Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Banking & Finance 28 (2004) 753–771
Journal of BANKING & FINANCE
www.elsevier.com/locate/econbase

Default correlation: An empirical investigation of a subprime lender

Adrian M. Cowan [a,*], Charles D. Cowan [b,c]

[a] *Department of Finance, Economics and Quantitative Methods, School of Business, University of Alabama at Birmingham, 1150 10th Ave. South, Birmingham, AL 35294, USA*
[b] *Analytic Focus LLC, Birmingham, AL 35203, USA*
[c] *University of Alabama at Birmingham, Birmingham, AL 35294, USA*

Abstract

In recent years, subprime lending has grown substantially as an important sector of the credit markets. This paper is concerned with the risk management of subprime loan portfolios and the importance of default correlation in measuring that risk. Using a large portfolio of residential subprime loans from an anonymous subprime lender, we show that default correlation is substantial for this lender. In particular, the significance of default correlation increases as the internal credit rating declines. Our results suggest that lenders and regulators would be well served investing in the understanding of default correlation in subprime portfolios.

JEL classification: G21; C10
Keywords: Credit risk; Default correlation; Mortgage; Subprime lending

1. Introduction

Subprime lenders include both large financial institutions that offer subprime loans as a subset of their portfolios and institutions that restrict their lending activities to such loans. Subprime loan portfolios generally have greater risks of default with higher credit spreads to compensate lenders for the higher risk. The subprime lending market has grown substantially in recent years. As noted by Scheessele (2002),

* Corresponding author. Tel.: +1-205-975-5490; fax: +1-205-672-7644.
E-mail address: acowan@business.uab.edu (A.M. Cowan).

0378-4266/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2003.10.005

total subprime lending grew from $90 billion in 1996 to $173 billion in 2001. In 2001, subprime loans represented a considerable 8.3% of the overall mortgage market. Given the growth of this market sector combined with the higher risks relative to other mortgage portfolios, understanding the credit risk of these portfolios is of critical importance both to the lenders themselves and to the regulators of these lenders. This paper focuses on one aspect of credit risk analysis; i.e., the importance of default correlation in measuring credit risk in subprime portfolios.

Significant advances are being made in the measurement and modeling of credit risk in lending portfolios. However, given inherent difficulties in necessary data acquisition, fewer advances have been made with respect to retail credit portfolios in general. This paper represents the first empirical investigation of default correlation within a totally subprime loan portfolio. Whereas Calem and LaCour-Little (2001) study mortgage portfolios of which subprime loans represent a subset, we are concerned with portfolios that consist exclusively of subprime loans. This is an important distinction as there are many lenders throughout the US who serve the subprime market exclusively. Using a proprietary data set of residential, subprime loans as a single case study, we are able to provide additional insight into the importance of default correlation in such portfolios.

Subprime lenders have historically used specified limits to manage credit risk exposure, such as a dollar limit established by borrower or a dollar limit established by geographic region. In fact, Carey (2000) observes that the monitoring of established credit limits has long been a part of examinations in the United States. Our results suggest that subprime lenders would be well served to develop more sophisticated credit measurement techniques. Despite the smaller exposures from subprime loans as compared with commercial portfolios, we find substantially larger default correlations than reported for commercial bonds and loans. This also has implications for capital requirements, as Gordy (2000) finds that capital requirements based on industry credit risk models vary considerably based on average default correlations in the portfolio. Despite the fact that we analyze the portfolio of a single lender, our results strongly support the investment in further understanding of default correlation both by regulators that evaluate internal risk models and by financial institutions developing models to manage risks.

The remainder of the paper is organized as follows. Section 2 discusses the previous literature associated with default correlation. Section 3 describes the data and methodology. We present our empirical results in Section 4. Section 5 concludes the paper.

2. Default correlation

Default correlation is a measure of the dependence among risks. Along with default rates and recovery rates, it is a necessary input in the estimation of the value of the portfolio at risk due to credit. In general, the concept of default correlation incorporates the fact that systemic events cause the default event to cluster. Coincident movements in default among borrowers may be triggered by common underlying factors.

Within the context of retail portfolios, systemic events might include the macroeconomic events of changes in the rate of unemployment or geographically specific events such as those modeled by Calem and LaCour-Little (2001). Default correlation is defined by Nagpal and Bahar (2001) as the relationship between default probabilities and joint default probabilities. They note that historical rates of default support the idea that credit events are correlated. This correlation is a critical factor in the estimation of the tails of the overall credit loss distributions. Failure to recognize the impact of shocks to the portfolio through default correlation will ultimately underestimate the measures of risk and economic capital required to manage that risk.

In contrast with other residential loan portfolios in which one would anticipate that default correlation can be very low, understanding default correlation is critical for lower credit quality subprime portfolios. Several authors have documented the relationship between the initial credit quality of the portfolio and default correlation in commercial portfolios. Generally, as credit quality declines, the importance of default correlation increases. For example, Zhou (1997) shows implied default correlations based on Z-values that are almost zero for highly rated firms but substantial for lowly rated firms even over short time horizons. Using corporate bond and loan portfolios, Lucas et al. (2001) provide numerical results showing that for a given correlation, a higher portfolio quality lowers extreme credit loss quantiles. Similarly, Loffler (2003) finds that correlation uncertainty is a more significant factor for portfolios rated B as compared with portfolios rated BBB for uncertainty in the 1% value-at-risk (VaR). Although these studies deal with commercial loan and bond portfolios, the management of a subprime loan portfolio is analogous to the management of a non-investment-grade bond portfolio. Our results are consistent with these previous studies. The significance of default correlation increases as the internal ratings of the lender decline. Thus, it is likely that ignoring default correlation in the development of credit risk models for subprime portfolios would lead to considerable model risk.

There are several methodologies currently employed in the development of default correlations within portfolios as discussed in Zhou (1997). For example, Loffler (2003) estimates default correlations based on the joint distribution of asset values. As discussed in Crouhy et al. (2000), equity prices are often used as a proxy to estimate asset correlations, given that asset values are not directly observable. One commonly employed method is the identification of a benchmark for the purpose of developing asset return correlations and then mapping these into default correlations. The approach requires making assumptions about the relationship between asset prices and default. However, this approach is not applicable within a retail context as there is no asset price for the individual borrower. Alternatively, default correlations can be inferred from historical default volatilities as described generally within Appendix F of J. P. Morgan's CreditMetrics Technical Document (1997). We adopt a similar approach in this paper. A default correlation coefficient is estimated based on the assumption that all loans within the risk class have identical default rates. The application of such an approach to subprime portfolios assumes that internal credit rating assignments are consistent. Specifically, we follow the exchangeable models framework of Frey and McNeil (2002).

3. Data and methodology

The data for this paper are the detailed loan history file of a large subprime lender. This lender identifies itself as a subprime lender and is identified as such on the most recent annual list of subprime lenders by the Department of Housing and Urban Development. [1] The protection of the anonymity of the lender precludes the provision of detailed descriptive statistics of the data.

Our sample consists of monthly data on residential loans from July 1995 through December 2001. We restrict our analysis to the 30-year fixed-rate loan portfolio, and the analyzed portfolio consists exclusively of loans on one- to four-family residential properties. [2] Although multiple factors contribute to the classification of a loan as subprime, the subprime nature of the analyzed portfolio is clearly reflected in the low Fair Isaac Corporation (FICO) credit scores at the time of origination. The lower the FICO score, the lower the creditworthiness of the borrower. A median FICO score of 587 in our sample is significantly below the national median score of 725, as reported by Calem and LaCour-Little (2001).

As noted by Jarrow and Turnbull (2000), credit risk time horizons are commonly one year. Initially, we considered examining all of these loans on an annual basis, but this gave us only six full years worth of data. Because credit events are typically less rare events for subprime lenders as compared with prime lenders, we take advantage of this characteristic to reduce the credit time horizon. Monthly rates are volatile and not generally reflective of overall economic activity. Thus, we calculate rates on a semiannual basis, breaking each year into January through June and July through December. Using a 6-month event horizon allows us to benefit from a more reasonable number of time periods while at the same time guarding against spurious volatility that may disguise any real correlation with external economic factors.

We use two separate measures of the default event for comparative purposes. First, we define default to occur at foreclosure when the lender takes the collateral as real estate owned (REO). Second, we employ a less stringent measure of 90 days or more delinquent to define default. This serves to expand our event sample, since approximately twice as many loans are 90 days or more delinquent as compared to foreclosed. This difference reflects the fact that not all delinquencies ultimately lead to foreclosure. Many delinquent loans eventually "cure" and return to current status or prepay because of sales.

Using the loan histories from this lender, we summarize the loans by characteristics of the borrower and within 6-month time periods. We then compute, for each time period, the rate of default for the overall pool of loans and also for subgroups of loans defined by the borrower characteristics. A number of discrete variables are

[1] The Department of Housing and Urban Development (HUD) annually identifies a list of lenders who specialize in either subprime or manufactured home lending. The listings for 1993–2001 are available from the HUD web site at www.huduser.org/datasets/manu.html.

[2] We retain the two- to four-unit property loans in the analysis as this is consistent with the permitted Basel II definition of residential mortgage loans and consistent with the current categorizations of residential loans on bank and thrift regulatory reports.

used to define subgroups, including payment to borrower income ratio, occupancy, and internal risk group. [3]

We use deciles to identify subgroups when dealing with sequential variables, such as credit score. For sequential variables, we place loans in order from smallest to highest value and create deciles. As a result, each category in the tabulation contains 10% of the borrowers. For the low end (10th decile) and the high end (90th decile) of the deciles, we further subdivide by 5% cutoffs; i.e., 0–5%, $5 \pm 10\%$, 90–95%, $95 \pm 100\%$. This yields 12 categories in all, as seen in the tables in Section 4. This segregation of the portfolio is conducted for all numerically sequential variables relevant to the borrower, loan, and property.

A general model of default correlation can be summarized by thinking about the correlation between events on a loan by loan AND half-year by half-year basis. If we look at each loan-time combination, we can consider loans held by the subprime lender in the same time period and also loan performance across time periods. To examine default, we assume that default events are independent across time periods. In other words, not defaulting in any one time period is not predictive of whether the loan defaults in the next time period. This is necessary because loans that default are immediately removed from the loan set. This allows us to consider loan-time blocks. The correlation matrix is block diagonal and has only three possible entries: 1, 0, and, ρ the correlation coefficient. The block diagonal matrix takes the form

$$\Sigma = \begin{bmatrix} B_1 & 0 & 0 & \cdots & 0 \\ 0 & B_2 & 0 & \cdots & 0 \\ 0 & 0 & B_3 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & B_k \end{bmatrix}_{N \times N} \quad \text{where } B_j = \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \rho & \rho & \rho & \ddots & \rho \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix}_{n_j \times n_j}. \tag{1}$$

A block is the set of all loans that are held by the lender in any one time period (in this case, a 6-month period). The correlation rho (ρ) is assumed to be the same in all blocks. The number of blocks, k, is the number of time periods during which a loan could exist. Naturally, in any time period some new loans enter the pool as new originations are made and some loans leave the pool as a result of either being paid off or being foreclosed. This means that the blocks are different sizes, depending on the number of loans that exist in any one time period.

Across time periods the default event is considered to be independent so that off-diagonal blocks in the matrix Σ are filled with zeros. The default correlation is measured as the variability of the default events (indicated as zero/one dummy variables) from time period to time period relative to the variability of the default event if there were no variability from time block to time block.

[3] We also looked at the source of the loan (wholesale or retail) and whether the documentation was full or limited. These factors added no significant discrimination in terms of default correlation and therefore we do not present the results.

A richer but much more difficult model to examine would have off-diagonal blocks in the matrix Σ be measures of autocorrelation for lags of 1, 2, 3, ... time periods. It is difficult to know the full impact of consideration of potential autocorrelation. On the one hand, the autocorrelation may be positive, indicating that if there are a large number of default events in time period j, it is likely there will be a large number of default events in time period $j+1$, and so the overall variability of default events may be even larger. This would indicate that the effects of changes in economic conditions both predominate and linger. If the autocorrelation were negative, this would suggest that the borrowers respond to and recover from changes in economic conditions so that a large number of defaults in one time period would likely be followed by a smaller number the next period. This may also reflect activities on the part of lenders to respond to increases in the number of defaults by redoubling efforts at workouts and other such remedies. A subsequent paper considers both contemporaneous and lagged correlations in default and delinquency.

If we assume no correlation within a time block, we would expect the likelihood of default to be simply the average rate of default. However, under the default correlation model, the rate will appear to vary from time period to time period, since it does, conditional on changing economic conditions. We cannot observe the changing conditions – we may not even know which variables are important to consider. But we can model the situation by allowing the default events to be correlated within time periods while independent between time periods.

One especially compelling result of this formulation is to consider the variance of the total number of defaults under this model. If $\underline{x}$ is a vector whose entries are zero or one indicating default, with an x value for each loan-time period, then the summary statistics for the total number of defaults, D, and its variance, observed over a k period term is

$$D = \underline{1}^T \underline{x} \tag{2}$$

and

$$\mathrm{Var}(D) = (\underline{x} - \bar{x})^{\mathrm{T}} \Sigma (\underline{x} - \bar{x}), \tag{3}$$

where $\underline{1}$ is a unit vector.

If we rewrite the matrix B_j as

$$B_j = (1-\rho)I + \rho \underline{1}\underline{1}^{\mathrm{T}}, \tag{4}$$

then the variance of the total number of defaults, $\mathrm{Var}(D)$, reduces to approximately

$$\mathrm{Var}(D) \cong (1-\rho)\frac{1}{N}\sum_{i}^{k}\sum_{j}^{n_i}(x_{ij}-\bar{x})^2 + \rho\frac{1}{N}\sum_{i}^{k} n_i^2(\bar{x}_i - \bar{x})^2. \tag{5}$$

When ρ equals zero (no correlation between default events), then the variance is simply the usual variance measured from each observation to the overall mean across all time periods. When ρ is one, in any block if any loan defaults, they all default. Thus when ρ is one, the contribution to the variance of the total number of defaults is only the variability between (but not within) the blocks. This representa-

tion makes it easy to see that the default correlation is simply the relation of the variability of the default rate over time periods relative to the variability one would see if there were no default correlation.

4. Empirical results

The magnitude of default correlation for the total pool of loans is low, but an examination of subgroups shows that default correlation can be a significant problem for a subprime lender. This is true even for a stable portfolio of loans like fixed-rate 30-year loans. We present our results and show how they vary by subgroup and also present some of the default and delinquency rates discovered.

As can be seen in Table 1 and the accompanying Fig. 1, there is some variation for the entire sample in the default rate as measured by foreclosure but much more variation over time in the delinquency rate. However, it is relatively large compared with the incidence of foreclosures reflected in the Calem and LaCour-Little (2001) sample. We find 3.5% of the loans foreclosed over a six-and-a-half-year observation period compared with only 0.8% over a 5-year period exhibited in the sample of Calem and LaCour-Little (2001). This difference primarily reflects the mix of prime and subprime loans in their portfolio. In addition, it may reflect differences in sample periods, although the Calem and LaCour-Little (2001) period overlaps our sample

Table 1
Overall default and delinquency rates for all 30-year fixed-rate subprime loans regardless of internal risk classification, July 1995 through December 2001

	Foreclosed REO	90 or more days delinquent
1995:2	0.28%	4.61%
1996:1	1.15%	4.84%
1996:2	1.83%	6.66%
1997:1	2.57%	6.07%
1997:2	3.56%	6.89%
1998:1	1.25%	2.17%
1998:2	0.65%	1.43%
1999:1	0.69%	2.34%
1999:2	0.58%	2.21%
2000:1	0.86%	2.60%
2000:2	0.85%	3.45%
2001:1	1.23%	2.91%
2001:2	0.98%	3.12%
Total	0.98%	2.95%
Default correlation	0.0079	0.0109

The second column in this table presents the foreclosure rates for 6-month periods between the latter half of 1995 through the latter half of 2001, while the third column presents the rates at which loans became 90 or more days delinquent at least once during the life of the loan. The variation of the rates is an indication of increase or decrease in the frequency of the events caused by events' being correlated with one another within a time period. The rates displayed in both columns are for all loans in the loan portfolio.

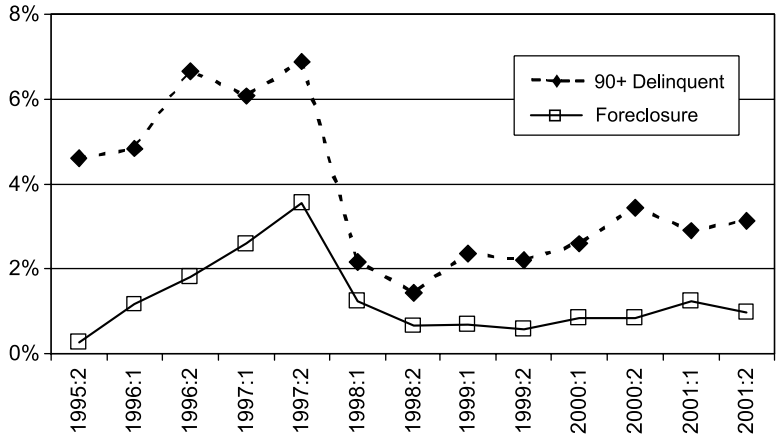

Fig. 1. Default and delinquency based on total 30-year fixed-rate subprime loan portfolio. This graph shows the steady increase in the delinquency and foreclosure rates from 1995 through 1997, followed by a decline in both rates to lower levels until the end of 2001. REO is Real Estate Owned, the foreclosure on the loan, while 90+ Del. is 90 or more days Delinquent.

period during the incidence of highest foreclosures and delinquencies. However, regardless of the higher incidence of default in our overall portfolio, default correlation is a minimal 0.0079 for foreclosures and a slightly higher 0.011 for delinquent loans. Neither outcome really demonstrates a very large default correlation.

More insight is available regarding default correlation once we investigate subsets of borrowers. When we examine the borrowers by subgroup as defined in the previous section, we find that some subgroups are subject to significant default correlation.

The first and most obvious variable to consider is the internal risk grade assigned to the borrower by the subprime lender. This rating would be comparable to the bond rating discussed above for commercial bonds. However, in the case of bonds, the bond ratings are independently assigned by rating agencies, they change over time, and they are tracked. A great deal of literature is devoted to the analysis and interpretation of these rating changes, and transition matrices can be used to better estimate the likelihood of default in the bond market. For mortgage loans, we observe only the risk grade at the time of origination, and we must assume that this is a sufficient indicator for the borrower. While we cannot measure how the risk grade might change over time for the borrower, the fact that the borrower started in a particular risk grade is still very indicative of the likelihood of default or delinquency. Default and delinquency by internal risk classification is presented in Table 2. There are two interesting results in this table. First, the credit rating categories appear to be relatively consistent as a risk measure. Second, default correlation reveals borrower differences when measured by foreclosure as compared with delinquency.

Internal risk classifications appear to accurately capture credit risk. The likelihood of delinquency and foreclosure monotonically increases as the risk category

Table 2
Default and delinquency rates of 30-year fixed-rate loans by internal risk classification

Risk grade	Foreclosed REO					90 or more days delinquent				
	AA	A	B	C	CC	AA	A	B	C	CC
1995:2	0.00%	0.00%	0.27%	0.99%	0.00%	0.00%	1.69%	3.23%	9.85%	11.70%
1996:1	0.00%	0.87%	0.76%	2.33%	2.42%	2.17%	1.92%	4.39%	9.73%	12.10%
1996:2	0.00%	0.34%	2.50%	3.54%	5.00%	0.65%	4.24%	6.54%	14.57%	10.00%
1997:1	0.00%	2.10%	3.06%	4.78%	7.29%	0.28%	3.30%	6.97%	15.65%	18.75%
1997:2	0.22%	2.07%	4.35%	11.19%	21.43%	1.09%	4.82%	8.70%	23.13%	21.43%
1998:1	0.00%	0.69%	2.06%	9.32%	7.27%	0.44%	2.33%	2.52%	9.32%	14.55%
1998:2	0.04%	0.66%	0.93%	5.15%	4.55%	0.42%	1.27%	2.05%	7.73%	13.64%
1999:1	0.30%	0.78%	0.70%	3.54%	2.63%	0.89%	2.15%	3.26%	8.86%	18.42%
1999:2	0.28%	0.50%	0.81%	2.92%	0.00%	1.14%	1.79%	2.96%	8.10%	11.11%
2000:1	0.32%	1.01%	0.99%	3.15%	4.93%	1.25%	2.10%	3.79%	9.73%	9.85%
2000:2	0.41%	0.74%	1.38%	2.61%	3.06%	1.87%	2.40%	5.39%	12.48%	15.31%
2001:1	0.75%	0.92%	2.07%	4.06%	4.97%	1.41%	3.08%	4.71%	10.39%	9.39%
2001:2	0.50%	0.83%	1.67%	5.27%	6.98%	1.77%	3.37%	5.48%	10.19%	11.05%
Total	0.41%	0.83%	1.42%	3.90%	4.53%	1.36%	2.51%	4.33%	10.74%	12.79%
Default correlation	0.0013	0.0040	0.0086	0.0197	0.0620	0.0030	0.0042	0.0084	0.0170	0.0123

Columns 2–6 (panel 1) in this table present the foreclosure rates for 6-month periods between the latter half of 1995 through the latter half of 2001, while the columns 7–12 (panel 2) present the rates at which loans became 90 or more days delinquent at least once during the life of the loan. The variation of the rates is an indication of increase or decrease in the frequency of the events caused by events' being correlated with one another within a time period. The rates displayed in a panel array the loans by the internal risk rating assigned by the lender (AA through CC).

decreases. In addition, the default correlation monotonically increases as the internal risk classification decreases. This suggests, as would be expected, that lower risk rated borrowers are more vulnerable to systemic events; i.e., these borrowers have more difficulty recovering from severe incidents. Thus, both the likelihood measures and the default correlation indicate that the internal risk classifications designed by the lender appear to capture the inherent risk characteristics of the loans.

The default correlation reveals an unexpected pattern by internal risk classification. The default correlation climbs from near zero to $\rho = 0.06$ as risk grade declines. The same measure applied to delinquencies shows the same trend, from very small to measurable, but for delinquencies the correlation is not as pronounced. We believe that this suggests that the less creditworthy borrowers reflect a pattern of habitual delinquency. For some borrowers, delinquency is a recurring state, and payment histories reveal a cycle of delinquency and recovery. Although this group of borrowers is more volatile relative to overall delinquency, the delinquency may not necessarily reflect some underlying systemic event for these individuals. Such a consistent pattern would reduce the correlation of default for delinquencies. At the same time, extraordinary events will tend to push lower grade borrowers into foreclosure in clusters as revealed in the higher measure of default correlation based on foreclosures. Thus, as would be expected, if the external event that occurs is sufficiently severe, it will lead to delinquency followed by default.

The likelihood of default, whether measured by foreclosure or delinquency, is much greater for properties where the owner is not the occupant, as shown in Table 3. The likelihood of default based on foreclosure is nearly three times larger for loans secured by non-owner-occupied properties as compared with owner-occupied. This is consistent with a great deal of the mortgage literature, which suggests that borrowers will not ruthlessly exercise the default option (e.g., Vandell, 1995). Borrowers will continue to remain current on their mortgages even after the value of the home falls below the outstanding mortgage balance. Although this failure to exercise the default option is attributed to many sources in the literature, our results are consistent with a utility explanation of behavior. This suggests that borrowers derive utility from remaining in their homes even after declines in market value. In contrast, no such utility would be found in non-owner-occupied dwellings.

Default correlation is highest for mortgages collateralized with second homes, which, in general, represent higher creditworthy borrowers as measured by initial FICO score or internal risk weighting. The default correlation for foreclosed second homes is 0.087 compared with only 0.026 for non-owner-occupied and a negligible 0.005 for loans on owner-occupied properties. This is consistent with the findings of Calem and LaCour-Little (2001) that the higher creditworthy borrowers *ex-ante* default primarily based on unusual circumstances. However, our results also suggest that it is not necessarily an event unique to the borrower.

Default rates by property type are depicted in Table 4. The results are consistent with those in Table 3, and the similar results are most likely due to the substantial overlap in the non-owner-occupied and 2–4 unit categories. The overall default and the default correlation are lowest for single-family detached residences and significantly higher for other property types. For properties with 2–4 units, the default

Table 3
Default and delinquency rates by type of occupancy

Occupancy	Foreclosed REO			90 or more days delinquent		
	Non-owner	Owner	Second home	Non-owner	Owner	Second home
1995:2	0.00%	0.33%	0.00%	5.52%	4.43%	7.14%
1996:1	1.53%	1.03%	5.88%	5.61%	4.79%	0.00%
1996:2	2.94%	1.64%	3.85%	8.82%	6.33%	7.69%
1997:1	5.05%	2.30%	0.00%	7.80%	5.77%	11.11%
1997:2	10.23%	2.72%	6.90%	10.23%	6.57%	3.45%
1998:1	1.81%	1.21%	0.00%	3.63%	2.05%	0.00%
1998:2	0.95%	0.63%	0.00%	4.95%	1.15%	0.00%
1999:1	2.57%	0.55%	0.00%	4.00%	2.21%	2.34%
1999:2	1.70%	0.50%	0.61%	4.36%	2.03%	3.07%
2000:1	2.70%	0.73%	1.09%	4.50%	2.49%	0.55%
2000:2	1.40%	0.82%	0.51%	6.70%	3.23%	2.53%
2001:1	2.97%	1.12%	0.00%	5.20%	2.77%	1.73%
2001:2	2.23%	0.91%	0.35%	3.80%	3.10%	1.40%
Total	2.37%	0.87%	0.63%	5.09%	2.80%	2.02%
Default correlation	0.0258	0.0054	0.0874	0.0080	0.0108	0.0556

Columns 2–4 (panel 1) in this table present the foreclosure rates for 6-month periods between the latter half of 1995 through the latter half of 2001, while the columns 5–7 (panel 2) present the rates at which loans became 90 or more days delinquent at least once during the life of the loan. The variation of the rates is an indication of increase or decrease in the frequency of the events caused by events' being correlated with one another within a time period. The rates displayed in a panel array the loans by the use of the property, as defined by whether the property is owner occupied or not, or if the property is a second home.

correlation is quite substantial, reflecting a very large increase in the default rate in both halves of 1997.

We next compute the default correlations based on groupings by credit score and present the sensitivity of default correlation to initial borrower FICO scores in Table 5. As would be expected, our results are similar to those reported by internal risk category, given that the credit score is an input to the final internal classification. The tendency for default correlation to be higher for foreclosures than delinquencies as exhibited for internal risk classifications in Table 2 is eliminated with the finer classifications provided by FICO score deciles. Whereas the use of deciles maintains a consistent number of loans across deciles, this is not the case for internal risk classifications. While default correlation, whether measured by foreclosure or delinquency, increases steadily as the internal risk classification declines, the same result does not hold when the loans are segregated by FICO score deciles. It is therefore interesting to note that this latter classification provides the greatest differences in default correlation. This finding offers a promising avenue for future research in evaluating the efficiency of internal versus external ratings.

We evaluate default correlation by mortgage payment to borrower income ratio. Similar to previous categories, this ratio is measured at time of origination. Deciles are created from the payment to income ratio, and borrowers are grouped from

Table 4
Default and delinquency rates by property type

Property type	SFR	Manufactured housing	2–4 Units	Condo	PUD
Foreclosed REO					
1995:2	0.34%	0.00%	0.00%	0.00%	0.00%
1996:1	1.26%	0.00%	0.93%	0.00%	0.00%
1996:2	1.77%	0.00%	3.30%	3.28%	0.00%
1997:1	2.39%	3.23%	6.48%	0.00%	3.77%
1997:2	3.39%	3.70%	10.71%	0.00%	1.79%
1998:1	1.38%	0.00%	1.49%	0.00%	0.56%
1998:2	0.72%	0.45%	0.89%	0.00%	0.00%
1999:1	0.68%	0.56%	1.04%	1.18%	0.22%
1999:2	0.56%	0.18%	1.52%	0.51%	0.33%
2000:1	0.91%	0.45%	1.23%	0.61%	0.30%
2000:2	0.84%	1.63%	1.04%	0.29%	0.55%
2001:1	1.21%	1.45%	2.52%	0.52%	0.49%
2001:2	0.98%	1.48%	1.80%	0.24%	0.31%
Total	0.98%	1.04%	1.76%	0.46%	0.40%
Default correlation	0.0067	0.0138	0.0461	0.0164	0.0258
90 or more days delinquent					
1995:2	4.93%	0.00%	4.60%	0.00%	3.45%
1996:1	4.68%	5.26%	8.41%	3.85%	2.22%
1996:2	6.37%	9.52%	14.29%	1.64%	6.00%
1997:1	6.16%	0.00%	10.19%	2.53%	3.77%
1997:2	6.78%	7.41%	8.33%	7.14%	7.14%
1998:1	2.28%	0.00%	2.49%	2.37%	0.56%
1998:2	1.33%	1.35%	4.44%	0.96%	0.57%
1999:1	2.32%	2.25%	4.57%	1.18%	1.51%
1999:2	2.19%	2.87%	4.43%	1.03%	0.33%
2000:1	2.56%	2.69%	5.04%	1.37%	1.50%
2000:2	3.44%	4.08%	5.52%	1.87%	1.94%
2001:1	2.92%	3.75%	3.77%	1.17%	2.55%
2001:2	3.21%	3.80%	3.42%	1.90%	1.74%
Total	2.96%	3.26%	4.71%	1.58%	1.70%
Default correlation	0.0105	0.0240	0.0225	0.0189	0.0237

Columns 2–6 (panel 1) in this table present the foreclosure rates for 6-month periods between the latter half of 1995 through the latter half of 2001, while the columns 8–12 (panel 2) present the rates at which loans became 90 or more days delinquent at least once during the life of the loan. The variation of the rates is an indication of increase or decrease in the frequency of the events caused by events' being correlated with one another within a time period. The rates displayed in a panel array the loans by the internal risk rating assigned by the lender (AA through CC).

lowest to highest. These results are presented in Table 6. Mortgage subgroups delineated by this ratio and credit score show the greatest differences overall in default correlation. One also finds that loans with higher ratios of mortgage payment to borrower income are more likely to default and the default correlation is larger. This

suggests that any disruption in household income flow will more likely lead to a default or delinquency as a larger portion of the borrower's income is going to the mortgage payment (and presumably less into savings). Thus, as expected, economic distress will lead to greater default correlation for these groups.

Finally, it is important to recognize that our limited sample period from July 1995 through December 2001 introduces potential selection bias. Default correlation captures systemic events, and thus, the time period over which it is measured naturally impacts the estimates. As noted by Carey (1998), data covering a very long period mitigate but do not eliminate this problem. There is an issue of how representative this time period is of general activity found in the subprime market. In general, the economy was enjoying expansion during much of our sample period, but there were underlying economic weaknesses as well. Similar to the Calem and LaCour-Little (2001) time frame, our sample period includes a period of primarily rising or stable house prices in most geographic markets, even subsequent to the beginning of the 2001 recession in March [4] of that year. According to the Federal Home Loan Mortgage Corporation's (Freddie Mac's) monthly mortgage rate survey, the sample period includes both periods of declining and rising 30-year mortgage rates, with a peak of 8.5% during the period in May of 2000 and a low of 6.6% in October of 2001. Despite the declining interest rate environment for much of 1997 and 1998, subprime borrowers are largely precluded from prepayment because of lockout periods and significant prepayment penalties. It is clear from both our overall portfolio and the categorized subsamples that surges occurred in defaults and delinquencies in the 1996–97 period. These surges coincide with the tremendous increase in personal bankruptcies over this same period. Personal bankruptcies increased 29% and 20% in 1996 and 1997, respectively, from prior years. Overall, our sample period represents a very mixed economic picture. Nevertheless, for any limited time period examined, there are always issues of selection bias to be considered.

5. Conclusion

This paper presents the first formal study of default correlation within an exclusively subprime mortgage loan portfolio. We find that default correlations for the specific portfolio studied are insignificant until the portfolio is segregated into appropriate risk groups. We analyze 6-month default correlation using both actual default (foreclosure) and a broader definition of delinquency that is consistent with previous literature (e.g., Calem and LaCour-Little, 2001). Contrary to our expectations, actual defaults generally result in higher default correlations than delinquencies. As anticipated, the magnitude of default correlation increases as the internally assigned risk grade declines.

[4] The National Bureau of Economic Research (NBER) identifies March 2001 as the beginning of the 2001 recession.

Table 5
Default and delinquency rates by credit score: Foreclosed REO (panel A) and 90 or more days delinquent (panel B)

Credit	0–5%	6–10%	11–20%	21–30%	31–40%	41–50%	51–60%	61–70%	71–80%	81–90%	91–95%	96–100%
Panel A												
1995:2	0.00%	7.69%	2.04%	0.00%	0.40%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
1996:1	0.00%	21.43%	3.33%	1.80%	1.42%	0.51%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
1996:2	0.00%	26.09%	3.08%	2.14%	2.45%	1.75%	0.38%	0.00%	1.59%	0.00%	0.00%	0.00%
1997:1	0.00%	16.67%	18.06%	3.59%	3.51%	0.74%	0.91%	0.32%	0.00%	0.00%	0.00%	0.00%
1997:2	50.00%	38.46%	14.52%	11.61%	4.40%	2.43%	1.72%	0.30%	0.48%	0.00%	0.00%	0.00%
1998:1	0.00%	27.27%	26.42%	7.59%	2.94%	0.37%	0.50%	0.00%	0.14%	0.00%	0.00%	0.00%
1998:2	0.56%	2.94%	3.54%	4.64%	1.51%	0.27%	0.11%	0.09%	0.09%	0.00%	0.00%	0.00%
1999:1	0.54%	2.11%	1.36%	0.84%	1.73%	1.58%	0.27%	0.44%	0.15%	0.17%	0.00%	0.00%
1999:2	0.17%	0.99%	1.07%	1.47%	1.22%	0.58%	0.73%	0.50%	0.23%	0.06%	0.00%	0.15%
2000:1	0.93%	0.79%	1.41%	1.72%	0.72%	1.27%	1.14%	0.31%	0.75%	0.55%	0.92%	0.13%
2000:2	1.65%	1.43%	1.72%	1.56%	1.09%	0.90%	0.79%	0.65%	0.41%	0.28%	0.68%	0.00%
2001:1	2.62%	2.18%	1.79%	1.94%	2.18%	1.37%	1.12%	0.78%	0.64%	0.71%	0.66%	0.37%
2001:2	1.66%	1.93%	1.97%	1.81%	1.89%	1.34%	0.68%	0.64%	0.69%	0.35%	0.32%	0.14%
Total	1.32%	2.14%	2.27%	2.11%	1.66%	1.06%	0.74%	0.47%	0.46%	0.34%	0.44%	0.14%
Default correlation	0.0110	0.0567	0.0435	0.0128	0.0027	0.0013	0.0009	0.0003	0.0006	0.0006	0.0027	0.0010

Panel B												
1995:2	0.00%	46.15%	18.37%	4.79%	4.78%	1.55%	2.94%	0.89%	0.00%	0.00%	0.00%	0.00%
1996:1	0.00%	57.14%	20.00%	7.66%	3.99%	3.06%	1.53%	0.53%	2.27%	0.00%	0.00%	0.00%
1996:2	100.00%	43.48%	36.92%	13.68%	5.52%	3.26%	1.88%	1.40%	0.00%	0.00%	0.00%	0.00%
1997:1	100.00%	54.17%	26.39%	10.31%	5.56%	5.17%	3.93%	0.96%	0.00%	0.00%	0.00%	0.00%
1997:2	100.00%	30.77%	37.10%	20.65%	7.60%	5.47%	2.41%	2.09%	0.95%	0.00%	0.00%	0.00%
1998:1	0.00%	36.36%	20.75%	13.10%	3.53%	2.43%	1.16%	0.68%	0.42%	0.00%	0.00%	0.00%
1998:2	0.00%	2.94%	1.77%	4.35%	4.87%	2.26%	1.46%	0.46%	0.47%	0.12%	0.00%	0.00%
1999:1	3.76%	3.51%	5.23%	5.39%	4.50%	2.96%	1.87%	1.45%	1.35%	0.69%	0.77%	0.00%
1999:2	3.28%	3.95%	4.14%	4.21%	2.76%	2.96%	2.20%	1.88%	1.15%	0.77%	0.90%	0.30%
2000:1	4.41%	4.74%	5.20%	3.99%	3.95%	2.81%	2.92%	1.63%	1.35%	0.73%	1.53%	0.64%
2000:2	8.38%	6.99%	5.53%	6.22%	5.10%	4.72%	2.20%	1.70%	2.10%	1.17%	1.22%	0.81%
2001:1	4.78%	5.39%	5.37%	5.05%	4.83%	3.94%	2.30%	1.84%	1.64%	1.53%	0.66%	0.65%
2001:2	5.97%	5.55%	4.77%	5.58%	5.79%	3.85%	3.51%	2.35%	1.59%	1.28%	1.37%	1.02%
Total	5.12%	6.19%	6.25%	5.98%	4.67%	3.49%	2.40%	1.64%	1.37%	0.98%	1.05%	0.65%
Delinquent correlation	0.0244	0.0701	0.0520	0.0138	0.0011	0.0013	0.0010	0.0009	0.0009	0.0009	0.0033	0.0023

Columns 2–13 (panel A) in this table present the foreclosure rates for 6-month periods between the latter half of 1995 through the latter half of 2001. The variation of the rates is an indication of increase or decrease in the frequency of the events caused by events' being correlated with one another within a time period. The rates displayed in a panel array the loans by the external risk rating determined by use of FICO scores. The credit score is the Fair Isaac FICO score obtained by the lender prior to loan origination. Each credit percentage indicates the decile or half-decile category created using credit score. The leftmost group (0–5%) has the lowest credit scores at the time the loan is made.

Columns 2–13 (panel B) in this table present the delinquency rates, 90 or more days at least once during the life of the loan, for 6-month periods between the latter half of 1995 through the latter half of 2001. The variation of the rates is an indication of increase or decrease in the frequency of the events caused by events' being correlated with one another within a time period. The rates displayed in a panel array the loans by the external risk rating determined by use of FICO scores. The credit score is the Fair Isaac FICO score obtained by the lender prior to loan origination. Each credit percentage indicates the decile or half-decile category created using credit score. The leftmost group (0–5%) has the lowest credit scores at the time the loan is made.

Table 6
Default rates by P&I to income ratio: Foreclosed REO (panel A) and 90 or more days delinquent (panel B)

P&I to income	0–5%	6–10%	11–20%	21–30%	31–40%	41–50%	51–60%	61–70%	71–80%	81–90%	91–95%	96–100%
Panel A												
1995:2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.12%	0.00%	1.00%	0.00%	1.72%	0.00%
1996:1	1.46%	1.19%	1.36%	0.60%	1.44%	0.00%	3.17%	0.69%	0.71%	1.23%	2.63%	0.00%
1996:2	1.32%	3.23%	0.00%	2.16%	0.00%	3.60%	3.31%	0.74%	1.39%	2.33%	2.20%	3.49%
1997:1	1.70%	0.00%	3.23%	1.96%	1.65%	2.99%	3.38%	2.86%	4.27%	2.31%	2.11%	5.13%
1997:2	4.85%	3.54%	2.00%	4.12%	2.94%	3.25%	3.08%	5.19%	2.13%	2.56%	5.26%	10.53%
1998:1	0.66%	1.24%	0.72%	1.48%	1.42%	0.30%	1.55%	1.17%	1.67%	1.66%	1.68%	3.06%
1998:2	0.00%	1.21%	0.26%	0.82%	0.89%	0.77%	0.16%	0.33%	0.51%	0.77%	1.65%	1.75%
1999:1	1.94%	0.71%	0.36%	0.39%	1.14%	0.32%	0.69%	0.58%	0.72%	0.55%	0.26%	0.85%
1999:2	1.06%	0.26%	0.81%	0.71%	0.15%	0.31%	0.65%	0.65%	0.67%	0.48%	0.70%	0.74%
2000:1	1.38%	0.95%	1.06%	1.03%	0.67%	0.49%	1.02%	0.57%	0.66%	0.94%	0.91%	0.95%
2000:2	0.84%	0.80%	0.71%	0.68%	0.50%	1.24%	0.82%	1.06%	0.95%	0.60%	0.68%	1.68%
2001:1	1.37%	1.03%	1.24%	1.08%	1.15%	1.32%	1.51%	0.61%	1.59%	1.05%	1.83%	1.24%
2001:2	1.63%	1.07%	0.77%	0.45%	0.76%	0.98%	1.31%	1.30%	1.02%	0.89%	1.04%	0.95%
Total	1.31%	0.94%	0.86%	0.86%	0.80%	0.91%	1.13%	0.90%	1.04%	0.90%	1.17%	1.31%
Default correlation	0.0103	0.0114	0.0084	0.0124	0.0075	0.0164	0.0108	0.0196	0.0090	0.0066	0.0127	0.0583

Panel B												
1995:2	3.23%	5.36%	4.21%	4.67%	4.49%	2.11%	5.62%	4.44%	3.00%	5.71%	8.62%	5.88%
1996:1	3.65%	1.19%	2.04%	3.57%	5.76%	1.40%	4.76%	7.59%	5.00%	8.02%	7.89%	7.77%
1996:2	6.62%	5.38%	4.07%	4.32%	4.83%	10.79%	6.61%	4.41%	6.94%	7.56%	8.79%	13.95%
1997:1	3.98%	5.00%	4.61%	7.84%	6.59%	3.59%	7.43%	6.86%	7.93%	4.63%	3.16%	15.38%
1997:2	4.85%	3.54%	5.00%	8.76%	5.88%	6.49%	9.23%	7.14%	4.96%	8.21%	12.28%	13.16%
1998:1	0.66%	1.24%	2.63%	0.74%	2.55%	1.18%	3.11%	2.33%	2.67%	3.05%	4.20%	4.08%
1998:2	3.13%	0.97%	1.28%	1.23%	2.08%	0.77%	1.47%	1.47%	1.02%	1.08%	1.23%	2.18%
1999:1	4.04%	3.02%	1.81%	1.96%	1.56%	2.16%	2.87%	1.97%	2.38%	2.22%	3.17%	2.56%
1999:2	2.47%	2.61%	1.61%	2.28%	1.62%	2.45%	2.03%	1.69%	2.26%	2.63%	2.64%	3.49%
2000:1	1.49%	2.38%	2.05%	1.87%	2.40%	3.14%	2.92%	2.48%	3.37%	3.18%	2.72%	3.47%
2000:2	4.10%	3.33%	2.70%	2.89%	2.69%	3.14%	4.15%	3.98%	3.27%	2.88%	5.19%	5.31%
2001:1	4.02%	2.78%	2.90%	2.22%	2.25%	1.93%	3.07%	3.52%	2.61%	2.79%	4.03%	4.70%
2001:2	3.61%	2.33%	2.76%	2.31%	2.39%	2.76%	3.11%	4.17%	2.95%	4.60%	3.03%	3.58%
Total	3.29%	2.65%	2.44%	2.44%	2.45%	2.62%	3.21%	3.24%	2.95%	3.34%	3.75%	4.47%
Delinquent correlation	0.0063	0.0081	0.0056	0.0230	0.0122	0.0261	0.0156	0.0129	0.0124	0.0159	0.0268	0.0454

Columns 2–13 (panel A) in this table present the foreclosure rates for 6-month periods between the latter half of 1995 through the latter half of 2001. The variation of the rates is an indication of increase or decrease in the frequency of the events caused by events' being correlated with one another within a time period. The rates displayed in a panel array the loans by the ordered principal and interest to income ratio. The P&I to income ratio is a measure of ability to pay. Each P&I to income ratio percentage reported as headers indicates the decile or half-decile category created using this ratio. The leftmost group (0–5%) has the highest ability to pay at the time the loan is made.

Columns 2–13 (panel B) in this table present the delinquency rates, 90 or more days at least once during the life of the loan, for 6-month periods between the latter half of 1995 through the latter half of 2001. The variation of the rates is an indication of increase or decrease in the frequency of the events caused by events' being correlated with one another within a time period. The rates displayed in a panel array the loans by the ordered principal and interest to income ratio. The P&I to income ratio is a measure of ability to pay. Each P&I to income ratio percentage reported as headers indicates the decile or half-decile category created using this ratio. The leftmost group (0–5%) has the highest ability to pay at the time the loan is made.

Briefly reviewing the results from actual defaults, we obtain a 6-month default correlation of 0.062 for CC-rated borrowers as compared with a minuscule 0.001 for AA-rated borrowers. When grouped by occupancy type, the default correlation increases to 0.087 for second-home loans and 0.026 for non-owner-occupied homes. In contrast, classification by property type results in a maximum default correlation of 0.046 for two- to four-unit properties.

If default correlations are very low within subprime portfolios, an expensive investigation of default correlations is not an efficient use of resources. However, our findings, combined with the findings of Loffler (2003) that lower grade portfolios are more sensitive to changes in default correlations, suggest that the industry should focus on this issue. Although our findings represent but one lender, they clearly provide sufficient insight to suggest directions for further investigation. If, as Carey (2000) suggests, bad tail loss rates are understated by estimating portfolio loss distributions by equally weighting events in each database year, then our results should compel both subprime lenders and regulators to further investigate the impact of default correlation.

Acknowledgements

We are grateful for the helpful comments received from participants at the Retail Credit Risk Conference at the Federal Reserve Bank in Philadelphia, April 2003, and from Lance Nail. We are also grateful to the editors, Loretta Mester and Mitchell Berlin. And we would especially like to thank an anonymous referee for constructive criticisms and comments.

References

Calem, P.S., LaCour-Little, M., 2001. Risk-based capital requirements for mortgage loans. Working Paper 2001–60. The Federal Reserve Board, Financial and Economics Discussion Series, November.

Carey, M., 1998. Credit risk in private debt portfolios. Journal of Finance 53, 1363–1387.

Carey, M., 2000. Dimensions of credit risk and their relationship to economic capital requirements. Working Paper 7629, National Bureau of Economic Research, March.

Crouhy, M., Galai, D., Mark, R., 2000. A comparative analysis of current credit risk models. Journal of Banking and Finance 24, 59–117.

Frey, R., McNeil, A., 2002. VaR and expected shortfall in portfolios of dependent credit risks: Conceptual and practical insights. Journal of Banking and Finance 26, 1317–1334.

Gordy, M., 2000. A comparative anatomy of credit risk models. Journal of Banking and Finance 24, 119–149.

Jarrow, R.A., Turnbull, S.M., 2000. The intersection of market and credit risk. Journal of Banking and Finance 24, 271–299.

Loffler, G., 2003. The effects of estimation error on measures of portfolio credit risk. Journal of Banking and Finance 27, 1427–1453.

Lucas, A., Klaassen, P., Spreij, P., Straetmans, S., 2001. An analytic approach to credit risk of large corporate bond and loan portfolios. Journal of Banking and Finance 25, 1635–1664.

Morgan, J.P., 1997. Credit Metrics^{TM} – Technical Document. J.P. Morgan & Co. Incorporated, New York.

Nagpal, K., Bahar, R., 2001. Measuring default correlation. Risk 14, 129–132.
Scheessele, R., 2002. Black and white disparities in subprime mortgage refinance lending. Working Paper HF-014, US Department of Housing and Urban Development, Housing Finance, April.
Vandell, K., 1995. How ruthless is mortgage default? A review and synthesis of the evidence. Journal of Housing Research 6, 245–264.
Zhou, C., 1997. Default correlation: An analytical result. Working Paper, Federal Reserve Board, Washington, DC.