such as other loan, property, neighborhood, or macroeconomic variables, raising questions about whether sustainability is a proxy for these other drivers of cash flow and value. Therefore, to separate the effects of sustainability features on default risk from these other factors, several control variables, suggested by prior research, were used in the models. The controls fall into four groups including loan, property, neighborhood, and economic characteristics.

Loan Characteristics

OLTV and ODSCR measure the LTV and debt service coverage ratios at loan origination. These are commonly used to predict default risk. Higher OLTV and lower ODSCR were expected to be associated with greater default risk. LOAN_SIZE_GP is an ordinal variable for the loan amount at origination. Esaki et al. (1999) found in a study of commercial loans that the smallest loans had the lowest default rate. LOAN_AGE_MONTHS is the number of months from the loan origination date to the observation date. Previous researchers have shown that default risk declines with age, though the pattern is nonlinear, increasing rapidly in the first few years and then declining (Snyderman 1991, Esaki et al. 1999, Archer et al. 2002). The same pattern was observed in this study sample. Linearity is not a requirement of the logistic regression model, however, and it was unnecessary to transform LOAN_AGE_MONTHS to obtain significant results. Some degree of non-linearity in the logit was detected for LOAN_AGE_MONTHS using the Box-Tidwell transformation (Menard 1995). Consequently, transformations of LOAN_AGE_MONTHS were tried in the model but they did not improve the results. ARM_FLAG is a dummy indicating whether the loan is adjustable or fixed.

Property Characteristics

NO_CONCERNS was a dummy indicating whether or not there were no substantial concerns about the property condition at the time of loan origination. This should reduce default risk by decreasing the need to divert cash flow to deferred maintenance. BUILT_YR is the year the property was built. Archer et al. (2002) found that default rates increased with building age, so BUILT_YR was expected to be inversely related to default risk (i.e., older buildings would default more often). This was the expectation for the nation as a whole, although it could be true that in some areas the historic or design qualities associated with older buildings may be desired and that could influence how age is related to default risk.

TOT_UNTS_CNT is the total number of units in the property. Smaller properties have been reported to experience more financial distress (Bradley et al. 2000). Archer et al. (2002), however, looked at unit count in a multivariate analysis and found that size (and value) was unrelated to default, although their univariate analysis showed that smaller properties had less default risk, contrary to Bradley et al. (2000). So the expected effect in this study was ambiguous.

Neighborhood Characteristics

Researchers have found that stress on properties is related to geographical effects. In fact, Archer et al. (2000) found geographical effects to be one of the most important dimensions for predicting default. Four control variables were created to control for these sorts of effects at the city and neighborhood level. PRINCIPAL_CITY is a dummy indicating whether the property is located in a Principal City, defined by the US Census as the largest incorporated or census designated place in a Core Based Statistical Area.

The goal was to control for whether or not the property was centrally located within a larger metropolitan or micropolitan area because many such areas have outperformed less central, suburban locations in the past decade and several of the sustainability features, such as Walkability, tend to be more common in central cities. Properties in Principal Cities were expected to have lower default risk. URB_RUR was also used to measure regional centrality. It was based on the 11 Urbanization Summary Groups available from the ESRI Tapestry™ Segmentation system, which groups locations into an urban-rural continuum from Principal Urban Centers to Small Towns and Rural places. The system also divides each urbanization group into places with higher and lower affluence; however, that element was ignored for URB_RUR. MEDHHINC000 is the median household income in the census tract as determined by the 2000 census. Higher income was expected to be linked with lower default rates. PROP_CRIME_MIL gives the annual number of property crimes per million persons at the city scale, reported by local police departments to the US Department of Justice. Higher crime was expected to increase default risk.

Regional and National Economy

Certain regional and national variables were included to control for difference in the national and regional economic context faced by the properties since loan origination. Nine dummies were created to indicate whether a property was located in each of the nine census divisions. Vandell et al. (1993) used a similar variable. Several other variables were also tested but found to be insignificant. These included state, metropolitan area, and city location. Additional variables designed to capture regional effects were whether the property was in one of the 25 largest cities (TOP25CITY), dummies for whether the property was located in New York City (NYC) or Washington, DC (DC), and changes in vacancy rates and prices in the metropolitan area in the most recent six-year period. USPRICE_CHANGE is a national indicator that captures the percent change in National Council of Real Estate Investment Fiduciaries (NCREIF) US Apartment Index that occurred between the time the loan was originated and the observation date (Q32011). AVG_PRICE_6 and AVG_OCC_6 were computed using the NCREIF Apartment Index for metropolitan statistical areas. They describe the average increase in apartment prices and the average occupancy rate in the metro area for each property over the last 6 years prior to the study observation date. Prior research updates LTV and DSCR over time on the theory that negative equity or cash flow will trigger default. Both are affected by NOI, which are in turn affected by vacancy rates and rental price indices. Therefore, changes in vacancy rates and rental price indices at the metropolitan scale can be used to capture changes in market conditions that strengthen or weaken mortgages over time (Goldberg and Capone 1998, Goldberg and Capone 2002).

Borrower Characteristics

Lenders consider borrower characteristics to be crucial to reducing default rates. Relevant variables include borrower character, experience, financial strength, and credit history. Unfortunately, data on these issues were not available for this study.

It is unlikely that the omission of borrower characteristics as controls weakened the results. In linear regression, omitted orthogonal variables (i.e., variables that are not correlated with the other independent variables) that are determinants of the dependent variable do not bias the parameter estimates. However, in logistic regression, Cramer (2007) showed that omitted orthogonal variables depress the estimated parameters of the remaining regressors toward zero. Since there is no reason to think that borrower characteristics would be correlated with the sustainability variables (i.e., since they are probably orthogonal), it is possible that the estimated effects of the sustainability variables on default risk reported below would be even larger if borrower characteristics could be included in the analysis.[8]

Collinearity

Correlation among the independent variables is indicative of collinearity. Collinearity can create modeling problems including insignificant variables, unreasonably high coefficients, and incorrect coefficient signs (e.g., negatives that should be positive). Collinearity will not affect the accuracy of a model as a whole, but it can produce incorrect results for individual variables. This makes it less of a concern for the tests of Hypothesis 1 than Hypothesis 2. Tolerance statistics, which check for a relationship between each independent variable and all other independent variables, were used as an initial check for collinearity and they raised no concerns (Menard 1995). However, a pairwise correlation matrix among the independent variables did point to some possible issues. LOAN_AGE_MONTHS and USPRICE_CHANGE were moderately correlated (.737) as were TOT_UNTS_CNT and LOAN_SIZE_GP (.662), both of which make intuitive sense. SUBWAY30 was also correlated with MIDATLANTIC (.684) and NYC (.601).[9] Correlations at this level do not automatically mean there will be collinearity issues, but they do raise the need for further tests, which were done and are reported below.

Results

Table 2 gives the three final models produced for the study. The first model predicts DEFAULT only using conventional explanatory variables unrelated to sustainability. The second model repeats the first but adds the sustainability variables. The third model is a reduced version of the second. It drops insignificant variables to produce a more parsimonious model to achieve the best fit with the fewest parameters. Using irrelevant variables increases the standard error of the parameter estimates and reduces significance (Menard 1995). Insignificant variables were kept in the second model so their effect on the sustainability variables could be considered.

Some variables were excluded from the 3 final models due to collinearity issues.

---

[8] Borrowers' behavior can also influence the sustainability outcomes of properties. For example, they can choose to use green-certified cleaning supplies and engage with tenants on sustainability practices. This discussion about the effect of omitting borrower characteristics for modeling default is not intended to indicate that borrowers are unimportant to property sustainability outcomes.

[9] Some readers may be surprised that DC was not highly correlated with SUBWAY30 because DC has a highly developed metro system. But the issue is not whether a high proportion of the DC properties received a 1 for SUBWAY30 but whether there are a disproportionate number of SUBWAY30 properties in the study that are located in DC such that the stronger DC economy is driving the findings for SUBWAY 30. For the whole sample, 0.6% of all cases were in DC and 1.7% of SUBWAY30 cases were in DC. So there is not a highly disproportionate number of SUBWAY30 in DC and the default rate of SUBWAY30 cases is not biased by the DC default rate. Compare this to NYC which has 2.8% of all cases but 29.3% of all SUBWAY30 cases. In this instance the NY economy and default rate could bias the default findings for SUBWAY30 cases. This is why there was a possible collinearity issue with NY in the model but not with DC.

LOAN_AGE was insignificant and had the wrong sign when it was included in the models with USPRICE_CHANGE. This was corrected when LOAN_AGE was used without USPRICE_CHANGE. LOAN_AGE was kept instead of US_PRICE CHANGE because it captured the information included in USPRICE_CHANGE (since they were correlated), and because it captures the effect of other forces that may have affected DEFAULT due to the seasoning of the loan and the market conditions when and since the loan entered the market.

TOT_UNTS_CNT was insignificant when included with LOAN_SIZE_GP; however, both were significant on their own. TOT_UNTS_CNT was used in the final models because it had more information content than LOAN_SIZE_GP, which was only ordinal rather than continuous.

The coefficient for SUBWAY30 was inflated when MIDATLANTIC and NYC were in the model, so both were excluded, which produced more conservative results (smaller effects) for SUBWAY30 in both cases. NYC was reintroduced in a robustness check, which is discussed below.

Since collinearity tends to produce unreasonably high regression coefficients, a rough indication of remaining collinearity in the models would be if any of the unstandardized coefficients (B) were greater than 2 (Menard 1995). This was not the case, indicating the steps taken to reduce collinearity were sufficiently effective.

In the final models, all the variables had the expected signs except for URB_RUR, however its results were insignificant. Also, larger properties had smaller default rates, supporting the findings by Bradley et al. (2000).

Goodness of Fit with and without Sustainability

The first hypothesis was that if certain sustainability features are added to a model of default risk, the accuracy of the model would be improved. This hypothesis was tested by comparing the goodness of fit of models with and without certain transportation-, location-, and housing-related sustainability features included.

Goodness of fit statistics are reported in the last four rows of Table 2. For all of the statistics except -2 Log likelihood, a higher value indicates a better fitting model. All four statistics show there was less discrepancy between the observed values for DEFAULT and the values produced by the model when sustainability features were included. That supports the acceptance of Hypothesis 1.

The Model Chi-square measures the total reduction for all the cases in default prediction errors that occurs when the independent variables are in the model, compared to when they are not in the model. Comparing the Model Chi-square across the models in Table 2 indicates that the models with the sustainability features predicted default more accurately than the model without them.

The -2 Log likelihood statistic is similar to the Model Chi-square since it is based on the total error made by the model in predicting default for all the cases combined. But -2 Log likelihood measures the total error made by the model with the independent variables included rather than the difference between the error with and without the independent variables. That is, it measures how poorly a model fits the data with all the independent variables in the equation. That is why a better model has a smaller -2 Log likelihood. Here again, comparing across the models in Table 2, the results show that the models with the sustainability variables more accurately predicted default than the model without them.

The Nagelkerke R-Square is a "pseudo R-square" that can be computed in logistic regression analysis. Pseudo R-squares are not analogous to the R-square in linear regression and are easily misinterpreted by readers familiar with linear regression models, according to Hosemer and Lemeshow (2000). For that reason, they recommend against reporting them. The Nagelkerke R-square is a measure of improvement from the null model to the fitted model (i.e., the improvement in each model produced by adding the independent variables). It is most useful for comparing multiple models predicting the same outcome with the same dataset, as in the present case. When used that way, the models with the higher R-squares are the ones that better predict the outcome. As Table 2 shows, the models with sustainability features included had higher Nagelkerke R-squares.

Goodness of fit was also tested using the Area Under the Receiver Operating Characteristic (ROC) Curve. This test measures the model's ability to discriminate between loans that do and do not default and is the likelihood that a loan that defaults will have a higher predicted probability than a loan that does not. If the result is equal to 0.5, the model is no better than flipping a coin. In the present study, ROCs were in the .83 to .84 range. These values indicate excellent discrimination (Hosemer and Lemeshow 2000). In other words, the models did an excellent job distinguishing between loans that will and will not default. Here again, the models with sustainability features did a better job than the model without.

According to all four goodness of fit measures, the results support the first hypothesis that if certain sustainability features are added to a model of default risk, the accuracy of the model improves.[10]

Interpretation of Sustainability Coefficients

The second hypothesis was that sustainability features will be associated with a lower risk of default. This hypothesis was tested by examining and interpreting the regression coefficients in the reduced model (Model 3).

According to the significance tests ("sig." in Table 2), all of the sustainability features were significantly related to default risk. This means it is highly improbable that the sustainability features would be this strongly related to default in a sample of this size if there really were no relationships. For example, there is less than a 0.01% chance that COMMUTE TIME, and just a 0.9% chance that PCTWALK, respectively, are unrelated to DEFAULT.

The size and direction of the relationships are indicated by the unstandardized coefficients (B). In Models 2 and 3, B gives the change in the risk of default associated with a 1-unit change in the sustainability variables, while the control variables are held constant. If B is positive, then default risk increases with each 1-unit increase in the sustainability variable. For example, in Model 3, a B of .041 for

[10] The Hosmer–Lemeshow test is another commonly recommended statistical test for goodness of fit in logistic regression models. However its assumption, that the expected frequencies are large, was violated because default was a relatively rare event, so its results would be invalid in the present study (Hosmer and Lemeshow 2000).

COMMUTE TIME indicates that the risk of default increases as the average commute time increases in the area where the property is located. If B is negative, then default risk decreases with each 1-unit increase in the sustainability variable. For example, in Model 3, a B of -.878 for SUBWAY30 indicates that the risk of default decreases when a property is located where 30% or more of the workers commute to work by subway or elevated train. And since these findings were estimated when the control variables were also in the model, we can say they are true regardless of neighborhood income (MEDHHINC000), whether or not the loan is adjustable (ARM_FLAG), the debt service coverage ratio at origination (ODSCR), etc.

Together, the significance tests and unstandardized coefficients in Models 2 and 3 indicate that when the sustainability features studied here are present, there is less risk of default, all else being equal. Note, however, that for some of the sustainability variables, a larger raw score indicates the property is *less* sustainable. That is the case for COMMUTE TIME and FREEWAY1000FT. For these variables, a positive B means that a lower value (which indicates less sustainability), such as a longer commute time, is associated with less default risk. For the other variables a negative B means that a higher value (which indicates more sustainability), such as more walking to work, is associated with less default risk.

The unstandardized coefficients can be used to obtain estimated odds ratios by exponentiating the coefficients (i.e., computing its base e anti-log). An odds ratio is the odds of an outcome in one group (e.g., the default rate for properties near a protected area) divided by the odds of an outcome in another group (e.g., the default rate for properties not near a protected area). It is analogous to relative risk (Grimes and Schulz 2008). The odds ratio associated with each independent variable is given as Exp(B) in Table 2. As explained by Menard (1995), the odds ratio is the number by which we would multiply the odds of default for each 1-unit increase in the independent variable. An Exp(B) greater than 1 indicates that the odds of default increase when the independent variable increases and an Exp(B) less than 1 indicates that the odds of default decrease when the independent variable increases. For example, a 1-unit increase in COMMUTE TIME (i.e., a one-minute increase, according to the definition column in Table 1) results in a 3.7% increase in the odds of default (the odds of default are multiplied by 1.037). Similarly, a 1-unit increase in SUBWAY30 results in 58.4% decrease in the odds of default (the odds of DEFAULT are multiplied by .584, which is .416 less than 1). [11]

Odd ratios can also be interpreted as relative risk when the outcome occurs less than 10% of the time, which is the case for DEFAULT in the study sample (Hosmer and Lemeshow 2000). Relative risk is the ratio of the probability of an event occurring in a group with and without a certain characteristic. We can say, for example, that in locations where commute time is 28 minutes (or about one minute above average), multifamily mortgages are 3.7% more likely to default than in locations with a 27-minute commute time. Similarly, in locations where 30.0% or more residents take a subway or elevated train to

[11] For a dummy variable, that can only be scored as 0 or 1, a 1-unit increase is equivalent to referring to all cases where the dummy variable has a score of 1 or Yes. So, for example, for SUBWAY30, if we say that a 1-unit increase decreases the odds of default by 58.4%, we are saying that when SUBWAY30 is scored 1 (i.e., when a property is located in a location where 30% or more of the workers commute by subway or elevated to work), the odds of default are 58.4% lower than in cases where the property is not located in such a location and SUBWAY30 is scored 0.

work, owners are 58.0% less likely to default in comparison to locations where fewer than 30.0% commute by subway or elevated. And, as noted above, the reductions in risk associated with the each of the sustainability features are unrelated to differences in regional location, neighborhood income, and the other control variables included in the models.

Table 3 summarizes the effects of a 1-unit change in the sustainability variables on default risk in relative risk terms. In every case, the effects are large, indicating that these sustainability features have a very significant effect on default risk, independent of other factors commonly used to predict default. In addition, the direction of each relationship indicates that in all cases more sustainability is related to less default risk. These findings confirm the second hypothesis that sustainability features are associated with a lower risk of default.

When dealing with continuous variables, such as COMMUTE TIME and PCTWALK, a 1-unit change is not always interesting. For example, a one-minute increase in commuting or a 1.0% increase in the percentage who walk to work may be too small to be considered important. Hosemer and Lemeshow (2000, p. 63) show that as in the case of a 1-unit increase, the effect from a multi-unit change in an independent variable on the odds of default can be determined from an estimated odds ratio. In the case of a multi-unit change, the odds ratio is estimated by exponentiating the product of the unstandardized coefficient (B) for a given variable times the number of units of change. If, for example, one were interested in how a 10-minute increase in COMMUTE TIME affected the odds of default, the odds ratio would be computed by exponentiating the product of 10 x B for COMMUTE TIME. Applying this method using the coefficients from Model 3 (Table 2), we find that for every 10-minute increase in the mean commute time for residents in a census tract, the risk of default increases by 45.0%. Similarly, for every 5-unit increase in the percent of workers in a tract who walk to work, the risk of default decreases by 15.0%.

As discussed above, NYC was excluded from the models. This was because SUBWAY30 was inflated with NYC in the model due to collinearity issues. However, it was important to know if the effects of SUBWAY30 were due to its association with NYC. To answer that question, Model 3 was rerun using the 33,733 cases that were not located in NYC. The unstandardized coefficient for SUBWAY30 (B) was -1.062 (.003) and the Exp(B) was .356. Compared to the results in Table 2, this indicates that the effect of SUBWAY30 outside of NYC was even stronger. When cases in New York City were excluded from the sample, the relative risk of default was 64.4% less for properties located where at least 30.0% of the tract residents commuted by subway or elevated, compared to 58.4% less when cases in New York City were included in the sample. This indicates that the effects of SUBWAY30 on default were not because a disproportionate share of the cases in subway-oriented locations were in NYC.

Discussion

The findings support the hypotheses. When certain transportation-, location-, and affordability-related sustainability features are included in a default probability model for multifamily housing, the model predicts default more accurately. Also, properties with the sustainable features studied are much less likely to default.

These findings suggest two important implications for practice, one pertaining to sustainability and the other to risk management.

First, certain aspects of sustainability can be fostered without increasing default risk by adjusting conventional lending standards for properties with the sustainability features studied here to achieve the same risk of default associated with less sustainable properties. For example, in the study sample, the mean OLTV and ODSCR were 0.61 and 1.52 respectively. However, according to Model 3, if a property is located in a subway-oriented census tract (where at least 30.0% of the workers commute by subway or elevated train), then the OLTV could have been raised to nearly 0.75 and the ODSCR could have been lowered to 1.28 (which maintains the normal ratio between the two found in the study sample) and the probability of default would have remained virtually equal to that found for properties not located near subway stations, all else being equal. According to the model, the extra risk produced by a higher OLTV and lower ODSCR would be offset by the lower risk produced by locating in a more sustainable location, leaving the total risk unchanged. Similar results can be produced with various other combinations of the studied sustainability features and conventional lending standards. Of course, lenders have programmatic constraints that set maximums for the OLTV and minimums for the ODSCR. Those constraints establish practical limits on how far lenders could go in adjusting for sustainability features unless they are willing to make exceptions to the programmatic constraints for that purpose.

If higher LTV ratios at origination could be obtained by borrowers for more sustainable properties, the borrowers would achieve a higher return on equity as long as positive leverage is possible (i.e., when the cost of debt financing is lower than the overall return generated by the property return on asset). *Ceteris paribus*, this should cause investors to prefer more sustainable investments, increase capital flow to more sustainable buildings, and foster transformation toward more sustainable cities.

The second implication of the findings is that lenders could improve their risk management practices by taking stock of whether a property has certain transportation-, location-, and affordability-related sustainability features when loans are originated. All the features studied here are based on existing federal datasets provided by the Census Bureau and other agencies. So it would be relatively easy to build an online address-based lookup tool that any lender can use to obtain certain sustainability information on a given property. In addition, recommended adjustments to OLTV and ODSCR ratios could be made available through a second online tool based on coefficients found in this study and confirmed by additional research. The results clearly show that traditional lending ratios have not fully recognized the presence or absence of certain sustainability features (by recognizing their effect on cap rates, values, or cash flows). The most common result is that these ratios have been based on an underestimation of the actual risk of default for properties without those sustainability features. This is because the “normal” rate of default expected for all properties is derived from the experience with more and less sustainable properties combined, without the effect of sustainability being recognized. In this situation, less sustainable properties, according to the findings in this study, should have a rate of risk higher than the norm. There is more risk associated with less sustainable properties (and less risk with more sustainable properties) than is being accounted for in traditional lending ratios. If better information could be made available on the size of these effects and on how lending ratios could be adjusted to offset the effects produced by the presence or absence of sustainability features, the risk of portfolios could be more effectively managed.

Unfortunately, this study is limited to those sustainability features for which data could be obtained. It is likely, however, that other sustainability features that were not studied here also have beneficial effects on default risk. The most likely examples are those pertaining to energy efficiency and green building certification because prior studies have shown they affect cash flow and valuations. Other sustainability features that could reduce default risk include rental unit flexibility, urban centrality, noise mitigation, water efficiency, childcare services, and school quality. This is expected because of their likely “materiality” to financial performance, rather than their positive effects on the public good (Pivo 2008). Further work on whether these features do or do not affect default risk would be most useful, but it will have to await for the development of better mortgage databases that account for them.

Other fruitful avenues for further research include repeating this study using other modeling methods (e.g., hazard models) and data sets. It would also be useful to conduct similar studies on other property types, including single-family homes and other commercial properties. And work should begin on a practical tool that helps lenders adjust conventional lending ratios based on sustainability information without increasing normal risk levels.

Conclusion

Perhaps the most important point of this study is that sustainability is just as much an issue with material consequences for investors as for those interested in social and environmental well-being. Seven years ago, Pivo and McNamara (2005) wrote about the common ground emerging between real estate investing and sustainability in the following passage:

> “It is probably apparent to anyone who thoughtfully considers real estate that it can both contribute to and be affected by many of the social and environmental issues that face the world’s societies. Until recently, however, most real estate investors would likely have said that while they are sympathetic, such issues are…not of direct concern to their investment practices. But today, a new view is emerging…that various social and environmental issues can have significant material consequences for their investment portfolios.”

Sustainability is financially consequential for property investors. Multifamily lenders can mitigate risk by gearing their portfolios toward more sustainable properties. Moreover, and for society this may be even more important, lenders can offer favorable financial terms to more sustainable properties without increasing risk, or as Benjamin Franklin once said, they can “do well by doing good”.

References

Archer WR, Elmer PJ, Harrison DM and Ling DC (2002), Determinants of Multifamily Mortgage Default, *Real Estate Economics* 30(3), 445-473.

Bradley DS, Cutts AC and Follain JR (2001), An Examination of Mortgage Debt Characteristics and Financial Risk among Multifamily Properties, *Journal of Housing Economics* 10, 482-506.

Bratt, RG (2002), Housing and Family Well-Being, *Housing Studies* 17(1), 13-26.

Cakmak, S, Mahmud M, Grgicak-Mannion A and Dales RE (2012), The influence of neighborhood traffic density on the respiratory health of elementary schoolchildren, *Environment International* 39(1), 128-132.

Cramer, JS (2007), Robustness of Logit Analysis: Unobserved Heterogeneity and Mis-specified Disturbances, Oxford Bulletin of Economics and Statistics 69(4), 545-555.

Ciochetti BA, Deng Y., Gao B., and Yao R (2002), The Termination of Commercial Mortgage Contracts through Prepayment and Default: A Proportional Hazard Approach with competing Risks, *Real Estate Economics* 30(4), 595-633.

Eichholtz P, Kok, N and Quigley JM (2009), *Doing Well by Doing Good? An analysis of the financial performance of green office buildings in the USA*, Royal Institute of Chartered Surveyors, London.

Ellison, L. and S. Sayce (2007), Assessing Sustainability in the Existing Commercial Property Stock: Establishing Sustainability Criteria Relevant for the Commercial Property Investment Sector, *Property Management* 25(3), 287–304.

Episcopos A, Pericli A and Hu J (1998), Commercial Mortgage Default: A Comparison of Logit with Radial Basis Function Networks, *The Journal Of Real Estate Finance And Economics* 17(2), 163-178.

Esaki H, L'Heureux S and Snyderman M (1999), Commercial Mortgage Defaults: An Update, *Real Estate Finance* 16(1), 80-86.

Federal Highway Administration (2012), *Health and Environmental Benefits of Walking and Bicycling*, fhwa.dot.gov.

Gauderman WJ, Vora H, McConnell R, Berhane K, Filliland F, Thomas D, Lurmann F, Avol E, Kunzli N, Jerrett M and Peters J (2007), Effect of exposure to t raffic on lung development from 10 to 18 years of age: a cohort study, *The Lancet* 369(9561), 571-577.

Goldberg L and Capone CA (1998), Multifamily Mortgage Credit Risk: Lessons From Recent History, *Cityscape* 4(1), 93-113.

Goldberg L and Capone CA (2002), A Dynamic Double-Trigger Model of Multifamily Mortgage Default, *Real Estate Economics* 30(1), 85-113.

Goodman J (2004), Determinants of operating costs of multifamily rental housing, *Journal of Housing Economics* 13(3), 226-244.

Grimes DA and Schulz KF (2008), Making sense of odds and odds ratios. Obstetrics and Gynecology 111(2), 423-426.

Grovenstein RA, Harding JP, Sirmans CF, Thebpanya S and Trunbull GK (2005), Commercial mortgage underwriting: How well do lenders manage the risks? *Journal of Housing Economics* 14, 355-383.

Haas, PM, Makarewicz C, Benedict A and Bernstein S (2008), Estimating Transportation Costs by Characteristics of Neighborhood and Household, *Transportation Research Record* 2077, 62-70.

Harrison, D., G. Smersh, and A. Schwartz. 2001. Environmental Determinants of Housing Prices: The Impact of Flood Zone Status, *Journal of Real Estate Research* 21: 3-20.

Hinterberger F, Luks F and Schjidt-Bleek F (1997), Material flows vs. 'natural capital' What makes an economy sustainable?, *Ecological Economics* 23, 1-14.

Hosemer DW and Lemeshow S (2000), *Applied Logistic Regression, 2nd Edition*, John Wiley & Sons, Inc, New York.

Hubbard, G (2009), Measuring Organizational Performance: Beyond the Triple Bottom Line, *Business Strategy and the Environment* 18, 177-191.

Kau JB, Keenan DC, Muller WJ, and Epperson JF (1990), Pricing commercial Mortgages and their Mortgage-Backed Securities, *The Journal of Real Estate Finance and Economics* 3, 333-356.

Krosinsky C and Robins N (2008), *Sustainable Investing: the Art of Long-Term Performance*, Earthscan, London.

Lee SH and Urrutia JL (1996), Analysis and Prediction of Insolvency in the Property-Liability Insurance Industry: A Comparison of Logit and Hazard Models, *The Journal of Risk and Insurance* 63(1), 121-130.

Lipman, B, *A heavy load: The combined housing and transportation burdens of working families*, Center for Housing Policy, October, 2006.

Lutzkendorf T and Lorenz D (2007), Integrating sustainability into property risk assessments for market transformation, *Building Research and Information* 35(6), 644-661.

McGreal S, Webb JR, Adair A, Berry J (2006), Risk and diversification for regeneration/urban renewal properties: evidence from the UK. *Journal of Real Estate Portfolio Management* 12(1): 1-12.

Menard, S (1995), Applied Logistic Regression, *Sage University Paper series on Quantitative Applications in the Social Sciences*, 07-106, Sage, Thousand Oaks, CA.

Miller N, Spivey J, and Florance A (2008), Does Green Pay Off? Journal of Real Estate Portfolio Management 14(4): 385-400.

Office of Health Hazard Assessment (2012), Health Effects of Diesel Exhaust, http://oehha.ca.gov.

Pearce, DW and Atkinson GD (1993), Capital theory and the measurement of sustainable development: an indicator of 'weak' sustainability, *Ecological Economics* 2, 103-108.

Pericli, Andreas, Hu, Jianxun and Masri, Joseph, The Prepayment of Fixed-Rate Mortgages: A Comparison of Logit with Proportional Hazard Models (June 25, 1996). Available at SSRN: http://ssrn.com.ezproxy2.library.arizona.edu/abstract=7982

Pivo G and McNamara P (2005) Responsible Property Investing, *International Real Estate Review* 8(1), 128-143.

Pivo, G (2007a), Responsible property investment criteria developed using he Delphi Method, *Building Research and Information* 36(1), 30-36.

Pivo, G (2007b), *Responsible Property Investing: What the Leaders Are Doing*. United Nations Environment Programme Finance Initiative, Geneva.

Pivo G and Fisher J (2010), Income, Value, and Returns in Socially Responsible Office Properties, *Journal of Real Estate Research* 32(3), 243-270.

Pivo G and Fisher J (2011), The Walkability Premium in Commercial Real Estate Investments, *Real Estate Economics* 39(2), 185-219.

Rauterkus SY, Thrall GI and Hangen E (2010), Location Efficiency and Mortgage Default, *Journal of Sustainable Real Estate* 2(1), 117-141.

Reed R, Bilos A, Wilkinson S and Schulte K-W (2009), International Comparison of Sustainable Rating Tools, *Journal of Sustainable Real Estate* 1(1), 1-22.

Rees WE (1992), Ecological footprints and appropriated carrying capacity: what urban economics leaves out, *Environment and Urbanization* 4(2): 121-130.

Roberts C (2009), Sustainable and Socially Responsible Investment, *Journal of Property Investment and Finance* 27(5), 445-446.

Shilling, J., J. Benjamin, and C. Sirmans. 1985. Adjusting Comparable Sales for Floodplain Location, *The Appraisal Journal* 53, 429-436.

Simons RA and Saginor JD (2006), A Meta-Analysis of the Effect of Environmental Contamination and Positive Amenities on Residential Real Estate Values, *Journal of Real Estate Research* 28(1), 71-104.

Slaper TF and Hall TJ (2011), The Triple Bottom Line: What Is It and How Does It Work?, *Indiana Business Review* 86(1).

Snyderman, M (1991), Commercial mortgages: default occurrence and estimated yield impact, *Journal of Portfolio Management* 18(1), 6-11.

Social Investment Forum (2011), *2010 Report on Socially Responsible Investing Trends in the United States*, Social Investment Forum, Washington, D.C.

Sridhar, K (2012), The Relationship between the Adoption of Triple Bottom Line and Enhanced Corporate Reputation and Legitimacy, *Corporate Reputation Review* 15 (summer), 69-87.

Stokes JR and Gloy BA (2007), Mortgage Delinquency Migration: An Application of Maximum Entropy Econometrics, *Journal of Real Estate Portfolio Management* 13(2), 153-160.

Teo AHL (2004), Delinquency Risk in Residential ARMs: A Hazard Function Approach, *Journal of Real Estate Portfolio Management* 10(3), 243-258.

Titman S and Torous W (1989), Valuing Commercial Mortgages: An Empirical Investigation of the Contingent-Claims Approach to Pricing Risky Debt, The Journal of Finance 59, 165-206.

Vandell KD (1984), On the Assessment of Default Risk in Commercial Mortgage Lending*, AREUEA Journal* 12(3), 270-296.

Vandell KD (1992), Predicting Commercial Mortgage Foreclosure Experience, *Journal of the American Real Estate and Urban Economics Association*, 20(1), 55-88.

Vandell, K.D. and Lane, J.S., 1989. The economics of architecture and urban design: some preliminary findings, *AREUEA Journal* 17(2), 235-260.

Vandell K, Barnes W, Hartzell D, Kraft D and Wendt W (1993), Commercial Mortgage Defaults: Proportional Hazards Estimation Using Individual Loan Histories, Journal of the American Real Estate and Urban Economics Association 4(21), 451-480.

Volk HE, Hertz-Picciotto I, Delwiche L, Lurmann F and McConnell R (2011), Residential Proximity to Freeways and Autism in the CHARGE Study, *Environmental Health Perspectives* 119(6), 873-877.

Williams BA, Mandrekar JN, Mandrekar SJ, Cha SS and Furth AF (2006), Finding Optimal Cutpoints for Continuous Covariates with Binary and Time-to-Event Outcomes, *Technical Report Series #79*, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, June 2006.

Wolch J, Jerrett M, Reynolds K, McConnell R, Chang R, Dahmann N, Filliland F, Su JG and Berhane K (2011), Childhood obesity and proximity to urban parks and recreational resources: A longitudinal cohort study, *Health and Place* 17(1), 207-214.

World Commission on Environment and Development (1987), *Our Common Future: Report of the World Commission on Environment and Development*, United Nations, New York.

Tables

TABLE 1 ▪ Definitions and summary statistics for variables in final models (n=37,386)

Variable	Definition	Min.	Max.	Mean	Std. dev.
DEPENDENT VARIABLE					
DEFAULT	Binary variable indicating whether loan was (1) or was not (0) in default. Default was defined as >90 days delinquent as of Q311.	0	1	.0086	0.092
SUSTAINABILITY VARIABLES					
COMMUTE TIME	Mean commute time in 2000 for residents in the census tract (minutes)	2.10	75.50	26.72	6.72
SUBWAY30	In census tract where at least 30% of workers use a subway or elevated train for work (1=yes, 0=no)	0	1	0.076	0.265
PCTWALK	Pct of workers in census tract that walked to work in 2000	0	100	4.92	7.40
RETAIL16	16 or more retail establishments in census block group	0	1	0.40	0.49
AFFORDABLE	Meets FNMA affordable housing standards (1=yes, 0=no)	0	1	0.10	0.30
FREEWAY1000FT	Within 1,000 feet of an interstate freeway (1=yes, 0=no)	0	1	0.40	0.49
PROTECTED1MILE	Within 1 mile of a protected area (1=yes, 0=no)	0	1	0.71	0.45
CONTROL VARIABLES					
OLTV	Loan to value ratio at origination	.59	100	61.30	16.29
ODSCR	Debt service coverage ratio at origination	.90	5.0	1.52	0.549
LOAN_AGE_MONTHS	Number of months since loan origination.	.00	468.00	73.15	52.91
LOAN_SIZE_GP	Original loan size; ordinal variable (1=<$3MM, 2=$3-5MM, 3=$5-25MM, 4=>$25MM	1	4	1.70	0.97
ARM_FLAG	Dummy for adjustable rate mortgage	0	1	0.31	0.46
NOCONCERNS	Dummy indicating absence (1) or presence (0) of concerns about physical condition at origination	0	1	0.28	0.40
BUILT_YR	The year the property was built	1800	2011	1967.83	26.25
TOT_UNTS_CNT	Total housing units in the property	2	3284	94.65	125.05
PRINCIPAL_CITY	Dummy for whether or not located in US Census Principal City	0	1	0.60	0.49
URB_RUR	Ordinal variable indicating whether property is in Principal Urban Center, Metro City, Urban Outskirts, Suburban Periphery, Small Town or Rural location.	1	7	1.92	1.16
MEDHHINC000	Median household income in census tract in 2000 (000 dollars)	0.0	200.00	42.70	16.94
PROP_CRIME_MIL	Annual number of property crimes per million persons in the city	0.0	2849.85	407.48	165.31
US_PRICE_CHANGE	Pct change in NCREIF US Apartment Index from quarter loan originated to Q32011	-13.54	137.89	16.26	27.22
NEWENGLAND	Dummy for New England region	0	1	0.03	0.17
MIDATLANTIC	Dummy for Mid-Atlantic region	0	1	0.14	0.35
ENCENT	Dummy for East North Central region	0	1	0.08	0.26

WNCENT	Dummy for West North Central region	0	1	0.04	0.19

TABLE 1 (cont.) ▪ Definitions and summary statistics for variables in final models (n=37,386)

Variable	Definition	Min.	Max.	Mean	Std. dev.
SOATLANTIC	Dummy for South Atlantic region	0	1	0.09	0.29
ESOCENTRAL	Dummy for East South Central region	0	1	0.02	0.14
WSOCENT	Dummy for West South Central region	0	1	0.08	0.27
MTN	Dummy for Mountain region	0	1	0.05	0.22
PACIFIC	Dummy for Pacific region	0	1	0.47	0.50
AVG_PRICE_6	Average apartment price change in percent from Q32005 to Q32011 in percent in the MSA	-50.30	20.79	-1.33	3.51
AVG_OCC_6	Average apartment occupancy rate in percent from Q32005 to Q32011 in percent in the MSA	33.43	100	91.04	3.66
TOP25CITY	Dummy for in one of 25 most populous US cities	0	1	0.230	0.42
NYC	Dummy for in New York City	0	1	0.028	0.164
DC	Dummy for in Washington, DC	0	1	0.057	0.075

TABLE 2: Logistic Regression Results for *DEFAULT*

	Model 1: Without Sustainability		Model 2: With Sustainability		Model 3: With Significant Variables Only	
	B (sig.)	Exp(B)	B (sig.)	Exp(B)	B (sig.)	Exp(B)
Loan						
OLTV	.041(.000)	1.042	.044 (.000)	1.045	.043 (.000)	1.044
ODSCR	-.868 (.003)	.420	-1.037 (.001)	.355	-1.043(.001)	.352
LOAN_AGE_MONTHS	-.005 (.002)	.995	-.004 (.019)	.996	-.004 (.020)	.996
ARM_FLAG	.578 (.000)	1.782	.468 (.001)	1.596	.477 (.000)	1.611
Property						
NOCONCERNS	-.902 (.000)	.406	-.820 (.000)	.440	-.827 (.000)	.437
BUILT_YR	-.015 (.000)	.985	-.016 (.000)	.985	-.016 (.000)	.984
TOT_UNTS_CNT	-.004 (.000)	.996	-.004 (.000)	.996	-.004 (.000)	.996
Neighborhood						
PRINCIPAL_CITY	.145 (.297)	1.156	.285 (.057)	1.330		
URB_RUR	.041 (.481)	.960	.020 (.745)	.980		
MEDHHINC000	-.028 (.000)	.973	-.033 (.000)	.968	-.035 (.000)	.966
PROP_CRIME_MIL	.001 (.001)	1.001	.001 (.000)	1.001	.001(.000)	1.001
Economy						
TOP25CITY	-.358 (.032)	.699	-.564 (.002)	.569	-.419 (.011)	.658
DC	-.1.158 (.115)	.314	-1.420 (.056)	.242	-1.247 (.091)	.287
REGION	unreported	unreported	unreported	unreported	unreported	unreported
AVG_PRICE_6	.004 (.790)	1.004	.005 (.742)	1.005		
Sustainability						
COMMUTE TIME			.041 (.000)	1.042	.037 (.000)	1.037
SUBWAY30			-821 (.014)	.440	-.878 (.008)	.416
PCTWALK			-.031 (.008)	.969	-.031 (.009)	.969
RETAIL16			-.417 (.002)	.659	-.421 (.002)	.656
AFFORDABLE			-.959 (.000)	.383	-.964 (.000)	.381
FREEWAY1000FT			.455 (.044)	1.576	.464 (.040)	1.590
PROTECTED1MILE			-.401 (.009)	.669	-.393 (.010)	.675
Constant	24.620 (.000)	4.926E10	25.841 (.000)	1.670E11	25.492 (.000)	1.178E11
n	37,382		37,382		37,382	
Goodness of Fit						
Model Chi-square	549.54 (.000)		625.55 (.000)		621.54 (.000)	
-2 Log likelihood	3097.149		3019.951		3023.967	
Nagelkerke R- Square	.157		.179		.178	
Under ROC Curve	.829		.841		.841	

TABLE 3: Summary of Sustainability Effects on Default Risk in Multifamily Mortgages

Variable	Definition	Effect on Relative Risk of Default
COMMUTE TIME	Per 1 minute increase in the commute time to work	3.7% more
SUBWAY30	Where ≥ 30% commute by subway or elevated	58. 4% less
PCTWALK	Per 1-unit increase in the percent who walk to work	3.1% less
RETAIL16	Where there are 16 or more retail establishments in the block group	34.4% less
AFFORDABLE	When property meets FNMA definition for affordable housing	61.9% less
FREEWAY1000FT	Where property is located within 1,000 feet of a freeway corridor	59.0% more
PROTECTED1MILE	Where property is located with 1 mile of protected open space	32.5% less