MANAGEMENT SCIENCE
Vol. 38, No. 11, November 1992
Printed in U.S.A.

COMPLETE PREPAYMENT MODELS FOR MORTGAGE-BACKED SECURITIES*

PAN KANG AND STAVROS A. ZENIOS
HERMES Lab for Financial Modeling and Simulation, Decision Sciences Department, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

The estimation of prepayment rates for pools of mortgages is a critical component in determining the value of mortgage-backed securities—MBS for short—and derivative products. This paper discusses the development of prepayment models for pools of fixed-rate mortgages. The models are complete: calibrated functional forms are given for all of the factors that determine prepayment rates. Hence, the models can be used as benchmarks against the simple models of the Public Securities Association, the Federal Housing Administration experience, or the variety of projected prepayment rates generated by proprietary industry models.

The key factors that determine prepayment rates are: (1) refinancing incentive, (2) seasonal variations, (3) seasoning of the mortgage pool, and (4) burnout effect. Each factor is modeled separately and is calibrated using historical data. A multiplicative relationship determines the prepayment rate of the mortgage pool. A novel feature of our model is the use of *basis* functions that capture the complex interactions between the control variables, i.e., interest rate differentials and time, and the response parameter, i.e., prepayment rates.
(MORTGAGE-SECURITIES; PREPAYMENTS; REGRESSION ANALYSIS; SECURITY PRICING)

1. Introduction

Mortgage-backed securities (abbreviated: MBS) hold a prominent position in the securities market. The outstanding mortgage debt in the U.S.—$3.5 trillion as of the second quarter of 1988—dwarfs the more established Government and Corporate debts. Since the '80s, mortgage debt has been financed in part by the issuance of mortgage-backed securities. Approximately 25% of the outstanding residential debt had been securitized as of 1988. These securities are actively traded in the secondary markets: trading volume reached $1.2 trillion in 1985. On top of the mortgage pass-through securities—where all the cashflow from the underlying mortgages, net a service fee, is passed on to the investors—there has been a proliferation of derivative products. Collateralized mortgage obligations (CMO), interest only (IO) and principal only (PO) are some of the derivative instruments.

Pricing MBS is a complex problem. First, a mortgage-backed security combines features of annuities and options: the homeowners' ability to prepay represents a *call* option. Second, this call option is not necessarily exercised optimally, according to the rational behavior prescribed by options pricing theory. For example, a homeowner may sell the property and prepay the mortgage even if the prevailing mortgage rates are higher than the contract rate of her mortgage. Employment changes, death, divorce and so on are factors that affect prepayment decisions.

To cope with uncertainty in pricing MBS—and to capture the complex relationships between the cashflows of a MBS and interest rate paths—researchers and Wall Street analysts resort to simulation methods. These methods combine models for the evolution of interest rates with models of prepayment for the mortgage pool. Naturally, estimating the prepayment rates of a pool is a critical input of pricing methodologies. Several authors argue that it is the weak link of current models for pricing and for the management of MBS portfolios.

* Accepted by William T. Ziemba; received March 1991. This paper has been with the authors 4 months for 1 revision.

0025-1909/92/3811/1665$01.25

In this paper we discuss the development of prepayment models for pools of fixed-rate mortgages. We consider the aggregate prepayment activity of groups of homeowners whose individual mortgages are pooled together in a—somewhat homogeneous—MBS.

The models are complete from two perspectives: First, they discuss in a common framework several key factors that affect prepayment rates. We concentrate, however, on factors that have a similar effect on all outstanding mortgage pools. For example, our model ignores the geographic location and other demographic effects. While we recognize that pools from different geographic locations will exhibit distinct prepayment behavior, our interest is in predicting prepayments of *generic* pools. As will be seen later, the factors that are included in the model have been sufficient for explaining historically observed prepayments on generic pools.

Second, our models present the precise functional forms used to estimate the prepayment rates. Hence, they can be used as benchmarks against simple models like the Public Securities Association, the Federal Housing Administration experience, or the variety of projected prepayment rates generated by proprietary models that are in widespread use by the market.

A novel feature of our model is the use of *basis* functions to capture the complex interactions between the control variables—interest rate differentials and time—and the response parameters—prepayment factors. Basis functions are simple functions (like the step function, or the ramp-like function) parametrized in such a way that a linear combination of these functions will provide an arbitrary level of accuracy in tracking some complex, nonlinear function. Basis functions are of common use in finite element or finite difference analysis, where the solution to some underlying differential equation (being an unknown complex function that cannot be specified analytically) is approximated by estimating the coefficients of basis functions. The choice of appropriate basis functions is more of an art than a science, and requires an understanding of the shape of the function that is being approximated. For the prepayment model we could easily identify appropriate basis functions, and this approach resulted in better estimation accuracy than currently employed methods. However, the calibration procedure is significantly more complex, and it requires techniques from large-scale optimization.

The paper starts with a critical review of existing literature. It proceeds in §3 to describe the modeling process, and give functional forms and explain the calibration techniques. Section 4 gives the complete prepayment models with a discussion of the statistical significance of their parameters. The use of these models with sample securities is illustrated in §5. Section 6 draws some conclusions and points out ways in which the models can be further improved.

2. Background

The literature on mortgage-backed securities is vast. A volume that covers several facets of mortgage-backed financing is Fabozzi (1989). Papers therein discuss pass-through securities and CMOs together with models for valuation and prepayment forecasting, management strategies, hedging, and related computing technologies.

Current standards for estimating the prepayment of mortgage-backed securities are the Public Securities Association (PSA) model and the Federal Housing Administration (FHA) experience. The PSA model assumes that the prepayment rates increase at a constant rate during the first 30 months of the life of the mortgage until they reach 6%. Then they remain constant. The FHA experience is summarized by a series of annual survival rates for various populations of FHA insured mortgages. Figure 1 illustrates both the PSA model and the 1985 FHA experience.

A severe limitation of both the PSA and the FHA models is that they are interest-rate invariant and assume a very simple aging process for mortgages. Hence, both models are

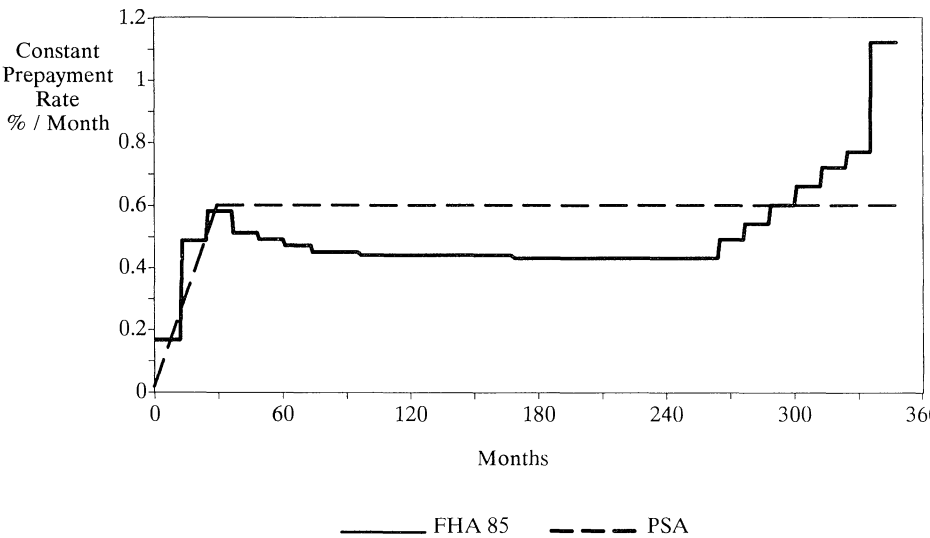


FIGURE 1. The Public Securities Association (PSA) model and the 1985 Federal Housing Administration (FHA) experience of mortgage prepayment rates.

of limited value within the context of pricing models where the evolution of interest rates is modeled explicitly.

One of the earlier successful attempts to model explicitly the effect of interest rates and other factors on prepayment is due to Curley and Guttentag (1974). Their work pointed out the severe limitations of models that assumed a fixed time period for the full prepayment of a mortgage (i.e., the 12-year rule of thumb used at that time). They argued that prepayment rates will depend on the age of the mortgages in the pool, the coupon rate and the current discount points. Their work opened the way for a series of models that explicitly accounted for the dynamic nature of prepayment activities.

Peters (1979) continued this line of inquiry in her dissertation. She identified more than a dozen explanatory variables that influence the prepayment behavior of different types of mortgages. Green and Shoven (1986) use a proportional hazards model to estimate the effect of variations of interest rates on prepayments. Their model indicates a very high correlation between the prepayment rate and the relation of the market rate to the mortgage contract. While it was successful in estimating prepayment rates for two pools of new mortgages, the proportional hazards model assumes that past attributes of the environment have no effect on present turnover. This assumption loses little realism for new mortgages. For mortgages that have experienced a long history of interest rate variations, this assumption will lead to gross overestimation of prepayment rates. A related study is due to Navratil (1985), that uses a logistic model to fit prepayment rates to interest rate differentials.

A model that added more key factors in estimating prepayments rates is due to Richard and Roll (1989). This appears to be at present one of the most comprehensive models for estimating prepayments. Their report gives details of the factors included in the model: seasonal variations, pure refinancing effects and aging effects. However, there is no discussion on how the factors are isolated from historical observations and no mention of the functional relationships that are used in the actual estimation. Nevertheless, sample results are presented, and their model appears to have high explanatory power. The Prudential-Bache prepayment model is outlined in Hayre, Lauterbach and Mohebbi (1988). They discuss the performance of their model, but the model itself is not revealed. Carron et al. (1988) discuss the prepayment models that are being used by First Boston. Their models include the same factors as those introduced by Richard and Roll, as well as a factor that takes into account housing activity, and hence, indirectly, economic growth. They present some experiences with their model, but neither the model nor its calibration procedures are revealed. A brief discussion of a prepayment model used in a generalized pricing framework is given in Jacob, Lord and Tilley (1987). The precise functional form of a simple model for GNMAs is given for the purpose of illustrations in their pricing framework.

The success of the aforementioned models is usually claimed to be quite high, as justified by very high correlations between actual and fitted prepayments. If a different model is fitted for each coupon and maturity date, some models achieve an R-squared of 99%. This claim is an exaggeration: the prepayment model will perform poorly in estimating prepayments under conditions that were not encountered in the sample of historical observations. Furthermore, R-squared gives little information on the accuracy of the model when used for extrapolations and in interest-rate simulation studies. Given also the noise in the observed data, such figures should be taken with a dose of skepticism. Jacob et al. are making the most sensible claims in this regard. Their simple model achieves an R-squared of 57% for GNMAs. This level of explanatory power is rather low, and pricing results obtained using this model are suspect. Nevertheless, according to the authors, it can be improved to 75–85% with more detailed models. Given the complexities of MBS and the noise in observed data, such explanatory power indicates that the model was successful in capturing the primary factors that affect prepayment.

3. Preview of the Modeling Process

The characteristics of a mortgage that provide indicators for the prepayment behavior of the mortgage borrower are: (1) the age of the mortgage, call it t, (2) the month of the year, m, and, (3) the ratio between the mortgage contract rate denoted by C, and the prevailing rate at which the mortgage can be refinanced, R. We also recognize that the prepayment activity of a pool will depend on the whole history of C/R ratios since its issuance. We use $(C/R)_t$ to denote the vector of these ratios for all time periods less than t.

Our models deal with the estimation of the prepayment activity of pools of mortgages, and not of individual mortgages. The formation of a pool, according to the legal specifications adhered to by the issuing agencies, creates MBS which are not perfectly homogeneous with respect to the aforementioned characteristics. For example, for an MBS with a given pass through rate, the underlying mortgage rates may be slightly different, as will be their maturities. Since our interest is in dealing with *generic* MBS we ignore this heterogeneity. Each pool is treated as a homogeneous collection of mortgages with contract rates equal to the WAC (weighted average coupon) of the pool and with maturity equal to its WAM (weighted average maturity). This treatment of the pools is consistent with current practices by the market, and our model loses little realism in this respect. Furthermore, detailed information on the individual mortgages that are in a given pool is usually unavailable.

Other characteristics of a mortgage pool may be relevant to the prepayment activity. Geographic location is particularly relevant in capturing demographic phenomena. Nevertheless, the three pool characteristics mentioned above are sufficient for the representation of the factors that we include in our models. The age of the mortgage and the vector $(C/R)_t$ are essential: we believe that the prepayment rates are not static, but change dynamically with the age of the mortgages in the pool. Evidence to this effect is seen in the FHA experience curve and is also assumed in the PSA model (see Figure 1). The month of the year captures seasonal variations, with high prepayment rates in the late summer and early fall, and low prepayment rates during the winter months. Our use of the C/R ratio, as an indicator of prepayment rates due to refinancing incentives, is in agreement with the analysis of Richard and Roll (1989). This measure differs from the widely adopted differential $C - R$, see Green and Shoven (1986), Hayre et al. (1988) or Navratil (1985).

The rationale for our choice is based on options pricing theory (Richard and Roll 1989). A homeowner has a financial incentive to prepay her mortgage if the cost of continuing the current payments is above the cost of payments of a new mortgage at the current rates (including mortgage origination costs). The present value of a 30-year mortgage, per unit of monthly payment, is given by

$$V_t = \frac{[1 - (1 + R)^{t-360}]}{R}. \tag{1}$$

The mortgage balance, per unit of monthly payment, is given by

$$B_t = \frac{[1 - (1 + C)^{t-360}]}{C}. \tag{2}$$

Mortgage owners will exercise the prepayment option if the ratio V_t/B_t exceeds some critical value that reflects the implicit costs of refinancing. It is easy to see that the ratio V_t/B_t is well approximated by C/R for a wide range of values of t. Hence we use this ratio as a proxy for the pure economic incentive to refinance the mortgage (see §7).

Our model for prepayment estimates contains four factors, each of which is in turn determined by the mortgage characteristics introduced above, and the prevailing refinancing rates. The factors, and their independent variables, are the following:

(1) *Seasonality effect*, $s(m, C/R)$. This factor captures the increase in house sales during the summer and early fall, and the decrease of sales in the winter months. This effect arises primarily due to housing activity, although an argument can be made that refinancing activity is higher following the end of a tax year.

(2) Pure economic *refinancing* incentive $\rho(C/R)$. This factor captures the homeowners' decision to refinance their mortgage if the prevailing mortgage rates R are lower than the coupon rate of their mortgage.

(3) Effect due to *seasoning* of the mortgage, $\sigma((C/R)_t, t, C/R)$. Prepayment rates rise in the early years of the life of a mortgage because homeowners do not take out mortgages with the expectation that they will quickly move. At the same time they may lack the energy and money needed to refinance a new mortgage. As the time from origination lengthens the probability of a move, leading to prepayment, increases. The rate with which prepayments increase depends on the history of $(C/R)_t$ ratios experienced by the pool since its issuance. Premium pools season faster than discount pools.

(4) *Burnout* effect for older mortgages, $\beta((C/R)_t, t, C/R)$. Prepayment activity tends to decrease for older mortgages. The shorter the period to maturity the larger the ratio C/R required to generate any level of savings due to refinancing. Furthermore, if $R < C$ some, but not all, borrowers will refinance. There are several reasons for this behavior. Some homeowners expect even more favorable opportunities in the future, some are just lethargic and others do not qualify for a new loan. Whatever the reason, with the passage of time those borrowers with the greatest sensitivity to C/R will prepay, leaving in the pool those with less sensitivity. The combined result of all these considerations is that prepayment activity tends to slow at the end of the life of the mortgage. However, pools that have been at discount since origination will experience less burnout than those that have been at par or at premium.

The last two factors together capture the effect of *aging* on the prepayment rates. The model combines these four factors to estimate the *constant prepayment rate* (CPR) for the specific mortgage characteristics:

$$\mathrm{CPR}(t, m, (C/R)_t) = s(m, C/R) \cdot \rho(C/R) \cdot \sigma((C/R)_t, t, C/R) \cdot \beta((C/R)_t, t, C/R) \cdot \epsilon. \tag{3}$$

ϵ is used to denote the error in the model which is assumed to be normally distributed with mean 1.0. The prepayment model is complex and highly nonlinear. The four factors depend on one or more of the independent variables, and most of the factors depend on the same variable, but not necessarily in the same way. The rest of this section is devoted to techniques for filtering out the four factors based on historical observations.

Historical data tabulate the outstanding balance for each observed pool for each month since its issuance, or since the beginning of the observation period. The observed outstanding mortgage balance, and the scheduled outstanding balance—if prepayments were not exercised—can be combined to estimate the monthly mortality rate of the mortgage. Hence, a series of constant prepayment rates are observed under known values of the independent variables. It is this series that provides the information for the calibration of our prepayment model. Details on the conversion of outstanding mortgage balance into a series of CPRs and other technical information can be found in Bartlett (1989).

3.1. *Modeling the Seasonal Variations*

Estimating the seasonal variations of prepayments is the easiest phase of the modeling process. First we cluster securities into cohorts: we group together securities that are in a prespecified age range (say, T_0 to T_1 months) and have the same (within a small tolerance) C/R ratio. For example, a 10% security that finds itself in a 10% refinancing environment at month m (say, January) when it is T_0 to T_1 months old is grouped together with an 11% security that finds itself in an 11% environment at the same month of the year, when it is also T_0 to T_1 months old. (Of course, the 10% and 11% interest rate environments were not realized in the same year, but in the same month of different years.) We used 20 intervals of C/R in the range [0.50, 1.60]. Within each cohort of homogeneous C/R the CPR factors are identified by month, and by the outstanding pool balance for each observation.

The seasonality adjustment factor for a given month, and a given C/R interval, is computed by scaling the weighted average prepayment activity of that specific month and C/R interval by the weighted average prepayment activity of all months for the same C/R interval. Weighted averages are computed based on the outstanding balance of each observed pool. Details of the mathematical model are given in Appendix A.

The seasonal adjustment is estimated for all C/R intervals, together with the standard deviation of these estimates. We then proceed to check if the seasonality factors generated for all cohorts of the same age range are within $\pm 2\sigma$ from each other. If the factors fail this test, the age range $[T_0, T_1]$ is adjusted and the process repeats. The rationale for this iteration is twofold. First, there is no reason to assume that seasonal variations are affected by C/R. Second, seasonal variations may not reach a steady state for very new or very old mortgages. New mortgages tend not to prepay (i.e., are not yet fully *seasoned*), while old mortgages suffer from *burnout*. None of these factors has yet to be eliminated from the observed data! Hence, it is important to find the age range for which these effects are insignificant. The iterative procedure explained here is used to identify the suitable age range. One iterative step with this method has been sufficient to produce acceptable seasonal variations. We found the age range [36, 168] months to give the best results for estimating both GNMA and FNMA seasonality factors.

3.2. *Modeling the Refinancing Incentive*

Having now estimated the effect of seasonal variations, we proceed to fit the remaining factors. First, the observed prepayment rates are seasonally adjusted by dividing each observation by the corresponding seasonality factor. The model of equation (3) becomes:

$$\hat{\text{CPR}} = \frac{\text{CPR}(t, m, C/R)}{\hat{s}(m)} = \rho(C/R) \cdot \sigma((C/R)_t, t, C/R) \cdot \beta((C/R)_t, t, C/R) \cdot \epsilon', \quad (4)$$

where $\hat{s}(m)$ is the seasonality adjustment factor for month m estimated by the seasonality model (and assumed to be identical for all C/R ranges). We now want to fit the refinancing incentive $\rho(C/R)$ to the seasonally adjusted observations $\widehat{\mathrm{CPR}}$. This has to be done in such a way that the residual effect, i.e., aging, satisfies

$$\sigma((C/R)_t, t, C/R) \cdot \beta((C/R)_t, t, C/R) \leq 1.$$

Of course, at this stage we have not yet isolated the aging effect. Hence, we proceed in an iterative way, as we did for the seasonality estimations.

We fit first a model using the seasonally adjusted CPRs in the age range 48–180 months. The problem is to estimate a functional form that closely tracks the CPR rates due to refinancing. We use a least-squares estimation procedure with suitable (shifted, ramp-like) basis functions to estimate the functional form of the refinancing effect. Details of the mathematical model are given in Appendix B.

Using this first estimate of the refinancing incentive (call it $\hat{\rho}_0(C/R)$), and the already

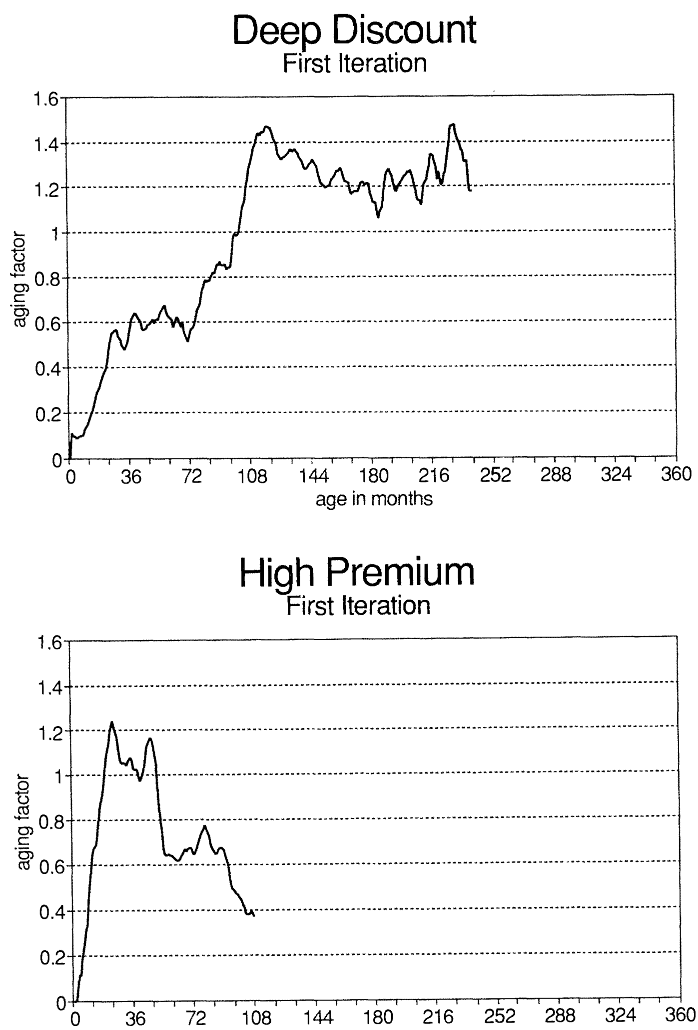
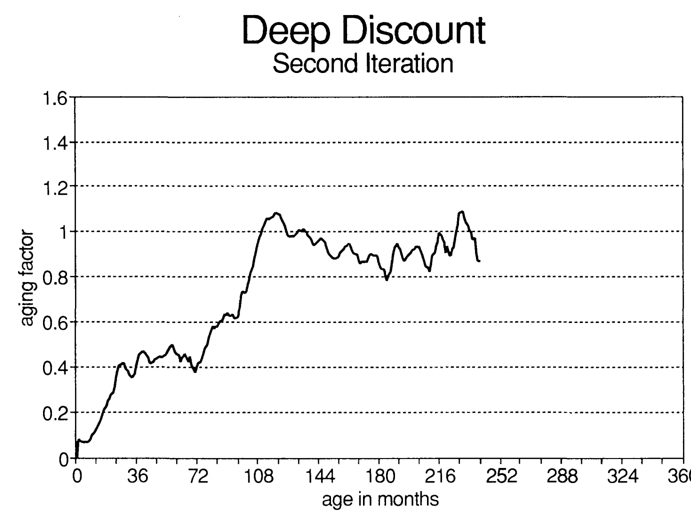


FIGURE 2. First approximation of the aging effect for GNMA pools. The aging effect exceeds 1.0 after the age of 84 months for the discount bonds. It exceeds 1.0 in the age range 24–48 months for the premium bonds. Hence, this approximation has failed to remove all seasonality and refinancing effect from the aging effect for the age ranges indicated.

estimated seasonality factors, we can remove both the seasonal variations and the refinancing incentive from the observed data. The remainder is due to the aging effect:

$$A_t = \frac{\mathrm{CPR}(m, C/R, t)}{\hat{s}(m)\hat{\rho}_0(C/R)} . \tag{5}$$

The aging effect A_t as a function of time is shown in Figure 2 for high premium and deep discount GNMA securities respectively. We observe from the two plots of Figure 2 that the aging effect exceeds 1.0 after 86 months for discount bonds, and in the range 24–48 for premium bonds. Hence, the assumption we made in simplifying equation (4) is not valid outside the indicated age ranges. Our current model still attributes some of the refinancing effect to the aging process. We, therefore, rerun the refinancing model of Appendix B using now different age ranges for premium and discount bonds: 24–48 months for premium bonds and 84–180 months for discount bonds. In essence, this is an attempt to isolate the seasoning effect using discount bonds, and the burnout effect using premium bonds. The model produces two refinancing factors, $\hat{\rho}_p$ and $\hat{\rho}_d$, for premium and discount bonds, respectively.



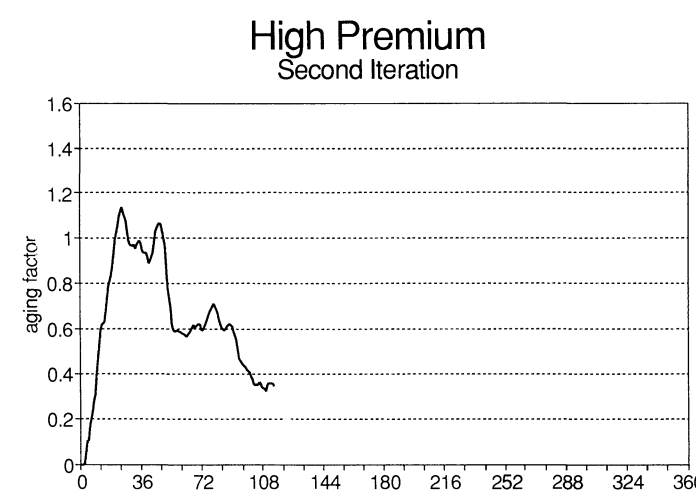


FIGURE 3. Second approximation of the aging effect for GNMA pools. The aging effect is almost always less than 1.0 for all age ranges for both premium and discount bonds. Hence, the model has successfully isolated the seasonality and refinancing effects from aging.

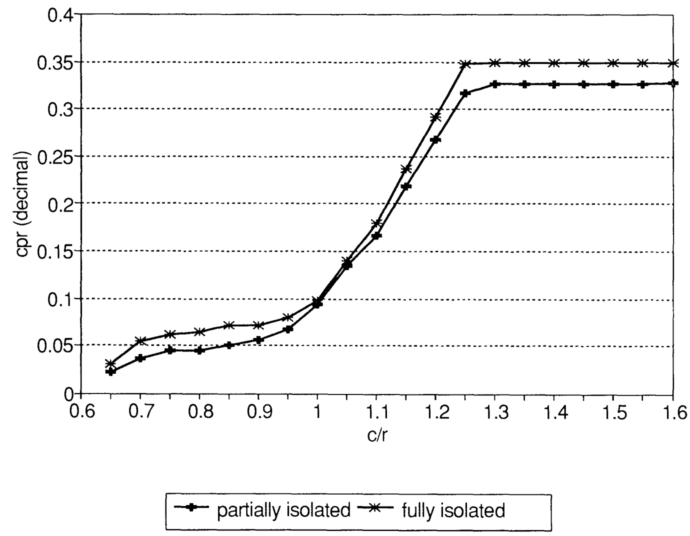


FIGURE 4. Refinancing effect for GNMA-I: With and without fully isolating the seasoning and burnout effects.

Using now these new models we isolate the aging factors, using equation (5) with $\hat{\rho}_p(C/R)$ and $\hat{\rho}_d(C/R)$ for premium and discount bonds respectively. The result is shown in Figure 3. We observe that the aging factor for discount bonds is correct when the observed data are adjusted using $\hat{\rho}_d(C/R)$. Similarly, the aging factor for premium bonds is correct when the data are adjusted using $\hat{\rho}_p(C/R)$. Combining the seasoning factors for premium and discount bonds we obtain the final model for the refinancing factor:

$$\hat{\rho}(C/R) = \begin{cases} \hat{\rho}_d(C/R) & \text{if} \quad C/R < 1, \\ \hat{\rho}_p(C/R) & \text{if} \quad C/R \geq 1. \end{cases} \tag{6}$$

Figure 4 illustrates the fitted model for the refinancing incentive of GNMA-I when a single refinancing model $\hat{\rho}_0$ was first employed for all bonds, and when the two different models $(\hat{\rho}_p, \hat{\rho}_d)$ were obtained for premium and discount bonds as appropriate.

3.3. *Modeling the Aging Effect*

Thus far we have modeled the seasonality and refinancing incentives. Hence, we may remove these two effects from the observed data, to obtain the simplified model in the remaining factors.

FIGURE 5. Weighted average of observed data for the aging of FNMA discount pools.

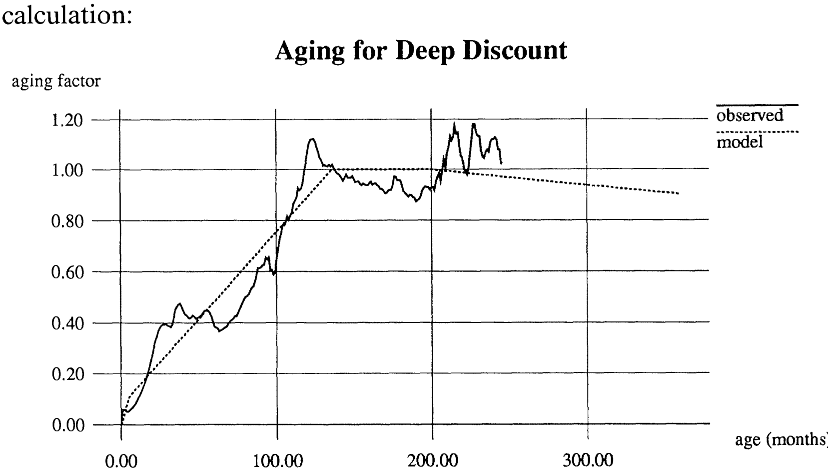
$$\hat{\text{CPR}} = \frac{\text{CPR}(t, m, C/R)}{\hat{s}(m)\hat{\rho}(C/R)} = \sigma((C/R)_t, t, C/R) \cdot \beta((C/R)_t, t, C/R) \cdot \epsilon''. \tag{7}$$

This model describes the aging process of the security. Figure 5 illustrates the weighted average of the CPR factors for FNMA discount pools, as observed from the historical data, and after they have been adjusted for seasonality variations and the pure refinancing incentive. This figure clearly illustrates the dynamic nature of prepayments as the mortgages in a pool age.

The two remaining factors—seasoning and burnout—are easily separated from each other: Seasoning occurs at the beginning of the life of a mortgage, while burnout occurs during the later part of its life. The next two sections are devoted to the estimation of these factors from the observed CPRs after they have been adjusted for the seasonal variations and the refinancing effects.

3.4. *Modeling the Seasoning Effect*

We recognize that the seasoning effect depends on the history of interest rates, as measured by the vector $(C/R)_t$. We approximate this dependence on the history of interest rates by separating pools into deep discount, discount, par, premium and so on. We then estimate the seasoning effect for each category separately. This is done by grouping together pools that had the same C/R at the same point in time. (Of course, this implies that different pools may be grouped together at different points in time.) Once we estimate the seasoning effect assuming that the vector $(C/R)_t$ is constant and equal to C/R, i.e., $\sigma(\{(C/R)_t = C/R\}, t, C/R)$, we can calculate the dynamic $\sigma((C/R)_t, t, C/R)$ by the simple calculation:



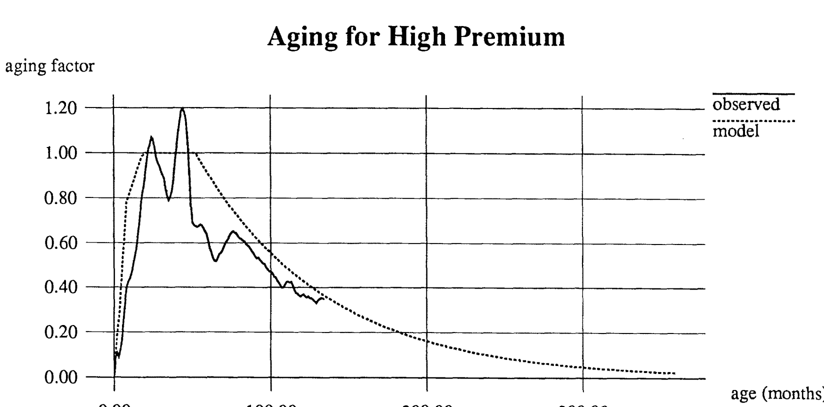


FIGURE 6. Weighted average of observed data and fitted model for the seasoning and burnout of deep discount and high premium GNMA bonds.

$$\sigma((C/R)_t, t, C/R) = \sum_{\tau=1}^{t} \sigma(\{(C/R)_\tau = C/R_\tau\}, \tau, C/R_\tau).$$

We use again constrained least-squares and appropriate basis functions to build a model that tracks the CPR due to the seasoning effect. We use in particular ramp-like basis functions to capture the nondecreasing nature of the seasoning effect as the mortgage ages after origination. Details of the mathematical model are given in Appendix C. The two diagrams in Figure 6 illustrate the fitted model and the weighted average observations of the seasoning factor for both deep discount bond and high premium GNMA bonds. As expected, high-premium bonds age much faster than the deep discount bonds. We also observe that the estimated model fits very accurately the observed data, an indication of the power of the basis functions we used in tracking the complex behavior of aging.

3.5. *Modeling the Burnout Effect*

The burnout effect is obtained in a fashion similar to that used in estimating the seasoning effect. However, we are now interested in the latter part of the life of a mortgage, when the burnout effect is having an impact. Furthermore, a different basis function is needed to capture the decay of prepayment rates as the mortgage matures. In this case we use shifted negative exponential basis functions. Details of the mathematical model for the burnout effect are given in Appendix D.

Figure 6 illustrates the fitted model and the weighted average of observed aging factors for both deep discount and high premium GNMA bonds. We observe that premium bonds burn-out before discount bonds, and at a faster rate. Our model tracks very well the observed data, an indication of the flexibility of the basis function approach.

3.6. *The Integrated Model*

The complete model is finally built by combining the four factors using the multiplicative relation:

$$\mathrm{CPR}(t, m, C/R) = s(m, C/R) \cdot \rho(C/R) \cdot \sigma((C/R)_t, t, C/R) \cdot \beta((C/R)_t, t, C/R).$$

The dynamic dependence of the prepayment activity on the history of interest rates is captured by observing the aging of a pool as it changes from being at par (at origination) to discount or premium (as interest rates move up or down).

It is also possible to obtain a confidence interval for the CPR estimate based on the estimates and standard deviations of the four factors. Let $\hat{\theta}_i$, $i = 1, 2, 3, 4$ be estimates of the four factors for a given age, month and C/R ratio. Let also $\hat{\sigma}_i^2$, $i = 1, 2, 3, 4$ be estimates of the variance for each one of the four factors. Then an $\alpha\%$ approximate confidence interval for $\mathrm{CPR}(t, m, C/R)$ can be obtained from large sample theory as:

$$\hat{\theta}_1 \cdot \hat{\theta}_2 \cdot \hat{\theta}_3 \cdot \hat{\theta}_4 \pm z_{1-\alpha/2} \sqrt{\sum_{i=1}^{4} \hat{\gamma}_i \hat{\sigma}_i^2}$$

where $\hat{\gamma}_i = \prod_{j=1, j \neq i}^{4} \hat{\theta}_j$.

4. The Calibrated Prepayment Models

We have now completed the discussion on the development and calibration of prepayment models. We build such models for GNMA-I, GNMA-II, FNMA and FHLMC securities. We used historical data in the form of monthly outstanding mortgage-balance provided courtesy of Financial Publishing Company, Boston, Massachusetts. The nonlinear least squares models were solved using the General Algebraic Modeling System (GAMS) of Brooke, Kendrick and Meeraus (1988).

As one may imagine, the development of the models is a complex and computationally intensive process. We used a distributed network of DEC stations at the HERMES Lab-

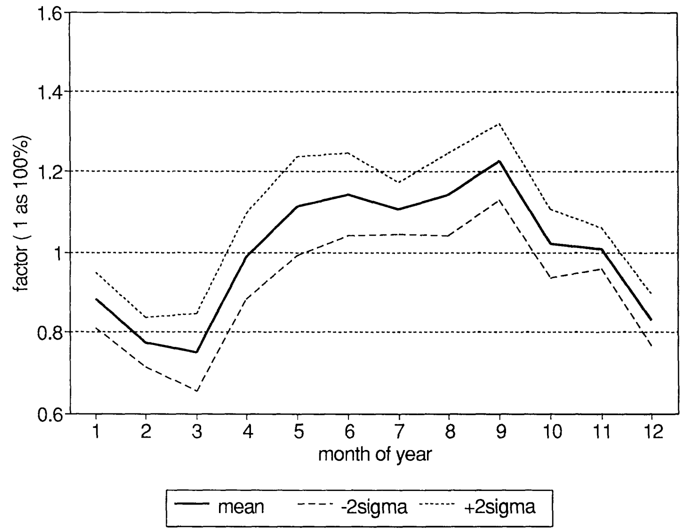


FIGURE 7. Seasonality variations of GNMA securities: mean value and 2-*sigma* confidence interval.

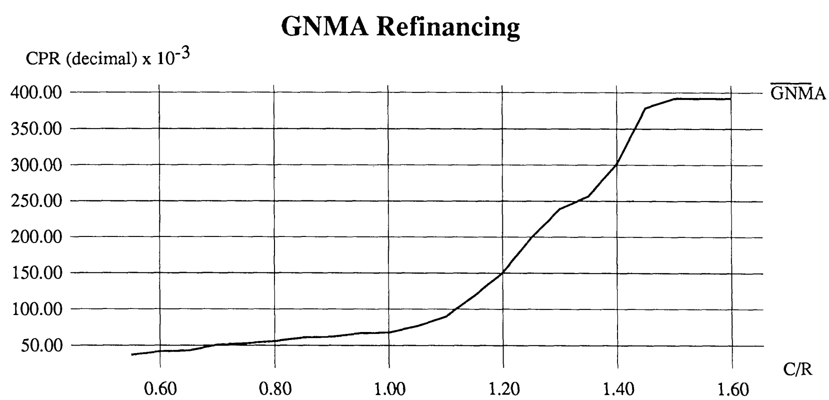


FIGURE 8. Pure refinancing incentive for GNMA securities.

oratory and a VAX 6400 mainframe. The analysis of the observed data was carried out by custom-made C-programs on the DEC stations. The same programs also set up the appropriate input files for the GAMS optimizers.

In this section we give details of the models we built for GNMA-I and FNMA.

4.1. *The GNMA-I Experience*

The model for various coupon rates of GNMA securities are given in the following figures: Figure 7 gives the seasonality factors for all C/R ranges, together with the confidence interval of 2σ. Figure 8 gives the pure refinancing incentive in the range $C/R = [0.5\text{–}1.6]$. Figure 9 gives the aging effect for both discount and premium bonds, combining both seasoning and burnout.

The model has an R-squared of 0.83 when compared against the aggregate data, aggregated by the number of observations in each cohort. A different R-squared is observed for different age ranges, and it is higher for the mid-life of the pool when the prepayment activity has been stabilized. The R-squared when the data in each cohort are aggregated by the outstanding mortgage balance—a more reasonable practice—is 0.85.

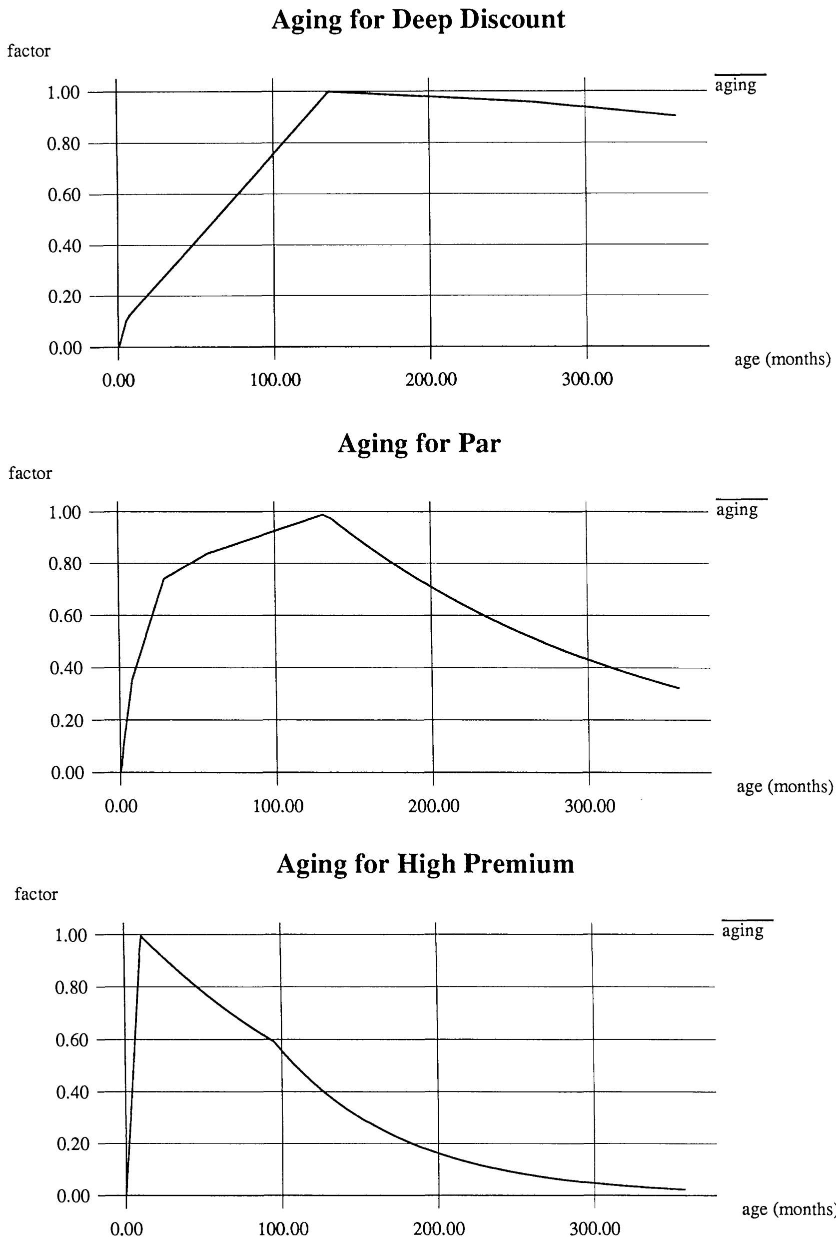


FIGURE 9. Seasoning effect of deep discount, par and high premium GNMA securities.

4.2. *The FNMA Experience*

The model for various coupon rates of FNMA securities are given in the following figures: Figure 10 gives the seasonality factors for all C/R ranges, together with the confidence interval of 2σ. Figure 11 gives the pure refinancing incentive in the range C/R = [0.5–1.6]. Figure 12 gives the aging effect for various C/R ranges.

The model has an R-squared in the range 0.78–0.81 when compared against the aggregate data, aggregated by the number of observations in each cohort. Different R-squared is observed for different age ranges, and is higher for the mid-life of the mortgage when the prepayment activity has been stabilized. The R-squared when the data in each cohort are aggregated by the outstanding mortgage balance—a more reasonable practice—is in the range 0.81–0.85.

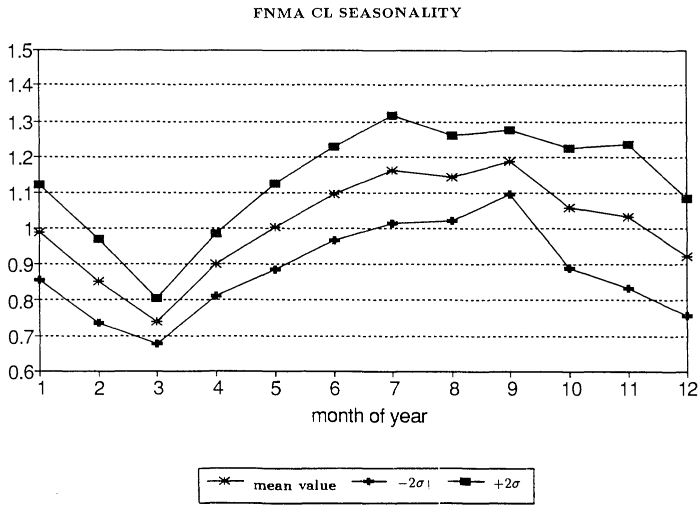


FIGURE 10. Seasonality variations of FNMA securities: mean value and 2-*sigma* confidence interval.

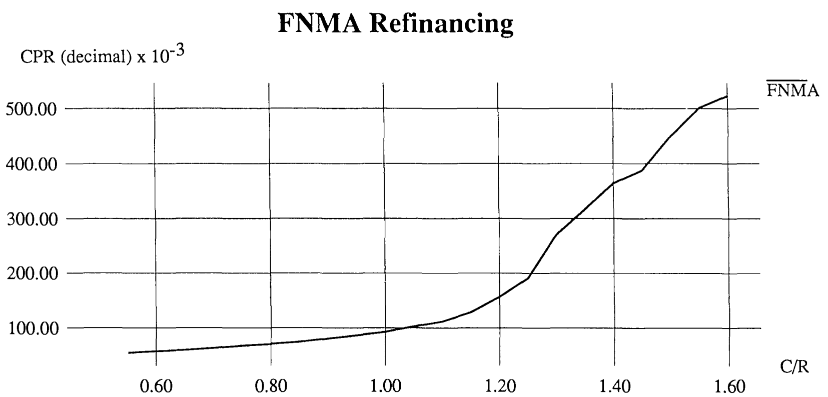


FIGURE 11. Pure refinancing incentive for FNMA securities.

5. Using the Models

We first used the model to observe its performance in explaining already observed interest rate patterns. We isolated a weighted average of GNMA-I securities of different coupon rates and ages from our database over the period 1980 to 1990. In this period we have observations for both the market mortgage rates R and the observed prepayment rates. We show in Figure 13 the prepayment rate as obtained from our model, against the empirically observed data. The model CPRs appear to track closely the observed data. The model is particularly accurate in the early part of the life of the pool and when interest rate changes are small. (See the results from the first 60–70 months of the 3-month-old GNMAs.) The model has also very good tracking behavior for mature pools and in more volatile interest rate environments. (Notice the behavior of the model after 7–8 years for the 9.9–10.1% GNMA.) The model has also been able to capture the burnout effects that are observed at months 100 and 110 when the prepayment activity has slowed down following the spike at month 90. The model appears to be less accurate as the mortgages age more. See, for example, the discrepancy at the later months of the aged GNMA. This behavior is expected since we have only limited observations for aged pools.

We then run the model to demonstrate the prepayment activity of a generic GNMA-I 8.9 assuming constant interest rate environment at 8.9%; see the first diagram in Figure

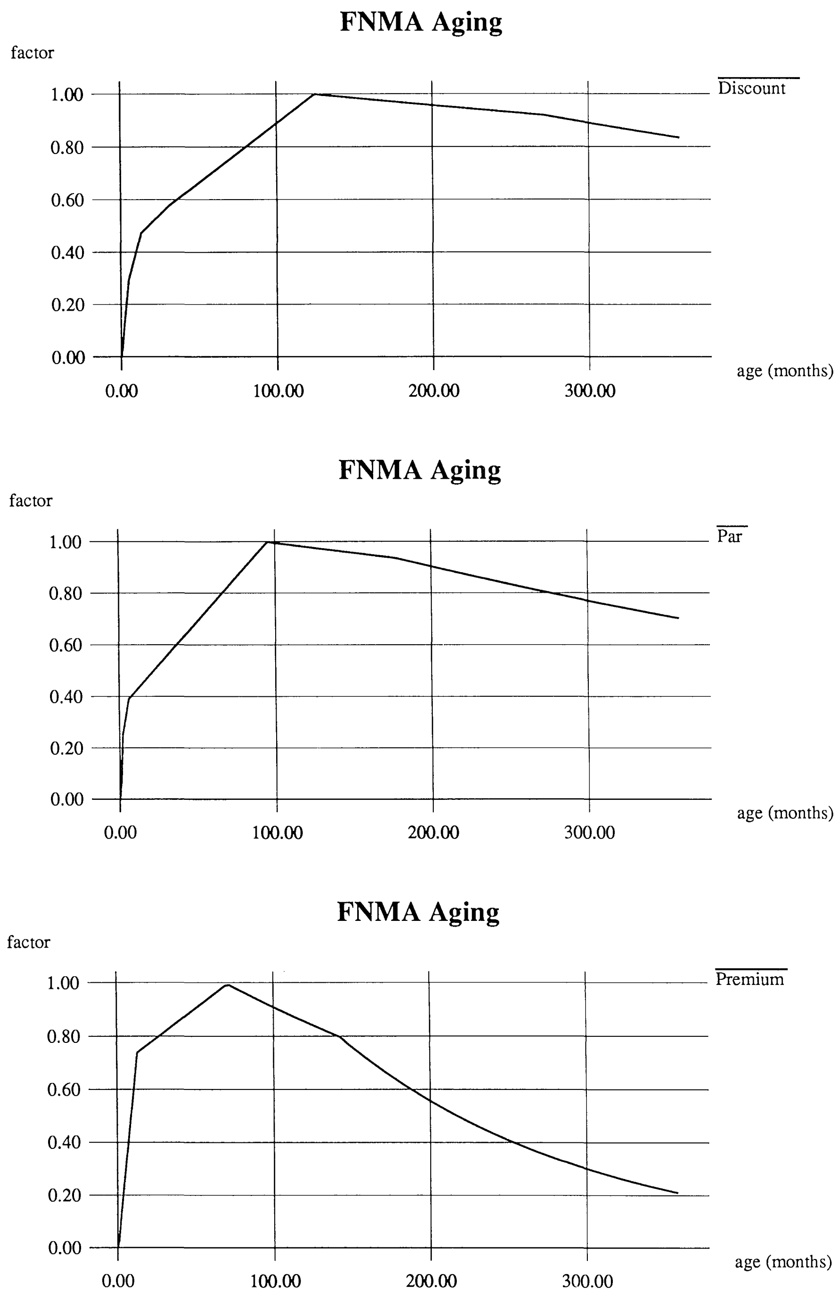
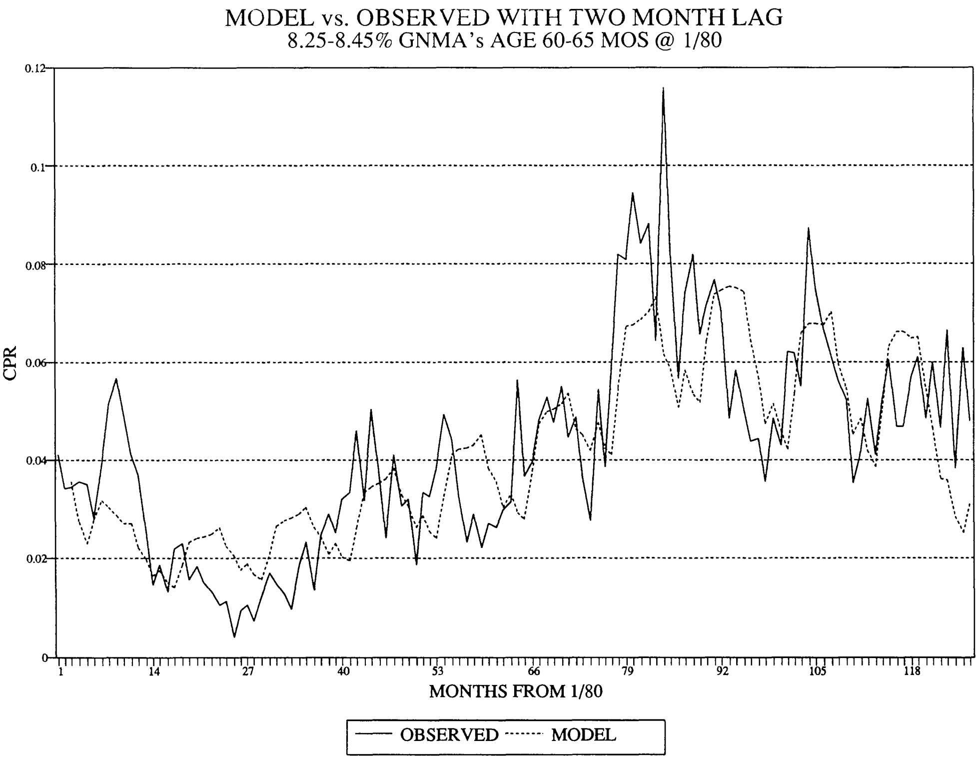


FIGURE 12. Aging effect of FNMA securities.

14. If interest rates were to drop by 200 basis points, however, the prepayment activity would rise more quickly and reach a higher level, and would also burn out sooner than in the constant interest rate environment. Further analysis and validation of the model are reported by Golub and Pohlman (1992).

6. Conclusions

We have presented in this study complete models for estimating the prepayment rates of mortgage-backed securities. Empirical observations for several hundred thousand securities over an eight-year period have been used to calibrate the model. Individual



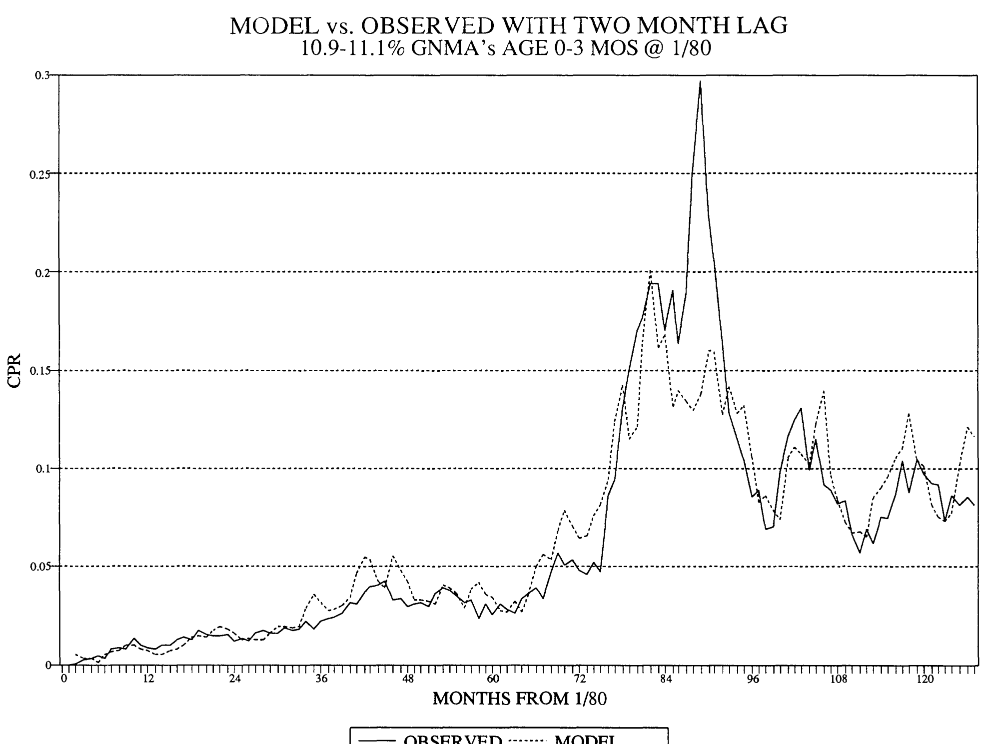


FIGURE 13. Validating the model: Observed and model predicted CPR for GNMA securities of different coupon rates and maturities.

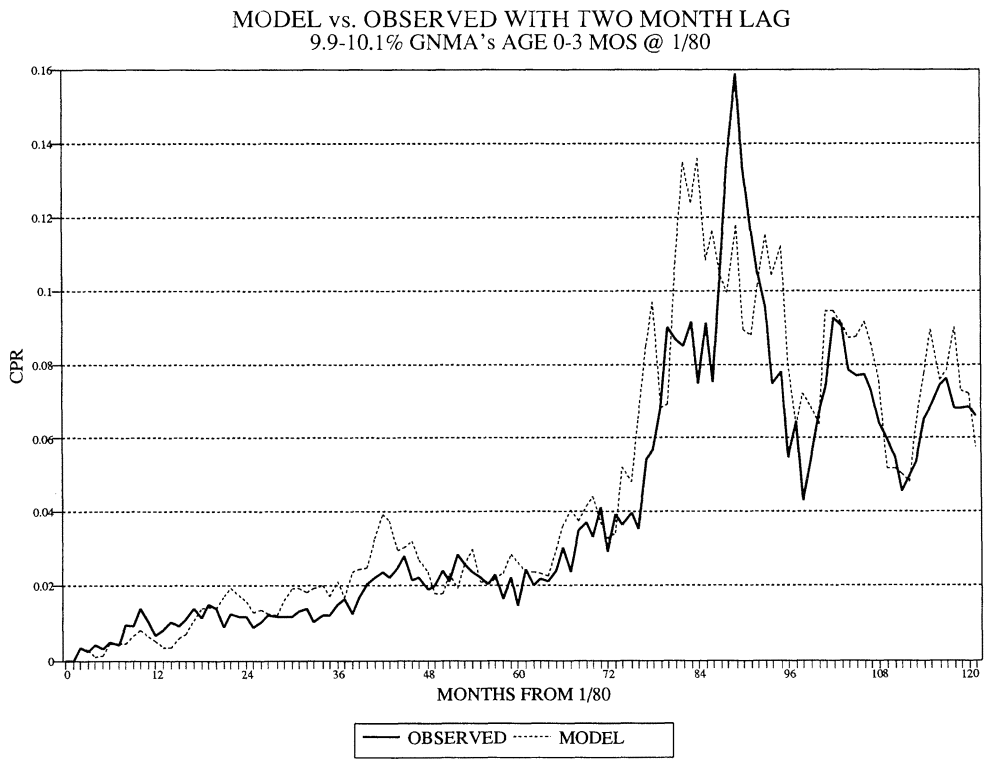


FIGURE 13 (*continued*).

components of the model fit the observed data extremely well: a manifestation of the flexibility of the basis functions we have used, departing from current practices. The overall model has also high explanatory power in estimating the prepayment activity of generic pools.

However, the model can not yet discriminate among differences in the prepayment activity of different securities that may belong in the same coupon/age cohort. This is an area where further work is needed. In particular, the current prepayment model can be used as a discriminant function within each cohort to separate securities that pay at higher speed than the model, from those that pay at a lower speed. The model itself, besides its many current uses, can serve as the basis for more detailed studies. We are now using this model in a joint project with the Blackstone Financial Management and the Federal National Mortgage Association to develop optimal portfolios of mortgage-backed securities. The model developed here, together with recent advances in parallel computing methodologies (Hutchinson and Zenios 1991), brings the search for interest-rate invariant portfolios within reach. A first attempt in this direction is reported in Zenios (1991).

7. Postscript

Since this paper was first completed B. Golub and L. Pohlman of BlackRock Financial Management conducted an empirical analysis of our model and compared it with the market prepayment projections as reported by the Bloomberg services, and a proprietary model developed by a Wall Street firm. One important shortcoming of our model, that emerged from this analysis, was that the use of C/R as a proxy for the ratio V_t/B_t is a poor approximation for mature mortgages. Changes in refinancing rates have an insignificant impact on the prepayment activity of a loan that approaches maturity. Following their analysis the model was recalibrated using V_t/B_t as the independent variable that

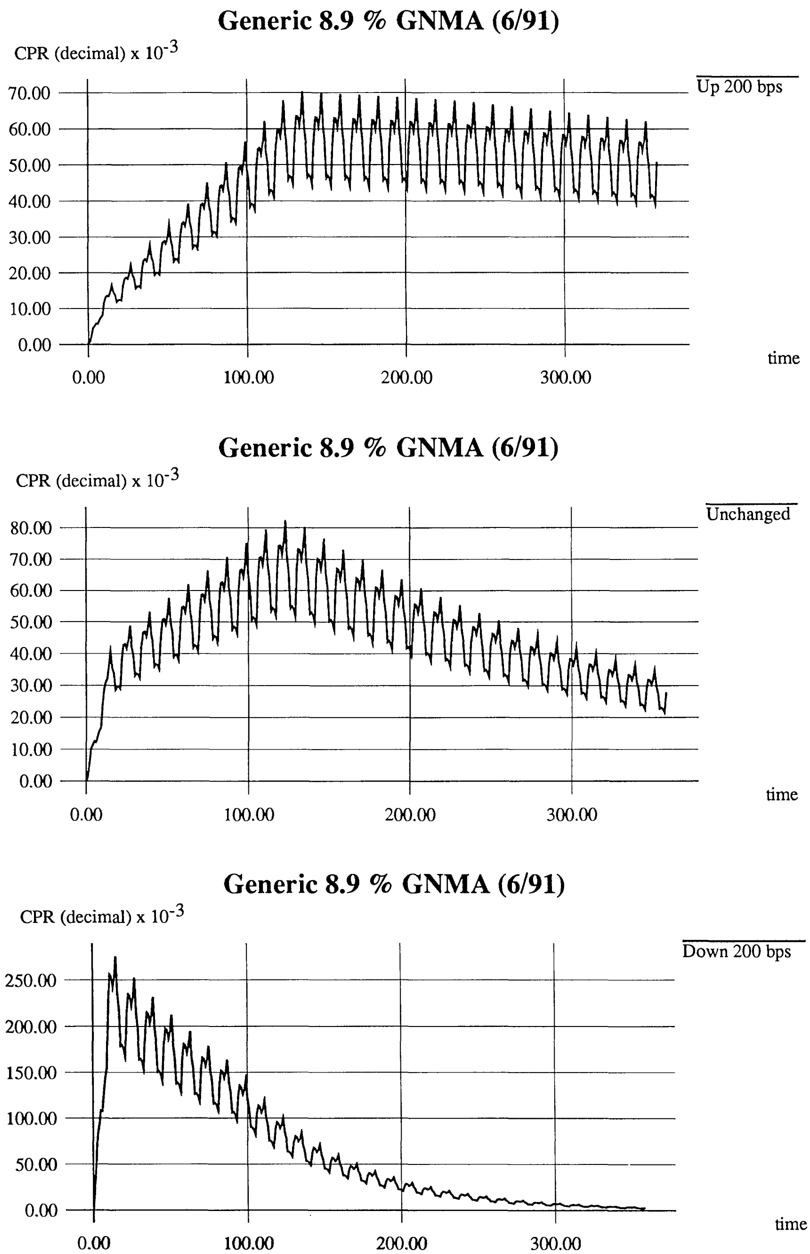


FIGURE 14. Projected prepayment activity of GNMA-I 8.9 generic pool under different interest rate projections.

determines the refinancing incentives. We found this recalibration procedure to yield much better results than those already reported here. The underlying modeling methodology and calibration procedures described in this report remain, nevertheless, unchanged. Additional details on the testing and validation of our prepayment model are given in the paper by Golub and Pohlman.[1]

[1] The work reported by Kang and Zenios was funded by the Decision Risk and Management Science Program of the National Science Foundation on University/Industry collaboration (SES-91-00216) with matching funds from Union Bank of Switzerland (UBS). The HERMES Laboratory for Financial Modeling and Simulation was established by an equipment grant from Digital Equipment Corporation.

The authors have benefited from the interaction with the Mortgage Group of UBS Securities and with Black Rock Financial Management (BFM). Discussions with Mr. Howard Hill (UBS), and Drs. Ben Golub and Larry Pohlman (BFM) are gratefully acknowledged. We would also like to thank Dr. Christian Gabathuler and Mr. Hossur Shrikantan from UBS (New York), and Prof. Dr. Rudolph Marti of the UBS Laboratory (Zurich). Financial Publishing Company, Boston, Massachusetts, kindly provided us with their mortgage database. Two anonymous referees provided insightful comments on the first draft of this paper. Their suggestions led to substantial improvements. We also benefited from the comments of J. Guttentag. The confidence interval analysis is due to A. Krieger whose suggestions we acknowledge without implicating him.

Appendix A. Mathematical Model of Seasonal Variations

Let $\mathrm{CPR}_m(i, C/R)$ be the observed factor for pool i within a given C/R range, in month $m = 1, 2, 3, \ldots, 12$. Let $i = 1, 2, 3, \ldots, n_m$, (i.e., there are n_m observations for month m). Let $w_m(i)$ be the outstanding balance for the ith observation during month m. The seasonal adjustment for month m is computed by

$$\hat{s}(m, C/R) = \frac{\sum_{i=1}^{n_m} w_m(i)\mathrm{CPR}_m(i, C/R)}{\sum_{\tau=1}^{12} \sum_{i=1}^{n_\tau} w_\tau(i)\mathrm{CPR}_\tau(i, C/R)}. \tag{8}$$

Appendix B. Mathematical Model of the Refinancing Incentive

We define the following:

y_j for $j = 0, 1, 2, 3, \ldots, N$ are point values within the range of C/R we want to model. This is the independent variable in modeling the purely economic incentive to refinance. Typically, $y_0 = 0.5$ and $y_N = 1.6$, and $N = 22$.

$\hat{\mathrm{CPR}}_i(x_j)$ for $j = 1, 2, 3, \ldots, N$ and $i = 1, 2, 3, \ldots, n_j$ indicates the ith observation of the seasonally adjusted prepayment rate for a given C/R in the range $[y_{j-1}, y_j]$. For each value of C/R in $[y_{j-1}, y_j]$, there are n_j such observations available from historical data.

w_{ij} for $i = 1, 2, 3, \ldots, n_j$, is the outstanding balance of the ith observation in the range $[y_{j-1}, y_j]$.

$g_j(x)$ is a family of shifted, ramp-like basis functions defined by:

$$g_j(x) = \begin{cases} 0, & x < y_j, \\ (x - y_j)/\delta & y_j \le x < y_{j+1}, \\ 1, & x \ge y_{j+1}, \end{cases} \tag{9}$$

where δ is the step size defined by $\delta = (y_N - y_0)/N$. The model for estimating the refinancing function then becomes:

$$\text{minimize} \quad \Phi(\alpha) = \sum_{j=1}^{N} \sum_{i=1}^{n_j} [w_{ij}\hat{\mathrm{CPR}}_i(x_j) - G(x)]^2 \tag{10}$$

subject to:

$$G(x) = \alpha_0 + \sum_{j=1}^{N} \alpha_j g_j(x), \tag{11}$$

$$\alpha_j \ge 0, \qquad \forall j = 1, 2, 3, \ldots, N. \tag{12}$$

If α^* denotes the solution of this model, then $G^*(x) = \alpha_0^* + \sum_{j=1}^{N} \alpha_j^* g_j(x)$ provides a first estimate of the refinancing incentive.

Appendix C. Mathematical Model of the Seasoning Effect

We define the following:

j for $j = 1, 2, 3, \ldots, N$ is an index for time period t_j during which seasoning is observed, and also for the basis function used in representing seasoning at this time period. Typically, we use 120 monthly intervals for discount bonds that season slowly. For premium bonds we have found 90 monthly intervals to work equally well.

$S_i(t_j)$ for $j = 1, 2, 3, \ldots, N$ and $i = 1, 2, 3, \ldots, n_j$ indicates the ith observation of the seasoning factor observed at some time instance that falls in the time period t_j. For each time intervale t_j there are n_j such observations available from historical data.

w_{ij} for $i = 1, 2, 3, \ldots, n_j$ is the outstanding balance of the ith observation for a given time intervals t_j. $g_j(t)$ is a family of ramp-like basis functions defined by:

$$g_j(t) = \begin{cases} 0, & t \leq 0, \\ t/t_j & 0 \leq t \leq t_j, \\ 1, & t \geq t_j. \end{cases} \tag{13}$$

The model for estimating the seasoning effect is:

$$\text{minimize} \quad \Phi(\alpha) = \sum_{j=1}^{N} \sum_{i=1}^{n_j} [w_{ij} S_i(t_j) - G(t)]^2 \tag{14}$$

subject to:

$$G(t) = \sum_{j=1}^{N} \alpha_j g_j(t), \tag{15}$$

$$\sum_{j=1}^{N} \alpha_j = 1, \tag{16}$$

$$\alpha_j \geq 0, \quad \forall j = 1, 2, 3, \ldots, N. \tag{17}$$

If α^* denotes the solution of this model, then $G^*(t) = \alpha_0^* + \sum_{j=1}^{N} \alpha_j^* g_j(t)$ provides the estimate of the seasoning effect.

Appendix D. Mathematical Model of the Burnout Effect

We use a set of negative-exponential basis function defined as

$$g_j(t) = \begin{cases} 1, & t \leq t_j, \\ \exp\{-\beta(t - t_j)\}, & t > t_j. \end{cases} \tag{18}$$

The least squares model for estimating the burnout effect is similar to the model of the seasoning effect. It can be defined as

$$\text{minimize} \quad \Phi(\alpha, \beta) = \sum_{j=1}^{N} \sum_{i=1}^{n_j} [w_{ij} S_i(t_j) - G(t)]^2 \tag{19}$$

subject to:

$$G(t) = \sum_{j=1}^{N} \alpha_j g_j(t), \tag{20}$$

$$\sum_{j=1}^{N} \alpha_j = 1, \tag{21}$$

$$\alpha j \geq 0, \quad \forall j = 1, 2, 3, \ldots, N, \tag{22}$$

$$\beta \geq 0. \tag{23}$$

References

BARTLETT, W. W., *Mortgage-Backed Securities: Products, Analysis and Trading*, New York Institute of Finance, 1989.

BROOKE, A., D. KENDRICK AND A. MEERAUS, *GAMS: A User's Guide*, The Scientific Press, 1988.

CARRON, A., E. FIRESTONE, R. GERBER, G. PATRUNO AND W. WELCH, *Prepayment Models for Fixed and Adjustable Rate Mortgages*, Technical Report, Fixed Income Research Report, First Boston, August 1988.

CURLEY, C. AND J. GUTTENTAG, "The Yield on Insured Residential Mortgages," *Explorations in Economic Research* (Summer 1974), 114–161.

FABOZZI, FRANK J. (Ed.), *The Handbook of Mortgage-Backed Securities*, Probus Publishing Company, Chicago, IL, 1989.

GOLUB, B. AND L. POHLMAN, "Mortgage Prepayments and an Analysis of the Wharton Prepayment Model," *Interfaces* (to appear).

GREEN, JERRY AND JOHN B. SHOVEN, "The Effects of Interest Rates on Mortgage Payments," *Journal of Money, Credit, and Banking* (February 1986), 41–59.

HAYRE, LAKHBIR S., KENNETH LAUTERBACH AND CYRUS MOHEBBI, *Prepayment Models and Methodologies: The Prudential-Bache Prepayment Model*, Technical Report, Prudential-Bache Financial Strategies Group, 1988.

HUTCHINSON, J. M. AND S. A. ZENIOS, "Financial Simulations on a Massively Parallel Connection Machine," *International Journal of Supercomputer Applications*, 5, 2 (1991), 27–45.

JACOB, DAVID P., GRAHAM LORD AND JAMES A. TILLEY, "A Generalized Framework for Pricing Contingent Cash Flows," *Financial Management* (Autumn 1987), 5–14.

NAVRATIL, FRANK J., "The Estimation of Mortgage Prepayment Rates," *Journal of Financial Research*, 8, 2 (Summer 1985), 107–117.

PETERS, HELEN F., "Termination Distributions of FHA Insured Residential Mortgages," Ph.D. thesis, The Wharton School, University of Pennsylvania, 1979.

RICHARD, SCOTT F. AND RICHARD ROLL, "Prepayments on Fixed-rate Mortgage-backed Securities," *Journal of Portfolio Management* (Spring 1989), 73–82.

ZENIOS, S. A., "Massively Parallel Computations for Financial Modeling under Uncertainty," In J. P. Mesirov (Ed.), *Very Large Scale Computing in the 21st Century*, SIAM, Philadelphia, PA, 1991, 273–294.

LINKED CITATIONS

- Page 1 of 1 -

You have printed the following article:

Complete Prepayment Models for Mortgage-Backed Securities
Pan Kang; Stavros A. Zenios
Management Science, Vol. 38, No. 11, Focused Issue on Financial Modeling. (Nov., 1992), pp. 1665-1685.
Stable URL:
http://links.jstor.org/sici?sici=0025-1909%28199211%2938%3A11%3C1665%3ACPMFMS%3E2.0.CO%3B2-%23

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

The Effects of Interest Rates on Mortgage Prepayments
Jerry Green; John B. Shoven
Journal of Money, Credit and Banking, Vol. 18, No. 1. (Feb., 1986), pp. 41-59.
Stable URL:
http://links.jstor.org/sici?sici=0022-2879%28198602%2918%3A1%3C41%3ATEOIRO%3E2.0.CO%3B2-X