Default Risk of Securitized Commercial Mortgages: Do Sustainability Property Features Matter?*

Xudong An
Department of Finance
San Diego State University
xan@mail.sdsu.edu

Gary Pivo
School of Landscape Architecture and Planning
University of Arizona
gpivo@u.arizona.edu

March 30, 2015
Revised April 16, 2015

Abstract: This is the first study to empirically examine the relationship between building sustainability features and performance of corresponding commercial mortgages across property types. We examine 22,813 loans in the CMBS universe and more than 664,000 quarterly observations of loan performance. Two types of sustainability features are considered: smart growth locations that are walkable, transit-oriented, near green infrastructure, or avoid traffic-related air pollution; and green buildings, including LEED and Energy Star properties. We find that several sustainability features are strongly associated with lower default risk after controlling for standard risk factors. The mortgage default risk of buildings within a quarter mile of fixed-rail transit stations is reduced by 30.1 percent compared to other locations, properties with a Walk Score of 77 (out of 100) have a 13.5 percent lower default risk than properties with a Walk Score of 45, Energy Star properties are 20 percent less likely to default than others and properties near protected open space are 3 percent less likely to default, *ceteris paribus*. On the other hand, some sustainability features are associated with increased default risk, including jobs-worker balance (for office and retail) and locations that avoid traffic-related air pollution (for office and multifamily housing), suggesting that market signals can conflict with sustainability goals. The results for green infrastructure are ambiguous. We also find the default model fits the data better when sustainability variables are included. When examined by property type we find that walkability has the most consistent benefit to default risk across property types, followed by transit and energy efficiency. Our results will be of interest to academics, professionals, and policy makers because they give new insights into the geography of default risk, new ways for lenders and debt investors to predict and manage risk, and a better understanding of the financial dynamics that support and deter sustainable urbanism.

* Paper presented at the 2015 RERI Research Conference, April 30, 2015. We are grateful to the Real Estate Research Institute (RERI) for its financial support and sincerely appreciate the helpful comments from our RERI mentors Paige Mueller of RCLCO and Douglas Poutasse of Bentall Kennedy. We are also grateful to Trepp, Real Capital Analytics, the US Green Building Council, the US Geological Survey, and Redfin for providing invaluable data. Any errors and omissions are solely the authors' responsibility.

Default Risk of Securitized Commercial Mortgages: Do Sustainability Property Features Matter?

Xudong An and Gary Pivo

1. Introduction

Sustainability has become an increasingly important theme in real estate development and investment. In addition to being socially responsible, going "green" can be economically beneficial. Existing research has found that "green" buildings or buildings in "smart growth" locations generate rent and price premiums (see, e.g., Eichholtz, Kok and Quigley, 2010, 2013; Pivo and Fisher, 2010, 2011; Fuerst and MacAllister, 2011; among many others). Some studies have shown increased investment return or profit associated with sustainable development or investment (see, e.g., Miller, et al., 2008; Pivo and Fisher, 2010; Deng, Li and Quigley, 2011; Deng and Wu, 2014). Following this line of research, an interesting question is whether real estate loans associated with sustainable properties have reduced default risk. Little research has been done on that question.

This paper provides the first systematic analysis of the relationship between sustainability property features and corresponding commercial mortgage default risk with a national sample of all the major commercial property types, including office, retail, apartment, and industrial. All loans in our sample are mortgages in the Commercial Mortgage-Backed Securities (CMBS) universe. The study was made possible by merging public use datasets from various federal and non-profit sources as well as proprietary datasets from the US Green Building Council (USGBC), Redfin, and Trepp Inc.

We consider two major dimensions of sustainability in real estate. The first dimension is green building, including properties that are green certified or energy star labelled because of their eco-efficient design or operational performance. Most, though not all, green building features pertain to the design and operation of the building, as compared to its location. The second sustainability dimension we consider is smart growth location. While green building emphasizes building design, smart growth emphasizes sustainable urban form by promoting real estate investment in walkable, less auto-dependent infill locations with protected landscapes and natural areas. Smart growth fosters sustainability by preserving green space, promoting health, protecting air and water quality, and reducing carbon footprints (EPA 2013).

We hypothesize that green properties and properties built in smart growth locations have lower default risk because they produce better income and value. We also hypothesize that the financial benefits of sustainability are not fully reflected in traditional default predictors. A simple example is that contemporaneous loan-to-value ratio (LTV) is one of the most commonly used variables in a default risk model, however, the current LTV is usually derived from regional or national indices that only measure the price appreciation/depreciation of an average property within a certain geography. Therefore, the property specific sustainability "premium" is not captured by the current LTV in a conventional default risk model. There are other reasons such as mortgage underwriters' underestimation of the potential benefits of green building and smart location features, which could cause those sustainability features to carry risk information that is not provided by the traditional default predictors. Consequently, there is a degree of inaccuracy in models that use standard covariates which would be reduced by adding sustainability features, resulting in better predictions of default risk, at least until valuation methods fully capture sustainability premiums.

Therefore, what we do in this paper is add various measures of green building and smart growth location variables to a conventional commercial mortgage default risk model and test whether they are significant and whether they help improve model fit. Our results show that sustainability property features are associated with statistically and economically significant reductions in CMBS loan default risk, and that they provide additional information to the conventional default predictors. For example, commercial properties with Energy Star labels are 20 percent less likely to default than those without Energy Star labels, all else being equal. The mortgage default risk of buildings within a quarter mile reach to public transit is reduced by 30.1 percent compared to that of buildings farther than a quarter mile from public transit, *ceteris paribus*. Properties with a Walk Score that is one standard deviation higher than others (e.g., 77 vs. 45 in our sample on a scale of 100) has a default probability that is 13.5 percent lower. Overall, walkability has the most consistent benefit to default risk across property types, followed by transit and energy efficiency. All of these results are in a default probability model where conventional predictors such as original LTV, contemporaneous LTV and debt service coverage ratio (DSCR), current occupancy rate, refinance incentives, macroeconomic conditions, MSA-fixed effects, and more, are already included and held constant. Moreover, the model fit is significantly improved when sustainability measures are included.

We also find that the impacts of sustainability features can vary across property types. For example, walkability helps reduce default risk significantly for office, retail and multifamily loans; however, it increases the default risk of industrial loans. Proximity to public transit has similar impacts on the default risks of office, multifamily, and industrial loans, but it has little or no impact on retail loans.

Our results on smart location effect echo those of Pivo's (2013, 2014). For example, based on the Fannie Mae multifamily building portfolio, Pivo found that *ceteris paribus*, defaults were 58% less likely for loans on multifamily properties in less auto-dependent locations where 30% or more of the workers living there commuted to work by subway or elevated train. Pivo's work, however, did not look at properties with securitized mortgages, which carry a higher underlying default rate than do the Fannie Mae loans. Nor did it examine default risk across multiple time periods and property types. Our paper fills these gaps. In addition, we take the first look at the impact of green building certification on mortgage loan default.

Our results also reinforce the work of others who have shown the importance of geographical effects on default risk. Archer et al. (2000) and An et al. (2013) have shown that geographical effects, such as zip code location and local unemployment rates, are important for predicting mortgage default when modeled with LTV and/or debt service coverage ratios.

Many groups are interested in the economics of green building and urban form because of their link to important issues including global warming, obesity, poverty, and housing affordability. However, virtually all the evidence being discussed by these groups deals with the real estate equity side of the topic drawing from studies on whether sustainability features affect rents, values and equity returns (e.g., Miller et al. 2008; Eichholtz, Kok and Quigley, 2010, 2013; Pivo and Fisher, 2010, 2011; Fuerst and MacAllister, 2011). This paper contributes to our understanding of green building economics and smart growth from the debt side.

At a practical level, the paper makes several contributions to the business community. It improves our understanding of how loans on sustainable real estate perform and deepens our understanding of the geography of default risk. It could help investors improve their ability to predict and manage default risk and indicate for investors, policy makers and other stakeholders whether lenders could offer more liberal terms for sustainable properties without increasing their exposure to risk. It could also help investors identify underpriced CMBS securities and possibly support the creation of new CMBS products that target green labelled properties in smart growth locations.

The rest of the paper is organized as follows: in the next section, we explain our data and methodology; section 3 describes our empirical results; our conclusions are in a final section.

2. Data and Methodology

2.1. Loan and Sustainability Data

Our CMBS loan data is from Trepp. Trepp gathers CMBS loan information from monthly master servicers' reports. The format of the report is laid out in the CRE Finance Council's Investor Reporter Package (IRP) and provides an internally consistent set of data across all CMBS loans. The initial dataset we received from Trepp includes 10,847,994 monthly observations of loan performance information. The loan performance information includes the status of the loan such as prepaid, delinquent, foreclosed or current in each month. It also contains updated loan balance, DSCR, occupancy rate and loss information if the servicer reports such information[1].

The nearly 11 million loan performance records in our database are for 89,865 CMBS loans from 658 CMBS deals. All the loans are for single properties so each loan can be tied to a specific location for analysis of locational features. Trepp also has data on loans for multiple properties that we did not study.

We have specific information for each loan such as origination date, original balance, actual rate (mortgage note rate adjusted by points), rate index for adjustable-rate mortgages (ARMs), margin and cap if an ARM, maturity term, amortization period, interest-only periods, property

[1] The IRP requires borrowers to provide regular updates of the current NOI, occupancy rate, and DSCR for each property. However servicers have been unwilling to enforce such a rule, resulting in some missing values in updated occupancy rates and DSCR in our data.

type, property rentable area, property year built, location of underlying property (address, longitude and latitude), prepayment provisions, originator, servicers (both master and special servicers), the date the loan was securitized (deal cutoff date), face value at the time of securitization, and LTV, net operating income, and DSCR at securitization. The dataset is comparable to that used in An, et al. (2013).

We focus on fixed-rate mortgage loans and exclude ARMs. ARMs are less than 2 percent of the sample. Given that we have to use the Real Capital Analytics (RCA) by-MSA and by-property type commercial real estate price index to help calculate contemporaneous LTV and that the RCA index is available for only a limited number of metropolitan statistical areas (MSAs), we focus on CMBS loans from the RCA MSAs. The 17 MSAs that we will show later are actually the top MSAs in terms of CMBS loan origination. For the same reason, we focus on the four major property types: multifamily, retail, office, and industrial. We also exclude loans originated before the year of 2000 because the RCA price index only starts from 2000. Further, we verify loan information on rate, LTV, and original balance at origination and exclude a few loans with invalid information on those variables. This leaves us with a final sample of 22,813 loans, including about 2 million monthly observations of loan performance information. Appendix Table 1 lists these various filters that were applied and their effect on the final sample size.

Table 1 gives the loans in our final sample by origination year from 2000 to 2012. The number of loans grows from 465 in 2000 to 4,581 in 2006. Then it declines in 2007 and drops to almost zero during the recent financial crisis. We only have 19 loans in 2008 and 6 loans in 2009 in our sample. It finally recovers to a few hundred in 2011.

In Table 2, we show the geographic distribution of our CMBS sample by MSA. New York has the highest number of loans among all the MSAs in our sample (16% of the sample), followed by Los Angeles (14%). Dallas, Houston, Austin and San Antonio combined have nearly 16% of our sample loans. Other major MSAs include Philadelphia (6%), Washington DC (6%), Atlanta (5%), and Chicago (5%). Table 3 contains property type distribution information. Retail loans represent 38% of our sample, office 28%, multifamily 24%, and industrial 10%.

We follow the existing literature to define default as the first time a loan enters into 60-day delinquency. Table 4 shows that among 22,813 CMBS loans in our final sample, 2,949 loans (or

13%) entered into default during our study periods of 2000 to 2013. This is a rather high default rate, and apparently part of it is because of the recent financial and real estate market crisis.

Table 5 contains descriptive statistics of the loan characteristics such as original balance, actual mortgage interest rate, maturity term, amortization term, age of the property, and LTV, occupancy rate, and DSCR information at securitization.

Our data on sustainability features was gathered from several public and private sources. The sources and the assembly, cleaning and matching efforts are discussed later when our focus variables are described in detail.

2.2. Hazard Model of Mortgage Default Probability

We estimate a default probability model to assess the relationship between sustainability property features and the default risk of CMBS loans. Note that default probability and loss given default (LGD) are the two components of default risk, but LGD is outside the scope of this paper.

The default probability model we estimate is a standard Cox proportional hazard model, which is widely used in the mortgage literature (see, e.g. Vandell, 1993; Seslen and Wheaton, 2010; An, et al., 2013). In a Cox proportional hazard model, the left-hand side (LHS) of the equation is the hazard rate of default, which is the probability of default for a loan at a certain age/period after origination given that there has been no default before that age/period (i.e. the conditional default rate). The hazard rate of default for a particular loan at a certain age/period is modeled as a function of a baseline hazard function, which is a function of the duration (age) of the loan, and a function of covariates, which are the default risk factors. The hazard model is convenient mainly because it allows us to work with our full final sample of loans despite some observations being censored when we collect our data.

Assume the hazard rate of default of a mortgage loan at period T since its origination follows the form:

$$h_i\left(T; Z_i\left(t\right)\right) = h_0\left(T\right)\exp\left(Z_i\left(t\right)'\beta\right), i = 1, \cdots n \qquad (1)$$

Here $h_0(T)$ is the baseline hazard function, which only depends on the age (duration), T, of the loan and is an arbitrary function that allows for a flexible default pattern over time; $Z_i(t)$ is a vector of covariates for individual loan i that include all the identifiable time-varying or time-invariant risk factors. Examples of covariates include LTV and DSCR. In the Cox proportional hazard model, changes in covariates shift the hazard rate proportionally without otherwise affecting the duration pattern of default. In this study, the covariates include sustainability property features and control variables that have been identified by the existing literature as default probability drivers. We will discuss the focus (sustainability) variables and control variables in detail in the next subsection.

The hazard model is estimated with the Maximum Likelihood Estimation (MLE) method using the event-history data of loans, which is constructed based on the loan performance records.[2] Given that many of the covariates are only available at a quarterly frequency, we construct a quarterly event-history for each loan. In the event history data for model estimation, the covariates could be time-varying or time-invariant (i.e. constant). For example, the size of the loan is used as a time-constant variable, while contemporaneous LTV and DSCR and occupancy rate are time-varying covariates, meaning they can be different in each observed time period. To construct contemporaneous LTV, we utilize the RCA price index (by property type and by MSA) to bring property value up-to-date, and calculate contemporaneous LTV as the ratio between current property value and remaining loan balance that is in the Trepp data. For contemporaneous DSCR and occupancy, if there is a quarter that the servicer did not report the current DSCR or occupancy rate we use the nearest quarter's value either before or after the missing quarter, as a proxy.

[2] See Clapp, Deng and An (2006) for details about the MLE estimation of the hazard model.

2.3 Focus Variables and Control Variables

Our focus variables are the sustainability property features. Sustainability is a multi-dimensional construct with multiple distinct but related dimensions treated as a single theoretical concept. As noted above, we focus on two key dimensions of sustainable building – green building and smart growth location.

The green building measures we consider include the Energy Star label issued by the US EPA and the Leadership in Energy and Environmental Design (LEED) certification issued by the US Green Building Council (USGBC). Certification information was provided by USGBC, which maintains a comprehensive database on green buildings including LEED-certified, Energy Star labelled, and other green properties. For our study, USGBC used the property address and geocodes (latitudes and longitudes) found in both the USGBC and Trepp datasets to determine whether properties in our Trepp data matched properties in the USGBC green buildings database, and if so whether the matching properties were LEED-certified or Energy Star labeled and the date that status was achieved.[3] Many of the records in our initial loans dataset were missing address or geocode data, leaving us with 72,357 (81%) of the 89,865 CMBS loans properties that could be checked for LEED or Energy Star status before further filters were applied (see Appendix Table 1). The USGBC matching algorithm used a combination of geocodes, mathematical proximity, and address text matching.

EnergyStar is a dummy for whether or not a property was labelled as Energy Star by the US Environmental Protection Agency during the study period. To be labeled, a building must be in the top quartile of energy efficiency when compared to other properties with similar operational characteristics (i.e. size, weather conditions, number of occupants, number of computers, and hours of operation per week). Since labels must be obtained by the owner annually, we scored a property as labelled in the first year it was officially recognized and every subsequent year even if it did not have an official label in the subsequent years, assuming it was more likely that the absence of subsequent labels was due to management choosing not to apply for the label rather than failing to maintain energy efficiency. There is a need to better understand the degree to which the green label itself operates as a distinct driver from energy- and eco-efficiency in real

[3] USGBC matched the Trepp data to both their public records available thru their Green Building Information Gateway and confidential records that identify LEED properties that are not publicly disclosed.

estate, so this assumption should be subjected of further analysis in future research. In the event-history data, the Energy Star covariate takes a value of 0 before the building was Energy Star labeled and 1 for every loan period when and after it is labeled.

The LEED variable we considered is LEED certified at any level (Certified, Silver, Gold or Platinum). The LEED program has different certification levels including “Certified”, “Silver”, “Gold”, and “Platinum”, and labeling standards are substantially more complex than those required for an Energy Star label. Additional points in the certification process are awarded for such factors as “site selection,” “brownfield redevelopment,” and the availability of “bicycle storage and changing rooms,” as well as energy performance (Eichholtz, Kok and Quigley, 2010).

Table 6 gives the frequencies of green building variables in our study sample. Energy Star properties are more prevalent than LEED properties; however they both comprise a small subset (i.e. 1 to 3 percent) of the sample.

Our smart growth location variables include measures that reflect the degree to which properties are located in less auto-dependent locations, proximate to green infrastructure, or exposed to traffic-related air pollution along primary roads. Table 7 contains descriptive statistics for all of the smart growth location variables for the 22,813 loans in our final sample.

The degree to which locations were less auto-dependent was measured using three separate metrics: Walk Score, whether the property was within walking distance (¼ mile) of a fixed rail transit station, and the degree to which jobs and workers at their place of residence are balanced in the census block group where a property is located.

WalkScore rates the walkability of an address on a 100 point scale by determining the distance to educational, retail, food, recreational and entertainment destinations. Studies show it to be a reliable and valid estimator of neighborhood features linked to walking (Carr, Dunsiger, and Marcus, 2010, 2011; Duncan, Aldstadt, Whalen, and Melly, 2011; Duncan et al., 2013). It is also a better predictor of walking for non-work trips than other related indices (Manaugh and El-Geneidy, 2011). Redfin produced Walk Scores for each property by using the geocodes given by Trepp in their CBMS loan database to map each property and then produced a Walk Score for that particular location. As already noted, the Trepp dataset had geocodes for 81% of the 89,865 properties, so Walk Scores could be generated for 72,357 of the properties in our initial dataset. The average Walk Score in our final sample was 46 out of 100. Walkability has been linked to various social and environmental benefits and increases with the number of desired destinations within walking distance of a property (Pivo and Fisher, 2011; Federal Highway Administration, 2012).

Transit_Quart_Mile is a dummy for whether the property is within ¼ mile of a fixed rail transit station. Transit station locations were obtained from The TOD Database maintained by the Center for Neighborhood Technology. It includes every fixed guideway transit station in the U.S. (as of October 2011). Sixteen percent of the properties in our final sample were located within ¼ mile of a fixed rail transit station. A number of health, environmental, social, and financial benefits are associated with properties being within walking distance of fixed rail transit (EPA 2013).

Jobs_Worker_Bal is the Household Workers per Job Equilibrium Index from the US EPA Smart Location Database (SLD)[4]. The index ranges from 0 to 1, with 1 indicating a balance (i.e. equality) in the number of resident workers and jobs in the census block group (CBG).[5] The average value for this variable in our final sample was 0.38, indicating that the typical property was in a job or worker rich CBG (a low score can indicate an imbalance of either too many or too few jobs relative to resident workers). The number of workers and jobs for this variable were derived by the EPA from the 2010 Census, the 2006-2010 American Community Survey, and US Census Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics. Job-worker balance indicates that housing and jobs are close to each other so people can work close to where they live. It is thought to produce several sustainability co-benefits including enhanced accessibility, reduced congestion, improved air quality and higher quality of life (Frank and Pivo 1994, Cervero and Duncan 2006, Stoker and Ewing 2014). The metric we used does not account for whether jobs and workers are matched in terms of incomes, which can influence the degree to which balance in simple numerical terms reduces commuting (Stocker and Ewing 2014). The SLD metric also was measured at the CBG level, although researchers have found that it may be a more accurate predictor of vehicle miles traveled when it is measured

[4] It is labeled D2c_WrEmIx in the SLD.
[5] Jobs_Worker_Bal = exp(-|(Workers/TotEmp) -1|), where exp = the exponential function (e [approximately 2.7] raised to the power of the number in parenthesis).

on a larger spatial scale, such as at the census tract or commuter-shed level (Stocker and Ewing 2014). Despite these concerns, this variable was available for the entire nation and addresses an important concern in sustainable urban form so it is used here with these caveats to the reader.

Green infrastructure in the vicinity of each property was measured by two separate metrics. One captures the distance to protected open space and the second captures tree canopy cover in the vicinity of each property.

Protected_Area_Quart indicates whether a property was located within a quarter-mile of a Protected Area according to the U.S. Protected Area Database (PAD). The PAD includes public lands at all government levels held for conservation and voluntarily provided privately protected areas. Thirty percent of the properties in our final sample were within ¼ mile of an open space mapped in the PAD. Protected open space helps sustain resource-based industry, recreation, wildlife, watersheds, and other ecosystem services such as greenhouse gas absorption and heat island mitigation. Access to parks and recreation has also been linked to lower childhood obesity and other social benefits (Wolch et al., 2011).

BlkGpTreeCover_Median indicates the median percent tree cover for the block group where each property is located as determined by its geospatial coordinates. This came to 11.9 percent for the properties in our final sample. Our tree cover data are from the 2011 National Land Cover Database produced by the US Forest Service. Research indicates net positive benefits from urban forestry and greening (Nowak et al. 2010; Hirokawa, 2011; Roy et al. 2012; Colding and Barthel, 2013; Gomez-Baggethun and Barton, 2013). This includes benefits to air quality, urban heat island mitigation, water quality, wildlife, energy conservation, and social and financial outcomes. The financial benefits are summarized by Roy et al. (2012) in their review of 115 studies, which reports that the most common economic benefit is increased property values. McPherson et al. (2005) studied street trees and found increased property value was the single largest economic effect and Joye et al. (2012) reported that shoppers are willing to pay higher prices for retail goods and entertainment in shopping areas with more tree cover. As with *Jobs_Worker_Bal,* we have some concern about whether the block group is the best scale to assess this feature, as far as effect on default is concerned. A more precise measure of tree coverage on the street within the immediate vicinity of the property may have a stronger effect on default risk, and we are working on developing such a variable for future studies.

in_500_Ft_Hwy_Buff indicates whether a property is located within 500 feet of a primary road. Three percent of our final sample properties were in this position. Primary road boundaries were obtained from the National Transportation Atlas Database, maintained by US Department of Transportation, and include divided, limited-access highways. There is evidence that living or working close to a freeway increases risk for autism, cancers, and respiratory disease (Gauderman et al., 2007; Volk, Hertz-Picciotto, Delwiche, Lurmann, and McConnell, 2011; Office of Health Hazard Assessment, 2012; Cakmak, Mahmud, Grgicak-Mannion, and Dales, 2012; Boehmer et al. 2013). We recognize that this last variable also measures convenience of location, as it is somewhat correlated with freeway and highway access. That could produce an effect on risk that is opposite to what might be caused by air quality problems and indeed it did give us somewhat ambiguous results that are described below.

Except in the case of Walk Scores, which were produced by Redfin using address data, in order for us to link our smart growth location data to the properties in our sample, we first had to map the Trepp properties using geospatial software. We did this using the geocodes provided by Trepp, which as noted above, was possible for 72,357 (81%) of the 89,865 CMBS loans in our initial dataset. Additional layers were then added to the map including boundaries for 2010 census block groups, the location of fixed rail transit stations, protected areas and primary roads, and the percent tree cover by 30 meter pixel. All of these layers were for the entire nation and came from the various sources described above plus the US Census Bureau for the census block group (CBG) boundaries. Once the loan properties were mapped, quarter mile buffers were automatically drawn around each Trepp property, allowing us to determine whether the properties were located within ¼ mile of a transit station or protected area. Additionally, a 500 foot buffer was drawn along each primary road in order to identify the loan properties located within 500 feet of a primary road. Next, each property was assigned the appropriate census block group geo-ID number by merging the property locations and CBG boundary layers. This allowed us to link data to the loan properties from the EPA Smart Location Database, which uses the CBG as its unit of analysis. Finally, summary statistics for the tree cover data pixels in the CBG around each property were computed using a zonal statistics procedure and each property was assigned a median tree cover value for the pixels in its CBG (the median being the better measure of central tendency given the distribution of the tree cover data in the CBGs).

The control variables we include in our model are essentially all of those identified by the existing literature as significant drivers of commercial mortgage default probability or their equivalents. These include the contemporaneous LTV and DSCR and occupancy rate we mentioned earlier, original loan balance (in log terms), original LTV (dummy for original LTV higher than 75%), refinance incentive (measured by percentage decline in market interest rate relative to the current note rate), age of the property (which also proxies to some degree for building class; for a discussion see Pivo and Fisher, 2010), prepayment restrictions (i.e. the presence of a prepayment lock out, prepayment penalty and yield maintenance in a particular loan quarter, which tend to limit refinancing and increase default risk), MSA unemployment rate innovation in a particular quarter (i.e. the change in the MSA unemployment rate over the prior four-quarter moving average, as a business cycle indicator), yield slope and corporate bond credit spread (as measures of macroeconomic conditions), and property type-fixed and MSA-fixed effects. Also, since mortgage default can be explained as the borrower's exercise of a put option, which increases with volatility, we include quarterly volatility of the 10-year Treasury rate and volatility of the RCA price index for each of the 17 MSAs and property types over the prior 12 quarters. For details on these control variables, please see An, et al. (2013) and An, Deng and Gabriel (2014).[6]

In Table 8, we report the sample statistics of our event history data. Each observation represents a loan record for a specific quarter. All continuous variables are standardized to zero mean and a standard deviation of 1 so the hazard ratios produced by the model for the continuous variables can be interpreted as the effect of one standard deviation change in the continuous variable on default risk.

Before proceeding to the modeling stage, we checked for potential collinearity issues among the focus and control variables. The only high correlations were between LEED and Energy Star. Due to the high correlation between these two variables we only include *EnergyStar* in the final model, though we do report some results that were produced when LEED was substituted for Energy Star in one of the models. *WalkScore* and *Transit_Quart_Mile* were strongly correlated (r = .53) but this did not create collinearity issues when they were used in the models.

[6] We also tried to include vintage-fixed effect and CMBS deal type-fixed effect (e.g. conduit, fusion, etc.). None are significant.

The original set of smart growth location variables available for all CBGs from the EPA Smart Location Database (SLD) is far greater in number than the single one we included in the final model and discussed above (*Jobs_Worker_Bal*). We do not use more of them in our final models due to their high correlations with the variables we include. The SLD includes street design variables that are thought to be correlated with walkability (e.g., pedestrian street network density and street intersection density), but they have strong or very strong positive correlations with Walk Score. SLD transit measures (e.g., aggregate frequency of transit service per square mile), demographic measures (e.g., percent of families with no cars), and destination accessibility measures (e.g., jobs or working age population within 45 minute travel time by car or transit), all of which are related to auto-dependence, are strongly correlated with *Transit_Quart_Mile*. SLD land use diversity measures, (e.g. employment and housing entropy), which also affect auto dependence, are strongly correlated with *Jobs_Worker_Bal.* It would have been feasible to use factor analysis to reduce these correlated measures into a smaller set of underlying factors, but we decided that using factors scores would make the model results difficult to interpret without adding much explanatory power. We expect, however, that several of the SLD variables which are not used in our models have similar effects to those that are used because of the correlations among them.

3. Results

Before presenting the hazard model results, let us first discuss some bivariate and correlation analyses that we conducted.

Table 9 compares the default rate of Energy Star/LEED labeled and that of non-Energy Star/LEED labeled properties. Here we count a property as labeled if during any time during our study period the property has an Energy Star label or LEED certificate. We see that of the 22,813 properties, Energy Star labeled properties have a default rate of 9.28% while non-Energy Star properties have a higher default rate of 13.05%. Similarly, LEED certified properties have a default rate of 5.58%, in contrast to the higher default rate of 13.01% for non-LEED certified properties. Interestingly, Table 10 shows that Energy Star labeled properties are associated with higher average DSCR on their CMBS loans – the average DSCR for Energy Star labeled property loans is 1.82, while the average DSCR for non-Energy Star labeled property loans is just 1.61. The DSCR comparison between LEED-certified properties and non-LEED certified properties reveals nearly the same difference. Given that DSCR is found by the existing literature to be a critical driver of commercial mortgage default, we suspect that part of the default risk difference in green vs. non-green property loans shown in Table 9 is through the impact of green building features on the property's operating performance and DSCR. However, we also notice there is no significant difference between green and non-green buildings in terms of occupancy rate. The financial benefits of green building may not be fully reflected by the conventional default risk factors such as occupancy rate but rather thru higher rents, lower utility bills, and lower cap rates (Pivo and Fisher 2011).

Table 11 contains the correlation results of *WalkScore* with average contemporaneous DSCR and occupancy rate. We see that *WalkScore* is positively correlated with both average DSCR and average occupancy rate. We expect this would mean lower default risk for properties with higher Walk Scores, consistent with prior research on the topic (Pivo 2014).

In Table 12, we compare the mean values of the smart growth location variables for defaulted and non-defaulted loans. For some smart growth variables, this bivariate analysis conveys the same message that we just discussed. For example, the average *WalkScore* of non-defaulted loans is about 48 (on a scale of 100) while the average *WalkScore* of defaulted loans is only about 37 (i.e. less walkable). The percentage of properties within a quartile mile of public transit is about 17% for non-defaulted loans, but only 8% for defaulted loans. The average *Protected_Area_Quart* is 0.31 for non-defaulted loans (i.e. 31% of the properties with non-defaulted loans are near protected open space) while it is only 0.25 for defaulted loans. However, for three other smart growth variables there is either *no difference* (as in the case of *Jobs_Worker_Bal)* or the non-defaulted loans are associated with *less sustainable* conditions (in the cases of *BlkGpTreeCover_Median* and *in_500_Ft_Hwy_Buff)*. These results suggest that sustainability variables can be positive, neutral, or negative for default risk, depending on which variables are examined.

Our main hazard model results are given in Table 13. For each parameter, the four statistics given are the regression coefficient and significance values, its standard error, and the hazard ratio (also known as relative risk). These results can be analyzed in a way similar to how we would analyze a multiple logistic regression model. The estimated coefficient gives the strength of the association with default risk, holding other variables constant. A positive coefficient indicates the variable is positively related to risk. It is the natural log of the hazard ratio. The hazard ratio is the predicted change in the risk of default produced by a change in the parameter while other variables are held constant. A hazard ratio greater than 1 indicates that the default risk increases when the variable increases and a hazard ratio less than 1 indicates that risk decreases when the variable increases. For the binary covariates the hazard ratio estimates the ratio of the risk of default in the loans with the feature to the default risk in the loans without the feature. For the continuous covariates, because they are standardized, the hazard ratio estimates the change in risk associated with a change of one standard deviation in the variable.

In column A, we report estimates of the baseline model, where only the control variables but not the sustainability variables are included. Our results are highly consistent with those found in the existing literature. Contemporaneous LTV (*curltv*) is highly significant and positively related to default probability – the higher the current LTV, the more likely the loan will be defaulted. Contemporaneous DSCR (*dscrnoi2*) and occupancy rate (*occrate2*) are also significant and negatively related to default probability – the higher the DSCR or occupancy, the lower the chance is that the loan will be defaulted. The refinance incentive (*refi_inc1*) is significant and negatively related to default probability, which is consistent with the competing risks argument – the more likely the loan is going to be refinanced, the less likely it is going to be defaulted (see, e.g., Deng, Quigley and Van Order, 2000). The high LTV (at origination) loans (*ltv75*) have higher probability of default, all else equal. Prepayment restrictions (*prep*), either lock out, or prepayment penalty, or yield maintenance (*yldm*) cause higher default risk. This is again consistent with findings in the existing literature (see, e.g., An, 2009; An, et al., 2013). MSA unemployment rate innovation (*unemp_msa_rla*) is positive and significant. This means when the local economy is bad, the chance of these CMBS loans being in default is higher, which makes perfect sense. Coefficients of the macroeconomic variables such as yield slope (*yldslope*) and the volatility of the Treasury rate (*tcm10ystd*) are both significant. There are significant property type-fixed effects (*cssaproptype*), e.g., comparing to retail loans, office and industrial loans have higher default risk, *ceteris paribus*. There are also significant MSA-fixed effects (*MSA-fixed effect*). The only surprise comes from RCA price index volatility (*rcacppi_vol*). Its coefficient is negative and significant.

In column B of Table 13, we report estimates of the new model offered by this paper, where we include the sustainability variables in addition to all the controls just discussed. The coefficient of *EnergyStar* is significant and negative, meaning that CMBS loans for properties that are Energy Star labeled have lower default risk than those for properties without the Energy Star label, *ceteris paribus*.[7] *WalkScore* is significant and negative, which means the more walkable the property, the less likely the CMBS loan is to be defaulted, everything else being equal. The coefficient of *Transit_Quart_Mile* and *Protected_Area_Quart* are also significant and negative, meaning that close proximity to public transit or protected areas (open space, parks, etc.) helps reduce default risk.

What is concerning is that *in_500_Ft_Hwy_Buff* is significantly and negatively associated with default probability. If we see *in_500_Ft_Hwy_Buff* as a variable for pollution, this result is the opposite of what we would expect if pollution was associated with greater default risk. A possible explanation is that this variable also measures highway access. It may be acting more like a traditional location variable measuring locational convenience; or at least the benefits of convenience are more fully internalized and outweighing the negative effects from air pollution, in terms of financial consequence.

The coefficient of *Jobs_Worker_Bal* is positive and significant (i.e. it increases default risk). In this sense, job-worker balance seems to work against commercial uses. This is plausible to us because we would expect commercial properties to do well in highly accessibility commercial centers where there is a high number of jobs relative to housing or in worker-rich bedroom communities where there is more affordable housing to support back-office and industrial workers and a large supply of residents to frequent shopping centers. This is a case where a more sustainable urban form associated with less vehicle miles of travel appears to be at odds with real estate market forces.

Finally, we find no significant effect from *BlkGpTreeCover_Median*, although the sign is in the expected direction. This could be because we measured tree cover at the block group scale instead of along the street in the immediate vicinity of each property, as we mentioned above.

[7] Given the high correlation between LEED and Energy Star, we used just Energy Star in the model. However, *LEED_any* (i.e. certified at any level) works similarly to EnergyStar. In the all property types model, LEED certified properties are about 30% less likely to default.

However, we did test tree cover within a quarter mile and 500 feet of each property and both were also insignificant. Tree cover may also be a proxy for suburban location because it has a moderate or modest negative relationship with variables that tend to be associated with urban locations such as *WalkScore* (r = -.375, p = .000) and *Transit_Quart_Mile* (r = -.181, p = .000). To further explore this issue, we added *Principle_City*[8] as a control for whether the property is in an urban or suburban location. Principle_City is correlated with *WalkScore* and *Transit_Quart_Mile* (r = .179, p = .000 and r = -.147, p = .000 respectively) and does not make *BlkGpTreeCover_Median* significant when it is added to the model. So the finding of no significant effect for *BlkGpTreeCover_Median* is likely the result of the measurement issue we have raised, the fact that tree cover does not affect default risk, or the possibility that its effects only occur above a certain threshold or cut-point we have not yet discovered. We suspect it is more likely a measurement or threshold problem because the research we cited above shows that urban tree cover is associated with financial benefits in commercial real estate.

To translate the hazard model coefficients into default probability differences we look at the hazard ratios given in Table 13. For example, commercial properties with Energy Star labels are 20 percent less likely to default than those without Energy Star labels, all else being equal (i.e. 1-.801 = 0.20). The default risk of buildings within a quarter mile reach to public transit is reduced by 30.1 percent compared to that of buildings farther than a quarter mile to public transit, *ceteris paribus* (i.e. 1-.699 = .301). Properties with a Walk Score that is one standard deviation higher than others (e.g., 77 vs. 45 in our sample on a scale of 100) has a default probability that is 13.5 percent lower (i.e. 1-.865 = .135). And properties within one quarter mile of a protected area are 3 percent less likely to default than those in other locations (i.e. 1-.970 = .03). Meanwhile, properties within 500 feet of a freeway are 14.8% *less* likely to default (1-.852 = .148) and properties in a CBG that is one standard deviation more balanced in terms of jobs and workers (e.g., .67 vs. the CBG mean of .38 on a scale of 0 to 1) are 3.4% *more* likely to default.

Notice that the significance of all aforementioned sustainability variables are in addition to the control variables we discussed including LTV and DSCR. This means that the benefits of green

[8] Defined by the US Census as "the largest incorporated place with a population of at least 10,000 in a core based statistical area (CBSA) or New England city and town area (NECTA), or if no incorporated place of least 10,000 population is present in the CBSA or NECTA, the largest incorporated place or census designated place (CDP) in the CBSA or NECTA. Additional places that meet specific criteria are also identified as principal cities."

building and smart location to CMBS loan default are not just through better LTV and DSCR measured in the model, as was suggested by the bivariate and correlation analyses we presented above. The sustainability variables are additional predictors of the CMBS loan default. We also notice that the inclusion of the sustainability variables changes the coefficients of a few control variables from insignificant to significant (comparing the baseline model and the current model). For example, log loan balance (*logbal*) and age (*age*) of the property are now significant, which is what one would expect in a default probability model. Moreover, the difference between multifamily and retail (the reference case) becomes significant, which is also what one would expect. Our explanation for this is that the sustainability variables are significant default risk predictors, and thus leaving out those variables would cause omitted variable bias (in the baseline model). We also notice that judged by the Akaike Information Criterion (AIC) and the -2 log likelihood (-2LogL), the model fit is improved significantly after we include the sustainability variables (lower AIC and -2LogL indicate better fit).

Also notice that in Table 13 we are not separating the effects of the sustainability variables by property type, so the coefficients we see in column B of Table 13 are average effects across all four property types. It is reasonable to hypothesize, however, that the impacts of sustainability features are different across property types. For example, walkability may be a benefit to multifamily properties but maybe not to industrial properties. This is exactly what we find in Table 14, where we interact the sustainability variables with property types. Here retail loans are the reference group.

In Table 14 we see that *WalkScore* is negatively associated with the default risk of retail, multifamily and office loans[9], however, it is positively associated with the default risk of industrial loans (coefficient of 0.24-0.125=0.115). Perhaps this is because it is more difficulty for trucks to maneuver in more walkable environments (Pivo *et al.*, 2002).

Proximity to public transit has similar impacts on the default risk of office, multifamily, and industrial loans, all being strongly negative as shown by the coefficients of *Transit_Quart_Mile*. However, the impact of proximity to public transit on retail loans is much smaller than on the

[9] Effects of sustainability feature on a property type are determined by adding the coefficient for the reference group (retail), as indicated by the coefficient estimates for the original terms, and the coefficient for the interaction term. For example, the effect of *WalkScore* on industrial is 0.24-0.125 = .115 while the effect of *WalkScore* on multifamily is -0.029-0.125 = -0.154.

aforementioned three property type loans and is marginally significant. These results are reasonable given that car-oriented shopping centers in the U.S. are probably just as successful as transit-oriented ones, however it is still remarkable to find that transit oriented shopping centers have less risk of default given the expectation that shoppers prefer to use a car for shopping trips.

There is also large variation in the coefficients of *in_500_Ft_Hwy_Buff* across the various property types. Being close to a major road has no significant impact on the default risk of retail loans and reduces default risk for office loans. But the strongest negative impact is on multifamily loans. This is a disturbing result because the public health issues caused by living near a freeway may be more acute than those caused by working or shopping there, yet the market seems to give the greatest reward to multifamily housing along freeways (in terms of lower default risk), perhaps because housing near a free is more affordable than in other locations precisely because of the associated disamenities.

Proximity to protected areas has no impact on retail and marginally increases default risk for offices. On the other hand it significantly lowers default risk for industrial and multifamily loans, as shown by the coefficients of *Protected_Area_Quart*. We suspect that multifamily properties may indeed be benefiting from the recreational and aesthetic amenities associated with being close to protected areas while the industrial properties are benefiting from the lower land prices, rents, and congestion, assuming the industrial properties near open spaces tend to be associated with suburban and urban fringe locations.

Tree cover significantly increases loan default risk for industrial, office and multifamily. However it significantly decreases risk for retail loan default, consistent with the prior research cited above on how urban forests benefit retail districts.

Job-worker balance significantly increases default risk for retail loans and has no significant effect on other property types. This suggests that the positive effect found in the all property types model was mainly caused by the effect on retail uses, suggesting that retail does better when proximate to concentrations of jobs or housing as opposed to more balanced locations.

Finally, the impact of Energy Star label is negatively associated with the default risk for both retail loans and office loans (since there is no multifamily Energy Star program, there are no Energy Star labeled multifamily properties in our sample).

We next run the hazard models separately for the different property types. This relaxes the assumption that different property type loans share the same coefficients of the control variables. Results are shown in Table 15. Compared to Table 14, we notice some changes in the property type results. Job-worker balance is now shown to have a significant positive impact on office and industrial loan default rather than having no effect, the effect of transit on retail has moved from barely significant to insignificant, and protected areas now flips to a barely significant negative effect on office default (instead of barely positive), while their effect on multifamily default risk has moved from negative to insignificant. Other results are consistent with what we discussed in the interaction model. Overall, the number of significant sustainability variables is the least in retail and multifamily loans, and walkability has the most consistent benefit to default risk across property types, followed by transit and energy efficiency. Meanwhile, jobs-worker balance appears to increase default risk and adjacency to a freeway tends to reduce default risk, indicating in both cases that sustainability and lower default risk do not always coincide. Finally, tree cover and proximity to open space, as indicators of green infrastructure, produce ambiguous results that vary depending on the model used (Appendix Table 2 summarizes the various results for the two green infrastructure variables).

4. Conclusions

We have demonstrated that certain sustainability features including building energy efficiency, walkability, and proximity to fixed rail transit significantly reduce default risk in CMBS loans. Both green building and smart location factors have been shown to be important. We have also demonstrated that adding these variables to a standard default risk model improves the accuracy of the model.

Our explanation for these results is that certain sustainability features can affect mortgage default risk through their impact on income and value and that those benefits may not be fully reflected in the conventional variables used in default risk models. In that case, adding sustainability features to the standard covariates improves the model and produces significant results. As long as the income and value estimates used to produce the standard covariates exclude the benefits of sustainability, it will remain important to include sustainability features in default risk models.

Some sustainability features did not have the expected effects. The results for green infrastructure were ambiguous, varying depending on the model used. Moreover, some sustainability features were associated with more default risk. Freeway adjacency lowers loan default risk for all property types except retail and jobs-worker balance increases default risk for office, industrial, and retail properties. Apparently, the desire to avoid default risk runs contrary to the goal of avoiding traffic related air pollution and reducing vehicle miles traveled thru greater jobs-worker balance. These results suggest that any discussion of sustainability and default risk needs to specify the property types and features under consideration because there are many sustainability issues and strategies for addressing them and not all of them have the same impact on real estate markets.

Of course, as with any first study of its kind, this project leaves room for much more research. There are other sustainability issues that should be examined, such as building water use efficiency and historic preservation, and other metrics for the issues studied here that should also be examined, such as different metrics for green infrastructure and urban forests. There may be additional controls that should also be considered, such as the characteristics of borrowers or submarket supply/demand trends. We would also be interested in looking at these relationships in other types of loan pools, such as loans held by lending institutions that are not sold into the CMBS market. And we would like to know more about how sustainability features affect origination terms and loan pricing in securitization. Finally, we want to check for nonlinearity and significant cutpoints in the observed relationships. It may well be that Walk Score or tree cover become significant above certain threshold levels and are insignificant below them.

In the meantime, the strength of these findings along with their consistency with prior work on related topics (Pivo 2013, 2014) raises the question of how underwriting tools and practices should be amended to take advantage of these results and whether lenders could and should offer more attractive terms to properties with certain sustainability features. After all, better models would help lenders better manage risk and better terms on sustainable properties could improve overall market efficiency and environmental outcomes without exposing lenders to greater risk.

References

An, X.. Macroeconomic Growth, Real Estate Market Conditions, and the Time Series Dynamics of CMBS Loan Default Risk. SSRN working paper, 2009.

An, X., Y. Deng and S. Gabriel. Pulling the Trigger: Default Option Exercise over the Business Cycle. SSRN working paper, 2014.

An, X., Y. Deng, J. B. Nichols, and A.B. Sanders. Local Traits and Securitized Commercial Mortgage Default. *Journal of Real Estate Finance and Economics,* 2013, 47:4, 787-813.

Boehmer, T.K., S.L. Foster, J.R. Henry, E.L. Woghiren-Akinnifesi and F.Y. Yip. Residential Proximity to Major Highways – United States, 2010, *Morbidity and Mortality Weekly Report* (MMWR), 2013, 62(03), 46-50.

Cakmak, S., M. Mahmud, A. Grgicak-Mannion, and R.E. Dales. The Influence of Neighborhood Traffic Density on the Respiratory Health of Elementary Schoolchildren. *Environment International*, 2012, 39:1, 128–32.

Carr, L.J., S.I. Dunsiger, and B.H. Marcus. Walk Score As a Global Estimate of Neighborhood Walkability. American Journal of Preventative Medicine, 2010, 39:5, 460–63.

——. Validation of Walk Score for Estimating Access to Walkable Amenities. *British Journal of Sports Medicine*, 2011, 45, 1144–48.

Cervero, R., & M. Duncan. Which reduces vehicle travel more: Jobs-housing balance or retail-housing mixing? *Journal of the American Planning Association*, 2006, 72, 475–490.

Clapp, J. M., Y. Deng and X. An. Unobserved Heterogeneity in Models of Competing Mortgage Termination Risks. *Real Estate Economics*, 2006, 34(2): 243-273.

Colding J and S. Barthel. The potential of 'Urban Green Commons' in the resilience building of cities. *Ecological Economics*, 2013, 86, 156-66.

Deng, Y., J. M. Quigley and R. Van Order. Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. *Econometrica,* 2000, 68(2): 275-307.

Deng, Y., Z. Li and J. Quigley. Economic Returns to Energy-Efficient Investments in the Housing Market: Evidence from Singapore. *Regional Science and Urban Economics*, 2012, 42(3): 506-515.

Duncan, D.T., J. Aldstadt, J. Whalen, S.J. Melly, and S.L. Gortmaker. Validation of Walk Score for Estimating Neighborhood Walkability: An Analysis of Four U.S. Metropolitan Areas. *International Journal of Environmental Research and Public Health*, 2011, 8, 4160–79.

Eichholtz, P., N. Kok and J. Quigley. Doing Well by Doing Good: Green Office Buildings. *American Economic Review,* 2010, 100(6): 2494–2511.

Eichholtz, P., N. Kok and J. Quigley. The Economics of Green Building. *Review of Economic and Statistics,* 2013, 95(1): 50-63.

EPA. *About Smart Growth*, http://www.epa.gov/dced/about_sg.htm, 2013.

EPA. *Our Built and Natural Environments: A Technical Review of the Interactions Between Land Use, Transportation, and Environmental Quality (2nd Edition)*, 2013.

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

Frank, L., & G. Pivo. Relationships between land use and travel behavior in the Puget Sound region. Seattle: Washington State Transportation Center, 1994.

Fuerst, F. and P. McAllister. Green Noise or Green Value? Measuring the Effects of Environmental Certification on Office Values. *Real Estate Economics*, 2011, 39(1): 45-69.

Gauderman, W.J., H. Vora, R. McConnell, K. Berhane, F. Filliland, D. Thomas, F. Lurmann, E. Avol, N. Kunzli, M. Jerrett, and J. Peters. Effect of Exposure to Traffic on Lung Development from 10 to 18 Years of Age: A Cohort Study. *The Lancet*, 2007, 369:9561, 571–77.

Gomez-Baggethun, E. and D.N. Barton. Classifying and valuing ecosystem services for urban planning, *Ecological Economics*, 2013, 86, 235-245.

Hirokawa, K.H. Sustainability and the Urban Forest: An Ecosystem Services Perspective. *Natural Resources Journal*, 2011, 51, 233-259.

Joye Y., K. Willems, M. Brengman, and K. Wolf. The effect of urban retail greenery on consumer experience: Reviewing the evidence from a restorative perspective. *Urban Forestry & Greening*, 2012, 9, 57-64.

Kau, J. B. and D. C. Keenan. An Overview of the Option-Theoretic Pricing of Mortgages. *Journal of Housing Research*, 1995, 6(2): 217-244.

Manaugh, K. and A. El-Geneidy. Validating Walkability Indices: How Do Different Households Respond to the Walkability of Their Neighborhood? *Transportation Research Part D*, 2011, 16, 309–15.

McPherson G., J.R. Simpson, P.J. Peper, S.E. Maco, and Q. Xiao. Municipal Forest Benefits and Costs in Five US Cities, *Journal of Forestry*, 2005, December, pp. 411-416.

Miller N., J. Spivey and A. Florance. Does Green Pay Off? *Journal of Real Estate Portfolio Management*, 2008, 12(1): 1-12.

Nowak D.J., S.M. Stein, P.B. Randler, E.J. Greenfield, S.J. Comas, M.A. Carr and R.J. Alig. *Sustaining America's Urban Trees and Forests*, GTR NRS-62, USDA, Forest Service, 2010.

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

Pivo, G. The Effect of Sustainability Features on Mortgage Default Prediction and Risk in Multifamily Rental Housing. *Journal of Sustainable Real Estate*, 2013, 5(1): 152-173.

Pivo, G. Walk Score: The Significance of 8 and 80 for Mortgage Default Risk in Multifamily Properties. *Journal of Sustainable Real Estate* 2014, 6(1), 187-210.

Pivo, G., Carlson D, Kitchen M and Billen D, Learning from Truckers: Truck Drivers' Views on the Planning and Design or Urban and Suburban Centers, *Journal of Architectural and Planning Research*, 2002, 19(1): 12-29.

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

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

Roy S., J. Byrne and C. Pickering. A systematic quantitative review of urban tree benefits, costs, and assessment methods across cities in different climatic zones. *Urban Forestry & Greening*, 2012, 11, 351-363.

Seslen, T. and W. Wheaton. Contemporaneous Loan Stress and Termination Risk in the CMBS Pool: How "Ruthless" is Default? *Real Estate Economics*, 2010, 38(2): 225-255.

Stoker, P. and R. Ewing. Job-Worker Balance and Income Matchin the United States, *Housing Policy Debate*, 2014, 24(2), 485-497.

Vandell, K. D. Handing Over the Keys: a Perspective on Mortgage Default Research. *Real Estate Economics*, 1993, 21(3): 211-246.

Volk, H.E., I. Hertz-Picciotto, L. Delwiche, F. Lurmann, and R. McConnell. Residential Proximity to Freeways and Autism in the CHARGE Study. *Environmental Health Perspectives*, 2011, 119:6, 873–77.

Wolch, J., M. Jerrett, K. Reynolds, R. McConnell, R. Chang, N. Dahmann, F. Filliland, J.G. Su, and K. Berhane. Childhood Obesity and Proximity to Urban Parks and Recreational Resources: A Longitudinal Cohort Study. *Health and Place*, 2011, 17:1, 207–14.

Yongheng D. and J.Wu. Economic Returns to Residential Green Building Investment: The Developers' Perspective. *Regional Science and Urban Economics*, 2014, 47: 35-44.

Table 1 Our CMBS Loan Sample by Origination Year

Origination year	Frequency	Percent	Cum. Pct.
2000	465	2.04	2.04
2001	1,726	7.57	9.6
2002	1,521	6.67	16.27
2003	2,778	12.18	28.45
2004	3,030	13.28	41.73
2005	4,378	19.19	60.92
2006	4,581	20.08	81
2007	3,815	16.72	97.72
2008	19	0.08	97.81
2009	6	0.03	97.83
2010	120	0.53	98.36
2011	292	1.28	99.64
2012	82	0.36	100
Total	22,813		

Note: These are all fixed-rate mortgage loans in the 17 MSAs listed in Table 2. Only loans for the four major property types are included.

Table 2 Geographic Distribution of our CMBS Loan Sample

MSA	Frequency	Percent	Cum. Pct.
Atlanta	1,219	5.34	5.34
Boston	570	2.5	7.84
Chicago	1,184	5.19	13.03
Washington DC	1,288	5.65	18.68
Denver	530	2.32	21
Riverside, San Bernardino-Ontario	713	3.13	24.13
Las Vegas	786	3.45	27.57
Los Angeles	3,292	14.43	42
Detroit	1,056	4.63	46.63
New York	3,718	16.3	62.93
Philadelphia	1,436	6.29	69.22
Phoenix	1,121	4.91	74.14
San Diego	711	3.12	77.25
Seattle	659	2.89	80.14
San Jose	297	1.3	81.44
San Francisco	677	2.97	84.41
Dallas-Houston-Austin-San Antonio	3,556	15.59	100
Total	22,813		

Table 3 Our CMBS Loan Sample by Property Type

Property type	Frequency	Percent	Cum. Pct.
Industrial	2,210	9.69	9.69
Multifamily	5,553	24.34	34.03
Office	6,303	27.63	61.66
Retail	8,747	38.34	100
Total	22,813		

Table 4 Performance of our CMBS Loan Sample

Default	Frequency	Percent	Cumulative Percent
0	19,864	87.07	87.07
1	2,949	12.93	100

Table 5 Descriptive Statistics of the Loan Characteristics

Variable	Mean	Std. Dev.	Minimum	Maximum
Original loan balance	14,316,155	31,596,708	100,000	806,000,000
Actual rate (%)	6.00	0.91	0.00	15.00
LTV at securitization (%)	65.64	16.13	0.42	95.80
DSCR at securitization	2.25	3.45	0.97	136.00
Occupancy rate at securitization (%)	95.81	6.13	2.30	100.00
Loan term (months)	116	25	35	360
Amortization term (months)	317	104	35	999
Age of the property (years)	34	24	2	113
Property rentable area (sqft)	223,718	12,793,857	25	1,690,000,000
Total number of loans	22,813			

Table 6 Frequencies of Green Building Variables

	Frequency	Percent	Cumulative Percent
LEED certified (any grade)			
0	22544	98.82	98.82
1	269	1.18	100
Energy Star label			
0	22048	96.65	96.65
1	765	3.35	100
Total number of loans	22,813		

Table 7 Descriptive Statistics of the Smart Growth Location Variables

Variable	Mean	Std. Dev.	Minimum	Maximum
WalkScore	46.29	32.60	0	100
Transit_Quart_Mile	0.16	0.36	0	1
in_500_Ft_Hwy_Buff	0.03	0.17	0	1
Protected_Area_Quart_Mi	0.30	0.46	0	1
BlkGpTreeCover_Median	11.93	20.82	0	97
Jobs_Worker_Bal	0.38	0.29	0	1
Total number of loans	22,813			

Table 8 Descriptive Statistics of the Event History Data

Variable	Mean	Std. Dev.	Minimum	Maximum
logbal	0.00	1.00	-3.72	4.27
ltv75	0.28	0.45	0.00	1.00
curltv	0.00	1.00	-2.57	7.98
dscrnoi2	0.00	1.00	-2.92	6.07
occrate2	0.00	1.00	-9.67	0.63
refi_inc1	0.00	1.00	-4.14	1.84
age	0.00	1.00	-1.42	3.60
lock	0.66	0.48	0.00	1.00
prep	0.00	0.06	0.00	1.00
yldm	0.10	0.30	0.00	1.00
WalkScore	0.00	1.00	-1.46	1.76
Transit_Quart_Mile	0.14	0.35	0.00	1.00
in_500_Ft_Hwy_Buff	0.03	0.17	0.00	1.00
Protected_Area_Quart_Mi	0.32	0.46	0.00	1.00
BlkGpTreeCover_Median	0.00	1.00	-0.59	4.25
Jobs_Worker_Bal	0.00	1.00	-1.34	2.19
estar_labl	0.02	0.13	0.00	1.00
unemp_msa_rla	0.00	1.00	-1.49	4.50
yldslope	0.00	1.00	-2.04	1.37
tcm10ystd	0.00	1.00	-1.70	2.00
rcacppi_vol	0.00	1.00	-1.44	6.31
crdspread	0.00	1.00	-1.02	4.03
Number of obs. (loan-quarter)	664,794			

Table 9 Green Building and Default

	Default		
	0	1	Total
Non-LEED certified	86.99	13.01	
LEED certified	94.42	5.58	
Non-Energy Star labeled	86.95	13.05	
Energy Star labeled	90.72	9.28	
Total number of loans	19,864	2,949	22,813

Table 10 Comparison of Average DSCR/Occupancy Rate between Green and Non-green Buildings

	N Obs	Variable	Mean	Std. Dev.	Minimum	Maximum
LEED						
0	22544	Average DSCR	1.61	0.54	0.09	5.00
		Average occupancy	93.41	7.32	9.09	100.00
1	269	Average DSCR	1.87	0.69	0.66	4.00
		Average occupancy	92.06	7.33	66.83	100.00
Energy Star						
0	22048	Average DSCR	1.61	0.53	0.09	5.00
		Average occupancy	93.43	7.32	9.09	100.00
1	765	Average DSCR	1.82	0.68	0.66	4.86
		Average occupancy	92.38	7.21	56.24	100.00

Table 11 Correlations Among Walk Score, Average DSCR and Average Occupancy Rate

	Walk Score	Average DSCR	Average Occupancy rate
Walk Score	1.00	0.13	0.08
		<.0001	<.0001
Average DSCR	0.13	1.00	0.18
	<.0001		<.0001
Average Occupancy rate	0.08	0.18	1.00
	<.0001	<.0001	

Table 12 Comparison of Means of Smart Growth Location Variables between Defaulted and Non-defaulted Loans

Default	N Obs.	Variable	Mean	Std. Dev.	Minimum	Maximum
0	19864	*WalkScore*	47.59	33.02	0.00	100.00
		Transit_Quart_Mile	0.17	0.37	0.00	1.00
		in_500_Ft_Hwy_Buff	0.03	0.17	0.00	1.00
		Protected_Area_Quart_Mi	0.31	0.46	0.00	1.00
		BlkGpTreeCover_Median	11.73	20.59	0.00	97.00
		Jobs_Worker_Bal	0.38	0.29	0.00	1.00
1	2949	*WalkScore*	37.43	28.02	0.00	100.00
		Transit_Quart_Mile	0.08	0.28	0.00	1.00
		in_500_Ft_Hwy_Buff	0.02	0.15	0.00	1.00
		Protected_Area_Quart_Mi	0.25	0.43	0.00	1.00
		BlkGpTreeCover_Median	13.28	22.28	0.00	96.00
		Jobs_Worker_Bal	0.38	0.30	0.00	1.00

Table 13 MLE Estimates of the Default Hazard Models

	A Baseline Model		B Model with sustainability features	
Parameter	Estimate (Standard Error)	Hazard Ratio	Estimate (Standard Error)	Hazard Ratio
logbal	0.008	1.008	0.043***	1.044
	(0.007)		(0.007)	
ltv75	0.404***	1.497	0.395***	1.484
	(0.013)		(0.013)	
curltv	0.15***	1.162	0.156***	1.169
	(0.008)		(0.008)	
dscrnoi2	-0.54***	0.583	-0.539***	0.583
	(0.009)		(0.009)	
occrate2	-0.262***	0.770	-0.263***	0.769
	(0.004)		(0.004)	
refi_inc1	-0.187***	0.829	-0.193***	0.824
	(0.01)		(0.01)	
age	-0.002	0.998	0.077***	1.08
	(0.007)		(0.008)	
lock	0.511***	1.667	0.462***	1.587
	(0.035)		(0.036)	
prep	0.194	1.214	0.201	1.222
	(0.111)		(0.111)	
yldm	0.145***	1.156	0.086*	1.090
	(0.04)		(0.041)	
WalkScore			-0.145***	0.865
			(0.008)	
Transit_Quart_Mile			-0.359***	0.699
			(0.026)	
in_500_Ft_Hwy_Buff			-0.164***	0.848
			(0.042)	
Protected_Area_Quart			-0.031*	0.970
			(0.015)	
BlkGpTreeCover_Median			-0.011	0.989
			(0.008)	
Jobs_Worker_Bal			0.034***	1.034
			(0.006)	
estar_labl			-0.222***	0.801
			(0.058)	
unemp_msa_rla	0.271***	1.311	0.263***	1.301

	(0.01)		(0.011)	
yldslope	-0.083***	0.921	-0.074***	0.928
	(0.009)		(0.009)	
tcm10ystd	0.044***	1.045	0.057***	1.059
	(0.009)		(0.01)	
rcacppi_vol	-0.07***	0.933	-0.072***	0.930
	(0.007)		(0.007)	
crdspread	-0.011	0.989	-0.012	0.988
	(0.009)		(0.009)	
cssaproptype IN	0.237***	1.268	0.14***	1.151
	(0.023)		(0.024)	
cssaproptype MF	-0.01	0.990	-0.035*	0.966
	(0.017)		(0.017)	
cssaproptype OF	0.582***	1.789	0.603***	1.827
	(0.014)		(0.015)	
MSA-fixed effect	Yes		Yes	
N	664,794		664,794	
-2LogL	572,091		537,843	
AIC	572,167		537,993	

Note: * for p<0.05; ** for p<0.01; and *** for p<0.001.

Table 14 MLE Estimates of the Default Hazard Model with Sustainability Variables Interacted with Property Type

Parameter	Estimate (Standard Error)
cssaproptype IN	0.341***
	(0.03)
cssaproptype MF	0.033
	(0.021)
cssaproptype OF	0.576***
	(0.019)
WalkScore	-0.125***
	(0.014)
Transit_Quart_Mile	-0.094*
	(0.047)
in_500_Ft_Hwy_Buff	0.081
	(0.07)
Protected_Area_Quart	0.023
	(0.024)
BlkGpTreeCover_Median	-0.077***

	(0.011)
Jobs_Worker_Bal	0.041***
	(0.009)
estar_labl	-1.04**
	(0.334)
*cssaproptype IN * WalkScore*	0.24***
	(0.033)
*cssaproptype MF * WalkScore*	-0.029
	(0.021)
*cssaproptype OF * WalkScore*	-0.068***
	(0.019)
*cssaproptype IN * Transit_Quart_Mile*	-0.389**
	(0.139)
*cssaproptype MF * Transit_Quart_Mile*	-0.255***
	(0.072)
*cssaproptype OF * Transit_Quart_Mile*	-0.315***
	(0.06)
*cssaproptype IN * in_500_Ft_Hwy_Buff*	-0.402**
	(0.155)
*cssaproptype MF * in_500_Ft_Hwy_Buff*	-0.603**
	(0.125)
*cssaproptype OF * in_500_Ft_Hwy_Buff*	-0.263**
	(0.097)
*cssaproptype IN * Protected_Area_Quart_Mi*	-0.375***
	(0.058)
*cssaproptype MF * Protected_Area_Quart_Mi*	-0.358***
	(0.042)
*cssaproptype OF * Protected_Area_Quart_Mi*	0.069*
	(0.031)
*cssaproptype IN * BlkGpTreeCover_Median*	0.201***
	(0.024)
*cssaproptype MF * BlkGpTreeCover_Median*	0.102***
	(0.015)
*cssaproptype OF * BlkGpTreeCover_Median*	0.08***
	(0.013)
*cssaproptype IN * Jobs_Worker_Bal*	-0.007
	(0.024)
*cssaproptype MF * Jobs_Worker_Bal*	-0.017
	(0.015)
*cssaproptype OF * Jobs_Worker_Bal*	-0.01

	(0.014)
*cssaproptype IN * estar_labl*	--
	--
*cssaproptype MF * estar_labl*	--
	--
*cssaproptype OF * estar_labl*	0.901**
	(0.339)
Other control variables	Yes
N	664,794
-2LogL	537,369
AIC	537,499

Note: * for p<0.05; ** for p<0.01; and *** for p<0.001.

Table 15 By Property Type MLE Estimates of the Hazard Model

	Estimate (S.E.)			
	Office	Multifamily	Retail	Industrial
WalkScore	-0.198***	-0.133***	-0.119***	0.112***
	(0.014)	(0.019)	(0.014)	(0.034)
Transit_Quart_Mile	-0.452***	-0.156**	0.007	-0.483***
	(0.04)	(0.058)	(0.048)	(0.133)
in_500_Ft_Hwy_Buff	-0.263***	-0.56***	0.096	-0.434**
	(0.068)	(0.106)	(0.071)	(0.143)
Protected_Area_Quart	-0.059*	0.067	-0.025	-0.404***
	(0.024)	(0.041)	(0.025)	(0.058)
BlkGpTreeCover_Median	-0.001	-0.085***	0.021	0.192***
	(0.013)	(0.017)	(0.014)	(0.03)
Jobs_Worker_Bal	0.041***	0	0.053***	0.049*
	(0.011)	(0.012)	(0.009)	(0.023)
estar_labl	-0.202***	--	-1.08**	--
	(0.061)	--	(0.334)	--
Other controls	Yes	Yes	Yes	Yes

Note: * for p<0.05; ** for p<0.01; and *** for p<0.001.

Appendix Table 1 Effect of Filters on Final Sample Size

Filter	Resulting Sample Size
Single Property Loans	89,865
Fixed-rate loans	88,239
Loans with Geocodes	72,359
Fixed-rate loans for the 4 major property types (office, retail, multifamily, and industrial)	71,446
Fixed-rate loans on 4 major property types covered by RCA price indices (in 17 MSAs and originated in 2000 or after)	35,785
Fixed rate loans on 4 major property types covered by RCA price indices (in 17 MSAs and originated in 2000 or after), with valid geocodes and other loan information	22,813

Appendix Table 2 Summary of Green Infrastructure Results

Property Group	*BlkGpTreeCover_Median*			*Protected_Area_Quart*			Summary
	Main Model	Interaction Model	By Type Models	Main Model	Interaction Model	By Type Models	
Office		increases risk	reduces risk		increases risk	no effect	ambiguous
Retail		no effect	no effect		reduces risk	no effect	ambiguous
MF		reduces risk	no effect		increases risk	reduces risk	ambiguous
Industrial		reduces risk	reduces risk		increases risk	increases risk	ambiguous
All types	no effect			reduces risk			ambiguous