Location-Efficient Mortgages: Is the Rationale Sound?

Author(s): Allen Blackman and Alan Krupnick

Source: *Journal of Policy Analysis and Management*, Autumn, 2001, Vol. 20, No. 4 (Autumn, 2001), pp. 633-649

Published by: Wiley on behalf of Association for Public Policy Analysis and Management

Stable URL: https://www.jstor.org/stable/3325776

Location-Efficient Mortgages: Is the Rationale Sound?

Allen Blackman
Alan Krupnick

Abstract

Location efficient mortgage (LEM) programs are an increasingly popular approach to combating urban sprawl. LEMs allow families who want to live in densely populated, transit-rich communities to obtain a larger mortgage with a smaller down payment than traditional underwriting guidelines allow. LEMs are premised on the proposition that homeowners in such "location-efficient" areas can safely be allowed to breach underwriting guidelines designed to prevent mortgage default because they have lower than average automobile-related transportation expenses and more income available for mortgage payments. This paper employs records of more than 8000 FHA-insured mortgages matched with data on various measures of location efficiency to test this proposition. The results suggest that it does not hold and that LEMs—like other low-down-payment mortgage programs—will raise mortgage default rates. This cost must be weighed against any potential anti-sprawl benefits LEMs may have. © 2001 by the Association for Public Policy Analysis and Management.

INTRODUCTION

Recent polls indicate that urban sprawl is now Americans' top local concern, edging out traditional issues such as crime and jobs (Pew Center, 2000). As apprehension about sprawl and attendant problems of traffic congestion, inadequate infrastructure, and auto emissions has grown, policymakers have responded with a host of purported remedies (Eggen, 1998), the location-efficient mortgage (LEM) among them. LEMs allow families who want to live in densely populated, transit-rich communities to obtain a larger mortgage with a smaller down payment than traditional underwriting guidelines allow. Advocates claim LEMs will curb sprawl by making homes in "location-efficient" communities more affordable to low- and moderate-income borrowers who would ordinarily be forced to live in less-expensive suburbs or exurbs. Moreover, they claim, even though LEMs breach underwriting guidelines designed to prevent mortgage default, they will not raise default rates because families living in location-efficient areas have lower than average automobile-related transportation expenses and therefore more income available for mortgage payments.

LEMs have a number of attractive features compared with conventional anti-sprawl policies such as altering local zoning codes, providing government funding to develop

Manuscript received April 2000; revise and resubmit recommended July 2000; revised August 2000; accepted April 2001.

Journal of Policy Analysis and Management, Vol. 20, No. 4, 633–649 (2001)

Published by John Wiley & Sons, Inc.

abandoned inner-city sites, and linking federal highway dollars to land use goals. While zoning codes involve some measure of government fiat, LEMs simply create economic incentives for desired land-use behavior but leave the ultimate decisions in the hands of the private sector. Also, if LEMs work as advertised and do not raise mortgage default costs, they will require virtually no additional expenditures by either the public or private sector.

LEMs appear to be well on their way to achieving widespread acceptance. Several federal agencies have funded the development of this new policy, and LEMs have figured prominently in national anti-sprawl and climate change initiatives.[1] In the summer of 1999, Fannie Mae, the nation's largest secondary mortgage institution, launched a $100 million pilot project making LEMs available in several large metropolitan areas including Chicago, Los Angeles, Portland, San Francisco, and Seattle. This initiative has been well-received in the national and local media (Allen, 1998; Woellert, 1999). LEM programs are under active discussion in several other cities (*Orlando Sentinel*, 1999; Rimer, 1999).

But will LEMs work as advertised? Ultimately, the effectiveness of LEMs in slowing sprawl will depend on their ability to redirect new development away from fringe areas to location-efficient areas. But even if LEMs do this, their viability as a policy instrument is questionable if they significantly raise default rates. If this happens, secondary mortgage institutions will be less likely to embrace LEMs, and primary mortgage lenders will be less likely to promote them. According to Avery and colleagues (1996, p. 641), even relatively small increases in default rates can make a mortgage program targeted to low- and moderate-income borrowers unprofitable.

Since LEMs have only recently been introduced, there are no empirical studies to provide guidance on their likely effect on default rates. LEMs have only just begun to receive attention in the academic literature. Danielsen, Lang, and Fulton (1999, p. 533) contains a general discussion of LEMs that effectively endorses the notion that they will not affect default rates. According to these authors:

> [LEMs] recognize that homebuyers who need only one car actually do have greater financial capacity, thus acknowledging that not all houses are alike even if they are "comparable" in traditional real estate appraisal terms. The Natural Resources Defense Council found that significant savings accrue from living in higher density neighborhoods that feature public transit and pedestrian access to everyday services (Benfield, Raimi, and Chen, 1999). Mortgage markets should consider these savings when calculating loan risk.

In a reply to this article, Easterbrook (1999) argues that LEMs are not likely to generate significant new demand for houses in location-efficient areas, but does not take up the issue of whether LEMs will affect mortgage default.

This paper presents the results of an empirical test designed to assess the proposition that LEMs will not exacerbate mortgage default. Since repayment histories for LEMs will not be available for several years, an indirect test is used based on the following logic. If it is true, as LEM advocates claim, that homeowners in location-efficient areas can safely be allowed to violate traditional underwriting guidelines because they have below-average transportation expenses and more funds available for mortgage payments, then historical records of repayment on conventional mortgages should show that on average, borrowers in location-efficient areas have lower default rates than similar borrowers with similar mortgages in other areas. In other words, there should be a statistically significant negative correlation between the location

[1] See for example a Vice Presidential speech on the "Livable Communities of the Twenty-first Century" initiative, available at www.smartgrowth.org/library/gore_pr9298.html.

efficiency of a home and the probability of mortgage default, all other things equal. A finding that no such relationship exists would suggest that allowing borrowers who live in location-efficient areas to breach conventional underwriting guidelines would have precisely the same effect as allowing randomly selected borrowers to breach these guidelines: It would raise default rates.

To test for a relationship between location efficiency and mortgage default, records of more than 8000 Federal Housing Administration (FHA)–insured mortgages originated in Chicago between 1988 and 1992 were matched with census tract–level data on location efficiency used by lenders issuing LEMs. Regression analysis indicates that, all other things equal, the relationship between location efficiency and the probability of mortgage default is not statistically significant. Although the analysis is not conclusive—it only indirectly tests the proposition that LEMs will not raise default rates, and it is based on a geographically and temporally limited sample of mortgage records—the authors believe their findings cast considerable doubt on the notion that LEMs will be a virtually costless anti-sprawl policy.

The empirical literature on the determinants of mortgage default underpins this analysis. A number of researchers have found an association between the probability of default and location, all other things equal, although none have looked specifically at the relationship between the probability of default and location efficiency. In an analysis of thousands of loans originated in Pittsburgh, Pennsylvania, von Fustenberg and Green (1974) find that suburban location reduced delinquency risks by as much as 47 percent compared with urban location, all other things equal, a result that appears to contradict the argument of LEM advocates. However, Mills and Lubuele (1994) conclude that residential mortgages on single-family properties in low- and moderate-income neighborhoods outperformed those in a locationally diverse national sample. Using a much larger data set, Van Order and Zorn (1996) find that neighborhood income is generally negatively related to default, all other things equal.

While the effect of locational characteristics on the probability of default remains something of an open question in the literature, the effect of loan-to-value ratios on the probability of default is not. Empirical studies consistently find that higher loan-to-value ratios (smaller down payments) significantly increase the probability of default. For example, empirical analysis of FHA– and Department of Veterans' Affairs (VA)–insured mortgages shows that loan-to-value ratio at the time the loan is originated is the most important predictor of default. When loan-to-value ratios are raised from 90 to 97 percent—the maximum ratio for LEMs—default rates increase sevenfold (Quercia and Stegman, 1992, p. 349; see also Berkovec et al., 1998; Deng et al., 1996; Van Order and Zorn, 1996). Higher than standard debt-to-income and housing-expense-to-income ratios have also been shown to exacerbate default risk (Avery et al., 1996, p. 644; Quercia and Stegman, 1992, p. 350).

BACKGROUND

In 1995, three nonprofit organizations—the Center for Neighborhood Technology in Chicago, the Natural Resources Defense Council (NRDC) in San Francisco, and the Surface Transportation Policy Project in Washington, DC—formed a consortium to develop LEMs.[2] The consortium was funded by the Department of Transportation, the Department of Energy, the Environmental Protection Agency, and several private foundations.

[2] The inspiration for LEMs was an NRDC-sponsored econometric study that found a strong negative relationship between both vehicle miles traveled and auto ownership on one hand, and residential density (the number of dwelling units per square mile) and transit accessibility on the other hand (Holtzclaw, 1994).

After several years of refinement, LEMs were unveiled in 1998. They are 15- to 30-year fixed interest mortgages of as much as $240,000 on a one-unit, owner-occupied house or condominium. They allow borrowers to "stretch" traditional lending guidelines that mandate a minimum down payment in the range of 5 to 20 percent of the appraised property value (or equivalently, a maximum loan-to-value ratio of 80 to 95 percent), a maximum housing-expense-to-income ratio of 28 percent, and a maximum debt-to-income ratio of 36 percent. LEMs allow a down payment as low as 3 percent (or equivalently, a loan-to-value ratio as high as 97 percent), a housing-expense-to-income ratio as high as 35 percent, and a debt-to-income ratio as high as 45 percent.[3]

The actual terms of individual LEMs are determined by computer models developed by the LEMs consortium for each city in which the new loans are made. For any given property in one of these cities, the model assesses the location efficiency of the property and then estimates the dollar savings in automobile-related expenditures a prospective owner would enjoy. This location efficiency value (LEV) is added to a mortgage applicant's income in calculating the housing-expense-to-income and debt-to-income ratios that determine the maximum mortgage amount. The end result is that borrowers who apply for a loan on a home in a location-efficient areas can "get a larger mortgage than possible with any other product now on the market" (CNT, 2000), presumably enabling them to afford such homes.

The LEMs consortium's promotional literature includes the following hypothetical example. A loan applicant with an income of $2100 per month, no long-term debt, and $6000 for a down payment wishes to purchase a $105,000 home in a location-efficient area. The conventional 28 percent limit on the housing-expense-to-income ratio implies that the maximum amount the loan applicant can borrow is $76,058, too little to afford the home. But according to the LEM computer model, living in this particular home would enable the borrower to save $653 per month in automobile-related transportation costs relative to living in a location-inefficient home. Adding this savings to the applicant's monthly income in calculating the housing-expense-to-income ratio enables the borrower to get a $115,611 mortgage, more than sufficient to purchase the home.

It is important to note that because LEMs enable borrowers with a fixed amount of funds available for a down payment to obtain larger loans, they have the effect of reducing the down payment as a percentage of the appraised property value (that is, raising the loan-to-value ratio). For example, in the hypothetical case described above, the down payment is 8 percent without a LEM and 5 percent with the LEM.

The LEM consortium's computer models calculate LEVs in three steps (see the Appendix). First, they use an econometric model (Holtzclaw, 1994) to predict both vehicle miles traveled and number of autos owned for a given residence using six independent variables. Two of these independent variables—household income and number of persons in the household—are specific to the loan applicant. The remaining

The study was seen as evidence that: (i) vehicle miles traveled, along with auto emissions and traffic congestion could be reduced by creating incentives for people to live in high-density, transit-rich areas, and (ii) families applying for a mortgage in location-efficient areas could safely be given larger loans than traditional underwriting guidelines allow, since they would have lower than average automobile-related transportation expenses. Note that Holtzclaw's findings are controversial. Schimek (1996) suggests that vehicle miles traveled are not very responsive to change in residential density except at very high densities (approaching 10,000 persons per square mile).

[3] LEMs are a new and still-evolving mortgage product. There has been some discussion of modifying them by, for example, requiring borrowers to purchase a long-term transit pass. This paper tests whether the bare-bones idea that underpins LEMs—location efficiency reduces default risk—stands on its own.

four independent variables—households per residential acre, households per total acre, pedestrian factor, and transit access—relate to the census tract in which the home is located. Second, given econometrically estimated vehicle miles traveled and auto ownership, auto expenses are calculated using FHA figures on the cost of owning and operating an automobile (FHA, 1992). Finally, automobile expenses for the applicant's household are subtracted from the base case—automobile expenses for a household of similar size and wealth in a neighborhood with relatively low density, poor transit access, and low pedestrian friendliness.

ECONOMETRIC MODEL AND DATA

Econometric models of mortgage default fall into two categories: those that adopt the lender's perspective and those that adopt the borrower's perspective (Quercia and Stegman, 1992). Both types of models use similar independent variables to explain the probability of default (for example, the loan-to-value ratio, the market value of the property, and the borrower's income). However, the former uses only information the lender has at the time the loan is originated, while the latter uses information available at the time of origination and in every payment period afterward. Not surprisingly, borrower's perspective models more accurately predict default and also better reflect how default decisions are really made.[4] Nevertheless, the lender's perspective model is employed here. This is appropriate because the purpose of this paper is not to understand the role location efficiency plays in the borrower's decision to default. Rather, it is to test whether a specific piece of information that the lender has at the time the loan is originated—the location efficiency of the property—has value as a predictor of default and should be taken into account in deciding whether to allow the borrower to breach traditional lending guidelines.

A simple logit regression is used to test for a relationship between the probability of default and various measures of location efficiency, holding constant as many mortgage, property, and borrower characteristics as the data permit. For each annual cohort of mortgages in the data, the following equation is estimated:

$$\mathbf{P} = \frac{\exp(\mathbf{X}\beta + \mathbf{Z}\gamma)}{1 + \exp(\mathbf{X}\beta + \mathbf{Z}\gamma)}$$

where,

- $\mathbf{P}$ is the probability of default,
- $\mathbf{X}$ is a vector of locational characteristics,
- β is a vector of parameters,
- $\mathbf{Z}$ is a vector, mortgage, property, and borrower characteristics, and
- γ is a vector of parameters.

[4] From a theoretical perspective, a consensus now exists that mortgage default is best viewed as a put option, that is, an option to "sell" the home to the lender for the value of the mortgage at the beginning of each payment period. Whether or not the borrower exercises this option in each period depends on contemporaneous information including: the market value of the house, the market value of the mortgage, the transactions costs associated with default, expectations about future market conditions that affect the value of waiting to exercise the option, and "triggering events" like unemployment, illness, and divorce (Avery et al., 1996, p. 623; Quercia and Stegman, 1992).

Separate regressions are run for each annual cohort to control for differences in the period during which default can be observed. One complication is that some of the locational characteristics included in the vector **X** may be correlated with each other. To control for collinearity, regression results are reported for a number of different specifications of the model.

The data consist of records of more than 8000 FHA-insured mortgages originated in the greater Chicago metropolitan area in 1988, 1989, 1991, and 1992, matched with census tract–level information on location efficiency in this area. Data from 1990 were not used because only 190 mortgage records from that year are available. Data availability also dictated focusing on a single city. At the time the analysis was conducted, the LEM consortium had only tabulated location-efficiency data for Chicago. In addition, matching mortgage data with location-efficiency data proved to be resource intensive. Spatially, the data are heterogeneous: The mortgage property sites are in more than 1000 different census tracts in urban and suburban Chicago.

The mortgage records—including detailed information on borrower, mortgage, and property characteristics—were obtained from the Department of Housing and Urban Development. The data on location efficiency were developed by the LEM consortium and come from the 1990 census, the Chicago Metropolitan Planning Organization, and the Chicago Transit Authority. Detailed information on the location-efficiency variables is provided in the Appendix. Table 1 defines the variables used in the regression analysis and Table 2 gives summary statistics.

The data on FHA-insured mortgages are well suited to this analysis for a number of reasons. First, as Berkovec and coauthors (1998) point out, they are particularly appropriate for analyzing the determinants of default since the incidence of default on FHA-insured mortgages is relatively high. The incidence of default in the pooled sample used here is 8 percent, with higher rates in earlier years and lower rates in later years (Table 2).[5] By contrast, the incidence of default is just 2 percent in a sample of more than 400,000 loans, originated between 1975 and 1983 and purchased by Freddie Mac (Van Order and Zorn, 1996). Second, underwriting guidelines for FHA-insured loans and LEMs are somewhat similar and as a result, one would expect the two programs to attract somewhat similar borrowers. Both types of mortgages allow a lower down payment and higher ratios of housing-expense-to-income and debt-to-income than do conventional mortgages. In addition, LEMs and FHA-insured mortgages have similar lending limits and target first-time homebuyers, among other borrowers.[6] Finally, the FHA-insured data are well suited to this analysis because they are extremely rich. As a result, it is possible to control for 18 borrower, mortgage, and property characteristics.

Summary statistics presented in Table 3 foreshadow the regression results. The table presents default rates for two subsamples of mortgages in each annual cohort: those that have an above-average LEV and LTV ratio—characteristics that make them roughly equivalent to LEMs—and those that have a below-average LEV and LTV

[5] The main reason more recent annual cohorts have lower default rates is simply because there are fewer years between the time the loan was originated and the time the data were collected (1997).

[6] FHA debt-to-income (DTI) and housing-expense-to-income (HEI) ratios are defined unconventionally: Income is gross income used to calculate conventional lending ratios less federal withholding tax; housing expenses include principal, interest, taxes, and insurance used to calculate conventional ratios plus expenses for utilities and maintenance; and debt payments include expenses for housing, plus any debt that requires at least 12 months of payments, revolving debt, city and state income taxes, and Social Security payments. By these definitions, the FHA HEI limit is 38 percent and the FHA DTI limit is 53 percent. The LTV limit is 97 percent. In our sample of FHA borrowers, the average HEI ratio is 24 percent, the average DTI ratio is 43 percent, and the average LTV ratio is 92 percent. More than three-quarters of the borrowers in our sample are first-time homebuyers. The maximum FHA loan amount in the Chicago primary metropolitan statistical area is $209,000.

Table 1. Definitions of variables.

Variable	**Definition**
Dependent	
DEFAULT	1 if borrower defaults, 0 otherwise
Location	
HH_DENS	Households per residential acre in census tract
LEV	Location efficiency value
PED_FACT	Pedestrian friendliness factor in census tract
RAW_DENS	Households per total acre in census tract
URBAN	1 if property is in an urban area, 0 otherwise
ZONAL_TR	Zonal transit access in census tract
Mortgage	
DTI	Debt-to-income ratio
HEI	Housing-expense-to-income ratio
LTV	Loan-to-value ratio
Property	
CONDO	1 if property is a condominium, 0 otherwise
HVAL	Appraised value of the property at time of purchase
HVAL2	HVAL squared
Borrower	
AGE	Borrower age
ASI_AMIN	1 if Asian or American Indian borrower, 0 otherwise
BLACK	1 if African American borrower, 0 otherwise
DEPNUM	Number of dependents (excluding borrower and co-borrower)
FIRSTBUY	1 if borrower is a first-time homebuyer, 0 otherwise
HISPANIC	1 if Hispanic borrower, 0 otherwise
INCOME	Total annual effective family income
INCOME2	Income squared
LQASS	Liquid assets available at closing
LQASS2	Liquid assets squared
NOCBINC	1 if no co-borrower or co-borrower income is zero, 0 otherwise
REPEATBUY	1 if borrower is not a first-time homebuyer, 0 otherwise
SINGLEF	1 if borrower is a female and there is no co-borrower, 0 otherwise
SINGLEM	1 if borrower is a male and there is no co-borrower, 0 otherwise
WHITE	1 if white borrower, 0 otherwise

ratio.[7] For each annual cohort, the mean default rate for the first subsample is significantly greater than that for the second. For example, for 1988, the LEM proxies have a default rate of 17 percent while mortgages with below-average LEV and LTV ratios have a default rate of 3 percent. This would seem to suggest that any desirable effect that greater location efficiency has on default rates is swamped by the undesirable effect of higher LTV ratios, so that loans with a high LEV and a high LTV ratio will likely raise default rates.

Although interesting, these summary statistics are merely suggestive. Since LEV and LTV ratios are correlated with other determinants of mortgage repayment, such as the borrower's liquid assets and number of dependents, the higher default rates on the LEM proxies (Table 3) may be driven by omitted variables. For example, persons living in high-LEV areas may have more dependents and may default more often as a result. The regression analysis below avoids this problem by assessing the independent effects on default of a wide range of factors.

[7] We credit an anonymous referee with suggesting this test.

Table 2. Summary statistics.

Variable	1988 n = 2337		1989 n = 1683		1991 n = 898		1992 n = 3225		All n = 8143	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Dependent										
DEFAULT	0.10	0.31	0.09	0.28	0.08	0.26	0.06	0.23	0.08	0.27
Location										
HH_DENS	11.17	54.06	9.34	19.21	13.22	72.36	11.59	49.38	11.18	49.58
LEV	114.15	94.26	109.69	92.81	113.65	94.67	114.20	98.99	113.19	95.91
PED_FACT	0.84	0.45	0.81	0.44	0.82	0.44	0.81	0.44	0.82	0.44
RAW_DENS	5.19	6.95	5.00	7.05	5.07	7.05	5.34	7.51	5.20	7.21
URBAN	0.33	0.47	0.30	0.46	0.32	0.46	0.30	0.45	0.31	0.46
ZONAL_TR	19.97	95.49	16.87	51.71	26.81	159.26	19.87	88.18	20.05	95.13
Mortgage										
DTI	50.66	8.29	52.00	6.99	34.97	6.35	34.92	6.42	42.97	10.82
HEI	22.44	6.52	23.82	6.14	23.85	5.47	24.26	5.79	23.60	6.09
LTV	92.00	6.76	91.78	6.38	92.66	7.55	92.58	7.15	92.26	6.94
Property										
CONDO	0.12	0.32	0.12	0.33	0.15	0.36	0.15	0.36	0.14	0.34
HVAL	69.11	20.25	75.70	23.36	81.50	25.79	91.15	26.57	80.57	25.89
HVAL2	5186.04	3086.87	6275.24	3842.72	7306.58	4796.90	9014.00	5021.30	7161.06	4576.83
Borrower										
AGE	33.52	8.97	33.47	9.23	33.51	8.98	33.75	9.37	33.60	9.19
ASI_AMIN	0.03	0.16	0.03	0.17	0.02	0.14	0.02	0.15	0.02	0.15
BLACK	0.28	0.45	0.23	0.42	0.25	0.43	0.22	0.41	0.24	0.43
DEPNUM	1.07	1.26	1.03	1.29	0.97	1.26	1.02	1.31	1.03	1.29
FIRSTBUY	0.78	0.41	0.80	0.40	0.81	0.39	0.75	0.43	0.78	0.42
HISPANIC	0.15	0.35	0.16	0.37	0.20	0.40	0.19	0.40	0.17	0.38
INCOME	39.81	14.40	41.64	14.32	41.87	14.15	44.99	111.31	42.47	70.95
INCOME2	1791.67	1595.62	1938.40	1653.89	1952.42	1404.65	14,409.35	698,436.40	6836.90	439,544.50
LQASS	8.39	9.83	9.39	10.44	9.38	12.19	9.80	11.01	9.27	10.72
LQASS2	166.48	593.19	196.63	885.48	235.94	1142.00	216.74	1013.80	200.28	902.31
NOCBINC	0.28	0.45	0.28	0.45	0.36	0.48	0.35	0.48	0.31	0.46
REPEATBUY	0.22	0.41	0.20	0.40	0.19	0.39	0.25	0.43	0.22	0.42
SINGLEF	0.12	0.32	0.12	0.32	0.15	0.36	0.17	0.37	0.14	0.35
SINGLEM	0.16	0.37	0.16	0.36	0.20	0.40	0.18	0.38	0.17	0.38
WHITE	0.55	0.50	0.58	0.49	0.53	0.50	0.56	0.50	0.56	0.50

Table 3. Mean default rates by year, and LEV and LTV (n).

	1988*	**1989***	**1991***	**1992***
Above-average LEV & LTV	0.165 (665)	0.133 (466)	0.100 (251)	0.080 (869)
Below-average LEV & LTV	0.030 (399)	0.035 (339)	0.032 (157)	0.028 (569)

*Difference in means statistically significant at 1 percent level.

RESULTS

Table 4 presents results for the four econometric models. In each model, the dependent variable is DEFAULT, a dummy variable that takes the value of 1 if the borrower defaults and 0 otherwise.[8] Model 1 is intended as a benchmark. It includes only the control variables—that is, the mortgage, property, and borrower characteristics—and excludes the location variables that are the focus of interest here. If a statistically significant regression coefficient in two or more annual cohorts is interpreted as implying an independent variable is generally correlated with the probability of default, then the results mirror those found in the literature (for example, Berkovec et al., 1998). With regard to mortgage characteristics, the probability of default is higher for mortgages with higher loan-to-value ratios. With regard to property characteristics, the probability of default is lower for properties with higher appraised property values. With regard to borrower characteristics, the probability of default is higher for borrowers who are African American, have more dependents, are first time homebuyers, have fewer liquid assets, and are single males.

Note that the probability of default is negatively correlated with the debt-to-income ratio and is ambiguously correlated with income. Although counterintuitive, such findings are not uncommon (e.g., Avery et al., 1996, p. 624; Vandell and Thibodeau, 1985, p. 309). They likely stem from the well-documented "U-shaped" effect of income on the probability of default (Deng et al., 1996, p. 274). In addition, the first result may stem from debt-to-income ratio thresholds written into the lending guidelines (Quercia and Stegman, 1992, p. 350).[9]

Model 2 regresses DEFAULT onto LEV—the transportation cost savings estimated by the LEM consortium's computer model—controlling for 18 mortgage, borrower, and property characteristics. In this model, LEV serves as an index of location efficiency that presumably captures residential density, transit access, and pedestrian friendliness. As Table 4 illustrates, LEV is significant in two annual cohorts (1989 and 1992) at the 10 percent level. However, the sign of the coefficient on LEV in both of these years is positive, implying that greater location efficiency is correlated with higher rates of mortgage default. This unexpected sign may arise because LEV acts as a proxy for a missing variable that is positively correlated with default. The most likely candidate is urban location. By definition, LEVs are generally higher in urban

[8] Our definition of default includes lender foreclosure as well as conveyance of the title to the lender or insurer in lieu of foreclosure. This definition of is commonly used in empirical studies of mortgage default (Quercia and Stegman, 1992, p. 343).
[9] It was not possible to control for these nonlinearities by using a series of dummy variables as in Berkovec et al. (1998) because there were too few observations.

Table 4. Logit regression result (dependent variable = DEFAULT).

Variable	Model 1				Model 2			
	1988 (n = 2337)	1989 (n = 1683)	1991 (n = 898)	1992 (n = 3225)	1988 (n = 2337)	1989 (n = 1683)	1991 (n = 898)	1992 (n = 3225)
Location								
LEV					0.00007 (0.073)	0.00193* (1.779)	−0.00205 (−1.136)	0.00183* (1.928)
HHDENS								
PED_FACT								
RAW_DENS								
URBAN								
ZONAL_TR								
Mortgage								
DTI	−0.03211***	−0.02175	−0.03953*	−0.00401	−0.03204***	−0.01941	−0.04066*	−0.00215
HEI	0.02472	0.07596**	0.08953	−0.03464	0.02477	0.07534**	0.08691	−0.03398
LTV	0.03734**	0.0452*	−0.00955	0.02792*	0.0374**	0.0466*	−0.01094	0.02864*
Property								
CONDO	−0.85981**	−1.15303**	−0.72759	−0.36466	−0.85986**	−1.21797***	−0.71806	−0.42474
HVAL	−0.00069	−0.05595**	−0.00391	−0.03284*	−0.00071	−0.05233**	−0.00645	−0.0322*
HVAL2	−0.00007	0.00021	−0.00003	0.00017**	−0.00007	0.00018	−0.00002	0.00017*
Borrower								
AGE	0.0059	0.00836	0.00652	0.00305	0.00583	0.00731	0.00807	0.00153
ASI_AMIN	0.28535	0.88152	0.69574	0.5552	0.28272	0.92348	0.78928	0.49909
BLACK	1.50553***	1.19737***	1.75441***	0.96857***	1.50195***	1.10681***	1.87963***	0.85648***
DEPNUM	0.19006***	0.166**	0.06072	0.21014***	0.18955***	0.14815**	0.07713	0.19273***
FIRSTBUY	−0.05011	−0.24002	−0.5947*	0.36865*	−0.05097	−0.27609	−0.58212*	0.34435
HISPANIC	−0.08513	−0.21497	0.2351	0.19418	−0.09071	−0.39273	0.43265	0.00946
INCOME	−0.03408	0.10108*	0.01769	−0.0295**	−0.03409	0.09476*	0.02667	−0.03048**
INCOMEZ	0.00021	−0.00096*	−0.00024	0.000004*	0.00021	−0.0009	−0.00035	0.000005*
LQASS	−0.1339***	−0.07729***	−0.10949***	−0.03502**	−0.13393***	−0.07813***	−0.11146***	−0.03589**
LQASS2	0.00174***	0.00054**	0.00069**	0.00028**	0.00174***	0.00054**	0.00069**	0.00029**
NOCBINC	−0.04847	0.21964	−1.13815**	−0.12258	−0.04767	0.26168	−1.167**	−0.08092
SINGLEM	0.63898**	0.14916	1.06484**	0.47322*	0.63906**	0.11422	1.06381**	0.47523*
Intercept	−3.6548*	−7.029**	−1.8046	−2.6641	−3.6656*	−7.3448**	−1.5675	−2.8495*
Pseudo R^2	0.1892	0.1549	0.1585	0.103	0.1892	0.158	0.1612	0.1056

* Statistically significant at 10 percent level. ** Statistically significant at 5 percent level. *** Statistically significant at 1 percent level.

Table 4. Logit regression results (continued) (dependant variable = DEFAULT). *continued*

Variable	Model 3				Model 4			
	1988 (n = 2337)	1989 (n = 1683)	1991 (n = 898)	1992 (n = 3225)	1988 (n = 2337)	1989 (n = 1683)	1991 (n = 898)	1992 (n = 3225)
Location								
LEV	–0.0006	0.0016	–0.0026	0.0007				
	(–0.542)	(1.180)	(–1.095)	(0.602)				
HHDENS					0.0006	0.0035	–0.0038	0.0010
					(0.125)	(0.818)	(–0.276)	(0.355)
PED_FACT					0.1027	0.4080	0.4923	0.3523
					(–0.329)	(–0.03)	(–1.287)	(–0.605)
RAW_DENS					–0.0074	–0.0007	–0.0794	–0.0114
					(0.401)	(1.294)	(0.88)	(1.206)
URBAN	0.1992	0.0854	0.1459	0.3106	0.1469	0.0649	0.0064	0.3320
	(0.958)	(0.324)	(0.363)	(1.359)	(0.736)	(0.254)	(0.016)	(1.498)
ZONAL_TR					–0.0011	–0.0003	0.0002	–0.0009
Mortgage					(–0.538)	(–0.137)	(0.054)	(–0.511)
DTI	–0.0323***	–0.0193	–0.0403*	–0.0027	–0.0323***	–0.0191	–0.0400*	–0.0028
HEI	0.0256	0.0760**	0.0857	–0.0347	0.0261	0.0764**	0.0962	–0.0324
LTV	0.0377**	0.0466*	–0.0111	0.0304*	0.0375**	0.0462*	–0.0120	0.0309*
Property								
CONDO	–0.8140**	–1.1974***	–0.6844	–0.3629	–0.7514**	–1.0807**	–0.4993	–0.1949
HVAL	0.0001	–0.0523**	–0.0054	–0.0304*	–0.0003	–0.0539**	–0.0070	–0.0301*
HVAL2	–0.0001	0.0002	0.0000	0.0002*	–0.0001	0.0002	0.0000	0.0002*
Borrower								
AGE	0.0053	0.0072	0.0075	0.0010	0.0056	0.0074	0.0082	0.0005
ASI_AMIN	0.2921	0.9262	0.7790	0.5198	0.2697	0.9685*	0.7229	0.5426
BLACK	1.4905***	1.0965***	1.8749***	0.8337***	1.4572***	1.0695***	1.7527***	0.7928***
DEPNUM	0.1968***	0.1507**	0.0827	0.2031***	0.1900***	0.1650**	0.0603	0.2066***
FIRSTBUY	–0.0725	–0.2802	–0.5788**	0.3437	–0.0671	–0.2603	–0.5793*	0.3375
HISPANIC	–0.1072	–0.4024	0.4337	–0.0090	–0.1392	–0.4290	0.3573	–0.0328
INCOME	–0.0326	0.0957*	0.0290	–0.0295**	–0.0328	0.1042*	0.0286	–0.0269*
INCOMEZ	0.0002	–0.0009	–0.0004	0.0000*	0.0002	–0.0010*	–0.0003	0.0000
LQASS	–0.1341***	–0.0782***	–0.1120***	–0.0363**	–0.1339***	–0.0799***	–0.1082***	–0.0360**
LQASS2	0.0017***	0.0005**	0.0007**	0.0003**	0.0017***	0.0006**	0.0007**	0.0003**
NOCBINC	–0.0586	0.2621	–1.1676**	–0.0930	–0.0467	0.2740	–1.1086**	–0.0949
SINGLEM	0.6389**	0.1077	1.0618**	0.4757*	0.6356**	0.0782	1.0293**	0.4784**
Intercept	–3.7437**	–7.3844**	–1.6259	–3.0792*	–3.8141**	–7.7401***	–1.9706	–3.4333**
Pseudo R^2	0.1898	0.1581	0.1615	0.1070	0.1900	0.1595	0.1645	0.1079

* Statistically significant at 10 percent level. ** Statistically significant at 5 percent level. *** Statistically significant at 1 percent level.

areas than in suburban areas and, as discussed in the introduction, empirical studies typically find that the probability of default is higher in urban areas than in suburban areas.

To test the hypothesis that LEV proxies for urban location, in model 3 a new control variable is added—a dummy called URBAN that takes the value of 1 if the property is located in an urban area and 0 otherwise (this variable is included in the FHA data set). Now, LEV is no longer significant in any of the four annual cohorts. This strongly suggests that the reason LEV is positively correlated with default in model 2 is its correlation with urban location. Correlation coefficients support this hypothesis. The simple correlation coefficient between LEV and URBAN is 0.66.

Because LEV and URBAN are strongly correlated, the results in model 3 may be affected by multicollinearity. To check this, URBAN was omitted from the model and urban location and controlled for by splitting the sample in each annual cohort. One subsample included only urban homes and a second subsample included only suburban homes. Model 2 was then estimated for each subsample. Qualitatively, the results (omitted for the sake of brevity) were the same as those reported in Table 4—LEV is not significant in any of the four annual cohorts.

While these results indicate that there is not a statistically significant relationship between the LEM consortium's index of location efficiency and the probability of default, it is still possible that one or more of the locational variables used to calculate this index are correlated with the probability of default. Model 4 tests this hypothesis. This model includes the four census tract–level locational variables used to calculate LEVs—households per residential acre (HH_DENS), pedestrian friendliness (PED_FACT), households per total acre (RAW_DENS), and transit access (ZONAL_TR)—along with 19 control variables. None of the locational variables are significant in any of the four annual cohorts.[10]

CONCLUSION

The results indicate that in this sample of mortgage records, the measure of location efficiency used by LEMs advocates—the LEV—is not significantly correlated with a lower probability of default, all other things equal. Moreover, neither are any of the components of this measure, including density, access to mass transit, and pedestrian friendliness. What accounts for these negative results? One explanation is that while homeowners in location-efficient areas may actually enjoy transportation costs savings, these savings are simply not large enough to affect their propensity to default. Indeed, there are a number of reasons to believe that the estimates of transportation cost savings generated by the LEMs computer model—often hundreds of dollars per month—are overstated.[11] A complimentary explanation is that homeowners rationally anticipate any transportation cost savings from location efficiency and allocate these saving to alternative uses, leaving their ability to repay housing debt not much improved.[12]

[10] To be certain that this result is not affected by multicollinearity stemming from correlations among the locational regressors, four auxiliary models that included just one locational variable along with the control variables were estimated. Qualitatively, the results (omitted for the sake of brevity) were the same as reported in model 4—none of the locational variables was negatively and significantly correlated with default. The same results were obtained when urban location was controlled for using the split sample discussed above.

[11] First, as discussed in the Appendix, the algorithm used to calculate transportation cost savings from location efficiency assumes that borrowers purchasing homes in location-efficient areas would otherwise live in particularly location-inefficient areas, an arbitrary assumption that clearly biases cost savings estimates upward. Second, compared with alternative models (Schimek, 1996), the econometric model used to calculate transportation costs (Holtzclaw, 1994) assumes that vehicle miles traveled are particularly responsive to household density.

[12] We are grateful to an anonymous reviewer for this explanation.

These findings probably come as no surprise to those who believe that mortgage lenders are reasonably well-informed profit maximizers. A statistically significant negative relationship between location efficiency and mortgage default would imply that lenders have overlooked opportunities to increase their profits by conditioning mortgage contracts on an easily observable determinate of loan repayment.

Since no demonstrable relationship was found between location efficiency and probability of default, the authors would argue that making low-down-payment loans available to borrowers in location-efficient areas is tantamount to making such loans available to a random sample of borrowers and will have the same result: It will raise default rates. This conclusion is subject to a number of caveats. First, it is based on an indirect test of the proposition that LEMs will not raise default rates. Second, this conclusion is based on a sample of mortgage records that is limited both geographically and temporally. It is conceivable—although highly doubtful—that the beneficial effects of location efficiency on mortgage repayment were for some reason dampened in Chicago in the late 1980s and early 1990s.[13] Research in a second city would be useful to confirm the generalizability of the results. (With the recent expansion of the LEM program, it is now possible to replicate this analysis using data from Los Angeles, Portland, San Francisco, and Seattle). Also, in a number of years, it will be possible to examine the historical record of repayment on LEMs.

Given these findings, should LEMs be dismissed as a viable policy? Not necessarily—if LEMs are effective in controlling sprawl or achieving other policy objectives, then their cost in terms of default losses would have to be weighed against their benefits, as well as against the costs of achieving these benefits by other means. The focus of this paper is not the possible benefits of LEMs. The main point is simply that just as policymakers recognize that low-down-payment loan programs like those operated by the FHA and VA involve tradeoffs between higher default costs and specific policy objectives (for example, expanded home ownership), they should also recognize that LEMs will entail a quid pro quo. Contrary to the claims of advocates, the findings presented here suggest that LEMs are unlikely to be a costless anti-sprawl policy.

This research was funded under EPA cooperative agreement CX 824429-01. We would like to thank: Joe Cook, Terrell Stoessell, and Deirdre Farrell Gabbay for excellent research assistance; William Schroeer (formerly of EPA) and Robert Noland of EPA's Policy Office for financial and intellectual support; William Shaw of the Department of Housing and Urban Development for providing access to and information about the FHA mortgage data; Kim Hoeveler and Peter Haas of the Center for Neighborhood Technology and John Holtzclaw of the Natural Resources Defense Council for providing data on location efficiency and for patiently explaining the LEMs Advisor; William Pizer for helpful comments; and four anonymous referees. The views in this paper are the authors' alone and not necessarily those of any of the above-named individuals or organizations.

ALLEN BLACKMAN is a Fellow in the Quality of the Environment Division at Resources for the Future.

ALAN KRUPNICK is a Senior Fellow in the Quality of the Environment Division at Resources for the Future.

[13] The fact that the mortgages in the sample were originated during a recession in which average default rates rose is not particularly germane to the analysis. All the borrowers in this sample were subjected to the same macroeconomic conditions. If it is true—as LEM advocates claim—that location efficiency mitigates default risk, then in times of economic distress borrowers in location-efficient areas should still have below-average default rates, all other things equal, despite the fact that average default rates for all borrowers have risen.

APPENDIX: LOCATIONAL VARIABLES

A. Sources of data

HH_DENS	1990 Census data and Chicago Metro. Planning Org.
LEV	LEMs Coalition
PED_FACT	LEMs Coalition
RAW_DENS	1990 Census data
ZONAL_TR	LEMs Coalition

B. Methodologies used by the LEMs coalition to derive PED_FACT, ZONAL_TR, and LEV[14]

PED_FACT is calculated by Travel Analysis Zones (TAZ), which are aggregations of census tracts. PED_FACT is the sum of three variables: STREET_GRID, YEAR_BUILT, and BONUSES. The variables are derived from Census and Chicago Area Transportation Study (CATS) Committee data.

STREET_GRID is simply the average block size in the TAZ. It is equal to the number of census blocks divided by the number of developed hectares (including areas zoned as residential, commercial, and industrial). The rationale for this variable is that smaller blocks are more pedestrian friendly.

YEAR_BUILT is an index based on the median vintage of the building stock, information derived from the 1990 Census. This variable purports to capture the safety of streets, the setback of buildings, and sidewalk weather protection. The index is:

median vintage	index
1939 or earlier	0.7
1940–1942	0.6
1943–45	0.5
1946–48	0.4
1949–50	0.3
1951–1952	0.2
1953–55	0.1
1956 or newer.	0.0

BONUSES is an ad hoc variable meant to give credit to areas that have implemented "traffic calming" policies like narrowing streets, widening sidewalks, and planting trees in median strips as well as areas that have built bike lanes. The variable ranges from zero to one for traffic calming credits and from zero to 0.5 for bicycle friendliness.

[14] This section is based on CNT, 1998 and 2000, as well as multiple personal communications with John Holtzclaw of the Natural Resources Defense Council and Peter Hass of the Center for Neighborhood Technologies.

ZONAL_TR purports to capture accessibility to mass transit in a given census tract. It approximates the average number of buses and train cars available to the population of a census tract each day. For each type of mass transit, the methodology involves four steps. For buses, first, a "buffer zone" is drawn around each bus stop in the area. Because of the grid layout of streets in Chicago, the buffer zone is shaped like a diamond with the distance from the center of the diamond to each point being one-quarter mile. The number of daily buses at each stop is determined from Chicago Transportation Authority route schedules. One bus arriving at one stop is counted as two buses: one inbound and one outbound. One bus that stops at two stops within the buffer zone is counted as four buses. Second, the number of buses per stop per day is multiplied by the population in the buffer zone to get the number of people served by each bus. Third, population served is attributed to census tracts as follows. If a buffer zone lies within a census tract, 100 percent of the population served in that buffer zone are assigned to the census tract. If only part of a buffer zone lies within the census tract, the census tract is assigned a fraction of the population served in the buffer zone. This fraction is the percentage of the buffer zone that lies within the census tract. The process is identical for trains, except that the buffer zone is one-half mile. Finally, the number of people served by each bus or rail car in each census tract is normalized by the number of people in each census tract.

LEVs are calculated using a three-step method. First, the LEMs coalition's econometrically estimated model is used to predict both vehicle miles traveled (VMTs) and number of autos owned per residence (VEH/HH). That model is,

$$\frac{\mathrm{VEH}}{\mathrm{HH}} = 1.721 \bullet \left(\frac{16.123}{9.955 + \mathrm{HH\_DENS}}\right)^{0.2797} \bullet \left(\frac{1 - \exp\left(-[0.000142 \bullet \mathrm{INCOME}\ ]^{1.2915}\right)}{0.98545}\right) \bullet \left(\frac{1 + 0.4893 \bullet \mathrm{HH\_SIZE}}{2.3660}\right) \bullet \left(\frac{11.753}{\mathrm{ZONAL\_TR} + 2.960}\right)^{0.0685}$$

$$\frac{\mathrm{VMT}}{\mathrm{VEH}} = 10655 \bullet \left(\frac{4.925}{0.1662 + \mathrm{RAW\_DENS}}\right)^{0.0547} \bullet \left(\frac{1 + 0.00653 \bullet \mathrm{HH\_SIZE}}{1.0178}\right) \bullet \left(\frac{1 - (0.0249 \bullet \mathrm{PED\_FACT})}{0.983}\right) - 0.0818(\mathrm{INCOME} - 22136)$$

$$\frac{\mathrm{VMT}}{\mathrm{HH}} = \left(\frac{\mathrm{VEH}}{\mathrm{HH}}\right) \bullet \left(\frac{\mathrm{VMT}}{\mathrm{VEH}}\right)$$

Note that two of these independent variables—household income and number of persons in the household—are specific to the loan applicant while the remaining four independent variables relate to the census tract in which the home is located. (For our regressions we constructed the variable HH_SIZE from FHA data using the formula, HH_SIZE = Borrower + Coborrower [if any] + number of dependents [if any]). Second, given econometrically estimated VMTs and auto ownership, auto expenses are calculated using Federal Highway Administration figures on the costs of owning and operating automobiles (FHA, 1992).

Auto expenses = (VMT/HH*$0.127 per mile) + (VEH/HH*$2207 per auto per year)

Costs are inflated to 1998 dollars using the consumer price index. Finally, automobile expenses for the borrower are subtracted from base case expenses, the difference being the LEV. The base case is meant to represent auto expenses for a household of similar size and wealth in a location-inefficient census tract, that is, one with relatively low density, transit access, and pedestrian friendliness. It is constructed as follows. Census tracts are grouped into quartiles for the entire Chicago region based on household densities per residential acre. The average household density in the lowest quartile is very close to one household per residential acre.

Using this density and the low values for zonal transit access and pedestrian factor a base cost is calculated for each household size and household income. The difference between the base case auto costs and the econometrically estimated auto costs is the LEV. The LEV will be negative if the census tract in question is less location efficient than the base case.

REFERENCES

Allen, J., (1998, July 26). Mortgages with a motive: A growing number of home loans serve a specific social purpose. Chicago Tribune,C1.

Avery, R., Bostic, R., Calem, P., & Canner, G. (1996, July). Credit risk, credit scoring & the performance of home mortgages. Federal Reserve Bulletin, 621–648.

Benfield, K., Rami, M., & Chen, D. (1999). Once there were greenfields: How urban sprawl is undermining America's environment, economy and social fabric. New York: Natural Resources Defense Council.

Berkovec, J., Canner, G., Gabriel, S., & Hannan, T. (1998). Discrimination, competition, and loan performance in FHA mortgage lending. Review of Economics and Statistics, 80, 241–250.

CNT [Center for Neighborhood Technologies] (1998). Summary of results of location and auto cost correlation study. Chicago: CNT.

CNT [Center for Neighborhood Technologies] (2000). Location-Efficient Mortgage project. http://www.cnt.org.

Danielsen, K.A., Lang, R. E., & Fulton, W. (1999). Retracting suburbia: Smart growth and the future of housing. Housing Policy Debate, 10, 513–540.

Deng, Y., Quigley, J., Van Order, R., & Freddie Mac (1996). Mortgage default and low downpayment loans: The costs of public subsidy. Regional Science and Urban Economics, 26, 263–285.

Easterbrook, G. (1999). Comment on Danielsen, Lang, and Fulton. Housing Policy Debate, 10, 541–547.

Eggen, D. (1998, October 28). A growing issue: Suburban sprawl, long seen as a local problem, emerges as a hot topic in state, national politics. Washington Post, A03.

FHA [Federal Highway Administration] (1992). The cost of owning and operating automobiles, vans, and light trucks. FHWA-PL-92-019. Washington, DC: GPO.

Holtzclaw, J. (1994). Using residential patterns and transit to decrease auto dependence and costs. San Francisco, CA: Natural Resources Defense Council.

Mills, E., & Lubuele, L. (1994). Performance of residential mortgages in low- and moderate income neighborhoods. Journal of Real Estate Finance and Economics, 9, 245–260.

Orlando Sentinel (1999, June 13). Loan will benefit buyers near transit. J6.

Pew Center for Civic Journalism (2000) Straight talk from Americans, 2000: National survey results. www.percenter.org/doingcj/research/r_ST2000nat1.htm.

Quercia, R., & Stegman, M. (1992). Residential mortgage default: A review of the literature. Journal of Housing Research, 3, 341–379.

Rimer, L.B. (1999, October 17). An anti-sprawl mortgage. The News and Observer (Raleigh, NC).

Schimek, P. (1996). Household vehicle ownership and use: How much does residential density matter? Transportation Research Record, 1552, 120–130.

Van Order, R., & Zorn, P. (1996). Income, location, and default: Implications for community lending. Mclean, VA: Federal Home Loan Mortgage Corporation.

Vandell, K. & Thibodeau, T. (1985). Estimation of mortgage defaults using desegregate loan history data. AREUEA Journal, 13, 292–316.

von Fustenberg, G., & Green, R. (1974). Home mortgages delinquency: A cohort analysis. Journal of Finance, 29, 154–158.

Woellert, L. (1999, December 6). Now the close-to-the-bus mortgage. Business Week, 6.