Institute of Business and Economic Research

Fisher Center for Real Estate and Urban Economics

PROGRAM ON HOUSING AND URBAN POLICY

WORKING PAPER SERIES

WORKING PAPER NO. W07-005

HOW CAN REGIONAL DIFFERENCES IN THE RISK-OF-FORECLOSURE BE EXPLAINED? EVIDENCE FROM SWEDISH SINGLE FAMILY HOUSING MARKETS

By

Roland Andersson
Mats Wilhelmsson

September 2007

UNIVERSITY OF CALIFORNIA, BERKELEY

How can Regional Differences in the Risk-of-Foreclosure be Explained? Evidence from Swedish Single Family Housing Markets

by

Roland Andersson[1] and Mats Wilhelmsson[2]

Royal Institute of Technology
Real Estate Economics
100 44 Stockholm, Sweden

September 2007

Word count: 10,218

1) rolanda@infra.kth.se and 2) Corresponding author: matsw@infra.kth.se. Funding from the Sweden-America Foundation and the Wenner-Gren Foundation is highly appreciated. I also thank the Haas School of Business and the Goldman School of Public Policy (UC, Berkeley) for their hospitality. Finally, we are grateful to Sandy Nixon for her editorial help with the manuscript.

How can Regional Differences in the Risk-of-Foreclosure be Explained? Evidence from Swedish Single Family Housing Markets

September 2007

Abstract

In Sweden, quite large differences in the risk-of-foreclosure for single-family housing exist between regions. The aim of this paper is to explain such differences, using data on foreclosures for all Swedish regions. In an option-based model, the risk-of-foreclosure is a function of such things as housing prices and incomes, as well as interest rate and housing price volatility. Instead of using housing prices and incomes to explain the risk of foreclosure, we use explaining variables in the labor market model. The main results indicate that the option-based model explains the variation in foreclosure rates. Specifically, interest rate – together with price volatility, price changes, price and rent level, income, and employment – explains around one-third of the total regional variation. Our extended option-based model explains slightly more. Specialization within the industrial sector seems to have a positive effect on foreclosure risk in that it, together with the educational level of the workforce, reduces the risk. Specifically, mortgage lenders and banks can reduce their risk by concentrating their business on dense regions with a high degree of employment within the manufacturing industry and with a higher educational level of the workforce.

Keywords: Foreclosure risk, default, spatial distribution, housing market.

1. Introduction

Mortgage defaults are important to the lending industry as well as to borrowers and investors. During the last few decades, research papers have focused increasingly on the pricing of credit risk in the mortgage market. In order to lower their credit risk, lenders could utilize such data-based research devoted exclusively to understanding the causes and consequences of mortgage default risk. However, mortgage defaults do not represent a definitive end of such an affair, at least not in Sweden. Not all mortgage defaults end up as a foreclosure sale. Negotiations can take place between lenders and borrowers that might lead to agreements about how to solve a situation of mortgage default, thereby stopping the default. Foreclosure is a more definite end of a credit story than a default. For example, in the U.S., the large number of so called subprime mortgages utilized for home ownership of single-family housing has recently led to a high frequency of foreclosure, with big losses for borrowers and often also for lenders. The prediction is that this situation will continue to occur in the future.

In Sweden, data show that quite large regional differences in the risk-of-foreclosure exist for owner-occupied single-family housing. For instance, the number of foreclosures in percent of turnovers has been considerably lower in some parts of Sweden than in other parts of the country during the whole period 1993-2006. This has occurred in spite of the lenders being aware of the larger risks of lending to some parts of the country. One would expect that a rational lending policy would result in more restrictive lending practices in such regions in order to diminish that region's risk-of foreclosures, such as giving a lower loan-to-value ratio. Therefore, since there are regional differences in the rate of foreclosures, it would appear they have to be explained by local variables. The aim of this paper is to explain such differences.

The main results indicate that the option-based model explains the variation in foreclosure rates very well. Specifically, interest rate together with price volatility and price changes, as well as price and rent level, income, and employment explain around one-third of the total variation. Our extended option-based model explains slightly more. Specialization within the industrial sector and educational level of the workforce seems to have a positive effect on the foreclosure risk.

The paper is arranged as follows: Section 2 below provides a selective review of the literature on the determinants of foreclosure and of mortgage default. Moreover, we present a brief literature review on housing and labor market literature. In Section 3, we will discuss the methods used and specifically the model specification. In Section 4, the data is presented and in Section 5, the econometric analysis. Section 6 ends the paper with a conclusion and a discussion of policy implications.

2. A literature review

We will review here papers from three different but related fields, covering mortgage defaults and labor markets.

2.1. Delinquents, mortgage defaults and foreclosure

A search on mortgage default gives you a large variety of research types. Defaults are analyzed from different perspectives such as: the borrower's point of view, the lender's point of view and the investor's perspective. Moreover, mortgage default is discussed in different stages of the process. For example, there is a huge difference between examining mortgage delinquencies and mortgage defaults. One way to describe the different choices for the borrower is as follows.

FIGURE 1 IN HERE.

At each payment period, the borrower has a number of choices or options. The borrower can pay interest and principal according to the schedule, pre-pay the mortgage balance or delay the payment and be delinquent. In the next stage after delinquency, the borrower can choose to restart the payment of the interest and principal or pre-pay the mortgage through a sale of the property. The borrower can also choose to default. However, even after the borrower has defaulted there is a possibility of selling the property and pre-paying the mortgage. Another option is to negotiate new loan terms with the bank or mortgage lender. If the borrower fails to renegotiate the loan terms, foreclosure does occur. However, even after the foreclosure, there is time for negotiation. However, if the negotiation fails, the sale is enforced (a foreclosure sale).

2.2. The mortgage default literature

The last decades have seen increasing number of papers presenting advances in understanding the risks of mortgage defaults foreclosures: from Jung (1962), to Foster and Van Order (1985) to recent articles by, for example, Miller and Peng (2006). A review of the early residential mortgage default literature can be found in Quercia and Stegman (1992) and Vandell (1995). Most of the analyzed data comes from the U.S., and it is mostly household mortgage data from the secondary mortgage market.

Jung's (1962) article is probably one of the first in which the determinants of mortgage default are studied. His results suggest that there is relationship between LTV and interest rate. His finding was that "there was a definite pattern of variability in the effective interest charged /…/systematically higher /…/ for loans with high LVRs [loan to value ratio]."

Twenty years later, Foster and Van Order (1984) investigate option-based explanatory variables, such as housing price volatility. Utilizing option models, they demonstrate the importance of interest rates affecting the borrower's decision to default. See also Vandell and Thibodeau (1985). Kau et al (1992) simulate default frequencies in a frictionless model, i.e. one with no transaction costs. They argue that the frequencies presented by this model are not much different from those actually observed. They also argue that the introduction of transaction costs into their simulation implies implausibly low default rates, and they therefore conclude that the frictionless model does a good job in explaining default behavior.

Based on the use of data on actual loan-loss severities in mortgages in the U.S., Lekkas et al (1993) performed a number of tests concerning the predictive ability of the frictionless model. They tested some propositions in line with this model, such as the proposition that the loan-loss severity of mortgages should be independent of initial Loan-to-Value (LTV) and should be the same in regions with high default frequencies as it is in regions with low default frequencies. Further, they said that his proposition should be the same for loans originated in both good and bad years. They found that the frictionless model does not do very well in these tests. Thus, their results show a strong relationship of mortgage defaults to LTV. They compared the loss severity of all LTV measurements as a percentage of mortgage balance for the country' with a subset of data from Texas during the same time. They found that the losses in Texas are substantially higher. The prediction of the frictionless model is not consistent with this empirical evidence.

In addition, Lekkas et al. found that people wait until the values of their houses have dropped by 20 to 30 percent before defaulting. An explanation might be that people get into trouble, for instance by losing a job, and have a limited liquidity for a limited period. Then, at the end of that period, they must sell either the house or default. Quigley and Van Order (1995) present further evidence concerning that a zero transaction cost model is not consistent with the data. They point out that cost of trading housing is an important transaction cost in the market.

Capozza et al (1997) examine the effects of three mortgage default trigger events: unemployment, moving rates, and divorce. As unemployment increases, borrowers encounter ability-to-pay problems leading to higher default rates, as well as increases in moving rates and divorces. They include in their model a rent-to-price figure, that is, a home rental cost divided by home price as a measure of the dividend or asset yield. A negative relationship with default was found. The higher the rent-to price ratio, the higher the value of delaying default and the lower the house-price-default boundary. They point out that a borrower who experiences a job transfer could be forced to default immediately, even if there were a more opportune time to default in the future. Their overall conclusion is that if house prices are low, defaults will be high regardless of whether trigger events occur or not. Earlier, Williams et al (1974) found that areas with a high unemployment rate also exhibit higher mortgage defaults rates.

In a later paper, Capozza et al (1998) focus on so-called "conditional probability of mortgage default over short horizon." This refers to the probability of a default conditional on the availability of all information that has arrived since the origin of the loan, such as the current LTV ratio and recent economic conditions. This is a very important distinction, since the loan-to-value ratio might have changed considerably from the original date of the loan up to date of the mortgage default, because of the amortization that has taken place, and because of changes in the market value of the housing. The authors demonstrate that some variables suggested by option pricing that matter unconditionally – such as the rental rate, interest-rate volatility, and interest reversion – are not of the same importance conditionally. Volatility of housing prices is found not to be very important conditionally as unconditionally. Their findings (1998) confirm the results of their (1997) study: "Trigger events, traditionally considered important precipitators of default, have relatively little influence" (1998, p 383).

Capozza et al (1998) assume that a default results in an immediate loss of the house in exchange for forgiveness of the debt. This is a simplification of the reality. Ambrose and Capone (1998) point out that mortgage default and foreclosure are two separate decisions or events and that foreclosure is but one outcome of a default episode. They investigate the influence of borrower characteristics, mortgage terms, and economic conditions on probabilities of various resolutions to find out under what conditions foreclosure is more likely to result from mortgage default. Lauria et al (2004) examine the variables that determine the time between default and foreclosure. They found, quite expected, that homes with lower LTV ratios have significantly longer periods before foreclosure proceedings are initiated, than do high LTVs. In addition, they found that the homes of borrowers who defaulted due to loss of job were foreclosed on at a significantly faster rate than on the homes of those employed, because the former were simply unable to afford the payments, also quite an expected result.

In a later study, Ambrose et al (2001) examine the timing of defaults, given home price and interest changes. They found that the cost of life events such as divorce, influence the mortgage default. On the other hand, Clapp et al (2001) found that current-loan-to-value ratios (CLTV) and borrower credit scores solely explain mortgage defaults. Danes and Cross (2005) have examined the impact of subprime lending on mortgage defaults, an issue of great concern in the U.S. as of today. In particular, they found that low credit scores and past delinquency rates significantly effect the likelihood of default by subprime borrowers.

Pennington-Cross (2003) found that the greater the housing price volatility, the greater the probability of negative home equity and more severe mortgage foreclose losses. Recently, Miller and Peng (2006) investigate the time variation of the volatility of single-family home value appreciation and ask why such a housing market seems more volatile in some periods than others. They also ask whether and how housing volatility affects the economy. In their model, the interactions between volatility of the home appreciation rate and the economy is analyzed, as measured by the per capita personal income growth rate, the population growth rate, the unemployment rate change, and the per capita gross metropolitan growth rate (GMP). One of their results was that shocks in the growth rate of per capita GMP and home value appreciation significantly affect the growth rate of per capita personal income and future volatility; see also Quigley (1999).

More recently, Bond (2007) investigated why Ohio has experienced such a significant increase in mortgage defaults in recent years. The Ohio unemployment rate was below the national average for most of the 1990's and increased dramatically above the U.S. level by 2005. In addition, Ohio's productivity grew at less than half the rate of the U.S. during the same period with a relatively high default rate. Besides, home prices in Ohio have grown at a substantially slower rate than in the U.S. as a whole. Such independent variables correctly relate to defaults, with relative unemployment being positively correlated and relative gross domestic product and home values being negatively correlated. All estimated parameters were highly significant.

2.3. Labor markets

An enormous number of papers exist about the labor market. Only a few papers of direct relevance to our interest in the development of local variables are reviewed: Roback (1982), Topel (1986), Rauch (1993), and Glaeser and Maré, (2001). Roback (1982) focuses on the regional wage differences, pointing out that if workers require a compensating wage differential to live in a big, polluted, or otherwise unpleasant city, the firms must have some productivity advantage to be able to pay the higher wage. In her study, she found that regional wage differences could be explained largely by local amenities. It seems reasonable if workers require a compensation for unpleasant conditions. However, positive amenities, such as more sunny days, might explain a wage differential in the contrary direction.

Utilizing data for the U.S., Topel´s analysis (1986) starts with observations about persistent differences between geographic areas concerning wages and income, unemployment, net migration flows and growth. He points out the disparity between levels and amplitudes of measured unemployment rates in rapidly growing so-called Sunbelt localities on the one hand, and the higher unemployment rates that occur in the industrial Northeast and Midwest regions on the other. Such disparities could be attributed to specific industrial composition and different changes in the local market conditions. Topel further points out that some areas have experienced an above-average employment growth supported by an in-migration of young and well-educated workers that has resulted in relative declines in local unemployment and increasing relative wages. He found that wage rates are quite sensitive to inter-area differences in market conditions and that wages are more flexible in response to transitory changes in local market conditions than to permanent ones. He also found that that the largest local wage adjustments occur among workers that exhibit the least geographic mobility, such as older workers with less education.

Rauch´s (1992) basic argument is that the average level of human capital could be considered a local public good. Thus, economically identical workers will tend to earn higher wages in regions that are rich in human capital than do workers in regions scarcer in human capital. His hypothesis is that cities with higher average levels of human capital should have higher wages and land values, and that migration to high wage areas leads to higher residential and commercial rents. He believes that such a migration will offset the higher wages and allow for a spatial equilibrium, where utility levels and production costs are equalized across metropolitan areas. He claims that his results reflect the productivity benefits of sharing ideas made possible by the geographic concentration of human capital externalities.

Glaeser and Maré (2001) want to explain why workers in cities in the U.S. earn 33 percent more than their non-urban counterparts do. They found evidence that a significant fraction of the urban wage premium accrues to workers over time and stays with them when they leave cities. Their conclusion is that a portion of the urban wage premium is a wage growth, not a wage level effect and that cities speed the accumulation of human capital. They conclude that this urban wage growth could be explained by the better coordination of labor markets or by the faster learning that occurs in cities.

3. Model Specification and Variables Used

3.1. The basic option-based model

The basic model in the present paper follows Capozzas et al's (1997) option-based model. In our case, key option model variables – such as interest rate, asset price, and price volatility in the underlying security – explain the regional variation in foreclosure rates. The model also bears some resemblance to Case and Shiller's (1997) state-level foreclosure model in which they use income, net migration, and employment rates as explanatory variables. The model is conditional in the sense that we are using information at the time of foreclosure and not at the time of origination of the mortgage.

We are using aggregated data, that is, we do not have any information on the underlying mortgages in the investigated labor market. Hence, we lack relevant information of such things as "expiration date" and "age of the mortgage." Our base model relates foreclosure rates with interest rate, price and rents, volatility in price, and a transaction-cost proxy namely income per capita. Lekkas et al (1993) and Capozza et al (1997) show a strong positive relationship between LTV and mortgage default and LTV and loss severity, respectively. As we do not have information about the average LTV rate in the labor market over time, we include the average change in home prices. Hence, we implicitly assume that if prices have decreased over the last five years, LTV will increase. Specifically, our base model has the following expression.

$$(1) \quad \begin{aligned} FCR_{i,t} = f(&\text{Interest rate}_{t}, \text{Sigma}_{i,t}, \text{Price}_{i,t}, \text{Rent}_{j,i,t}, \text{Rent - to - Price}_{i,t}, \\ &\text{Price changes}_{i,t}, \text{Income}_{i,t}, \text{Employment}_{i,t}) \end{aligned}$$

where *FCR* is foreclosure rate in labor market i and time t. In the current study, we use two different definitions of foreclosure rate. We define the first as the total number of foreclosed sales divided by the total number of sales (the turnover). However, a measure of mortgage loss is a better indication of mortgage risk than defaults rates based on the number of foreclosed sales. In the words of Quercia and Stegman (1992), "expected losses may be a better measure of default risk than default rates, because expected losses provide a more accurate basis for estimating mortgage insurance premiums, mortgage interest rate premiums, and the potential default costs of a government subsidy program." Hence, as an alternative, the second definition uses the value of the sales related to the total values of all sales, that is, in some sense the loss severity[1] (see Lekkas et al, 1993).

Interest rate is one of the key option-based variables (see e.g. Foster and Van Order, 1984). High interest rate means high option values and, therefore, the foreclosure rate should be high, *ceteris paribus*. In the present study, the interest rate only varies over time and not across labor markets, which in Sweden's case seems to be a realistic assumption. *Sigma* measures the volatility in price of the underlying asset. We have defined it as the standard deviation of the annual price changes over the previous five years. The anticipated effect is positive (Pennington-Cross, 2003). The *price* variable is the average house price and the *rent* variable is the average rental level in the rental-housing sector. Hence, we have not used quality-adjusted prices and rents. It is important to remember that apartment rents in Sweden are not market rents, clearing the housing demand and supply. Instead, rents in Sweden are a result of local negotiations between the tenant organization and the municipality housing company.

We also use the rent-to-price ratio as a proxy for owner-occupied dividends. Capozza et al (1997) utilized a similar variable. They interpret the variable rent-to-price ratio as the "dividend" of owner-occupied housing. The interpretation is that when the rents go up given the single-family home prices, the dividends increase, and it is less likely that the foreclosure option is "in the money." On the other hand, a low dividend indicates that the rents are lower

[1] We do not have the information about mortgage balance; on the other hand, we do have the sale price of the house. Hence, we do not define loss severity as in Lekkas et al (1993) where it is defined as the difference between mortgage loan balance and the value of the house. Our measure can be seen as an implicit assumption that the mortgage balance is equal to zero, which is not likely, suggesting that the loss is equal to the total value of the foreclosed sales.

than the single-family house prices. Thus, it indicates that the option is "in the money," and if the transaction cost is low enough, it is optimal to exercise it. However, low rents in relation to the "true" market rents also imply very low vacancy rates, well below the natural vacancy rate.[2] Hence, even if the foreclosure option is favorable, it is not possible to exercise it. For this reason, we define the variable as *expected rent-to-price* where rents are conditional on vacancy rates.

The price change over the last five years defines the *price changes* variable in the equation. Furthermore, the taxable income per capita in the region defines the *income* variable. We use the last variable as a proxy for transaction cost (or reputation cost). The argument is that high-income borrowers induce larger mortgage default and foreclosure losses (not only in terms of money). Hence, income should have a negative impact on foreclosure rates. Finally, the last variable included in Equation 1 is employment measured as the *employment rate*. As the employment rate increases, fewer households experience problems paying their mortgages, and hence, foreclosure rates are reduced. In the terminology of Capozzas et al and Foster and Van Order (1984), this is called a "trigger event variable."

3.2. The extended option-based model

Our objective was not only to estimate a model á *la* Capozzas et al but also to utilize aggregated data instead of individual mortgage data. Many of the variables (e.g. house price and income) in the option-based model are a function of other determinants. For example, other exogenous variables determine the income level in the region, such as its industrial composition. As has been said, there exist a vast number of articles about the determinants of labor income, and our objective is to merge a simple labor-market model into the option-based foreclosure model. The key objective with our examination is to analyze whether

[2] Rosen and Smith (1983) introduced the concept of natural vacancy as dependent on the issue of matching at the same time the demand for and supply of rental properties.

different aspects of agglomeration and diversity play a role in explaining the variation in foreclosure rates.

The merged model can be seen as an extended option-based model in which we explain the variation in foreclosure rates among the 100 labor markets in Sweden over the period 1994-2001 with a set of variables that explains income and housing prices, together with the more traditional option-based key variables. The former variables are, for example, *market size*, *density*, *diversity* and *specialization*. All of these variables are indicators of agglomeration economies. In the agglomeration model, labor productivity increases – in the words of Hoover (1937) – "when there are external economies of localization and urbanization." The reason for these external economies, according to Marshall (1925), is that in an agglomeration, firms are able to share labor and other inputs (better matching); and in an agglomeration, knowledge is spread more efficiently (see Quigley, 1998). An increase in labor productivity increases production and causes the demand for labor to increase, which results in higher wages, suggesting a relationship between indicators of agglomeration economies and foreclosure rates. The hypothesized linkage between economic diversity and house prices and foreclosure rates follows from Jane Jacob's (1961, 1969) verbal insights about economic growth and urban heterogeneity, as well as from a more recent work quantifying the linkage between economic diversity and economic growth (e.g. Wagner and Deller, 1998).

Furthermore, we use variables such as *specialization* and *human capital* (see e.g. Rauch, 1992). Results presented by Glaeser et al (1995) and Andersson et al (2004) seem to indicate that those areas with high levels of human capital experience large increases in income per capita, linking human capital to rate of foreclosure. Similar variables describing the industrial composition of the labor market used in this study are also utilized in Andersson et al (2005) and economic diversity and the creation of innovations (Andersson et al 2005).

4. Data and Descriptive Statistics

4.1. The panel

The utilized data set is a cross-sectional time-series panel data set consisting of 100 labor markets in Sweden over the period 1994 to 2001.[3] The Swedish Labor Ministry defines the labor markets based on commuting patterns and use methods analogous to those used to define MSAs (Metropolitan Statistical Areas) in the United States. Most, but not all, of Sweden's 100 labor-market areas contain a central city and a number of surrounding jurisdictions.

4.2. Foreclosures in Sweden

Foreclosure data for all Swedish municipalities during the period 1994-2001 aggregated up to the labor-market level are used. The Swedish Enforcement Authority has the responsibility of enforcing foreclosures in Sweden. This authority sells the real estate via public auction where the highest bid is accepted. The condition for a foreclosed sale is that the expected price is larger than the mortgage balance. After the auction, the creditors distribute the purchase sum among themselves. As in the majority of the U.S. states, Sweden allows for deficiency judgements (Clauretrie, 1989); that is, if "the proceeds from the foreclosure sale are insufficient to satisfy the loan balance /…/, a lender (insurer) can proceed to recover directly against the mortgagee's personal assets."

An early study by Bjork (1994) examines the experience of repossession. She states that it takes one to two years from mortgage delinquency to foreclosure sales. However, in order to speed up the procedure, the legal process was changed in early 1992. Today, the estimated average time in Sweden to foreclose is only four to six months (Sanders, 2005). Thus, the foreclosure proceedings are less than one year (compared to five years in Italy and less than

[3] The data comes primarily from NAI Svefa, Statistics Sweden (SCB) and Föreningssparbanken - SPINTAB.

one year in some states in the U.S.), which means that there is most likely no time lag between foreclosure rates and the independent variables used in the empirical models.

According to Jappelli and Pagano (1994), the legal cost of a mortgage foreclosure for the mortgage lender is about 5 percent of the selling price. The total number of enforced sales amounted to 14,000 over the period 1993 to 2006, compared to 600,000 sales in total. The highest number of foreclosures occurred in 1994 with almost 3,000 enforced sales out of around 46,000 sales all together. The selling price of all foreclosed sales totaled SEK 7,458 million. Hence, totalling the legal cost only, amounted to SEK 373 million. Thus, investigating the determinants of foreclosure is of great economic importance.

It is important to remember that not all mortgage defaults end up as foreclosure sales, as we illustrate in Figure 1. The foreclosure rates that we have estimated, and are using, are an underestimation of the mortgage defaults, which are an underestimation of all mortgage delinquencies. Furthermore, less than 10 percent of all applied foreclosures end up in a foreclosure (see Bjork, 1994).

Figures 2 and 3 illustrate the foreclosure rates over time and across labor markets. It is apparent that foreclosure rates have declined over the years, with a maximum number occurring around 1994/95. At its maximum, the foreclosure rate in Sweden was around 8 percent, a very high number in comparison to other countries (e.g. Bond, 2007, reports that the U.S. rate increased from 0.9 percent in 1994 to around 1.5 percent in 2004). The foreclosure rates started to decline in 1996 and by the year 2004, they were lower than 1 percent. Regardless whether we measure foreclosures as number of sales or the value of the properties, the foreclosure rates are, on average, of the same magnitude. Why have they been so high? The extremely high-observed foreclosure rates can be explained by the financial crises that Sweden experienced in the beginning of the 1990's with interest rates as high as 12 percent on average in 1995.

FIGURE 2 IN HERE:

The average foreclosure rate over the whole period is close to 2 percent. Out of 100 labor markets, 11 have an average foreclosure rate over the period below 1.5 percent. Half of the labor markets show a number that is higher than 3 percent, and as many as 21 labor markets have a foreclosure rate above 4.5 percent. The spatial distribution of foreclosures is not uniform across the labor markets in Sweden, and this is further highlighted in Figure 3 (larger dots indicate higher foreclosure rates). Some markets exhibit a larger number of foreclosed sales compared to the total number of sales in the housing market, while other have had a very low number of foreclosed sales in the whole investigated period.

FIGURE 3 IN HERE

The spatial pattern is very scattered and not very clear, but it seems that the foreclosure rates have been lower in the middle and the south parts of Sweden, while they have been higher in the inland parts of the north. Of course, this pattern is very similar to the inverse spatial distribution of, for example, income per capita.

4.3. The explanatory variables

Besides data on foreclosure, we have supplemented the data set with a number of other variables. The first set of data is some key variables in the option-based models. The variables that are included are *income per capita*, *interest rate*, *rent-to-price ratio*, *employment rate*, price volatility (*sigma*) and the price change (*delta-price*) in the previous five years.

We utilize the five-year fixed-mortgage interest rate. As for rents, the data available are based on a national survey conducted by Statistics Sweden for the average rent level in a number of municipalities (45–66 out of 290, depending on the year) and in all counties. We use a measure of the rent at the municipal level if possible, otherwise at the county level. We have subsequently aggregated the municipality data up to the labor market level.

The second type of data we utilize is data concerning the labor markets. As a proxy for agglomeration, we define the *Density* as employment per squared kilometer. Moreover, the numbers employed in the non-agricultural sector defines the market *size* variable, and the inverse of the Hirfindahl-index defines the *Diversity* variable.[4] Finally, the share of employment within the manufacturing sector measures the *Specialization* in the labor market, and the measure of *Human capital* is equal to the average number of years in the school system. The table below describes the dependent variable and the independent variables used in the econometric analysis.

TABLE 1 IN HERE

Sweden has experienced a major decrease in foreclosure rates in recent years. However, the foreclosure rates are still higher than in other countries. Foreclosures average 4.2 percent of the turnover over the period 1994-2001 (3.7 percent if we measure foreclosure as the value instead of the number of sales).[5] However, the standard deviation is high suggesting that the spatial distribution is of significance. At the same time, housing prices have increased over the period 1990-1997, and the price volatility has been high. As anticipated, the variation around the average rent level has been small due to non-market rents. The coefficient of variation is, for example, only 0.05 regarding the rent level, compared to 0.38 regarding the owner-occupied housing market. That is, the variation has been much higher on the more competitive housing market.

The average interest rate (five-year fixed mortgage rate) has been as high as 8 percent with a standard deviation of almost 2 percent. A credit expansion in the late 1980s led to a property boom. The interest rates went up to around 12 percent in the early 1990s, caused by a

---

[4] $H_i = \sum_{s=1}^{S_i} \left(\frac{e_{si}}{e_i} \right)^2$, where S is the total number of industries (24 in our case) within region i and e is the numbers employed.

[5] We are using a shorter period due to lack of data concerning some of the independent variables, that is, we are investigating the period 1994-2001 and not 1993-2006.

tightening credit market (see Allen, 2001). Subsequently, the property prices plunged, and as the lending was based on inflated property values, many borrowers obtained a negative LTV.

All labor-market characteristics vary considerably across the labor markets. Especially, market size and density varies. For example, some labor markets are very small, and others are very large with a very low density. However, the exception is diversity. The variation around the average is very low, indicating that the variable may be a poor explanatory variable. As we are using the geographical-level labor market, most of these are large and rather well diversified in respect to industrial composition.

5. Econometric Results

5.1. The basic option-based models

The table below presents the results from the option-based model. The first three models use foreclosure rates based on the number of foreclosed sales related to all sales as a dependent variable, and the last three models use foreclosure rates based on the value. In an attempt to test whether we have a problem with spatial dependency in our empirical models that explain the foreclosure rate, we have used the test statistic Moran's I. If present, we are estimating spatial autoregressive models (SAR) and spatial error models (SEM). [6]

We have estimated one OLS (ordinary least square) model and two different types of spatial econometric models (SAR and SEM). Independent variables used are price volatility (*sigma*), price changes (*Delta-price*), interest rate, income per capita, employment rate, and rent-to-price ratio. We have defined the spatial weight matrix as the inverse of the squared distance

[6] See Anselin (1988) for an in-depth discussion about spatial econometrics. Using longitude and latitude coordinates (squared and cross products) [suggested by Galster et al (2004)] do not reduce the spatial dependency in our data set.

between the geographical middle points of the labor markets per year.[7] That is to say, we do not allow any lagged spatial dependency lagged in time.

TABLE 2 IN HERE

The spatial variation in foreclosure rates can be highly explained by the option-based model. The included variables explain around 30 percent of the variation, a number comparable with, for example, Case and Shiller's (1996). As anticipated, price volatility relates positively to foreclosure rates, that is, high foreclosure rates can be observed in high-price-volatility labor markets. Pennington-Cross (2003) also found that the greater the housing volatility, the more severe the mortgage foreclosure losses. The economic interpretation is straightforward. If price volatility increases by 10 percent, the foreclosure rate will increase by 0.3-0.8 percent.

On the other hand, price increases over the past five years reduce the foreclosure risk. However, this is not true if we take into account the spatial autocorrelation. Moreover, as income and employment increase, foreclosure risk is reduced. Furthermore, as expected, interest rate has a positive impact on the risk, and the parameter seems to be robust. If interest rate increases by 1 percent, the foreclosure rate is expected to increase by 0.4-0.9 percent. Even after taking care of the spatial autocorrelation problem, the parameter is of the same magnitude (slightly lower t-values).

Moreover, the housing-rent level in the labor market seems to increase the risk. The parameter concerning the dividend variable (expected rent-to-price) of owner-occupied housing is negative as in Capozza et al (1997). That is to say, if the dividend goes up, foreclosure risk goes down. The parameter concerning the expected rent-to-price ratio is remarkably robust. Moreover, an income increase by SEK 20,000 will reduce the risk by 0.3-0.5 percent, and an

[7] We have also used the inverse of the distance. However, the results are not altered, and the same conclusions can be drawn.

increase in the employment rate by 1 percent decreases the risk by almost 0.1 percent. The presence of spatial dependency does not affect the estimate.

Surprisingly, many of the variables do not have statistically significant parameters if we instead analyze the value of the foreclosed sales. The exceptions are interest rate, income and dividends. It is also much harder to explain the regional variation in foreclosure rates with the included variables.

The conclusion regarding spatial dependency is that it causes only a minor problem. Therefore, we will henceforth estimate all the models with OLS, but we will also estimate and present Moran's I.[8]

5.2. Robustness test

There is a major difference regarding foreclosure rates between the first four years in our study period and the last four years. To test whether our parameters are robust or not, we have split the data set into two (1994-1997 and 1998-2001) and have estimated the models separately. Moreover, it is important to recognize the problem of timing. We have information about the year of the foreclosure sale. Despite the fact that foreclosure proceedings take a short time in Sweden, it may be quite likely that the mortgage default or even the mortgage delinquencies appeared a number of years earlier. In an attempt to test for this, we have estimated the models with a time lag. Even if a multiple-year lag structure would have been more accurate, we have only tested a one-year and two-year lag due to lack of data.

Finally, in an attempt to test for robustness in our estimated parameters, we have tested for temporal autocorrelation. It is problematic to test the hypothesis of a first-order autoregressive process AR(1), as the distribution concerning the test statistics is complicated (Wigren and

[8] All spatial models are available upon request.

Wilhelmsson, 2007). In the option-based model above we have estimated an AR(1) model on the disturbance and have tested the hypothesis that $\rho>0$, that is, no temporal autocorrelation.

TABLE 3 IN HERE

As said, there is a large difference between the foreclosure rates the first four years (6.0 percent in average) and the last four years (2.4 percent in average). As anticipated, the included variables relate differently in a high-foreclosure rate regime than in a low-foreclosure rate regime. For example, interest rate does not explain the reduction in risk the last four years, but is very important explaining the rise and decline in the first investigated years. Naturally, the reason for this difference is that the decline in risk of defaults has been very modest the last four years. Price volatility is important in the second period, but not at all in the first. The reverse is true when it comes to price changes. Case and Shiller (1997) use the same type of data that we have, namely, aggregated data, and their results also imply that high default rates strongly follow house price declines (in their case between 1975 and 1993). The rent and price level variable is more important in the first high-foreclosure rate period.

As anticipated and judging by R^2, the one and two-year lag models did not seem to be superior to the non-lagged model. However, it appears that price volatility, price change and employment rate relate more strongly to foreclosure rate in the lagged models (higher t-values). As the overall result is not significantly better, we will continue to use the non-lagged representation of the model.[9]

All models exhibit temporal (and spatial) autocorrelation. We can reject the null hypothesis of no autocorrelation and hence the disturbance follows an AR(1) process due to, for example, non-stationary in the data. If we include an AR-representation in the model, the explanation power increases substantially. The parameter concerning AR(1) is highly significantly

[9] It may be more reasonable to expect that the relationship is temporal lagged the first examined years and not lagged at all the last years, as the legal process has gradually improved. The empirical tests that we have performed weakly indicate that (not presented).

different from zero (and positive). However, the parameters concerning the option-based variables appear to be rather robust and the economic interpretation only moderately changes this.

5.3. The extended option-based models

In the table below the results from the more exploratory models are presented. As dependent variables are the following: price volatility, price change, interest rate, income per capita, employment, and the rent-to-price ratio, and they are used, respectively. As independent variables, different measures characterizing the labor market are used. These variables measure the size of the market, the density, diversity and the degree of specialization of the labor market. We have also tested whether we have a problem with spatial autocorrelation.

TABLE 4 IN HERE

One word of caution, all the models exhibit spatial and temporal autocorrelation, and some of the variables are not stationary. In other words, it is hard to interpret the results. Nevertheless, the results indicate that the labor market characteristics are all highly important in explaining the variation in the underlying key option-model variables. Labor market size seems to relate positively to: price volatility, income per capita, and employment rate. For example, the variation in house prices seems to be higher in large labor markets, *ceteris paribus*. The price volatility, income per capita, and rent-to-price ratio, all relate negatively to employment density. A strong inverse correlation between income and density has also been found in, for example, Mills and Tan (1980), as well as in Kurban and Persky (2007). In our case, the negative correlation could be a result of a high correlation between market size and density, that is, we have some kind of a multicollinearity problem. The correlation between labor market size and density is 0.7.

The diversity of the industrial sector relates positively to price changes and employment rate. Diversity is negatively related to price volatility and, surprisingly, to rent-to-price ratio and income per capita. The latter implies that Jane Jacobs' hypothesis does not hold. That is to say, our result is in contradiction to Glaeser et al (1992), who found that diversity has a positive effect on economic output. Specialization of the labor market seems to have a mixed result on the key option variables. The same is true of human capital. Both variables increase the income per capita (as in Andersson et al, 2004) and increase house prices, but lower the housing price volatility. The former results confirm Glaeser et al's (2006) results. In the table below, the foreclosure rates are explained by the labor market characteristics.

TABLE 5 IN HERE

As before, we have a problem with spatial autocorrelation. However, the spatial autoregressive model and the spatial error model (not presented) show that the result is robust. Of the labor market characteristics, density, specialization, and human capital seem to have an effect on the foreclosure rates. Controlling for the option-based variables' specialization and human capital still explain some of the variation in foreclosure rates. This is especially true if we examine the results using foreclosure rates based on the number of foreclosed sales and not on the value of them.

The foreclosure rate increases by 0.4 percent if employment density increases from 10 employees per square kilometers to 20, hence, a very small effect. However, the risk reduces by 0.8-1.5 percent if specialization increases by one standard deviation. If the average educational background increases by one year, the model predicts that the risk drops by as much as 6 percent.

6. Conclusions

Mortgage defaults are important to the lending industry as well as to borrowers and investors. During the last two decades, research papers have focused increasingly on the pricing of credit risk in the mortgage market. The key objective with our examination is to analyze whether different aspects of agglomeration and diversity play a role in explaining the variation in foreclosure rates.

To summarize the literature review, most researchers have concluded that the loan to value ratio (LTV) is an important determinant of the probability of a mortgage default. Little equity, in addition to situations with relatively low or decreasing house prices, increases the likelihood of a mortgage default. Some researchers have found that life events such as job losses, declines in income, moving rates, divorce, illness, and death, etc, trigger mortgage defaults, while evidence presented by Capozza et al (1997 and 1998) give an argument for playing down the importance of such explanatory variables. However, it might be important to distinguish between regions that are growing, as measured by increasing income per capita, decreasing unemployment rates and increasing house prices, on the one hand, and regions with contrary development on the other. Such a distinction might illuminate why fundamental explanatory variables, such as the life events mentioned above – on average for a country – do not seem to determine very much of the mortgages defaults and foreclosures. They might for regions that are not doing so very well. For example, a divorce experienced in a growing region with increasing home prices need not lead to a mortgage default or a foreclosure. Divorced owners could simply sell the houses, and that is it, while a divorce could be fatal in a region with decreasing house prices.

Moreover, what may explain decreasing house prices? That could be due to the development of abovementioned fundamental explanatory variables that then trigger mortgage defaults and foreclosures. The aim of this paper is to test such a hypothesis.

The basic model is an option-based model in which the foreclosure risk is explained by variables such as interest rate, price volatility, price changes, rents and prices, together with transaction or reputation cost variables like income and trigger variables such as employment rate. We extend this option-based model by including agglomeration variables such as information about the industrial composition of the labor market. To help us, we are utilizing a panel data set, which include 100 labor markets in Sweden over the period 1994-2001. Our data has information about foreclosure rates and option based variables, together with the variables describing the industrial composition of the labor markets.

The most consistent result seems to be that a higher educational level of the workforce reduces the foreclosure risk. Another result is that a higher degree of employment in the manufacturing industry reduces the risk. Hence, specialization has a lowering effect on foreclosure risk. Moreover, a third result is that density and diversity have a very small (or no effect at all) on the risk. Furthermore, the option-based model can explain foreclosure risk measured as the number of foreclosed sales related to the total number of sales. However, foreclosure risk measured in value of the sales is much more difficult to explain.

What are the policy implications? For banks and mortgage lenders, it is of great interest to be able to determine whether a labor market falls in a certain risk category. In an attempt to illustrate and categorize the labor markets according to their risk, we examine labor markets with an above average foreclosure risk. Here, we define a risky labor market as a labor market where the foreclosure rate is above the national average. Hence, our response variable is a binary variable equal to one, if the labor market is more risky than the national average, and otherwise, it is equal to zero. We can categorize more than 36 percent of the labor markets as risky.

By estimating a probability model, we are able to classify the labor markets. In the current paper the classification is carried out by a logit regression model because for probabilistic events such as above average risk, outcomes must fall in the [0,1] interval. In the table below, we exhibit the result from the logistic regression.

TABLE 6 IN HERE

The first model uses all explanatory variables, that is, both the option-based variables and the labor market characteristics. Very few of the option-based variables can explain the probability of being an above-average risky labor market. Only the price-level variable has some explanatory power. A higher price level seems to reduce the risk. Out of the labor market variables, only specialization and human capital can explain the probability of being a risky labor market.

In order to determine whether we can utilize the model to classify the labor market as more or less risky, three different statistics are used. The first statistic is *sensitivity*, which measures the probability of classifying the labor market as a risky market. The second statistics are *the specificity*, which estimates the probability that we classify the market as non-risky when it is non-risky. A third measure is the statistics, *false,* that estimates the probability that the market is non-risky on the condition that it is risky. Finally, the measure *correctly classified* gives an answer to the question of what percentage of the labor markets we correctly forecasted.

As can be observed in the table, the first model, using all variables, correctly classifies 67.1 percent of all labor markets. The percentage is higher for *specificity* than for *sensitivity*. The probability of classifying a labor market as non-risky when it is risky equals 63 percent.

The model using only the option-based variables performs much worse. Only 64.9 percent of the labor markets are correctly classified, and the probability of classifying a labor market as non-risky when it is risky has dropped to 59 percent.

The last model, using only labor market characteristics, outperforms the other models. In its simplicity, almost 70 percent of the labor markets are correctly classified. The variables that explain the likelihood of being an above-average risky labor market are density, as well as specialization and human capital. All of them relate negatively to risk. Specifically, a dense labor market with a high degree of its employment within the manufacturing industry, and with a more educated workforce, reduces the risk of foreclosure being above average. Thus, mortgage lenders and banks can reduce their risk by concentrating their business in regions with the above characteristics.

References

Allen, F. (2001). Financial Structure and Financial Crisis. *International Review of Finance*, Vol.2, 1-19.

Ambrose, B. and Capone, C. (1996) "Cost-Benefit Analysis of Single Family Foreclosure Alternatives", *Journal of Real Estate Finance and Economics*, 13 (2), 105-120.

Ambrose, B. and Capone, C. (1998), Modeling the Conditional Probability of Foreclosure in the Context of Single-Family Mortgage default Resolutions, *Real Estate Economics*, Vol. 26(3), 391-429.

Ambrose, B., Capone, C and Deng, Y. (2001), Optimal Put Exercise. An Empirical Examination of Conditions for Mortgage Foreclosure, *Journal of Real Estate Finance and Economics,* Vol. 23(4), 213-.

Andersson, R., Quigley, J. and Wilhelmsson, M. (2004), University Decentralization as regional policy: The Swedish Experiment. *Journal of Economic Geography*, Vol.4, 371-388.

Andersson, R., Quigley, J. and Wilhelmsson, M. (2005), Agglomeration and the spatial distribution of creativity, *Papers of Regional Science*, Vol. 84(3), 445-464.

Anselin, Luc (1988), *Spatial Econometrics: Methods and Models.* Dordecht: Kluwer Academic Publishers.

Bjork, M. (1994). Investigating the Experience of Repossession: A Swedish Example. *Housing Studies*, Vol.9(4), 511-529.

Bond, M. (2007), Examining Mortgage Default Rates in Ohio, Paper presented at the 23rd Annual Meeting of the American Real Estate Society, San Francisco, California, April, 11-14, 2007.

Capozza, R. D., Kazarian, D. and Thomson, T. (1997), Mortgage Default in Local Markets, *Real Estate Economics*, Vol. 25(4), 631-655.

Capozza, D., Kazarian, D. and Thomas, D. (1998) The Conditional probability of Mortgage Default, *Real Estate Economics*, Vol. 26(3) 359-389.

Case, K.E. and Shiller, R.J. (1996). Mortgage Default Risk and Real Estate Prices: The Use of Index-Based Futures and Options in Real Estate. *Journal of Housing Research*, Vol. 7(2), 243-258.

Clapp J., Goldberg, G.M., Harding, J.P. and LaCour-Little, M. (2001), Movers and Shuckers: Interdependent Prepayment Decisions. *Real Estate Economics*, Vol.29(3), 411-450.

Clauretie, T.M., (1989). State Foreclosure Laws, Risk Shifting and the Private Mortgage Insurance industry. *The Journal of Risk and Insurance*, Vol.56(3), 544-554.

Danis, M. and Cross, A. (2005), A Dynamic look at Sub prime Loan Performance, *Journal of Fixed Income*, Vol. 15, 1, 28-39.

Foster, C. and Van Order, R. (1985). FHA Terminations. A prelude to Rational Mortgage Pricing, *AREUEA Journal*, Vol. 13, 3, 273-291.

Galster, G., Tatian, P. and Pettit, K. (2004). Supportive Housing and Neighborhood Property Value Externalities. *Land Economics*, Vol.80(1), 33-54.

Glaeser, E., Gyourko, J. and Saks, R. (2006). Urban Growth and Housing Supply. *Journal of Economic Geography*, Vol.6, 71-89.

Glaeser, E., Kallal, H. D., Scheinkman, J. and Shleifer, A. (1992). Growth in Cities. *The Journal of Political Economy*, Vol.100(6), 1126-1152.

Glaeser, E., Scheinkman, J. and Shleifer, A. (1995). Economic Growth in a Cross-Section of Cities. *Journal of Monetary Economics*. Vol.36, 117-143.

Glaeser, L. E. and Maré, C. D. (2001). Cities and Skills, *Journal of Labor Economics*, Vol. 19, no 2.

Hoover, E.M. (1937). *Location Theory and the Shoe and Leather Industries*. Cambridge, MA: Harvard University Press.

Jacobs, J. (1662). *The Death and life of great American cities*. Random House, New York.

Jacobs, J. (1969). *The economy of cities*. Random House, New York.

Jappelli, T. and Pagano, M. (1994). Saving, Growth, and Liquidity Constraints. *The Quarterly Journal of Economics*, Vol.109(1), 83-109.

Jung, A. (1962). Terms on Conventional Mortgage Loans on Existing Homes. *Journal of Finance*, Vol.17, 432-43.

Kau, J. B., Keenan, D. C., and Kim, T. (1992 and 1994), Default Probabilities for Mortgages, *Journal of Urban Economics,* 35 (3): 278-296.

Kurban, H. and Persky, J. (2007). Do Metropolitan Areas with Rish Central Cities Experience Less Sprawl? *Economic Development Quarterly*, Vol.21(2), 179-184.

Lauria, M., Baxter, V., Bordelon, B. (2004), An investigation of the time between mortgage default and foreclosure, *Housing Studies*, Vol. 19, 4, 581-600.

Lekkas, V., Quigley, M. J. and Van Order, R. (1993), Loan Loss Severity and Optimal Mortgage Default, *Journal of the American Real Estate and Urban Economics Association*, V21, 4: 353-371.

Marshall, A. (1925). Principles of Economics, 8th edn. London: Macmillan.

Miller, N. and Peng, L. (2006), Exploring Metropolitan Housing price Volatility, *Journal of Real Estate Finance and Economics,* 33: 5-18.

Mills, E.S. and Tan, J.P. (1980). A Comparison of Urban Population Density Functions in Developed and Developing Countries. Urban Studies, Vol.17(3), 313-321.

Pennington-Cross, A: (2003), Subprime and prime mortgages: loss distributions, OFHEO, Working paper 03-1.

Quercia, R.G. and Stegman, M.A. (1992). Residential Mortgage Default: A Review of the Literature. *Journal of Housing Research*, Vol.3(2), 341-379.

Quigley, J.M. (1998). Urban Diversity and Economic Growth. *The Journal of Economic Perspective*, Vol.12(2), 127-138.

Quigley, J.M. (1999), Real Estate Prices and Economic Cycles, *International Real Estate Review*, 20: 325-345.

Quigley, J.M. and Van Order, R., (1995), "Explicit Tests of Contingent Claims Models of Mortgage Default, *Journal of Real Estate Finance and Economics*, 11(2): 99-117.

Rauch, J. (1992), Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities, *Journal of Urban Economics*, 34: 380-400.

Roback, J. (1982), Wages, Rents and the Quality of Life, *Journal of Political Economy*, 90: 1257-78.

Rosen, K. and Smith, L. (1983). The price adjustment process and the natural vacancy rate. *American Economic Review. 73, 779–786.*

Sanders, A. (2005). Barriers to homeownership and housing quality: The impact of the international mortgage market. *Journal of Housing Economics*, Vol.14(3), 147-152.

Topel, H. R., (1986), Local Labor Markets, *Journal of political Economy*, Vol 94, no 3: 111-143.

Vandell, K. (1995), How ruthless is Mortgage Default? A Review and Synthesis of the Evidence, *Journal of Housing Research*, 6(2): 245-264.

Vandell, K. D. and Thibodeau, T. (1985), Estimation of Mortgage Defaults Using Disaggregate Loan History Data, *AREUEA Journal,* Vol.13(3): 292-316.

Wagner, J.E. and Deller, S.C. (1998). Measuring the effects of economic diversity on growth and stability. *Land Economics*, Vol.74(4), 541-560.

William, A.O., Beranek, W. and Kenkel, J. (1974). Default Risk in Urban Mortgages: A Pittsburgh Prototype Analysis. *AREUEA Journal*, Vol.2(2). 101-112.

Wigren, R. and Wilhelmsson, M. (2007). Housing Stocks and Price Adjustments in 12 West European Countries between 1976 and 1999. *Housing, Theory and Society*, Vol.24(2), 133-154.

Figure 1. From Mortgage Origination to Enforced Foreclosure Sale

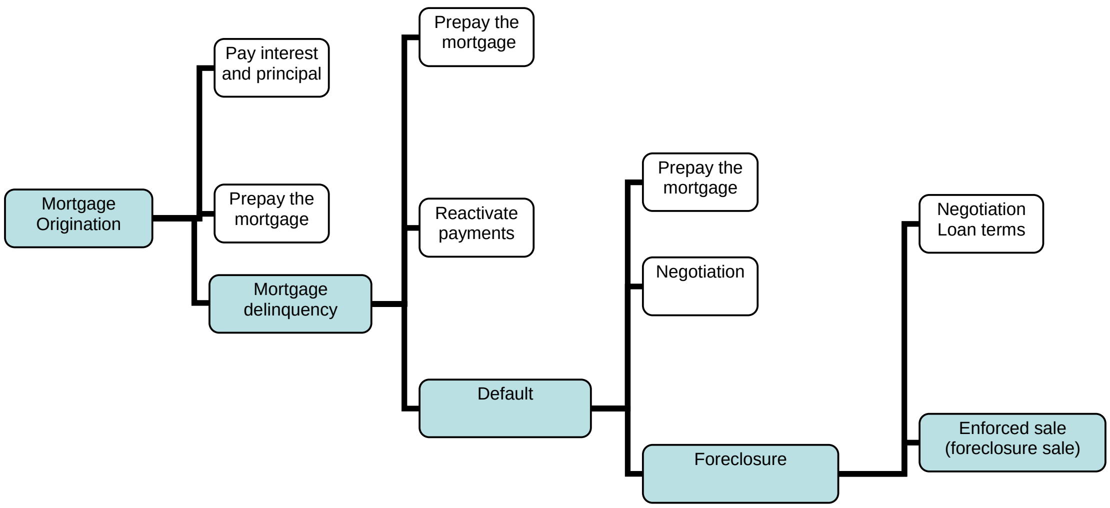

Figure 2. Foreclosure rates in Sweden over time, 1994-2001.

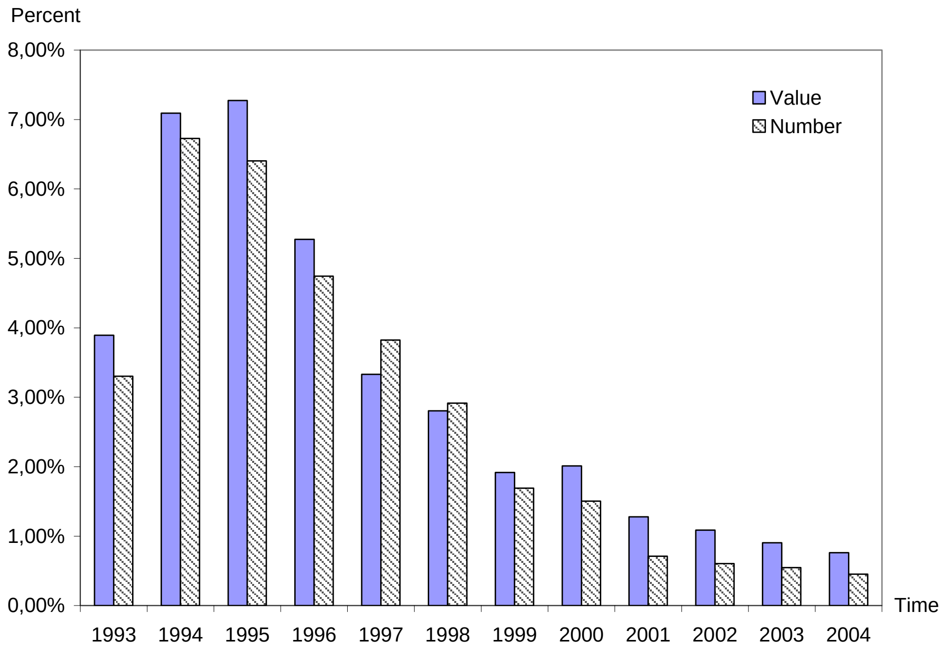

Figure 3. Spatial Distribution of Foreclosures in Sweden, 1995 and 2001.

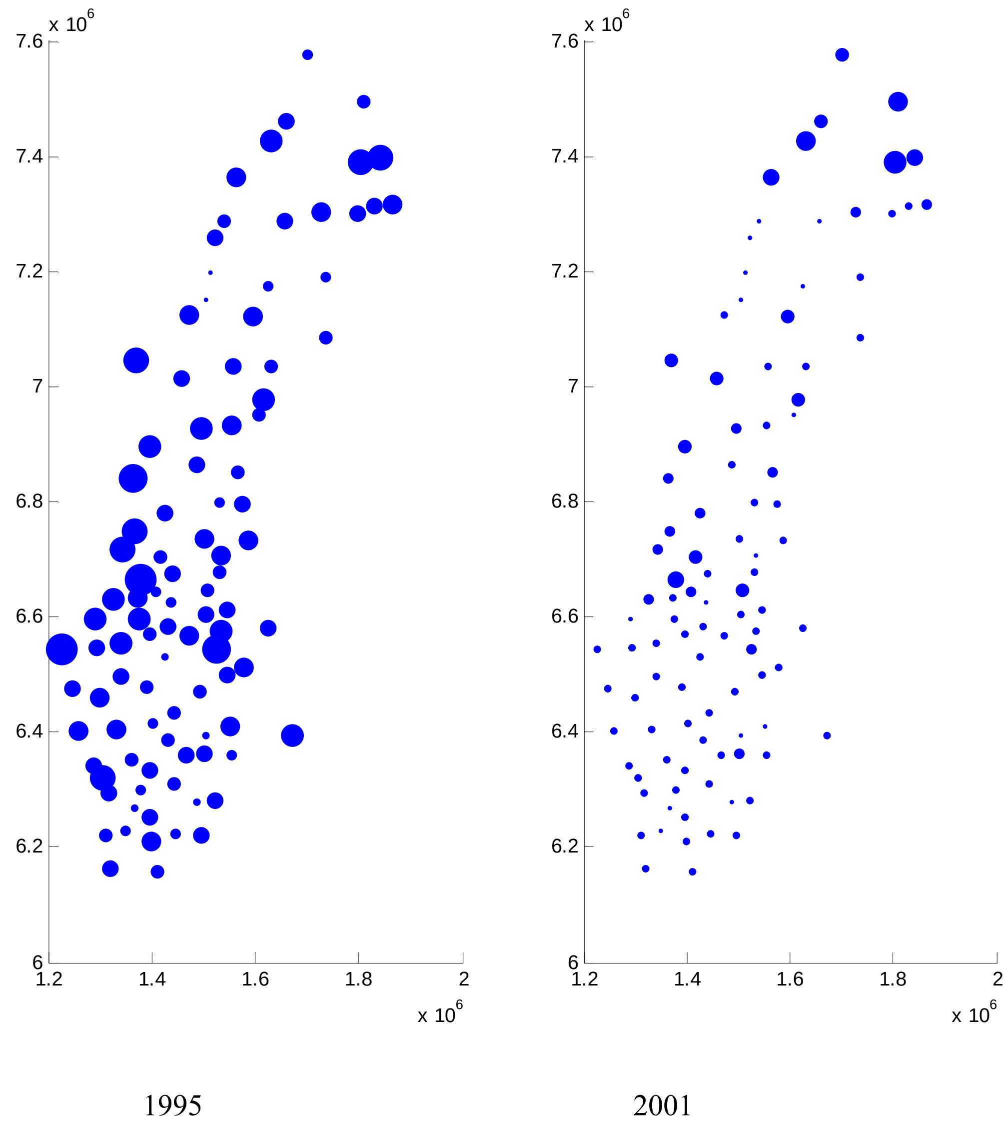

Table 1. Descriptive Statistics.

Variable	Unit	Average	Standard deviation
Foreclosure rate - number	Percent	0.0423	0.0407
Foreclosure rate - value	Percent	0.0367	0.0449
Sigma	Percent	0.0924	0.0481
Price change	Percent	0.0917	0.1564
Rent	SEK	658.5020	36.5691
Price	SEK (000)	485.7503	183.6670
Rent-to-price	Ratio	1.5054	0.4568
Expected Rent-to-price	Ratio	1.5054	0.3710
Interest Rate	Percent	8.1337	1.9631
Income per capita	SEK (000)	121.7672	17.2364
Employment rate	Percent	0.4106	0.0417
Size	Employment (0000)	3.7073	10.7115
Density	Employment per square kilometer	10.4918	13.2739
Diversity	Inverse of Herfindahl-index	0.8827	0.0219
Specialization – manufacturing	Percent	0.2241	0.1022
Human Capital	Years	10.7309	0.3227

Table 2. Option Based Model (OLS and Spatial Econometric Models).

	FC-number			FC-value		
	OLS	SAR	SEM	OLS	SAR	SEM
Sigma	0.0742	0.0700	0.0804	0.0347	0.0346	0.0379
	(2.60)	(2.55)	(2.78)	(1.03)	(1.03)	(1.13)
Delta-Price	-0.0213	-0.0139	-0.0126	0.0049	0.0068	0.0089
	(-1.95)	(-1.32)	(-0.92)	(0.38)	(0.54)	(0.66)
Interest Rate	0.0078	0.0047	0.0092	0.0076	0.0061	0.0086
	(7.48)	(4.26)	(6.74)	(6.07)	(4.92)	(6.61)
Income per capita	-0.0003	-0.0002	-0.0001	-0.0005	-0.0004	-0.0004
	(-2.80)	(-2.17)	(-1.02)	(-4.07)	(-3.39)	(-2.81)
Employment rate	-0.0834	-0.0719	-0.0923	-0.0001	-0.0089	-0.0192
	(-2.32)	(-2.07)	(-2.45)	(0.00)	(-0.21)	(-0.44)
Expected Rent-to-Price	-0.0325	-0.0369	-0.0370	-0.0295	-0.0321	-0.0328
	(-2.53)	(-2.97)	(-2.43)	(-1.93)	(-2.13)	(-2.10)
Rent	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	(1.91)	(23.48)	(9.22)	(0.95)	(12.73)	(11.74)
Price	-0.0001	-0.0001	-0.0001	-0.0001	-0.0000	-0.0001
	(-2.52)	(-2.87)	(-0.95)	(-0.88)	(-1.28)	(-1.09)
Constant	0.0385	0.0233	-0.0046	0.0440	0.0307	0.0112
	(0.88)	(0.88)	(-0.17)	(0.84)	(0.95)	(0.34)
Rho	-	0.3974	-	-	0.2278	-
		(5.43)			(3.08)	
Lambda	-	-	0.4460	-	-	0.2420
			(6.86)			(7.42)
R-square	0.3168	0.3441	0.3587	0.2050	0.2176	0.2175
Moran's I	6.7417	-	-	3.2404	-	-

Note: t-values within parentheses.

Table 3. Option Based Model (split data set, AR(1) and 1-2 years lag).

Dependent variable: Foreclosure rates – number of sales

	All	1994-1997	1998-2001	1 Year lag	2 Year lag	AR(1)
Sigma	0.0742	0.0045	0.1476	0.0799	0.0808	0.0597
	(2.60)	(0.10)	(4.72)	(2.69)	(2.93)	(2.12)
Delta-Price	-0.0213	-0.0516	0.0172	-0.0349	-0.0436	-0.0278
	(-1.95)	(-2.85)	(1.34)	(-3.08)	(-3.98)	(-2.59)
Interest Rate	0.0078	0.0071	-0.0018	0.0044	0.0003	0.0021
	(7.48)	(3.85)	(0.51)	(4.12)	(0.32)	(2.01)
Income per capita	-0.0003	-0.0004	-0.0001	-0.0003	-0.0001	-0.0002
	(-2.80)	(-1.96)	(-1.25)	(-2.49)	(-1.14)	(-1.72)
Employment rate	-0.0834	-0.1109	-0.0830	-0.1069	-0.0947	-0.0828
	(-2.32)	(-1.71)	(-2.22)	(-2.87)	(-2.62)	(-2.34)
Expected Rent-to-Price	-0.0325	-0.0508	-0.0116	-0.0296	-0.0407	-0.0171
	(-2.53)	(-2.25)	(-0.88)	(-2.23)	(-3.21)	(-1.36)
Rent	0.0001	0.0003	-0.0001	0.0001	-0.0001	-0.0001
	(1.91)	(3.28)	(-1.48)	(0.05)	(-0.98)	(-0.76)
Price	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(-2.52)	(-2.35)	(-1.89)	(-1.36)	(-2.51)	(-0.56)
AR	-	-	-	-	-	0.3051
						(9.18)
Constant	0.0385	0.0035	0.1553	0.1260	0.1823	0.1144
	(0.88)	(0.05)	(3.09)	(2.84)	(4.37)	(2.72)
ρ-AR(1)	0.3025	0.2246	0.2897	0.2631	0.2912	-
	(9.40)	(4.55)	(6.97)	(8.14)	(7.62)	
R-square	0.3168	0.1891	0.1379	0.2853	0.2217	0.3621
Moran's I	6.7417	3.9772	6.1209	7.8951	5.1273	7.81
Obs	800	400	400	700	600	700

Note: t-values within parentheses.

Table 4. Labor Market Models (OLS).

	Dependent variables Sigma	Delta-price	Interest rate	Income per capita	Empl. rate	Price	Rent	Exp.Rent-to-price
Size	0.0010	-0.0007	0.0141	0.4644	0.0003	4.4112	0.1135	-0.0003
	(4.73)	(-0.96)	(1.63)	(7.32)	(2.04)	(6.70)	(0.71)	(-0.55)
Density	-0.0006	0.0006	0.0387	-0.2864	-0.0001	3.7893	0.0679	-0.0023
	(-3.31)	(0.97)	(5.14)	(-5.20)	(-0.02)	(6.64)	(0.49)	(-5.61)
Diversity	-0.0469	0.1961	0.7729	-10.5081	0.0611	270.1195	-15.8429	-0.1562
	(-3.22)	(4.15)	(1.32)	(-2.45)	(5.55)	(6.07)	(-1.47)	(-4.91)
Specialization	-0.1021	0.3630	-3.8335	67.0489	0.2407	443.7672	58.1260	0.1175
	(-5.37)	(5.90)	(-5.01)	(11.99)	(16.76)	(7.65)	(4.15)	(2.83)
Human Capital	-0.0546	0.1829	-3.6456	36.3283	0.0555	119.2855	54.6159	0.0267
	(-8.21)	(8.50)	(-13.64)	(18.57)	(11.04)	(5.88)	(11.15)	(1.84)
Constant	0.7400	-2.1080	47.0577	-273.682	-0.2872	-1158.82	70.5201	1.3371
	(10.33)	(-9.09)	(16.32)	(-12.97)	(-5.30)	(-5.30)	(1.34)	(8.55)
R-squared	0.1627	0.1682	0.1852	0.4341	0.3637	0.4638	0.2121	0.1390
Moran's I	5.0791	27.5482	59.2790	25.5079	8.5049	10.6155	39.5044	10.0670

Note: t-values within parentheses.

Table 5. Foreclosure models with labor market characteristics.

	No. Foreclosure		Value. Foreclosure	
Sigma	-	0.0621	-	0.0267
		(2.17)		(0.77)
Delta-price	-	-0.0201	-	0.0014
		(-1.84)		(0.11)
Interest rate	-	0.0075	-	0.0069
		(6.75)		(5.19)
Income per capita	-	-0.0003	-	-0.0005
		(-2.39)		(-4.00)
Employment rate	-	0.0073	-	0.0423
		(0.19)		(0.91)
Exp. Rent-to-price		-0.0136		-0.0166
		(-1.04)		(-1.05)
Rent		0.0001		0.0001
		(1.57)		(0.59)
Price	-	-0.0001	-	-0.0001
		(-2.52)		(-1.67)
Size	0.0092	0.0002	0.0001	0.0003
	(0.93)	(1.30)	(0.48)	(1.49)
Density	0.0004	0.0001	0.0005	0.0001
	(2.46)	(0.64)	(2.63)	(0.37)
Diversity	-0.0403	0.0784	0.0719	0.1386
	(-0.66)	(1.39)	(1.02)	(2.03)
Specialization	-0.1530	-0.0874	-0.1016	-0.0333
	(-10.14)	(-5.37)	(-5.79)	(-1.69)
Human Capital	-0.0597	-0.0177	-0.0491	-0.0055
	(-10.76)	(-2.89)	(-7.60)	(-0.74)
Constant	0.7486	0.1247	0.5179	-0.0215
	(9.54)	(1.34)	(5.67)	(-0.20)
R-squared	0.1670	0.3446	0.0713	0.2120
Moran's I	15.7132	7.1867	10.3245	3.5252

Note: t-values within parentheses.

Table 6. Above Average Foreclosure risk – logistic regression.

	All variables		Option based		Labor Market	
	Odds ratio	t-value	Odds ratio.	t-value	Odds ratio.	t-value
Sigma	3.637	0.69	5.166	0.87	-	
Delta-price	0.209	-1.93	0.175	-2.34	-	
Interest rate	0.911	-1.25	0.863	-2.02	-	
Income per capita	0.995	-0.57	0.988	-1.43	-	
Employment rate	2.484	0.35	0.032	-1.44	-	
Exp. Rent-to-price	0.294	-1.44	-0.109	-2.76		
Rent	1.003	1.02	1.003	0.86		
Price	0.997	-3.36	0.997	-4.36	-	
Size	1.011	0.61	-		1.006	0.52
Density	0.992	-0.68	-		0.976	-2.04
Diversity	10772.71	1.78	-		265.232	1.10
Specialization	0.004	-4.96	-		0.001	-6.47
Human Capital	0.421	-2.11	-		0.313	-3.52
Sensitivity	0.3686		0.2696		0.4061	
Specificity	0.8462		0.8679		0.8442	
False	0.6314		0.7304		0.5939	
Correctly classified	0.6713		0.6488		0.6838	
R-squared	0.1095		0.0759		0.0876	

Note: The logit regression model is comparable to a linear regression model but is appropriate to models where the dependent variable is dichotomous. The restriction of 1/0 boundaries can be solved by the logistic transformation:

$$P_i^* = \left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_n X_{n.i} + \varepsilon$$

where the dependent variable P is the probability of above average risk ranging from 0 to 1. The relative weight of the factors in predicting the outcome is given by the coefficients. The interpretation of the coefficient is as a multiplicative effect on the odds ratio. The method of maximum likelihood is utilized to determine estimates of the parameters $\beta_1, \ldots, \beta_k$ and associated t-values.