

# Location Efficiency and Mortgage Default

---

**Authors** Stephanie Y. Rauterkus, Grant I. Thrall, and Eric Hangen

**Abstract** Using a sample of over 40,000 mortgages in Chicago, Jacksonville, and San Francisco, we model the probability of mortgage default based on differences in location efficiency. We used two proxy variables for location efficiency: 1) vehicles per household scaled by income and 2) Walk Score™. We find that default probability increases with the number of vehicles owned after controlling for income. Further, we find that default probability *decreases* with higher Walk Scores in high income areas but *increases* with higher Walk Scores in low income areas. These results suggest that some degree of greater mortgage underwriting flexibility could be provided to assist households with the purchase of location efficient homes, without increasing mortgage default. They also support the notion that government policies around land use, zoning, infrastructure, and transportation could have significant impacts on mortgage default rates.

The recent housing crisis has created a renewed interest in the drivers of mortgage default behavior. According to Haughwout, Peach, and Tracy (2008), “Historically, four key characteristics (“risk factors” or “underwriting criteria”) have been thought to determine the probability that a mortgagor will default. Those factors are the loan-to-value ratio (LTV), the debt service-to-income ratio (DTI), the mortgagor’s credit score, and the extent to which the mortgagor’s income and assets have been verified by third party sources such as employers, tax returns, and bank account statements.” A rich literature on mortgage default behavior exists that describes these predictors of default, and that explores the role of life events such as job loss, illness and divorce as well as market events, particularly home price declines (e.g., Foster and Van Order, 1984; Thibodeau, 1985; Quercia and Stegman, 1992; Kau, Keenan, and Kim, 1993; Vandell, Kau, Keenan, and Kim, 1994; Kau, Keenan, Muller, and Epperson, 1995; Vandell, 1995; and Elul, 2006). More recently, a literature has examined the terms and originators of the loans themselves, particularly subprime and broker-originated loans, as predictors of default risk (e.g., Alexander, Grimshaw, McQueen, and Slade, 2002; and Ding, Quercia, Ratcliffe, and Lei 2008).

One possible factor that may also impact mortgage default risk is location efficiency. Location-efficient homes are located in areas that enable lower automobile ownership. The most important determinants of location efficiency are the compactness of residential development (number of housing units per net acre of residential development) and transit access (number of buses per hour at bus stops within ¼ mile walk of the home) (Holtzclaw, Clear, Dittmar, Goldstein, and

Haas, 2002). Such neighborhoods typically are close to schools, shopping, workplaces, and other amenities.

People living in these homes are able to drive less because they have other options for many trips: to walk, bike or use public transit. As a result, they may save significantly on the cost of driving, driving for fewer miles and potentially reducing the number of cars they would otherwise need to own. In turn, these expense reductions should create an economic buffer for location-efficient homeowners that may reduce their propensity to default compared to otherwise similar homeowners who must lay out a substantial portion of their household budget for automobile transportation.<sup>1</sup> Furthermore, location-efficient homeowners may be buffered against increases in gasoline costs such as those observed in the summer and fall of 2008. Even before this spike in gas prices, transportation costs were the second-largest expenditure for the typical American household (Brookings Institution, 2006), averaging \$8,500 per year.<sup>2</sup> Therefore, one might well expect that households could afford to purchase a more expensive home if that home enabled them to reduce transportation expenses (e.g., by eliminating one car, or simply by driving less because they can now walk or use transit more). This realization has led to the development of a new housing affordability index that takes into account the expected transportation costs associated with the home's location.<sup>3</sup> Applying this index in place of more traditional measures of housing affordability shows that many car-dependent areas of metropolitan regions turn out to be much less affordable than initially thought, once transportation costs are factored in.<sup>4</sup>

These research results raise the question: If housing affordability is impacted by location efficiency, shouldn't mortgage underwriting take it into account? Environmental advocates have long promoted the idea of a "Location Efficient Mortgage<sup>®</sup>" that would reward homebuyers of location-efficient homes with more flexible mortgage underwriting terms.<sup>5</sup> For lenders to offer such flexible mortgage products, they must be convinced that location efficiency does indeed alter the risk profile of a mortgage. The specific hypothesis that must be tested is that after holding traditional underwriting factors constant, mortgages on location-efficient homes will have superior performance to those of non-location-efficient homes. If the hypothesis is shown to be true, then presumably underwriting guidelines should account for the location efficiency of a home. For example, underwriters could set a higher debt-to-income ratio for loans to location-efficient homes (reflecting the household budget savings on transportation in such homes) relative to the debt ratio allowed for location-inefficient homes.

If location efficiency is indeed predictive of lower mortgage default, it also provides additional justification for government policy interventions to promote a more location-efficient built environment. The harm of mortgage foreclosures is well documented (and now something that a great many communities nationwide are experiencing firsthand). Both families and lenders incur significant costs when a foreclosure happens.<sup>6</sup> However, the costs do not end there. The resulting REO property reduces the community's tax base, encourages arson and crime,<sup>7</sup> and contributes to significant loss in property value of neighboring properties.<sup>8</sup> Policies to promote a more location-efficient built environment generally fall under the

rubric of “Smart Growth”<sup>9</sup> and include zoning tools and urban growth boundaries to promote denser development, targeting of infrastructure investments to built areas of the region, enhancements to transit and pedestrian infrastructure, open space preservation programs, and other policy tools.

---

## Review of the Literature

Few studies to date have directly addressed the relation between mortgage default and location efficiency. Holtzclaw, Clear, Dittmar, Goldstein, and Haas (2002) aimed to find support for the development of the location efficient mortgage (LEM) by studying 2,820 transportation analysis zones (TAZs) in Chicago, Los Angeles, and San Francisco. The authors use auto ownership from 1990 U.S. Census data and vehicle miles traveled from odometer readings as their measures of location efficiency. They find that auto ownership and annual driving distance are a function of the density of the owner’s neighborhood. They also find a relationship, though weaker, between annual driving distance and the pedestrian and bicycle friendliness of the driver’s neighborhood.

In a preview of the 2002 study, Holtzclaw (2000) related the value of smart growth to location efficiency. In this study, Holtzclaw used preliminary location efficiency results to explain how the smart growth principle of creating walkable neighborhoods would in turn reduce automobile dependence and thereby reduce motor vehicle pollution.

Using a cohort of 8,000 FHA-insured loans made in the Chicago area from 1988 to 1992, Blackman and Krupnick (2000) conclude that location efficiency does not improve the risk characteristics of a mortgage. However, this study is somewhat limited in that it looks at only one metropolitan area and limited vintages of loans. A number of other studies have looked at whether certain neighborhood characteristics are predictive of future home value trends or of default risks, although most of these studies have focused on income and demographic characteristics as opposed to location efficiency per se (e.g., Goetzmann and Spiegel, 1997).

---

## Data

We use data from a number of sources. First, we obtain loan-level mortgage data from LPS Applied Analytics. LPS, formerly McDash Analytics, owns the largest loan-level database of mortgage assets with loan-level data for more than 39 million active first and second mortgage loans, including portfolios serviced by nine of the top 10 mortgage servicers in the nation. This database represents approximately two-thirds of the mortgage market. The dataset has both static and dynamic variables that relate to characteristics of the loan and the borrower at origination (static) and at specific points in time subsequent to loan origination (dynamic). Since this data does not have property addresses, we obtain property transfer records for counties in each metropolitan area and match these records to the mortgage data in order to identify the location of each property.

We obtain a sample of loans outstanding as of December 2008 in each of three metropolitan areas: Chicago, IL; Jacksonville, FL; and San Francisco, CA. We choose these cities in order to examine potential regional differences in mortgage default. All of these regions are areas that have struggled with foreclosures. Additionally, each has a distinct development history and urban form; significant effects of location efficiency in these cities would suggest a greater likelihood that the findings are applicable to other communities nationally. To ensure that our analysis included both the central city and outlying areas, we obtain property transfer records for multiple counties in each metropolitan area. In Chicago, we selected properties located in Cook, Dupage, Kane, Kendal, Lake, McHenry, and Will counties. Exhibit 1 shows a map of our Chicago data. In Jacksonville, we selected properties in Clay, Duval, Nassau, and St. Johns counties. Exhibit 2 shows a map of our Jacksonville data. In San Francisco, we obtain property transfer data

**Exhibit 1** | Chicago, IL





Exhibit 3 | San Francisco



(ESRI) for each block group identified in our data, which include an estimate of vehicles per household.<sup>10</sup> Then, to normalize for differences in income, we divide this value by income.<sup>11</sup> The result is our first key location efficiency variable—*vehper000inc*. This variable measures the number of vehicles the typical resident in a given block group owns for every \$1,000 of income. For example, across all three of our samples, this value ranges between 0.02 and 0.04. Using 0.03 as an average value, this means that a person who earns \$20,000 per year would own 0.60 vehicles while a person who earns \$200,000 per year would own 6.0 vehicles.

We use Walk Score as an alternative measure of location efficiency in our study.<sup>12</sup> Walk Score rates the walkability of a specific address on a scale from 0 to 100, by compiling the number of nearby stores, restaurants, schools, parks, etc. within a one-mile radius from the subject location. Higher scores are indicative of more walkable locations and therefore should generally correlate with more walkable

neighborhoods whereas locations with scores below 50 are considered to be car dependent. After mapping the property addresses to obtain geographic coordinates for each location, we apply the Walk Score algorithm (defined below) to obtain the Walk Score for each location. The Walk Score algorithm awards points based on the distance to the closest amenity in each category and no points are awarded for amenities located more than one mile from the subject address.<sup>13</sup> However, there are some key factors that affect walkability that are not accounted for in the Walk Score. If there is topography such as steep hills, freeways or bodies of water within the one-mile radius, they are ignored when calculating the Walk Score. Also, if the location is prone to extreme weather conditions, these are not reflected in the Walk Score. That is, the Walk Score is strictly based on the distance between a location and all amenities within a one-mile radius. Also, the distances used are “as the crow flies” distances as opposed to the actual distance walked along street grids. In effect, Walk Score assumes that residents can walk to each amenity using a straight path. In addition, Walk Score does not consider proximity to public transit. All three of the study cities appear on the Walk Score list of 40 most walkable cities. San Francisco tops the list with an overall score of 86; Chicago ranks fourth with a score of 76 and Jacksonville ranks 40<sup>th</sup> with a score of 36. Note that these Walk Scores refer to the city jurisdiction and not the metropolitan area.

Exhibit 4 describes the variables used in the study. *DTI\_orig* is a static variable from the LPS Applied Analytics database. This measures the debt-to-income (“back end”) ratio of the borrower at origination of the mortgage as reported by the servicer.<sup>14</sup> DTIs for mortgage borrowers typically range between 36% and 41%. Our sample observations tend to have DTIs at the lower end of this range. This statistic was not available for the Jacksonville sample. Also, a large number of loans in Chicago (37%) and San Francisco (39%) were missing this statistic.

*FICO\_orig* is also a static variable from LPS. This measure of credit worthiness created by the Fair Isaac Corporation is commonly used to measure an applicant’s credit risk. Scores range from 300 to 850 and scores above 680 are typically considered to be very good or “prime” borrowers. This data is more complete than the DTI variable, but still a number of the FICO scores are also missing in our sample.<sup>15</sup> Across all three subsets of our data, mean credit scores are well above 680, indicating that the average borrower would likely qualify for prime borrowing terms. In San Francisco, however, credit scores are even higher. With a mean FICO score of 745, the average borrower in our San Francisco sample has excellent credit.

We obtain loan-to-value ratios (LTVs) from LPS. LTV expresses the amount of a first mortgage lien as a percentage of the total appraised value of real property. Conforming loans according to Fannie Mae and Freddie Mac standards are those with LTVs less than or equal to 80%. The mean LTV across all of our samples is less than 80% and as low as 71% in San Francisco. This indicates that the average mortgage in our study is a conforming loan and not subject to private mortgage insurance requirements.

We determine the age of each mortgage by calculating the number of days between the loan closing date and December 31, 2008. This age is reflected in our

**Exhibit 4** | Descriptive Statistics

Variable	N	Mean	Std. Dev.	Min.	Max.
Panel A: Chicago, IL					
<i>DTI_orig</i> (%)	11,834	36.97	10.79	1.00	99.00
<i>FICO_orig</i>	16,542	695.74	65.74	387.00	844.00
<i>LTV_ratio</i> (%)	18,607	79.36	11.58	4.72	107.50
<i>Mort_age</i> (days)	18,735	995.30	43.98	930.00	1,080.00
<i>Orig_amt</i> (\$)	18,735	222,036.40	119,196.70	15,000.00	1,900,000.00
<i>PCI</i> (\$)	18,735	33,731.98	21,626.74	0.00	154,098.00
<i>Popgwth00CY</i> (%)	18,735	0.45	4.72	-3.90	96.40
<i>PNonwhite</i> (%)	18,735	43.43	30.19	0.00	100.00
<i>VehperHH</i>	18,735	1.47	0.44	0.05	4.00
<i>Vehper000HHinc</i>	18,735	0.03	0.01	0.00	0.76
<i>Walk_Score</i>	18,735	64.53	19.53	0.00	100.00
Panel B: Jacksonville, FL					
<i>FICO_orig</i>	18,188	690.85	71.96	307.00	900
<i>LTV_orig</i> (%)	22,844	78.60	20.42	3.13	212.00
<i>Mort_age</i> (days)	22,848	1,479.59	1,068.13	92.00	11,202
<i>Orig_amt</i> (\$)	22,848	128,016.60	89,270.37	5,000.00	2,680,000
<i>PCI</i> (\$)	22,848	27,699.41	12,259.82	6,400.00	116,161.00
<i>Popgwth00CY</i> (%)	22,848	2.48	3.67	-1.30	23.30
<i>PNonwhite</i> (%)	22,848	38.57	26.02	1.70	99.90
<i>Unemprt</i> (%)	22,848	6.82	4.89	0.00	40.00
<i>VehperHH</i>	22,848	1.67	0.31	0.20	2.49
<i>Vehper000HHinc</i>	22,848	0.04	0.01	0.01	0.09
<i>Walk_Score</i>	22,848	32.56	20.20	2.00	97.00
Panel C: San Francisco, CA					
<i>DTI_orig</i> (%)	9,789	36.10	12.55	1.00	99.00
<i>FICO_orig</i>	14,916	744.60	45.71	478.00	828.00
<i>LTV_orig</i> (%)	15,880	71.13	13.17	6.00	101.00
<i>Mort_age</i> (days)	15,945	779.35	403.92	1.00	1,460.00
<i>Orig_amt</i> (\$)	15,945	607,150.40	256,258.90	43,000.00	6,300,000.00
<i>PCI</i> (\$)	15,945	55,616.93	34,803.86	8,901.00	176,894.00
<i>PNonwhite</i> (%)	15,945	52.72	20.89	7.57	99.49
<i>Unemprt</i> (%)	15,945	6.86	5.24	0.00	81.00
<i>VehperHH</i>	15,945	1.39	0.47	0.03	3.44
<i>Vehper000HHinc</i>	15,945	0.02	0.01	0.00	0.40
<i>Walk_Score</i>	15,945	84.34	12.00	28.00	100.00

*mort\_age* variable. In Chicago, our loans are 30 to 36 months old. In Jacksonville, we have a much wider age range with loans ranging from 3 months to 30 years old. The age of the San Francisco loans falls between one day and four years old.<sup>16</sup>

We do not include the origination amount in our analysis, but we show it to here to emphasize the differences between these three geographic locations. Mean origination values in San Francisco are nearly three times the mean values in Chicago and five times the mean values in Jacksonville.

The demographic makeup of the three locations varies also. Similar to the disparity in origination values, per capita income varies greatly across the cities at the block group level. Per capita income in San Francisco is double the income measure in Jacksonville. Residents in the Chicago area earn slightly more but mean per capita income in Chicago is still only 60% of the mean value in San Francisco. Population growth since 2000 (*popgwth00CY*) at the block group level in Jacksonville is more than five times the rate of growth in Chicago. The minority share of the population at the block group level is also much greater in San Francisco (53%) than in Jacksonville (39%) and Chicago's minority share falls between the two.

Walk Score, as described previously, measures walkability on a 100-point scale with more walkable areas earning higher scores. In Chicago, Jacksonville, and San Francisco, our samples have mean Walk Scores of 65, 33, and 84 respectively. These values are consistent with those reported on the Walk Score listing of "America's Most Walkable Neighborhoods." That listing reports values of 76, 36, and 86 for our sample cities of Chicago, Jacksonville, and San Francisco. According to the Walk Score website, Chicago would be classified as "Somewhat Walkable" according to our sample statistics, Jacksonville would be considered "Car Dependent," and San Francisco would be considered "Very Walkable." Recall that our samples include observations located in counties that surround these cities, whereas the Walk Scores cited in this paragraph are for the city jurisdiction. Also, these are mean values; parts of the city may differ greatly.

*Vehperhh* is quite similar across cities with the mean value ranging from a low of 1.39 in San Francisco to a high of 1.67 in Jacksonville. We tested the correlation between Walk Score and vehicles per household and found that they are significantly negatively correlated. The correlation coefficient for the Chicago sample,  $\rho_C$ , is  $-0.6102$ ; the Jacksonville correlation coefficient,  $\rho_J$ , is  $-0.5726$ ; and the San Francisco correlation coefficient,  $\rho_{SF}$ , is  $-0.6395$ . These values indicate that the higher the Walk Score for the neighborhood where one lives, the fewer cars one tends to own.

---

## Methodology

We test the null hypothesis that borrowers with homes located in location efficient areas are no more likely to default on their mortgages than borrowers with homes located in car-dependent areas. In other words,

$H_0$ : The degree of location efficiency in a neighborhood has no impact on mortgage performance in that neighborhood.

$H_A$ : Mortgage performance is related to location efficiency.

We use a probit regression analysis to model the probability that borrowers will default on their mortgages. We use the model  $default = a_0 + b_1LEFF + b_2CONTROL + \varepsilon$ , where *default* is a binary variable set to one if the mortgage is in default and zero otherwise, *LEFF* is a variable measuring location efficiency, and *CONTROL* is a vector of control variables that are widely seen as driving foreclosure risk.<sup>17</sup> These control variables include three of the four most commonly used default predictors: back-end debt-to-income ratio (DTI), loan-to-value ratio (LTV), and FICO score. While the LPS database includes a variable for loan documentation type (no, low, full or unknown), few of the observations in our sample included data for this variable.

We find that income tends to be highly (positively) correlated with Walk Score. Therefore, we do not include per capita income in our vector of control variables. Instead, we stack our three location-based datasets into one combined dataset. Here, we create two sets of binary variables and one set of interaction variables. The first set of binary variables, *CHI*, *JAX*, and *SF*, are set to one if the observation relates to a property located in Chicago, Jacksonville, or San Francisco respectively and are set to zero otherwise. The second set of binary variables, *qrt1*, *qrt2*, *qrt3*, and *qrt4*, are set to one if the observation relates to a property located in a block group whose per capita income falls in the first, second, third, or fourth quartile with regard to the combined dataset.<sup>18</sup> We then create a set of interaction variables, *citywalkqrt#* in all possible combinations. This gives us four variables (*chiwalkqrt1*, *chiwalkqrt2*, etc.) for each city and 12 total interaction variables. For example, *chiwalkqrt1* can be interpreted as the Walk Score conditional on the property being in Chicago and located in a block group with a low per capita income. If a particular property does not meet both of these conditions (city and per capita income quartile), this value will be zero. With these new location efficiency interaction variables now also controlling for income, we use the combined dataset to estimate our original equation  $default = a_0 + b_1LEFF + b_2CONTROL + \varepsilon$  except now, *LEFF* is vector of location efficiency interaction variables and we use FICO score and binary variables indicating whether or not the block group is made up of primarily white or primarily minority residents as control variables.<sup>19</sup>

---

## Results

The results of our probit regressions are shown in Exhibit 5. In all cities, *lnvehper000inc* is positive and highly significant ( $p < 0.01$ ) before controlling for other factors and the pseudo  $R^2$  is approximately 3%. In Models 2 through 4 we add additional control variables to the model. In all models, *lnvehper000inc* remains positive and highly significant ( $p < 0.01$  in all but one model) across all specifications.

In Model 2, we add key factors commonly found to be predictive of mortgage default—DTI, FICO, score and LTV.<sup>20</sup> All three of these control factors are

**Exhibit 5** | Probability of Default and Vehicles per Household

Variable	Model 1	Model 2	Model 3	Model 4
Panel A: Chicago, IL				
<i>lnvehper000inc</i>	0.6761 (16.23)***	0.3069 (4.83)***	0.2969 (4.66)***	0.2574 (3.97)***
<i>lnDTI_orig</i>		0.4021 (7.16)***	0.4036 (7.19)***	0.3966 (7.06)***
<i>lnFICO_orig</i>		-3.1686 (-16.37)***	-3.1624 (-16.32)***	-3.1961 (-15.68)***
<i>lnLTV_orig</i>		0.5138 (3.87)***	0.5160 (3.87)***	0.33753 (3.09)***
<i>lnMort_age</i>			0.5106 (1.29)	0.6575 (1.64)
<i>Popgwth00CY</i>			-0.0114 (-1.02)	-0.0133 (-1.22)
<i>Minblkgrp</i>				0.1213 (2.57)**
<i>Ownocc</i>				-0.0119 (-0.16)
<i>Refi</i>				-0.2584 (-6.71)***
<i>Whiteblkgrp</i>				-0.1364 (-3.08)***
Constant	1.1946 (8.15)***	16.7104 (11.30)***	13.0952 (4.21)***	12.9345 (4.13)***
Pseudo R <sup>2</sup>	0.0218	0.0834	0.0839	0.0954
N	18,632	10,372	10,372	10,372

**Exhibit 5** | (continued)  
Probability of Default and Vehicles per Household

Variable	Model 1	Model 2	Model 3	Model 4
Panel B: Jacksonville, FL				
<i>lnvehper000inc</i>	0.9654 (15.91)***	0.7121 (9.83)***	0.5563 (6.62)***	0.3845 (4.29)***
<i>lnFICO_orig</i>		-2.43 (-17.28)***	-2.3828 (-16.73)***	-2.5219 (-16.23)***
<i>lnLTV_orig</i>		0.6425 (9.01)***	0.6673 (9.22)***	0.6258 (7.46)***
<i>lnMort_age</i>			-0.1595 (-6.02)***	-0.1701 (-5.89)***
<i>Popgwth00CY</i>			0.0015 (0.27)	-0.0044 (-0.79)
<i>lnUnemprt</i>			0.1225 (4.39)***	0.0105 (0.35)
<i>Fixed</i>				-0.6711 (-19.21)***
<i>Minblkgrp</i>				0.3340 (6.94)***
<i>Ownocc</i>				-0.2590 (-6.68)***
<i>Refi</i>				-0.1806 (-4.53)***
<i>whiteblkgrp</i>				-0.1485 (-3.51)***
Constant	1.5611 (8.04)***	13.7274 (13.76)***	13.7016 (13.01)***	15.2165 (12.86)***
Pseudo R <sup>2</sup>	0.0257	0.0853	0.0922	0.1593
N	22,707	18,066	17,941	17,941

**Exhibit 5** | (continued)

Probability of Default and Vehicles per Household

Variable	Model 1	Model 2	Model 3	Model 4
Panel C: San Francisco, CA				
<i>Lnvehper000inc</i>	0.7813 (10.34)***	0.5025 (4.39)***	0.4752 (3.80)***	0.3843 (2.41)**
<i>lnDTI_orig</i>		0.4156 (3.81)***	0.5348 (4.65)***	0.5581 (4.82)***
<i>lnFICO_orig</i>		-4.9592 (-10.38)***	-4.4795 (-9.10)***	-4.3745 (-8.43)***
<i>lnLTV_orig</i>		2.8739 (6.20)***	2.8841 (5.23)***	2.6579 (4.34)***
<i>lnMort_age</i>			0.5777 (6.05)***	0.4486 (4.51)***
<i>lnUnemprrt</i>			0.0325 (0.72)	0.0231 (0.50)
<i>Fixed</i>				-0.5899 (-5.10)***
<i>Minblkgrp</i>				0.1848 (1.75)*
<i>Ownocc</i>				-0.1187 (-0.80)
<i>Refi</i>				-0.0838 (-1.01)
<i>whiteblkgrp</i>				0.0170 (0.11)
Constant	0.9680 (3.38)***	18.5375 (4.85)***	10.8706 (2.55)**	11.8320 (2.53)**
Pseudo R <sup>2</sup>	0.0353	0.1712	0.2023	0.2254
N	15,883	9,088	8,973	8,973

Notes: This table reports the results of various probit model specifications. The dependent variable in all specifications is the binary variable *default* which is set to one if the mortgage is in default and zero otherwise. Location efficiency is measured by Vehicles per Household as adjusted by income. Z-scores are shown in parentheses.

\* Significant at the 10% level.

\*\* Significant at the 5% level.

\*\*\* Significant at the 1% level.

significant as has been found in the prior literature. In addition, the explanatory power of the model grows to as much as 17% in San Francisco after adding these factors.

In Model 3, we add mortgage age, population growth, and unemployment rate. Neither mortgage age nor population growth explains the probability of default in Chicago. However, mortgage age is a significant factor in Jacksonville and San Francisco. In Jacksonville, default probability increases as the age of the mortgage decreases, but in San Francisco default probability increases as the age of the mortgage increases. In addition, the default probability rises with the unemployment rate in Jacksonville but is not a significant explanatory factor in San Francisco.

In Model 4, we add binary control variables. Here again, the results vary by city. In fact, none of these variables are significant in all cities. In San Francisco, neither mortgages initiated to refinance an existing mortgage nor properties located in block groups that are predominantly white are more likely to default; however, these are significant factors in Chicago and Jacksonville. Again, the explanatory power of our model is as high as 23% in San Francisco.

The coefficient for income-normalized vehicle ownership is positive and significant—even in models that include standard default risk factors such as DTI, FICO, and LTV and additional variables that explain regional differences in foreclosure likelihoods. We interpret these results to mean that income-normalized vehicle ownership increases the probability of mortgage default.

The following example explains our interpretation of these results. Assume that we wish to estimate the probability that an average borrower in San Francisco will default on their mortgage. According to our descriptive statistics, an average borrower in San Francisco lives in a neighborhood averaging 0.02 vehicles per thousand dollars of income, and has a DTI of 36.1%, a FICO score of 745, an LTV of 71.1%, a mortgage that is 779 days old, and lives in a block group where the unemployment rate is 6.86%.<sup>21</sup> Assume also that this particular borrower originated the mortgage for a new home purchase, does not live in the home, and has an adjustable interest rate and lives in an integrated block group. For this particular borrower, our model indicates that the probability of default is:

$$F[11.8320 + 0.3843 \times \ln(0.02) + 0.5581 \times \ln(36.1) - 4.3745 \\ \times \ln(745) + 2.6579 \times \ln(71.13) + 0.4486 \times \ln(779) \\ + 0.0231 \times \ln(6.86)],$$

where  $F$  is the cumulative distribution function of the standard normal. According to our model, the probability that this borrower will default is 1.2%.

We use Walk Score as an alternative measure of location efficiency. The results of our probit regressions for the individual cities using  $\ln\text{Walk\_Score}$  as our key

location efficiency variable are shown in Exhibit 6. Panel A presents the results for Chicago, Panel B presents Jacksonville, and Panel C shows the results in San Francisco. The first model in all cities tests how well Walk Score alone predicts the likelihood of default. In all cities, *InWalk\_Score* is significantly different from zero when modeled without controlling for other factors. However, the explanatory power of the models as a whole is very low as indicated by pseudo  $R^2$  values close to zero. The signs of the coefficients are negative in both Chicago and San Francisco but positive in Jacksonville. This suggests that walkability decreases the probability of default in walkable cities but *increases* the probability of default in car-dependent cities.

Model 2 in all cities includes Walk Score plus key risk factors commonly related to mortgage default. These are DTI, FICO score, and LTV.<sup>22</sup> In these models, the common risk factors are all significantly different from zero, indicating that they are predictive of default. However, in the Chicago sample, Walk Score loses significance when these factors are added to the model. In addition, the overall explanatory power of this model rises significantly across all cities to as much as 17% in San Francisco as indicated by the pseudo  $R^2$ .

In Model 3, we add additional discrete control variables. Specifically, we add the age of the mortgage, population growth, and unemployment rate. With the exception of Chicago, the sign and significance of our key variable, *InWalk\_Score*, is consistent across all model specifications. With respect to the control variables, here again we find regional differences in the results. The age of the mortgage is insignificant in Chicago, negative and significant in Jacksonville, and positive and significant in San Francisco. These results indicate that the vintage of the mortgages in default as of December 2008 varies greatly across cities. In Jacksonville, newer mortgages are more likely to be in default, in San Francisco, older mortgages are more likely to be in default, and in Chicago there is no discernible trend with respect to mortgage age. There is less variability in the results regarding population growth and the unemployment rate. The probability of default increases with increases in the block group unemployment rate. Population growth in the block group decreases the probability of default but is a less significant factor.

In Model 4, we add binary control variables. These variables are ‘flags’ indicating whether or not the interest rate is fixed (*Fixed*), the property is located in a primarily minority block group (*Minblkgrp*), the property is owner-occupied (*Ownocc*), the purpose of the loan is for refinancing an existing mortgage (*Refi*), and whether or not the property is located in a block group where the population is primarily white (*Whiteblkgrp*). While *InWalk\_Score* remains stable, the additional control variables exhibit regional differences. Specifically, whether or not the property is owner-occupied is only a significant factor in mortgage default in Jacksonville. In addition, despite a pseudo  $R^2$  of nearly 23% in San Francisco, four of the 11 factors included in the model are insignificant. Two of these, *Refi* and *Whiteblkgrp*, are negative and highly significant in both Chicago and Jacksonville.

In order to estimate our model while controlling for income, we combine the observations from all three cities and create an interaction variable that flags the

**Exhibit 6** | Probability of Default and Walk Score

Variable	Model 1	Model 2	Model 3	Model 4
Panel A: Chicago, IL				
<i>InWalk_Score</i>	-0.2094 (-6.36)***	-0.0047 (-0.09)	0.0018 (0.03)	-0.0184 (-0.33)
<i>InDTI_orig</i>		0.4132 (7.39)***	0.4141 (7.40)***	0.4065 (7.26)***
<i>InFICO_orig</i>		-3.3874 (-17.64)***	-3.3659 (-17.49)***	-3.3196 (-16.24)***
<i>InLTV_orig</i>		0.5513 (4.08)***	0.5532 (4.08)***	0.3954 (3.22)***
<i>InMort_age</i>			0.6259 (1.58)	0.7610 (1.90)*
<i>Popgwth00CY</i>			-0.0173 (-1.79)*	-0.0193 (-2.05)**
<i>Minblkgrp</i>				0.1371 (2.90)***
<i>Ownocc</i>				0.0082 (0.11)
<i>Refi</i>				-0.2544 (-6.60)***
<i>Whiteblkgrp</i>				-0.1661 (-3.77)***
Constant	-0.3365 (-2.49)**	16.8658 (11.37)***	12.3698 (3.98)***	12.0527 (3.85)***
Pseudo R <sup>2</sup>	0.0029	0.0806	0.0820	0.0949
N	18,677	10,412	10,412	10,412

**Exhibit 6** | (continued)

Probability of Default and Walk Score

Variable	Model 1	Model 2	Model 3	Model 4
Panel B: Jacksonville, FL				
<i>InWalk_Score</i>	0.1552 (8.96)***	0.1299 (6.32)***	0.0911 (4.01)***	0.0609 (2.54)**
<i>InFICO_orig</i>		-2.6104 (-18.89)***	-2.4725 (-17.51)***	-2.5701 (-16.62)***
<i>InLTV_orig</i>		0.6779 (9.34)***	0.6918 (9.43)***	0.6408 (7.57)***
<i>InMort_age</i>			-0.1592 (-6.05)***	-0.1699 (-5.92)***
<i>Popgwth00CY</i>			-0.0005 (-0.09)	-0.0064 (-1.16)
<i>InUnemprrt</i>			0.1740 (6.62)***	0.0309 (1.04)
<i>Fixed</i>				-0.6646 (-19.08)***
<i>Minblkgrp</i>				0.3601 (7.60)***
<i>Ownocc</i>				-0.2628 (-6.80)***
<i>Refi</i>				-0.1816 (-4.56)***
<i>whiteblkgrp</i>				-0.1841 (-4.39)***
Constant	-2.0574 (-34.62)***	12.0247 (12.05)***	11.9973 (11.62)***	13.9957 (12.06)***
Pseudo R <sup>2</sup>	0.0081	0.0784	0.0889	0.1581
N	22,848	18,184	18,059	18,059

**Exhibit 6** | (continued)  
Probability of Default and Walk Score

Variable	Model 1	Model 2	Model 3	Model 4
Panel C: San Francisco, CA				
<i>lnWalk_Score</i>	-0.8513 (-6.87)***	-0.6898 (-3.47)***	-0.6633 (3.20)***	-0.5191 (-2.17)**
<i>lnDTI_orig</i>		0.4323 (3.97)***	0.5420 (4.73)***	0.5573 (4.82)***
<i>lnFICO_orig</i>		-5.0750 (-10.71)***	-4.5492 (-9.30)***	-4.4011 (-8.50)***
<i>lnLTV_orig</i>		2.9607 (6.27)***	2.9074 (5.29)***	2.6494 (4.37)***
<i>lnMort_age</i>			0.5720 (6.04)***	0.4427 (4.48)***
<i>lnUnemprrt</i>			0.0883 (2.06)**	0.0548 (1.20)
<i>Fixed</i>				-0.5921 (-5.13)***
<i>Minblkgrp</i>				0.2450 (2.53)**
<i>Ownocc</i>				-0.1228 (-0.84)
<i>Refi</i>				-0.0812 (-0.98)
<i>whiteblkgrp</i>				-0.0613 (-0.40)
Constant	1.7178 (3.16)***	20.2040 (5.24)***	12.2428 (2.85)***	12.8369 (2.71)***
Pseudo R <sup>2</sup>	0.0134	0.1659	0.1995	0.2250
N	15,945	9,124	9,009	9,009

Notes: This table reports the results of various probit model specifications. The dependent variable in all specifications is the binary variable *default* which is set to one if the mortgage is in default and zero otherwise. Location efficiency is measured by Walk Score. Z-scores are shown in parentheses.

- \* Significant at the 10% level.
- \*\* Significant at the 5% level.
- \*\*\* Significant at the 1% level.

city where the property is located and the per capita income range of that block group along with the Walk Score for the property. The results of our probit regression for the combined dataset are shown in Exhibit 7. Here we find that 10 of the 12 Walk Score variables are significantly different from zero. Further, we find that the sign of the coefficient changes from positive to negative as income increases. That is, the signs of the coefficients for all three quartile one variables (*Chiwalkqrt1*, *Jaxwalkqrt1*, and *SFwalkqrt1*) are positive. Also, the coefficients decrease in size and significance (as indicated by lower Z-scores) as the income levels increase. In Jacksonville and San Francisco, this trend continues to the point that the signs of the coefficients for the quartile four variables (*Jaxwalkqrt4* and *SFwalkqrt4*) are negative. In other words, higher Walk Scores appear to increase

**Exhibit 7** | Income and Walkability

Variable	Model 1	Model 2
<i>Chiwalkqrt1</i>	0.1712 (10.35)***	0.1435 (8.49)***
<i>Chiwalkqrt2</i>	0.1305 (7.66)***	0.1230 (7.16)***
<i>Chiwalkqrt3</i>	0.0857 (4.89)***	0.0895 (5.07)***
<i>Chiwalkqrt4</i>	0.0382 (2.15)**	0.0479 (2.67)***
<i>Jaxwalkqrt1</i>	0.0838 (4.48)***	0.0583 (3.07)***
<i>Jaxwalkqrt2</i>	-0.0120 (-0.56)	-0.0119 (-0.56)
<i>Jaxwalkqrt3</i>	-0.0121 (-0.54)	0.0057 (0.25)
<i>Jaxwalkqrt4</i>	-0.0552 (-1.66)*	-0.0343 (-1.02)
<i>SFwalkqrt1</i>	0.0537 (2.74)***	0.0091 (0.45)
<i>SFwalkqrt2</i>	0.0428 (2.17)**	0.0052 (0.26)
<i>SFwalkqrt3</i>	-0.0316 (-1.70)*	-0.0422 (-2.24)**
<i>SFwalkqrt4</i>	-0.0873 (-4.88)***	-0.0878 (-4.87)***
<i>lnFICO_orig</i>	-3.3622 (-35.41)***	-3.2068 (-33.33)***
<i>Minblkgrp</i>		0.2317 (9.13)***
<i>Whiteblkgrp</i>		-0.1119 (-4.21)***
Constant	20.2153 (32.73)***	19.2171 (30.69)***
Pseudo R <sup>2</sup>	0.1378	0.1427

Notes: This table reports the results of a probit model specification to estimate the probability of default. The dependent variable is the binary variable default which is set to one if the mortgage is in default and zero otherwise. Location efficiency is measured by Walk Score. Z-scores are shown in parentheses. We use the interaction term *citywalkqrt#*. Each *citywalkqrt#* variable identifies the city in which the property is located, the Walk Score for that unique property and the income quartile for the block group in which that property is located such that quartile one contains properties located in the lowest income areas and quartile four contains properties located in the highest income areas. We use *lnFICO\_orig* to control for borrower credit worthiness. In Models 1 and 2, *N* = 49,597.

\* Significant at the 10% level.

\*\* Significant at the 5% level.

\*\*\* Significant at the 1% level.

the probability of default in low income areas but decreases the probability of default in high income areas.

Our results can best be interpreted with an example. Assume that we wish to estimate the probability that an average borrower in a low income Chicago neighborhood will default. Based on the descriptive statistics for our Chicago sample, this borrower would have a FICO score of 696, a Walk Score of 65, and the home will be located in a block group with a per capita income of less than \$22,118.50. Based on the results of our model, the probability of default is:

$$F[20.2153 + 0.1712 \times \ln(65) - 3.3622 \times \ln(696)],$$

where  $F$  is the cumulative distribution function of the standard normal. This results in a default probability of 14.08%. Alternatively, we can also estimate the default probability for an average borrower in a high income San Francisco neighborhood. According to the descriptive statistics for our San Francisco sample, this borrower would have a FICO score of 745, have a home in a block group, a per capita income of greater than \$41,551, and a Walk Score of 84. According to our model, the probability of default is:

$$F[20.2153 + 0.1712 \times \ln(84) - 3.3622 \times \ln(745)].$$

This results in a default probability of 10.35%. Thus, our results indicate that significant differences in default probabilities across locations can be partially explained by income, Walk Score, and credit worthiness.

---

## Conclusion

Our findings indicate that location efficiency matters. Taken together, our results have significant policy implications with respect to lending practices, land use, and transportation planning.

We test our hypotheses regarding vehicle ownership and mortgage default and find that the probability of default does increase with vehicle ownership. Moreover, the impact of location efficiency (as measured by income-normalized vehicle ownership) on default is nontrivial. Consider the following example, of a borrower in Chicago with a FICO score of 680, a debt-to-income (back end) ratio of 41%, and an LTV of 80%. If the borrower is purchasing a home of average location efficiency for the Chicago metro area (0.0299 *vehper000 inc*), our model return default probability is 9.9%. A second borrower could have exactly the same underwriting characteristics but purchase a location-efficient home (0.0172 *vehper000 inc*). The default probability for this borrower falls to 7.2%. A third borrower with the same credit score and LTV could have a remarkably higher

debt-to-income ratio, of 62.5%, and, if they purchased the location-efficient home, the model would return a default probability of 9.9%—the same probability as for our first borrower. Our results also suggest that increases in *vehper000 inc* could compensate for substantial increases in LTV or decreases in FICO score. The results strongly suggest that researchers developing and refining automated mortgage underwriting models should seek to test the impact of location efficiency variables within their models.

It is interesting to note that there is an additional possible explanation for why mortgages for location-efficient homes perform better, beyond the household budgetary savings of lower vehicle ownership. This explanation is that location-efficient homes might hold their value better than other homes, and therefore better enable borrowers to avoid foreclosure through alternative measures (such as selling or refinancing the home) if they fall behind on payments or need to manage payment shock from an adjustable rate mortgage. Since the variable we use to measure default is whether the mortgage is in the foreclosure process or in REO status, we cannot be sure of the degree to which each of these two forces (household budget savings and home value performance) is contributing to the superior performance of location efficient mortgages. Future research on the linkages between vehicle ownership and default should explore these relationships more deeply.

We find highly nuanced results when we use Walk Score as an alternative measure of location efficiency. We find that it is closely related to income. Therefore, we find that its impact on mortgage performance differs across income categories. In low-income neighborhoods, Walks Score is positively related to mortgage default. We interpret this to mean that low-income, high default rate areas tend to be quite walkable (e.g., in urban areas within the central city). At the other extreme, we find that in high-income neighborhoods, Walk Score is negatively related to mortgage default. We interpret this as more of a location efficiency effect, with the causality flowing in the opposite direction as in low-income neighborhoods. That is, in high-income, walkable neighborhoods, borrowers have lower transportation costs and thus have more funds available to cover debt service thereby lessening the likelihood of default.

It is not clear from these results that it would be appropriate to utilize the Walk Score in mortgage underwriting, however. These results suggest that such a move could potentially reward borrowers in high income neighborhoods while punishing borrowers in low income neighborhoods if the interaction between income and walkability as measured by Walk Score were scored during the underwriting process. However, utilizing income-normalized vehicles per household as a proxy measure for location efficiency in underwriting criteria might avoid this issue, since it has a consistent effect in lowering default risk regardless of the income level of the neighborhood.

Moreover, our results provide support for policies supporting smart growth development and urban revitalization. That is, designing neighborhoods in such a way that reduces transportation needs is beneficial to borrowers as well as the environment.

Finally, we conclude that differences in location efficiency also help to explain the regional differences in mortgage performance often cited in academic literature. We examine three cities that vary greatly with respect to location efficiency, average borrower qualifications, and demographics. While significant empirical evidence exists to explain the various borrower qualifications and neighborhood characteristics that increase and decrease the probability of mortgage default, these results provide evidence that the degree of location efficiency present in a city or region may help to explain differences in mortgage performance.

---

## Endnotes

- <sup>1</sup> These costs include the cost to lease or purchase the car, maintenance, gas, insurance, and parking.
- <sup>2</sup> Center for Neighborhood Technology, on the Internet at [http://www.cnt.org/repository/heavy\\_load\\_10\\_06.pdf](http://www.cnt.org/repository/heavy_load_10_06.pdf). Additionally, the U.S. Bureau of Labor Statistics 2006 Consumer Expenditure Survey estimates annual household transportation costs at about \$8,500 per year, \$8,000 of which are automobile related.
- <sup>3</sup> The Affordability Index: A New Tool for Measuring the True Affordability of a Housing Choice. Center for Transit Oriented Development and Center for Neighborhood Technology. Brookings Institution, Market Innovation Brief, January 2006.
- <sup>4</sup> For detailed data and maps of the Housing-Transportation Affordability Index, visit the Center for Neighborhood Technology website at: <http://www.cnt.org/tcd/ht>.
- <sup>5</sup> See <http://www.locationefficiency.com/> for information about a pilot project in four U.S. cities (Seattle, San Francisco, Los Angeles, and Chicago) that has had several participating lenders.
- <sup>6</sup> The Joint Economic Committee estimates the average cost of a foreclosure nationally at \$79,000. Sheltering Neighborhoods from the Subprime Foreclosure Storm. Special report from the Joint Economic Committee. April 2007. Lenders are estimated to incur costs of an average of 26% of the loan amount in a foreclosure (Parisi and Uppuluri, 2008). The Anatomy of Loss Severity Assumptions in US Subprime RMBS. Standard and Poors.com, May 7, 2008.
- <sup>7</sup> See the National Vacant Properties Campaign (2005). Vacant Properties: The True Costs to Communities. [www.vacantproperties.org](http://www.vacantproperties.org).
- <sup>8</sup> One study estimates that each foreclosure within 1/8 mile of a home reduces that home's value by between 0.9% and 1.136% (Immergluck and Smith, 2005).
- <sup>9</sup> Smart Growth principles include creating a range of housing opportunities and choices; creating walkable neighborhoods; encouraging community and stakeholder collaboration; fostering distinctive, attractive communities with a strong sense of place; making development decisions predictable, fair, and cost effective; mixing land uses; preserving open space, farmland, natural beauty, and critical environmental areas; providing a variety of transportation choices; strengthening and directing development toward existing communities; and taking advantage of compact building design. See the Smart Growth Network at <http://www.smartgrowth.org/about/principles/default.asp>.

- <sup>10</sup> Census block groups are the smallest geographic level at which detailed Census data is available; they generally contain between 600 and 3,000 people, with an optimum size of 1,500. Block group boundaries are usually drawn with the input of local governments. See [www.census.gov](http://www.census.gov).
- <sup>11</sup> We used median household income as the income variable and divided this value by 1,000 in order to avoid generating a very small number as the resulting ratio.
- <sup>12</sup> For a detailed description of Walk Score, visit [www.walkscore.com](http://www.walkscore.com).
- <sup>13</sup> For specific information about the Walk Score Algorithm methodology, see [www.walkscore.com/rankings/ranking-methodology.shtml](http://www.walkscore.com/rankings/ranking-methodology.shtml).
- <sup>14</sup> Commonly, DTI is expressed as a pair of ratios X/Y with the first ratio representing housing-related debt and the second ratio representing all debt payments. LPS provides only the ratio of debt payments related to the subject loan to borrower income.
- <sup>15</sup> Twelve percent of the FICO scores in Chicago, 20% in Jacksonville, and 6% in San Francisco are missing in our sample.
- <sup>16</sup> The difference in the ages the three samples—especially Jacksonville as compared to Chicago and San Francisco—is largely due to the process that we used to match the mortgage data to the property transfer data. In Jacksonville, there were fewer total observations in the property transfer database and we were therefore able to match the entire property transfer database to our loan database; however, in Chicago and San Francisco, there were many more property transfers. We therefore chose to match our mortgage sample to property transfers between January 2004 and December 2008. We chose this timeframe because LPS Analytics' coverage is limited prior to 2004.
- <sup>17</sup> We code a mortgage as in default if it is in foreclosure or REO in order to analyze extreme cases of default. We attempted an OLS regression with default rates by block group as the dependent variable. After eliminating all block groups with less than 30 observations (in order to maintain statistical significance) our sample was reduced by more than 80%, making this analysis infeasible.
- <sup>18</sup> The median per capita income for the combined dataset is \$28,951. We assign each observation to quartiles accordingly such that quartile 1 observations are located in block groups with per capita income up to \$22,118.50, quartile 2 is between \$22,118.51 and \$28,951, quartile 3 is between \$28,952 and \$41,550, and quartile 4 is block groups with per capita income greater than \$41,551. There is approximately the same number of observations in each quartile.
- <sup>19</sup> The *minblkgrp* variable is created by coding *minblkgrp* to 1 if the percentage of the population in that block group that is non-white is 75% or more and 0 otherwise. The *whiteblkgrp* variable is created by coding *whiteblkgrp* to 1 if the percentage of the population in that block group that is non-white is less than 25% and 0 otherwise.
- <sup>20</sup> Recall that DTI was unavailable for the Jacksonville sample.
- <sup>21</sup> In our sample, the median household income is \$64,579.85, so this borrower would own approximately 1.29 vehicles.
- <sup>22</sup> Recall that DTI was not available for the Jacksonville sample.

---

## References

Alexander, W.P., S.D. Grimshaw, G.R. McQueen, and B.A. Slade. Some Loans are More Equal than Others: Third-Party Originations and Defaults in the Subprime Mortgage Industry. *Real Estate Economics*, 2002, 30:4, 667–97.

Blackman A. and A. Krupnick. Will Location Efficient Mortgages Raise Default Rates? Resources for the Future Discussion Paper 99-49-REV, 2000.

Brookings Institution. The Affordability Index: A New Tool for Measuring the True Affordability of a Housing Choice. Center for Transit Oriented Development and Center for Neighborhood Technology. Brookings Institution, Market Innovation Brief, 2006.

Ding, L., R. Quercia, J. Ratcliffe, and W. Lei. Risky Borrowers or Risky Mortgages: Disaggregating Effects Using Propensity Score Morels. Working Paper, Center for Community Capital, 2008.

Elul, R. Residential Mortgage Default. Federal Reserve Bank of Philadelphia Business Review, 2006, Third Quarter, 21–30.

Foster, C. and R. Van Order. An Option-Based Model of Mortgage Default. *Housing Finance Review*, 1984, 3:4, 351–72.

Goetzmann, William and Matt Spiegel (1997). “A Spatial Model of Housing Returns and Neighborhood Substitutability.”

Haugwout, A., R. Peach, and J. Tracy. Juvenile Delinquent Mortgages: Bad Credit or Bad Economy? Federal Reserve Bank of New York Staff Reports, No. 341, August 2008.

Holtzclaw, J. Smart Growth—As Seen from the Air: Convenient Neighborhood, Skip the Car. Unpublished Working Paper, 2000.

Holtzclaw, J., R. Clear, H. Dittmar, D. Goldstein, and P. Haas. Location Efficiency: Neighborhood and Socio-Economic Characteristics Determine Auto Ownership and Use—Studies in Chicago, Los Angeles and San Francisco. *Transportation Planning and Technology*, 2002, 25, 1–27.

Immergluck, D. and G. Smith. There Goes the Neighborhood: The Effect of Single-Family Mortgage Foreclosures on Property Values. Chicago: The Woodstock Institute, 2005.

Kau, J., D. Keenan, and T. Kim. Transaction Cost, Suboptimal Termination, and Default Probabilities for Mortgages. *Journal of the American Real Estate and Urban Economics Association*, 1993, 21: 247–63.

Kau, J., D. Keenan, W. Muller, and J. Epperson. The Valuation at Origination of Fixed Rate Mortgages with Default and Prepayment. *Journal of Real Estate Finance and Economics*, 1995, 11, 5–36.

Francis P. and S. Uppuluri. The Anatomy of Loss Severity Assumptions in US Subprime RMBS. Standard and Poors.com, May 7, 2008.

Quercia, R.G. and M.A. Stegman. Residential Mortgage Default: A Review of the Literature. *Journal of Housing Research*, 1992, 3:2, 341–76.

Vandell, K. How Ruthless is Mortgage Default? A Review and Synthesis of the Evidence. *Journal of Housing Research*, 1995, 6:2, 245–64.

Vandell, K., J. Kau, D. Keenan, and T. Kim. Default Probabilities for Mortgages. *Journal of Urban Economics*, 1994, 35:3, 278–96.

Vandell, K. and T. Thibodeau. Estimation of Mortgage Defaults Using Disaggregate Loan History Data. *Journal of the American Real Estate and Urban Economics Association*, 1985, 13:3, 292–316.

*The authors acknowledge the support of the Natural Resources Defense Council. Rauterkus acknowledges the generous support of the Federal Reserve Bank of Atlanta*

*for this project. The views expressed in this paper are those of the authors and do not necessarily represent those of the Natural Resources Defense Council, the Federal Reserve Bank of Atlanta or the Federal Reserve System. The authors are especially grateful to Norman Miller for helpful comments and suggestions.*

Stephanie Y. Rauterkus, University of Alabama at Birmingham, Birmingham, AL 35294-4460 or [srauter@uab.edu](mailto:srauter@uab.edu).

Grant I. Thrall, University of Florida, Gainesville, FL 32611 or [grant@thrall.us](mailto:grant@thrall.us).

Eric Hangen, I<sup>2</sup> Community Development Consulting, Inc., Cranston, RI 02905 or [ehangen@i2community.org](mailto:ehangen@i2community.org).

