

Housing Prices and the (In)stability of Mortgage Prepayment Models:

Evidence from California

Joe Matthey and Nancy Wallace¹

July 27, 1998

¹Matthey is a research officer at the Federal Reserve Bank of San Francisco. Wallace is Associate Professor at the Haas School of Business, U.C. Berkeley and a visiting scholar at the Federal Reserve Bank of San Francisco. This paper presents the authors' own views, not those of the Federal Reserve System.

Abstract

Most empirical models of mortgage terminations emphasize refinancing incentives related to interest rate movements. We consider three sources of risk that lead to observed mortgage payment terminations: interest-rate related refinancing, default, and moving. We estimate models that identify the relative importance of regional risk factors leading to termination, using data on mortgage and housing market activity in fifteen California counties from 1992 through 1996. As expected, we find that the time-series dynamics of interest rates and house prices are important determinants of the exercise of the refinancing and default options across regional markets. We also find that mobility effects differ significantly across regions and have an appreciable effect on overall mortgage termination activity. Our results suggest that standard methods of mortgage-backed-securities valuation could be improved by explicitly modeling the dynamics of housing prices and by modeling how house prices affect mortgage terminations.

1 Introduction

The market for securities backed by fixed-rate residential mortgage debt is large and growing. Proper valuation of these securities hinges on the ability to model mortgage termination risk. Most empirical models of mortgage terminations emphasize refinancing incentives related to interest rate movements. Recent contributions to the literature also have modeled defaults related to housing price declines and individual financial distress. Housing price declines also can be a barrier to homeowner mobility. However, even though a substantial portion of mortgage terminations are due to home sales by households who choose to move, not to refinance or to default, empirical mortgage prepayment models generally do not incorporate information about how housing prices and other factors affect home sales volumes.

The purpose of this paper is to empirically test for the importance of regional housing price effects in determining observed heterogeneity in mortgage terminations. We consider three sources of risk that lead to observed mortgage payment terminations: interest-rate related refinancing, default, and moving. We develop and estimate models that are intended to identify the relative importance of regional risk factors leading to termination. The data set for this analysis reflects mortgage and housing market activity in fifteen California counties from 1992 through 1996. As expected, we find that the time-series dynamics of interest rates and house prices are important determinants of the exercise of the refinancing and default options across regional markets. We also find that mobility has a strong regional component which is related to house prices; the regional dynamics of mobility turn out to be an important offset to the effect of defaults on total terminations. On the whole, our results suggest that standard methods of mortgage-backed-securities valuation could be improved by explicitly modeling the dynamics of housing prices and

by modeling how house prices affect mortgage terminations.

The paper is organized into five additional sections and some concluding remarks. In the following section, we consider the importance of exogenous mortgage termination and heterogeneous refinancing transactions costs and decision frequencies in both one- and two-factor rational models of mortgage valuation. We argue that by treating exogenous mortgage termination and differences in refinancing determinants as fixed effects which are not necessarily related to regional economic conditions, both types of valuation models fail to account for all the heterogeneity observed in payment terminations across pools of mortgages or mortgage-backed securities. We conclude the section with a discussion of the remaining empirical issues in accurately modeling mortgage terminations. In Section 3, we specify and estimate empirical models of mortgage defaults, home sales, and refinancing; these empirical models explain much of the regional variation in mortgage terminations as a correlate of the path of housing prices, not as fixed effects. Section 4 compares the implications of these estimated models for mortgage termination probabilities. An estimated dynamic model of housing prices implies a time-path for each region's probabilities of mortgage termination, given a particular interest rate environment. In Section 5, we illustrate some of what this model of geographic variations in mortgage termination probabilities implies for mortgage pool valuations. Our concluding remarks suggest directions for future research on how methods of mortgage-backed-securities valuation could be improved by explicitly modeling the role of housing price dynamics.

2 Approaches to Modeling Mortgage Terminations

2.1 The Rational Valuation Paradigm

Much of the academic mortgage valuation literature models mortgage terminations as a function of optimal call policies (see Dunn and McConnell, 1981; Brennan and Schwartz, 1985) or optimal call policies with transaction costs (See Stanton, 1995, 1996; Timmis, 1985; Schwartz and Torous, 1989; 1993). In addition, most of the early models consider the mortgage-refinancing termination option in isolation from default, with interest rates as a sole stochastic factor determining the value of this option.

For example, in the one-factor model framework of Stanton (1996), which is a generalization of Dunn and McConnell (1981), termination predictions can be written as:

$$(1) \quad \pi_{it} = \lambda + \rho I(r_t \leq r_{it}^*) + v_{it}$$

In this particular model, the hazard function governing termination takes on the value λ if $r_t > r_{it}^*$ and the value $\lambda + \rho$ if $r_t \leq r_{it}^*$, where r_t is the risk free rate and r_{it}^* is a time- and state- dependent threshold value where the call is optimally exercised. Here, π_{it} is the predicted termination rate for mortgage i at time t , and λ is the mean Poisson arrival rate of exogenous termination. I is an indicator variable for whether the spot interest rate, r_t , has fallen below the critical level, r_{it}^* , at which refinancing becomes optimal, and v_{it} is the error term. The parameter ρ governs the frequency with which refinancing decisions are made. The critical interest rate r_{it}^* depends on the expected future evolution of interest rates, the parameters (ρ, λ) , and the level of transactions costs faced by this individual.

Stanton's (1996) approach to modeling mortgage terminations at the more aggregate level of a mortgage

pool is to assume that the function which specifies how individual mortgage holders react to evolving interest rates, at a given transactions cost, is common across mortgage pools, but the distributions of transactions costs themselves differ across pools. In particular, the distributions of transactions costs across individuals within the k th pool are assumed to be given by a Beta probability distribution with parameters α and β_k , so that the mean transactions cost is $\alpha/(\alpha + \beta_k)$ of the outstanding mortgage balance. Overlaid on this is an additional “decision-frequency” impediment to refinancing; with probability $(1 - \rho)$, the decision to refinance is not even considered, and with probability ρ the optimal exercise rule is followed. Since neither ρ nor λ differ across pools in Stanton (1996), this model incorporates heterogeneity in termination predictions only through differences in the critical interest rates, r_{kt}^* .

In this paper, we are interested in the related problem of valuing groups of mortgages which are collateralized by houses in a known geographic location, which we index by the subscript j . The value V_{jt} to a lender at time t of the mortgages in group j are the expected present discounted value of the cash flows to be received, $X_{j,t+\tau}$, between the present and the termination date of the mortgages, T . If time is discretized, this present value can be written as :

$$V_{jt} = E_t \sum_{\tau=0}^{T-t} \beta_{j,[t,t+\tau]} X_{j,t+\tau} .$$

The discount factors $\beta_{j,[t,t+\tau]}$ should be treated as stochastic functions which depend on the evolution of interest rates and on the price of interest rate risk. If the Poisson process governing additional terminations is independent of the interest rates, then the present value relation may be written in the one-factor model

context as:

$$\begin{aligned}
(2) \quad V_{jt} = & \sum_{\tau=0}^{T-t} [\lambda X_{j,t+\tau}^p E_t \beta_{j,[t,t+\tau]} + (1-\lambda) X_{j,t+\tau}^p \rho E_t \beta_{j,[t,t+\tau]} I(r_{t+\tau} \leq r_{j,t+\tau}^*) \\
& + (1-\lambda) X_{j,t+\tau}^s (1-\rho) E_t \beta_{j,[t,t+\tau]} I(r_{t+\tau} \leq r_{j,t+\tau}^*) \\
& + (1-\lambda) X_{j,t+\tau}^s \rho E_t \beta_{j,[t,t+\tau]} (1 - I(r_{t+\tau} \leq r_{j,t+\tau}^*))]
\end{aligned}$$

Here, X^p denotes the cash flow upon prepayment and X^s denote the cash flow for a scheduled payment, amounts which we assume are known in advance¹. The first line of equation (2) shows two reasons for cash flows at the prepayment level, X^p : an exogenous prepayment or an interest-rate related prepayment. The second and third lines show two reasons for cash flows remaining at the scheduled level, X^s , when no exogenous prepayment was triggered: failure to exercise an optimal call or failure of the interest rate to be in the exercise region.

Researchers such as Dunn and McConnell (1981) and Stanton (1996) consider the continuous-time counterpart to this valuation problem with a particular assumption about how the term structure of interest rates evolves and use methods for solving partial differential equations to solve for V_{jt} . In this solution method, the critical interest rates $r_{j,t+\tau}^*$ that define the optimal exercise regions for prepayment are “calculated” at least implicitly.

A similar valuation paradigm also motivates MBS pricing practice in the capital markets. It is common to estimate MBS passthrough valuations by Monte Carlo integration of the expected present discounted

¹We abstract from the possibility of partial prepayments.

value of cash flows. As in the rational valuation paradigm, predicted cash flows are discounted according to a shifted Treasury spot curve, with discount factors and prepayment cash flows both linked to realizations of interest rates, as in equation (2) above². However, empirical valuation models often do not impose theoretical constraints of optimizing theory on the empirical specification of the relationship between interest rates and termination behavior. Rather, simplified empirical proxies tend to be used for the refinancing incentives, with Monte Carlo simulations used to capture the covariation between the evolution of the discount factors and the interest-related prepayment responses. The freedom from the difficulties of embedding rational prepayment allows more modeling effort to be devoted to explaining time-series and cross-section variation in non-interest-related prepayments³.

2.2 Introducing Heterogeneity in One-Factor Models

Single-factor rational prepayment models that include a Poisson arrival rate λ for exogenous terminations can be generalized to let that Poisson process vary over time and over conditioning variables. In principle, the mean arrival rate of the Poisson process can depend on exogenous observables which vary over time (Dunn and McConnell (1981), Kau et. al. (1992)). Alternatively, mortgages may be stratified by some

²Limitations of the rational prepayment valuation paradigm emerge in this market context. The Treasury spot curve shift factor is called the “option-adjusted spread” (OAS). Comparisons of the implied OAS for different MBS securities are used by the market as a guide to relative value. However, the single-factor rational prepayment paradigm seems to imply that such an OAS must be due to agency credit risk, as future values on the Treasury spot curve already incorporate interest rate risk, and the theory elaborated in Dunn and McConnell (1981), for example, implies that this interest rate risk is the only factor to be priced in the absence of systematic credit risk. In practice, the implied OASs for agency passthroughs tend to be much larger than could be reasonably ascribed to agency credit risk, suggesting that the market price of prepayment risk is not just interest rate risk (Cheyette (1996)) but includes additional unaccounted for factors.

³For example, the Wall Street models of MBS prepayment activity of Patruno (1994) and Hayre and Rajan (1995) include submodels for separate, unobservable pieces of aggregate prepayment activity, such as home relocation proxies, refinancing, and proxies for default.

time-invariant observable, such as their pool affiliation or geographic location, and the variation in λ can be modeled as a fixed effect.

In this paper, we follow this latter fixed-effects route for incorporating heterogeneity in the single-factor model. We modify Stanton (1996) by assuming no cross-section or time-series variation in the Beta transactions costs distribution parameters α and β , but we allow for variation across regions in the exogenous prepayment parameter λ_j . Furthermore, when we want to introduce heterogeneity in refinancing incentives in the single-factor model, we let the decision-frequency parameter ρ_j take a geographic subscript (j) and vary across place. This variant of the single-factor model can be written as

$$(3) \quad \pi_{ijt} = \lambda_j + \rho_j I(r_t \leq r_{ijt}^*) + v_{ijt}$$

for an individual mortgage i within group j . For the group as a whole, let $F_j(r_t \leq r_t^*)$ denote the proportion of individuals for whom transactions costs are low enough for refinancing to be optimal. In principle, this distribution function F_j can differ across groups j , reflecting differences in underlying transactions costs distributions. However, we assume common transactions costs distributions and write the group-level single-factor model as

$$(4) \quad \pi_{jt} = \lambda_j + \rho_j F(r_t \leq r_t^*) + v_{jt}$$

2.3 Geographic Factors and Two-Factor Models

Although the bulk of empirical mortgage termination research is based on one-factor models with interest rates as the underlying stochastic process, the more theoretical work of Kau and his co-authors (Kau et al., 1992; 1994), summarized in Kau and Keenan (1995), introduces a geographic dimension by emphasizing

the evolution of house values as a determinant of default. In these models, default is a competing risk to other sources of termination, such as refinancing. Generalizing the above, we have

$$(5) \quad \pi_{jt} = \lambda_j + \rho_j F_R(r_t \leq r_t^*) + F_D(h_{jt} \leq h_{jt}^*) + v_{jt}$$

Here, as before, F_R is a distribution function for the proportion of mortgage holders for whom the spot interest rate has fallen below the critical level, r_t^* , at which refinancing becomes optimal. Similarly, F_D is a distribution function for the proportion of mortgage holders for whom the house price, h_{jt} , has fallen below the critical level, h_{jt}^* , at which default becomes optimal.⁴ In this case, λ_{jt} would be the exogenous termination probability due to the relocation behavior of homeowners who have not defaulted.

The market mortgage valuation paradigm is centered on the one-state-variable model driven off a stochastic interest rate. However, we also introduce the two-state variable model with stochastic interest rates and house prices, equation (5), because it formalizes one channel for house prices to affect terminations (through defaults) and allows us to distinguish between three different types of mortgage termination: defaults, refinancing, and mobility-related terminations. In the following empirical work, we will generalize this two-factor model and allow house prices to affect refinancing and mobility-related terminations⁵.

⁴In general, both critical values, r_t^* and h_{jt}^* , depend on both the interest rate and house price state variables, so, strictly speaking, there are critical regions (in interest rates and house prices), not critical values (in a single factor), determining the exercise regions. However, for our heuristic purposes, it is useful to push this theoretical dependence on the other factor into the background.

⁵As with other empirical models of default (Quigley and Van Order, 1995) or of the competing risks of prepayment and default (Deng et al., 1996; Quigley, et al., 1994), we let the theory of rational prepayment and default guide model specification but do not fully impose the theoretical constraints of rational option valuation on the specifications.

2.4 Limitations of the Findings in the Extant Literature

Empirical termination model specifications usually are constrained by unavailability of data. The most serious dataset constraint in previous studies is the lack of separate identification of mortgage terminations due to refinancing, mobility, and defaults. Stanton (1995) and others who have modeled prepayment activity for MBS pools work with data in which all three of these sources of terminations—defaults, refinancings, and purchases—appear as an aggregate rate of mortgage termination⁶. Researchers who study individual mortgage loans also usually have not had complete identification of the types of terminations. For example, Schwartz and Torous (1993) used an individual loan dataset in which defaults and prepayments are separately identified, but no distinction was made between prepayments due to refinancing and prepayments due to other reasons, such as sale of the home or household portfolio rebalancing. Archer, Ling, and McGill (1996; 1997) studied the behavior of individual housing units tracked through the American Housing Survey (AHS), a dataset which did not permit them to track the activity of households who prepaid in order to move. Many datasets also have lacked the geographic information needed to model the role of local housing market conditions in mortgage terminations. For example, MBS market participants seldom know the evolution of the value of the houses which collateralize the mortgages, because the location of the houses is not disclosed in sufficient detail.

Despite these limitations, substantial evidence suggests that geographic housing market factors significantly affect refinancing and mobility-related terminations, in addition to defaults. For example, for

⁶With pool-level data that does not distinguish relocations from refinancings, switching-regime models offer a promising means of identifying the separate parameters governing these types of prepayment behavior; see, for example, Kau and Springer (1992).

mortgage pools, Beckett and Morris (1990) use the location of the operations of the originator-servicer as a proxy for the geographic location of the collateral and document substantial variation in overall prepayment speeds by state from 1982 to 1988. Evidence from the early 1990s also has corroborated the importance of geographic factors to prepayments. As noted by Monsen (1992) and demonstrated formally by Caplin et al. (1993), home prices declined in much of the Northeast over the 1990-92 period, and the reduction in collateral depressed prepayment activity there relative to other states. Caplin et al. (1993) attribute this depressed level of prepayment in the Northeast to lower refinancing activity, but their data does not allow them to actually distinguish between prepayments related to refinancing and prepayments related to home purchases. Using loan-level data on refinancings, Peristiani et al. (1996, 1997) were able to document a large effect of low home equity on the propensity to refinance, but they were not able to address the issue of how much regional economic conditions affect the propensity to prepay for other reasons, such as home purchases. Separately, using another dataset, Engelhardt (1998) shows that house prices have had a strong effect on household mobility within metropolitan areas.

2.5 The Importance of Parameterization to Valuation

Incorporating heterogeneity via group-specific rates of exogenous termination (λ_j), decision-frequency (ρ_j), or transactions costs (α, β_j) can be important to mortgage valuation. Over a moderately sized range for actual values of λ_j , the effect on mortgage value can be economically significant, depending on the mortgage contract (coupon) rate and the slope of the term structure⁷. Second, implied mortgage values also are quite

⁷See, for example, Dunn and McConnell (1981) and Archer and Ling (1993).

sensitive to the assumed levels of transactions costs and decision-frequency; high transactions costs (or low decision-frequency) reduce the value to the borrower of the option to refinance and hence increase the value of the mortgage to the lender.

In principle, termination heterogeneity may arise either from the characteristics of the individual mortgage holders or from the characteristics of the regional housing and labor markets in which they participate. Termination models which emphasize individual mortgage holder characteristics are costly to develop and maintain; we believe that for the purpose of MBS valuation, this modeling effort is better directed to understanding the regional component of prepayment risk. Also, the optimal portfolio diversification strategies for pools of mortgages and mortgage-backed securities differ depending on whether the underlying heterogeneity is individual in nature or common within a regional market. Accordingly, it is important to determine whether regional sources of heterogeneity account for a significant amount of the cross-sectional variance in observed termination behavior⁸.

In this paper, our primary empirical objective is to marshal evidence on how to parameterize mortgage pool valuation models. We examine the empirical frequencies of defaults, mobility-related mortgage terminations, and refinancings across a range of geographic areas (California counties) which have experienced quite different housing price environments. We consider both one-factor and two-factor specifications for mortgage termination risk with the objective of identifying the primary time-series and cross-section determinants of this risk. In most specifications we allow some of the cross-section determinants of terminations

⁸Other recent research, such as Archer, Ling, and McGill (1997), also recognizes the need to identify the relative importance of geographic-specific (versus individual-specific) factors in mortgage terminations. We extend this line of work in several dimensions, including considering whether the spatial correlation in rates of exogenous prepayment is likely to persist for relatively long periods of time.

to remain static over time, but in our preferred (unrestricted two-factor) specification we allow all three types of mortgage termination probabilities to evolve along projected paths for housing prices in each county.

To tie these results on termination frequencies back to the valuation literature, we calculate the effects on the value of county-level mortgage pools of differing parameterizations for λ_{jt} and ρ_{jt} , using a standard one-factor valuation procedure. In one specification, we take ρ to be common across areas and λ_j to be fixed effects estimated from total termination frequencies. Other specifications take advantage of data on terminations by type and progressively drop restrictions on housing prices as determinants of the evolution of defaults, refinancing, and mobility.

3 Modeling Loan Purchase Originations, Refinancing, and Defaults

We now turn to the task of building empirical models of mortgage defaults, home sales, and refinancing, so that we can quantify the benefits of incorporating geographic factors in mortgage pool valuation models. This section provides a brief introduction to our data and discusses model specification issues and basic empirical results. The appendix describes the specifics of our dataset construction.

In brief, we transform data into measures of the frequency of mortgage termination for all reasons, π_{jt} , and decompose this into three types of termination:

$$(6) \quad \pi_{jt} = \pi_{jt}^D + \pi_{jt}^M + \pi_{jt}^R.$$

Here, π_{jt}^D is the frequency of default, π_{jt}^M is the frequency of paying off a mortgage to move, and π_{jt}^R is

the frequency of prepaying without moving, which we take to be refinancing. We let the index j denote a county, and we focus on activity in fifteen of California's largest counties, a group which currently accounts for more than three-quarters of California's total population.⁹

The time-series pattern of refinancing activity in California was similar to that in the overall U.S. in the 1992-96 period which we study (Chart 1)¹⁰. Refinancing activity in California peaked at about 950,000 loans in 1993, when the mortgage interest rate on conventional thirty-year loans fell to a low of about 7-1/4 percent. In 1994, when the conventional mortgage rate increased about 100 basis points, refinancing activity dropped substantially. In 1995 and 1996, as mortgage rates drifted down only slightly, refinancing activity remained relatively low.

Mortgage defaults in California increased from a low of a little over 12,000 in 1990 to about 92,000 in 1996, as the level of home prices in the state continued to fall throughout this period (Chart 2). California single-family home sales volumes, the primary measure we use to estimate mobility-related terminations, remained near 300,000 per year throughout most of the 1992-96 sample period, with slight increases in 1994 and 1996 to about 330,000 purchase-related moves (not shown). Our method of converting corresponding county-level statistics on refinancing, defaults, and home sales volumes into estimates of the frequency of mortgage termination is explained in the data appendix.

⁹The fifteen counties include nine from the San Francisco Bay Area (Alameda, Contra Costa, Marin, San Francisco, San Mateo, Santa Clara, Santa Cruz, Solano, and Sonoma), five from the Los Angeles Area (Los Angeles, Orange, Riverside, San Bernardino, and Ventura), and Sacramento.

¹⁰The index of refinancing in California shown in Chart 1 is from our HMDA data source, and the U.S. refinancing index is from the Mortgage Bankers Association.

3.1 Estimates of the Single-Factor Model

Turning to estimation, we start by describing the fit of a simplified form of the single-factor model of mortgage termination. As discussed above, the basic rational prepayment mortgage valuation theory assumes that the observed frequency of prepayment for a group of mortgages j may be written as $\pi_{jt} = \lambda_j + \rho_j F(r_t \leq r_t^*) + v_{jt}$. Although in general the decision-frequency parameters ρ_j might differ across groups j , we first consider shutting down this source of variation by assuming $\rho_j = \rho$ for all j . In this case, all of the cross-section variation in mortgage termination probability derives from default and moving behavior, whereas all of the time-series variation derives from refinancing behavior. Accordingly, we can estimate the single-factor model by a fixed effects method, with county-effects picking up the cross-sectional variation in λ_j and year-effects picking up the time-series variation in $\rho F(r_t \leq r_t^*)$.

Table 1 shows the results. First, when data on total terminations are used as the endogenous variable, the time-series dimension accounts for the vast majority (86.8 percent) of the full-panel variance in terminations. However, there is a wide range of mean termination frequencies across counties, from a low of below 12 percent per annum in San Bernardino to a high of almost 18 percent per annum in Marin County. Accordingly, grouping the data by county affiliation explains a lot of the cross-section variance in termination frequencies; of the 13.2 percent of the full-panel variance in terminations which is not explained by year fixed-effects, 9.7 percent of the variance is explained by introducing county-specific means. The data on termination by type reveal that the variance in refinancing rates is dominating the variance in total terminations. For refinancing, as with total terminations, most of the variance is in the time-series dimension. However, for defaults and moving, most of the variance is in the cross-section dimension.

With just data on total terminations, the unknown parameters λ_j , ρ , and $F(r_t \leq r_t^*)$ are not separately identified under the fixed effects estimation approach. Rather, we can identify how much an individual county's λ_j deviates from the average across counties of the λ_j , and we can identify how much an individual year's $F(r_t \leq r_t^*)$ deviates from the average across years of the $F(r_t \leq r_t^*)$. We achieve identification by equating the model's implication for the average rate of refinancing with the sample mean rate of refinancing, which is 9.75 percent per annum. Then, we infer estimates of λ_j from the differences between this full-sample average rate of refinancing and the county group means for total terminations shown in table 1. For example, the average rate of total terminations in San Bernardino was .1154, so the estimate of λ_j for this county is .1154-.0975=.018. The complete estimation results for λ_j using the county group means for total terminations are given in the first column of table 2. The implied λ_j range from the low of .018 for San Bernardino to a high of .081 for Marin County¹¹.

We also consider another method for resolving this identification issue, using a fixed-effects approach on the data on terminations by type. In this case, we assume that the theoretical exogenous prepayment construct λ_j is the underlying population mean of the frequency of termination for default and mobility, $\lambda_j = E[\pi_{jt}^D + \pi_{jt}^M]$, and we estimate this population mean by the sample average frequency of default and moving (calculated from the entries such as those shown in the lower rows of table 1)¹². For example, the average rate of default plus moving in San Bernardino was .0213+.0281=.0494, so the estimate of λ_j for

¹¹Note that the results shown in the first column of table 2 do not depend on whether we use year fixed-effects as the proxy for refinancing incentives or some other observable, such as an interest rate process, which varies across time but not across place.

¹²Harding (1997), is among those who have interpreted the exogenous prepayment rate λ in the single-factor model as the sum of mortgage default and prepayment for home sale.

this county rounds to .049. The complete results using this latter method are shown in column two of table 2. With estimates derived from terminations by type, the minimal λ_j is .041 for Sacramento County and the maximal λ_j is .056 for Riverside County, only a 150 basis point spread, which is narrower than the 630 basis point spread in λ_j using total terminations, shown in column one of table 2.

The narrower spread arises because the cross-sectional variations in default rates tend to be negatively correlated with the cross-sectional variations in mobility. Rates of moving for reasons other than default were low in the major counties of the Los Angeles area, particularly San Bernardino, Riverside, and Los Angeles County, and defaults were high in these areas. Major counties in the San Francisco Bay Area were at the high end of the range of non-default moving frequencies and at the low end of the range for default frequencies.

The wider spread for total terminations arises because cross-sectional differences in refinancing rates (inadvertently) spill over into these estimates of λ_j . Refinancing frequencies generally were low in the major counties of the Los Angeles area and high in the San Francisco Bay area. The total terminations estimation method “interprets” the low rates of refinancing in Los Angeles area counties as low rates of exogenous termination and “interprets” the high rates of refinancing in the San Francisco Bay area counties as high rates of exogenous termination.

With data on terminations by type, heterogeneity in refinancing rates can be incorporated in the calibration of the single-factor model. To do this, we set the cross-county average of the frequency-of-decision parameter, ρ_j , at the .483 value estimated by Stanton (1996, table 4)¹³ Notationally, we write

¹³Stanton estimates a value of .66 for a ρ^c which is the continuous-time Poisson arrival rate expressed at an annual rate;

$\rho_{.92-96} = .483$. For the single-factor model, we incorporate across-county heterogeneity in refinancing behavior by choosing $\rho_j = \rho_{j92-96}$ so that the ratio of county-specific to average decision frequency, $\rho_{j92-96}/\rho_{.92-96}$, matches the historical average relative rate of refinancing in this county, $\pi_{j92-96}^R/\pi_{.92-96}^R$. As shown in column three of table 2, the implied ρ_{j92-96} range from a low of .327 in the Los Angeles area county of San Bernardino to a high of .639 in the San Francisco area county of Marin.

3.2 Estimates of the Two-Factor Model with Fixed Effects for Mobility

In the standard rational two-factor model, which we represented as equation (5), the exogenous probability of termination, λ , is assumed to be due to home resales not related to defaults and is represented by fixed effects. As with the single-factor model, the magnitude of these fixed effects may be assumed to be common across geographical regions or to differ across geographical regions, depending on whether the modeler has the data needed to stratify geographically and chooses to do so.

Estimates of a regression function analogue to equation (5) are shown in table 3, using, alternatively, data on total terminations and data on terminations by type. We use a single proxy for the value of the refinancing prepayment option, a measure of the spread between the current mortgage rate on new issues and the contract mortgage interest rate on outstanding loans¹⁴. In our simple model, defaults are primarily determined by the fraction of homeowners in the county who have negative housing equity¹⁵. For most

the discrete-time analogue of $\rho^d=.483$ is derived from this as $\rho^d = (1 - \exp(-\rho^c))$.

¹⁴The current mortgage rate is measured by the annual average of the monthly average commitment rate on 30-year fixed rate mortgages reported in the Freddie Mac Primary Mortgage Market Survey. The contract rate on the outstanding stock is measured by the yield on mortgages in portfolio as reported in Freddie Mac Investor-Analyst reports.

¹⁵Capozza, Kazarian, and Thomson (1997) are among those who have found a correlation between default frequencies and the local incidence of trigger events such as unemployment and divorce. We discuss in a later section of this paper whether the main implications of our parsimonious model of defaults is robust to including proxies for trigger events.

counties in our sample, a peak in home prices was reached in 1990, and we include in the model the ratio of current home prices to the level of home prices in 1990.¹⁶ Finally, county fixed effects are used to pick up differences across counties in mobility rates.

Using data on the frequency of total terminations as the dependent variable, OLS estimation results (column 1, table 3) reveal a sizable and statistically significant effect of the mortgage rate spread on terminations. However, on balance, higher housing prices actually increase total terminations, not decrease them as predicted by the simple two-factor model. (This coefficient on house prices in the total terminations model is not, however, statistically distinguishable from zero.) Finally, the county fixed effects parameter estimates range from a low of -.054 in San Bernardino to a high of .004 in San Mateo, a range of about 580 basis points (columns 2-4, table 3).

Estimation of the model using data on terminations by type shows that the effect of the mortgage rate spread on refinancing is basically the same magnitude as the effect of the spread on total terminations. Also, house prices are shown to be negatively correlated with the rate of default, as implied by the two-factor model but not seen in the results using data on total terminations. Finally, the range of estimated fixed effects for mobility narrows substantially relative to the results using total terminations; however, the range still is relatively wide, from a low of .028 in San Bernardino to a high of .051 in Sonoma.

¹⁶The county-level home price index is a repeat-sales measure which was constructed by Experian.

3.3 Estimates of the Unrestricted Two-Factor Model

Refinancing incentives are primarily determined by the mortgage rate spread. However, in principle, the path of home prices also can affect the refinancing decision, particularly if house prices fall enough to give rise to a binding collateral constraint. Also, the literature on household mobility recognizes that collateral constraints are a possible impediment to terminating a mortgage in order to relocate. Furthermore, as emphasized by Stein (1995) and Mayer and Genesove (1997), the volume of sales in the market for existing homes is likely to be linked to the amount of trade-up equity available to current homeowners. Hence, we also consider models which relax the restrictions that exclude the housing price variable from the equations explaining refinancing and mobility.

In the upper panel of table 4 we show the results of estimating an unrestricted two-factor model in which the mortgage rate spread and house price variables (but not fixed effects) are used to explain terminations by type. As suggested by the broader theory, an increase in house prices tends to increase the frequency of refinancing. Also, house prices are positively correlated with mobility; the coefficient on house prices in the non-default moving equation is positive and quite distinguishable from zero.

Our measure of mortgage terminations is constructed from empirical frequencies observed for default, refinancing, and relocations. By construction, the fraction of homeowners who do not terminate their mortgages is one minus the fraction who terminate, $\pi_{jt}^N = 1 - \pi_{jt}^R - \pi_{jt}^D - \pi_{jt}^M$. In the data, these frequencies satisfy the definitional constraints of being bounded by zero and one and of summing to unity. However, the levels forms of regressions shown in table 3 and the upper panel of table 4 do not impose these boundary or adding up restrictions on the fitted values in the regressions. Accordingly, in principle, the model could

predict that total terminations (or a particular type of termination) exceeds unity or becomes negative¹⁷. To avoid this, we refine the unrestricted two-factor model by using the logarithm of odds ratios as the dependent variable in the estimating equations. That is, we use as dependent variables the three log relative termination probability variables $\ln(\pi_{jt}^R/\pi_{jt}^N)$, $\ln(\pi_{jt}^D/\pi_{jt}^N)$, and $\ln(\pi_{jt}^M/\pi_{jt}^N)$. The log relative form results shown in the lower panel of table 4 are qualitatively similar to the estimates in levels form; the mortgage rate spread is important to predicting refinancing, and house prices have a noticeable effect on all three types of termination.

4 Comparison of Model Implications for Termination Probabilities

The regression estimates of the two-factor models can be used to calibrate sequences of mortgage termination probabilities. For mortgage valuation purposes, these implied termination probabilities can be expressed in terms of values for the parameters λ_{jt} and ρ_{jt} that appear in the single-factor model, equation (4).

4.1 Effects of Elimination of Negative Equity

We consider the implications of the unrestricted two-factor model estimated in log relative form. In this model, implied termination probabilities (and hence, λ_{jt} and ρ_{jt}) are quite sensitive to the assumed level of house prices. As shown in Chart 2, California home prices declined substantially in the early 1990s.

¹⁷In practice, the boundary constraints are not violated by the fitted values of the dependent variables within sample. However, out-of-sample projections for defaults become negative under the model estimated in levels form.

To illustrate the model's implications for the dependence of termination frequencies on house prices, we first present an easily-understood scenario, the elimination of the negative equity in house prices (i.e., the return of each county house price to its 1990 level). Table 5 shows the changes in terminations implied by this perturbation of the house price explanatory variables. Among the counties, house prices declined the most in Los Angeles County, so the effects of a reversion to previous peak house price levels are largest there: total terminations in Los Angeles County are implied to jump by about 5 percentage points. The size of the implied change in termination probabilities also is large for the other Los Angeles area counties, which suffered relatively large declines in house prices.

Among the various channels for housing prices to affect mortgage terminations, the refinancing channel appears to be the strongest. For the Los Angeles area counties, the implied increase in refinancing activity generally exceeds 3 percentage points, and the model also implies that the more moderate declines in Bay area house prices held down refinancing there by between 1 and 2 percentage points per annum. The effects of a rebound in house prices on defaults and moving without default also are quite noticeable, but the largest of these effects (on moving) is only about one-third the magnitude of the effect on refinancing.

4.2 Effects along an Expected Path for Housing Prices

The scenario discussed above, that house prices in each county return exactly to 1990 levels, is not the most plausible scenario for the near-term evolution of house prices beyond the 1992-96 sample period. To help identify a more plausible scenario, we need to augment our model of mortgage terminations with a model of the evolution of housing prices in California counties. We present here a relatively simple model of

California housing prices to illustrate the basic dynamics. Also, we confirm that the dynamics of California housing prices appear to be similar to the dynamics of housing prices in other areas, which lets us draw from results in the broader literature on housing price dynamics.

To a certain extent, the time-series variation in housing prices is an asset market outcome which defies complete explanation. However, for most states and metropolitan areas, there is a clear long-run relationship between real house prices and fundamental determinants of value, such as the number of people in the area, real income levels, and the costs of adding to the stock of housing. Error-correction models such as those estimated by Abraham and Hendershott (1996) and Meese and Wallace (1994) suggest that after a boom period in which actual prices have grown to exceed fundamental values, there will tend to be periods of very intense declines in real housing prices before prices start to recover.

Our particular model of California house prices uses the error-correction form employed in the Capozza, Mack, and Mayer (1997) study of house price dynamics in sixty-two U.S. metropolitan areas. Letting Δp_{jt} denote the logarithmic change in real housing prices in area j in year t , we assume

$$(7) \quad \Delta p_{jt} = \xi \Delta p_{jt-1} + \delta(p_{jt-1} - p_{jt-1}^*) + \gamma \Delta p_{jt}^* + \epsilon_{jt}$$

The term p_{jt}^* is the underlying fundamental value of housing. Actual prices, p_{jt} , may temporarily deviate from this current fundamental value, but in the long-run actual prices tend to revert to the fundamental values. The coefficients ξ , δ , and γ and the realizations of the residual shocks ϵ_{jt} govern the speed and extent of this adjustment process.

Capozza, Mack, and Mayer (1997) find that in their annual sample of sixty-two U.S. metropolitan areas from 1979 to 1995, OLS estimation yields fitted values of $\hat{\xi} = .48$, $\hat{\delta} = -.24$, and $\hat{\gamma} = .18$, given their

estimates of fundamental values (p^*). When we use a parsimonious model of fundamental housing values for California counties, we find similar-sized estimates of the parameters governing the residual dynamics, $\hat{\xi} = .50$, $\hat{\delta} = -.35$. However, at $\hat{\gamma} = 2.22$, we estimate a larger current-period response of actual prices to changes in fundamental values.

The parsimonious model of fundamental housing values for California counties we use here takes the rate of net migration of people to a CMSA, $y_{m(j)t}^{net}$ to be the primary determinant of changes in housing demand in that CMSA.¹⁸ These population flows to a metropolitan area are assumed to have a differential impact on prices in the counties in the metropolitan area, depending on supply-side factors in the counties. For example, housing supply in counties within CMSAs is constrained to differing degrees by natural topographic features and land use regulations.¹⁹ We subsume these factors into a single index, which we label S_j , and estimate from data on the extent to which job growth in a county also tends to lead to an increase in the housing stock in that county, rather than an increase in commuting from other counties in the area.²⁰

The level of the fundamental value of housing in an area also reflects the extent of amenities in that area. We attempt to capture such differential amenities by fixed effects parameters A_j . We proceed as if other CMSA-wide demand-side determinants of fundamental values, $D_{m(j)t}$, interact multiplicatively with

¹⁸Here $m(j)$ is an index representing the CMSA containing county j .

¹⁹See Rose (1989) for evidence that topographical features and land use regulations are important to explaining the variation across metropolitan areas in land prices.

²⁰Specifically, the supply factor index, S_j , for county j is the exponential of negative one times the average value of the residual in OLS estimates of an equation explaining the growth rate of the single-unit housing stock in that county by the growth rate of employment (by place of work) in that county. Such an equation was estimated using annual data from 1987 to 1996, pooling across nineteen California counties and one rest-of-California geographic area.

the supply-side factors S_j :

$$(8) \quad p_{jt}^* = A_j + \theta S_j D_{m(j)t}$$

Thus, given our assumption that the rate of net migration of people to a CMSA, $y_{m(j)t}^{net}$ is the primary determinant of changes in housing demand in that CMSA, $\Delta D_{m(j)t} = y_{m(j)t}^{net}$, the implied equation for Δp_{jt} , solving out for p^* , is:

$$(9) \quad \Delta p_{jt} = -\delta p_{j0}^* + \xi \Delta p_{jt-1} + \delta p_{jt-1} - \delta \theta S_j \sum_{\tau=1}^{t-1} y_{m(j)\tau}^{net} + \gamma \theta S_j y_{m(j)t}^{net} + \epsilon_{jt}$$

OLS estimation of this equation yields the above-mentioned parameter estimates of $\hat{\xi} = .50$, $\hat{\delta} = -.35$, $\hat{\gamma} = 2.22$, and also $\hat{\theta} = 1.28$.

To see what these estimated values of the coefficients imply about the dynamics of housing prices it is useful to examine the difference equations which govern these dynamics. In particular, $a(L) = b(L)^{-1} = (1 - (1 + \xi + \delta)L + \xi L^2)^{-1}$ is the lag polynomial which governs the way residual shocks ϵ_{jt} pass through to actual prices, p_{jt} . Shocks to fundamental values, p_{jt}^* , are transferred to actual prices according to the product of this lag polynomial governing residual shocks and $\gamma - (\delta + \gamma)L$. At the estimated values of $\hat{\xi} = .50$ and $\hat{\delta} = -.35$, the roots of $b(L)$ are complex and imply that the infinite-order lag polynomial $a(L)$ has the damped, oscillating coefficients shown by the dashed line in Chart 3. A one-unit transitory innovation in ϵ_{jt} increases the logarithmic level of real house prices in the current and three subsequent periods, but thereafter the effect dampens down to close to zero. As shown by the solid line, a permanent change in the fundamental path of house prices is estimated to have two to three times as large an impact on actual prices in the short-run, but most of the convergence to the long-run unitary impact takes place

by the fifth year of adjustment.

Even given these dynamic response functions, the implied dynamics for actual prices still depend on two unknowns, the dynamics of the fundamental values and the size of residual shocks to the system. The typical residual shock is relatively small. For example, Chart 4 shows the actual and predicted values of logarithmic real house price changes for Los Angeles County; the difference between the actual and predicted values, the residual ϵ_{jt} , is small relative to the overall variation in actual price changes.

The estimated current effects of changes in fundamental prices, $\hat{\gamma}\Delta p_{jt}^*$ shown by the thin solid line in Chart 4, generally explain a moderate proportion of the variation in actual price changes. For example, the current effect of changes in fundamental prices for Los Angeles County swung from about 2 percent in the mid-1980s to about -10 percent in the early 1990s. Furthermore, because changes in fundamental prices in one direction tend to be followed by changes in fundamental prices in the same direction, and the bulk of the cumulative responses to these shocks are spread over about five years (as shown in Chart 3), the changes in fundamental prices contribute a lot to the overall explanatory power of the model.

With regard to the dynamics of the fundamental prices themselves, our model implies that these are completely determined by the dynamics of net migration from the CMSA, amplified to a lesser or greater extent by the (time-invariant) supply-side factors S_j . Although the heterogeneity in dynamics induced by the supply-side amplification effects is notable—the estimated supply-side index S_j ranges from a low of about .2 in Riverside and Placer counties to highs of 2.3 in Marin County and 2.7 in San Francisco and San Mateo counties—the main driving variable for fundamental housing values is the rate of net migration from the CMSA.

Following the stylized facts developed in Blanchard and Katz (1992) and the more directly applicable empirical results of Gabriel, Matthey, and Wascher (1995,1997), we calculate projected paths for net migration, $y_{m(j)t}^{net}$, beyond 1996 by using a simplified version of these migration models in which changes in the net migration rate depend only on changes in the differential between a CMSA's unemployment rate and the unemployment rate in other areas (Chart 5). With unemployment rate differentials for California CMSAs assumed to continue to revert to their means over the 1997 to 2000 period, the model implies that net migration rates also move back towards historical averages. The effects are largest for the Los Angeles CMSA, for which a roughly 2 percentage points projected narrowing of the unemployment rate differential implies about a 1 percentage point slowdown in the rate of net outmigration.

Given the estimates of equation (9) and the actual history of real housing prices and migration through 1996, the migration projections imply that in most counties real housing price changes turn positive in 1997 and jump noticeably in 1998, similar to actual recent experience. Implied changes in termination probabilities along these projected paths for housing prices under the two-factor model without fixed effects also have been calculated and show responses similar to those illustrated in table 5, albeit with a specific time profile over which each county's house price reverts to the previous peak level and the implied changes in termination probabilities accumulate to the values shown in table 5.²¹ For example, for Los Angeles County, where the 1996 house price was about 25 percent below the 1990 peak, the 1990 peak is projected to be re-attained in the year 2000; the projected drop of about 1 percentage point in the default rate, pickup of almost 2 percentage points in the mobility rate, and several percentage point increase in

²¹The conversion from real to nominal house price is made by assuming that CPI inflation is steady at 3 percent. For these out-of-sample simulations, the interest rate spread variable is held constant at the 1996 end-of-sample value.

refinancing rates is spread over the 1997 to 2000 period. In contrast, for the Bay Area county of Alameda, where the 1996 house price was only about 8 percent below the 1990 level, the 1990 house price level is projected to be surpassed in 1998; for this county, projected changes in termination probabilities occur soon after the end of 1996 sample period, but the effects are smaller than for Los Angeles area counties.²²

5 Effects on Mortgage Pool Valuation

We now turn to an illustration of the effect of geographic variations in mortgage termination probabilities on the valuation of mortgage pools, assuming that the groups of mortgages in the pools are drawn from individual geographic areas and all mortgages are fully insured. As noted in our discussion of equation (2) above, valuing the expected cash flows from a mortgage pool requires specification of both a discount rate process and a decision rule which associates each interest rate state with either continuation of the mortgage or prepayment. In this section of the paper, we use the empirically-derived refinancing, default, and mobility-related termination probabilities discussed in the previous section as alternative calibrations of parameters in a generalization of the Stanton (1995) rational prepayment mortgage pricing model. As

²²It would be useful in future work to explore more fully the robustness of the estimated responses of termination probabilities to house price changes. Relative to various alternative specifications we did explore and do not otherwise report here, the estimated coefficients on house prices shown in table 4 (the unrestricted two-factor models) were toward the middle of the range. For example, one could argue that the empirical termination model should include a house price level variable—in addition to the included variable which only picks up the percentage changes in house prices since a reference year—because some of the transactions costs of refinancing, moving, or defaulting are lump-sum in nature, not just a proportion of the house price or mortgage balance. Indeed, adding the 1992-96 average level of housing prices as an explanatory variable to this model does help predict relatively higher rates of refinancing in counties with high average house prices, for example, and the estimated coefficient on the “house price relative to 1990” variable in the refinancing equation drops from 3.5 to 2.5. However, adding other additional purely cross-sectional explanatory variables tends to increase the estimated magnitude of the effect of house price changes on terminations. For example, in the extreme case of adding county fixed effects to the refinancing equation, the estimated coefficient on the “house price relative to 1990” variable is 5.8, quite a bit larger than the 3.5 value shown in table 4.

in Stanton (1995, 1996), we assume a Cox, Ingersoll, and Ross (CIR) term structure process to derive the discount factors and to compute rational exercise of the refinancing option, we use a Poisson(ρ) parameter for frequency-of-refinancing-decision, and we use a Beta(α, β) distribution for transactions costs. We generalize this model in allowing the Poisson arrival rate λ_{jt} for exogenous (default and mobility-related) terminations to vary over time and place in some specifications. Also, we allow the Poisson arrival rate for the transactions cost impediment to exercising the refinancing option, ρ_{jt} , to vary over time and place. Given our assumptions about transactions costs and the exogenous prepayment process λ_{jt} , we calculate the value of the mortgage using standard numerical methods for solving partial differential equations²³.

The effects on mortgage values of variation in the underlying parameters ($\alpha, \beta, \lambda_{jt}, \rho_{jt}$) are sensitive to the slope of the term structure at the time of valuation and to whether the mortgages under consideration have high or low coupons relative to prevailing mortgage rates. Mortgage values in the case of low coupon, discount mortgages are shown in table 6, and the table 7 results pertain to premium mortgages. In the results of both table 6 and table 7, the mortgage values pertain to a relatively flat term structure environment²⁴.

²³We use the Crank-Nicholson implicit finite difference method for solving the appropriate PDE on a discrete grid of time and (transformed) interest rate values; see Stanton (1995) for an exposition of this procedure. In defining the cash flows, we let the borrower's "optimal" exercise decision determine the timing of refinancing; such exercise can be impeded by the transactions costs, which we include, but we do not let the mortgage pool investor receive any transactions costs paid by the borrower. Upon default, investors are assumed to receive the full amount of outstanding principal; this assumption is consistent with how investors in passthrough MBS see cashflows upon default, as mortgage insurers guarantee return of principal.

²⁴We use the Stanton (1995) CIR parameters, originally due to Pearson and Sun, of $\kappa = .29368$ for the speed of adjustment, $\sigma = .11425$ for interest rate volatility, and $q = -.12165$ for the price of interest rate risk. However, we drop the long-run mean instantaneous spot rate to $\mu = .04935$. From this family of curves, we choose a relatively flat term structure that has a current spot rate of 7.1 percent and a slight rise to a thirty-year bond yield of 7.4 percent. Given our baseline assumptions for the distribution of transactions costs ($\alpha = 2.96, \beta = 3.154$), the frequency of decision $\rho = .483$, and rate of exogenous prepayment ($\lambda = .050$), this implies that the coupon on a par mortgage would be 7.23 percent. To define the discount and premium mortgages, we subtract and add 150 basis points to this par coupon rate.

We use the Stanton (1996) estimates for the parameters of the transactions cost Beta distribution ($\alpha = 2.96$, $\beta = 3.154$) and take this same Beta distribution to apply to all counties. This implies that the mean transaction cost faced by the representative borrowers is about 48 percent of the mortgage balance at exercise of the prepayment option. Mortgage “pool” values for each county are calculated by drawing a sample of 100 borrowers from this transactions cost distribution, calculating the mortgage value at each level of transactions cost, and averaging over the 100 representative mortgages.

Heterogeneity across counties is introduced into the model by calibration of λ_{jt} and ρ_{jt} to our empirical results on termination probabilities. We calibrate the fixed-effect estimates of the single-factor model using the table 2 values of λ_j and ρ_j . To calibrate λ_{jt} and ρ_{jt} to the results from the two-factor model, we proceed as if 1996 is the first year of a 30-year mortgage term and start by equating the λ_{jt} and ρ_{jt} for 1996 to values implied by the within-sample fit of the model. Then, we roll the calibrated λ_{jt} and ρ_{jt} forward using the implied changes in termination probabilities along the dynamic house price path which was described in the preceding section²⁵.

With a low coupon, the mortgages are valued at a discount relative to new issues, and the option to refinance has virtually no value. Almost all of the cross-sectional variation in mortgage values is implied by the alternative assumptions about the rate of exogenous termination from default and moving, λ_{jt} ²⁶. High values of λ_{jt} imply more early return of the principal from the low-coupon mortgage, increasing the

²⁵More specifically, we set λ_{jt} equal to the predicted sum of default and moving probabilities, $\pi_{jt}^D + \pi_{jt}^M$, and we set ρ_{jt} so that the time-varying ratio of county-specific to average decision frequency, $\rho_{jt}/\rho_{.t}$, matches the projected relative rate of refinancing in the county, $\pi_{jt}^R/\pi_{.t}^R$. After the year 2000, we hold the model projections at the projected values for the year 2000.

²⁶In table 6 and table 7, we define the option value as the excess over actual mortgage value of an amortizing bond with the same coupon rate and same rate of exogenous terminations, but no other (endogenous) early terminations.

value of the mortgage to the pool investors. Under the parameterization from the fixed-effect single-factor model using total terminations, the range of implied mortgage pool values is quite wide, with valuation lows in San Bernardino and other Los Angeles area counties where rates of mobility (and refinancing) were relatively low and valuation highs in Marin and other San Francisco area counties where rates of mobility (and refinancing) were relatively high.

However, this fixed-effects total terminations calibration method mistakenly “interprets” the low rates of refinancing in Los Angeles area counties as low rates of exogenous termination and mistakenly “interprets” the high rates of refinancing in the San Francisco Bay area counties as high rates of exogenous termination. Using data on terminations by type, the single-factor model implies a narrower, but still sizable range of mortgage pool values in the discount case. Also, the position of some counties in the cross-section distribution of values differs significantly depending on whether the calibration is to total terminations or to terminations by type. For example, the calibration to terminations by type shows that discount mortgages from Riverside County should be less heavily discounted than one would expect from the history of total terminations; much of the low rate of total terminations in Riverside County owed to low rates of refinancing, not to high rates of “exogenous” terminations.

The final two columns of table 6 show implied mortgage values with the parameters λ_{jt} and ρ_{jt} varying over time, as implied by the two-factor models and projections of house prices. In most counties, projected increases in housing prices lower default probabilities. In the calibration to the restricted two-factor model, this is the only channel for house prices to affect terminations. Accordingly, beyond the first few years of the mortgage term the λ_{jt} from the restricted two-factor model tend to be below the fixed λ_j from

the single-factor model using terminations by type. Therefore, in the discount case the state average mortgage pool value is lower under the restricted two-factor model than under the single-factor model with terminations by type. Also, across counties the range of values is quite wide, with lows again in the Los Angeles area counties of Riverside and San Bernardino, where rates of mobility were relatively low; the restricted two-factor model assumes that these low mobility rates persist throughout the mortgage term.

Under the calibration to the unrestricted two-factor model, mobility rates in counties such as Riverside and San Bernardino are projected to increase significantly over time as house prices rebound. Also, other things equal, the probability of refinancing increases as the rebounding house prices ease collateral constraints. In the discount case, only the former mobility effect is quantitatively significant, so the state average mortgage value is higher under the unrestricted two-factor model than under the restricted two-factor model; also, incorporating the dynamic mobility effect substantially narrows the range of variation in values across counties. In this sense of implying a relatively narrow range of mortgage pool values, the results of the preferred unrestricted two-factor model are similar to the results of the single-factor model using terminations by type.

In a qualitative sense, the results reported in Table 7 for the premium mortgage, high-coupon case are symmetric to those for the discount mortgage, low-coupon case. Under the maintained assumption that the unrestricted two-factor model is the closest specification to the “truth”, we also see in the premium case that using either a fixed-effect single-factor model of total terminations or a dynamic, but restricted two-factor model using terminations by type would lead to substantial errors in estimating mortgage pool values. However, now that the considered coupon is high, the failures in these models to incorporate the effects of

house prices on the rate of mobility-related “exogenous” termination leads to an overestimation of mortgage pool values in areas with historically low mobility and house prices, rather than the underestimation of values seen in the discount case.

An additional feature arises in the premium case, making the quantitative results not fully symmetric to the discount case; with a high coupon, the option to refinance has non-negligible value. Accordingly, the alternative calibrations of λ_{jt} and ρ_{jt} also introduce heterogeneity across models in the county-specific option values. (High values of λ_{jt} reduce option values and high values of ρ_{jt} increase option values.) On a state-average basis, the net effects on estimated option value of choosing one of the considered models over another are small. Also, under the preferred unrestricted two-factor model the range of variation in option values across counties is relatively small. However, among the “misspecified” models, the range of option values (not shown) is wide if the calibrations are to either the single-factor model with total terminations or to the restricted two-factor model. In general, although the mortgage valuation errors from model misspecification are somewhat smaller in the premium case than in the discount case, the option valuation errors from model misspecification are larger in the premium case than in the discount case. For premium mortgages, appropriate model specification appears to be particularly important to correctly gauging the interest-rate-sensitive component of the mortgage pool value.

6 Conclusion

We have considered the importance of exogenous mortgage terminations and heterogeneous refinancing decision frequencies in a rational model of mortgage pool valuation. Our findings suggest that two types

of model misspecification are likely to lead to particularly large valuation errors in some contexts. First, total mortgage terminations are likely to be depressed in areas which recently have experienced declines in house prices, and (single-factor) models which relate total mortgage terminations only to interest rates are prone to mistakenly interpret these low total mortgage terminations as persistently low rates of borrower mobility and default. Second, in such weak markets (two-factor) models which relate one individual type of mortgage termination (refinancing) primarily to interest rates, another individual type of mortgage termination (default) primarily to house prices, and treat the remaining mobility-related terminations as exogenous are prone to make valuation errors and get the sign wrong on the sensitivity of total terminations to changes in house prices. Although declines in house prices tend to increase defaults, this only partly offsets the reductions in mortgage terminations from lower rates of refinancing and moving. Accordingly, a generalization of the two-factor model to include the influence of house price collateral constraints on refinancing and moving is recommended.

Also, additional research is needed to more fully understand the role of regional housing and labor markets in generating regional variations in mortgage terminations. In this paper we have emphasized that regional employment demand shocks can lead to labor market disequilibria which generate between-area migration, corresponding movements in house prices, and ultimately affect refinancing and within-area mobility rates, in addition to default rates. In future research using regional data, we believe that it would be particularly useful to investigate the robustness of the finding that collateral constraints are quite important to refinancing behavior. An alternative view embedded in much of the literature emphasizes the correlation between individual borrower characteristics and their propensities to refinance, as if some

borrowers face much higher transactions costs of refinancing than others, irrespective of the value of the collateral. A nested test of these competing explanations for heterogeneity in refinancing behavior would be useful.

7 References

- Abraham, Jesse M. and Patric H. Hendershott (1996). "Bubbles in Metropolitan Housing Markets," *Journal of Housing Research*, Vol. 7, Issue 2, 191-207.
- Archer, Wayne R. and David C. Ling (1993). "Pricing Mortgage-Backed Securities: Integrating Optimal Call and Empirical Models of Prepayment," *Journal of the American Real Estate and Urban Economics Association*, 21, 4, 373-404.
- Archer, Wayne R., David C. Ling, and Gary A. McGill (1996). "The Effect of Income and Collateral Constraints on Residential Mortgage Terminations," *Regional Science and Urban Economics*, 26(3-4), June, 235-61.
- Archer, Wayne R., David C. Ling, and Gary A. McGill (1997). "Demographic versus Option-Driven Mortgage Terminations," *Journal of Housing Economics*, 6, 2, June, 137-63.
- Becketti, Sean and Charles S. Morris (1990). "The Prepayment Experience of FNMA Mortgage-backed Securities," Federal Reserve Bank of Kansas City research working paper 90-01, July.
- Blanchard, Olivier, and Lawrence Katz (1992). "Regional Evolutions," *Brookings Papers on Economic Activity*, 0(1), 1-61.
- Brennan M., and E.S. Schwartz (1985). "Determinants of GNMA Mortgage Prices," *Journal of AREUEA*, 13, 209-28.
- Caplin, Andrew, Charles Freeman, and Joseph Tracy (1993). "Collateral Damage: How Refinancing Constraints Exacerbate Regional Recessions," NBER Working Paper No. 4531, November.
- Capozza, Dennis R., Dick Kazarian, and Thomas A. Thomson (1997). "Mortgage Default in Local Markets," *Real Estate Economics*, Vol. 25, No. 4, 631-55.
- Capozza, Dennis R., Charlotte Mack, and Christopher J. Mayer (1997). "The Dynamic Structure of Housing Markets," mimeo.
- Cheyette, Oren (1996). "Implied Prepayments," *Journal of Portfolio Management*, Fall, 107-15.
- Deng, Yongheng, John M. Quigley, and Robert Van Order (1996). "Mortgage default and low down-payment loans: The costs of public subsidy," *Regional Science and Urban Economics*, 26 263-285.

Dunn, K. and J. McConnell (1981). "Valuation of GNMA Mortgage Backed Securities," *Journal of Finance*, 36, 599-616.

Engelhardt, Gary V. (1998). "House Prices, Collateral, and Household Mobility," mimeo, Dartmouth.

Gabriel, Stuart, Joe Matthey, and William Wascher (1995). "The Demise of California Reconsidered: Interstate Migration over the Economic Cycle," *Economic Review*, Federal Reserve Bank of San Francisco, No. 2, 30-45.

Gabriel, Stuart, Joe Matthey, and William Wascher (1997). "Jobs, Housing, and the Quality-of-Life: Intrametropolitan and Out-of-State Moves among California Households," manuscript, December.

Harding, John P. (1997). "Estimating Borrower Mobility from Observed Prepayments," *Real Estate Economics*, 25:3, 347-71.

Hayre, Lakhbir and Arvind Rajan (1995). "Anatomy of Prepayments: The Salomon Brothers Prepayment Model," *Salomon Brothers*, June.

Kau J., D. Keenan, W. Muller III, and J. Epperson (1992). "A Generalized Valuation Model for Fixed-Rate Residential Mortgages," *Journal of Money, Credit and Banking*, 24:3, 279-99.

Kau, James B., Donald C. Keenan, and Taewon Kim (1994). "Default Probabilities for Mortgages," *Journal of Urban Economics*, 35, 278-96.

Kau, James B. and Donald C. Keenan (1995). "An Overview of the Option-Theoretic Pricing of Mortgages," *Journal of Housing Research*, 6:2, 217-44.

Kau, James B. and Thomas M. Springer (1992). "The Prepayment of Fixed-Rate Mortgages Underlying Mortgage-Backed Securities: A Switching Regimes Analysis," *Journal of Housing Economics*, 2, 257-275.

Mayer, Christopher and David Genesove (1997). "Equity and Time to Sale in the Real Estate Market," *American Economic Review*, 87(3), June, 255-69.

Meese, Richard and Nancy and Wallace (1994). "Testing the Present Value Relation for House Prices: Should I Leave my House in San Francisco?" *Journal of Urban Economics*, 35(3), May, 245-66.

Monsen, Gordon (1992). "The New Thinking on Prepayments," *Mortgage Banking*, October, 48-56.

Patrino, Gregg N. (1994). "Mortgage Prepayments: A New Model for a New Era," *The Journal of Fixed Income*, December, 42-56.

Peristiani, Stavros, Paul Bennett, Gordon Mosen, Richard Peach, and Jonathan Raiff (1996). "Effects of Household Creditworthiness on Mortgage Refinancings," *FRBNY Research Paper No. 9622*, August.

Peristiani, Stavros, Paul Bennett, Gordon Mosen, Richard Peach, and Jonathan Raiff (1997). "Credit, Equity, and Mortgage Refinancings," *FRBNY Economic Policy Review*, July, 83-99.

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

Quigley, John M., Robert Van Order, and Yongheng Deng (1994). "The competing risks for mortgage termination by default and prepayment: A minimum distance estimator," mimeo.

Rose, Louis A. (1989). "Urban Land Supply: Natural and Contrived Restrictions," *Journal of Urban Economics*, vol. 25(3), May, pp. 325-45.

Schwartz, Eduardo S. and Walter N. Torous (1989). "Prepayment and the Valuation of Mortgage-Backed Securities," *Journal of Finance*, 44(2), 375-92.

Schwartz, Eduardo S. and Walter N. Torous (1993). "Mortgage Prepayment and Default Decisions: A Poisson Regression Approach," *Journal of the American Real Estate and Urban Economics Association*, 21:4, 431-49.

Stanton, Richard (1995). "Rational Prepayment and the Valuation of Mortgage-Backed Securities," *The Review of Financial Studies*, Fall, Vol. 8, No. 3, 677-708.

Stanton, Richard (1996). "Unobservable Heterogeneity and Rational Learning: Pool-specific versus Generic Mortgage-Backed Security Prices," *Journal of Real Estate Finance and Economics*, 12(3), May, 243-63.

Stein, Jeremy C. (1995). "Prices and Trading Volume in the Housing Market: A Model with Down-payment Effects," *Quarterly Journal of Economics*, May, 379-406.

Timmis, G. C.(1985). "Valuation of GNMA mortgage-backed securities with transaction costs, heterogeneous households and endogenously generated prepayment rates," working paper. Carnegie-Mellon.

Appendix: Mortgage Termination Data Sources

To measure refinancing in California by county, we use the only relatively comprehensive, publicly available source of information: statistics on loan originations collected under the Home Mortgage Disclosure Act (HMDA) and disseminated by the Federal Financial Institutions Examination Council (FFIEC). These statistics show for given geographic areas, as detailed as the Census Tract, the number of mortgage loan applications originated by purpose, including those for refinancing and home purchase. For each California county, the frequency of refinancing is measured as the ratio of HMDA loan originations for refinancing to the estimated total stock of mortgages on owner-occupied dwellings in the county.

The HMDA dataset does not report mortgage *terminations* triggered by home purchases (sales), it only reports mortgage *originations* for home purchases. We do not have access to a comprehensive dataset on mortgage terminations triggered by home purchases. Available proxies for the geographic distribution of mortgage terminations for home purchases within California include the HMDA data on mortgage *originations* for home purchases and data on the actual volume of sales of new and existing homes.

The time-series pattern of overall home sales in California is similar to the pattern of loan originations for home purchases. Both measures of sales activity picked up in 1994, dropped back in 1995, and picked up again in 1996. In the computations reported in this paper, we use the home sales data instead of the home loan for purchase data, but the results are not very sensitive to this measurement choice.

To create a county-level proxy for mobility-related loan terminations triggered by home sales, we subtract from the annual level of sales of existing homes the number of defaults on home mortgage loans. This is equivalent to assuming that each default triggers a home sale and that there is a one-to-one correspondence between home sales and non-refinancing mortgage terminations. Our main conclusions are robust to using another measurement assumption, that the ratio of non-refinancing mortgage terminations to home sales is equal to the ratio of mortgages outstanding to the stock of owner-occupied homes.

We measure mortgage defaults in California by the number of trustees deeds recorded. For counties within the Los Angeles Area and San Diego, such estimates were published in the *Real Estate and Construction Report* of the Real Estate Research Council of Southern California, as provided to them by TRW-Experian. Comparable figures for the multi-county groupings of the San Francisco Bay Area, Sacramento Valley, Central Valley, and Central Coast were obtained from the Experian news release “California Foreclosures Highest Level Since 1989, Experian Reports”, March 18, 1997. We distribute these multi-county totals to individual counties within these areas by the distribution of notices of default on home loans recorded in these years; the data on notices of default on home loans by county was obtained from various issues of the *Real Estate and Construction Report* of the Real Estate Research Council of Northern California.

The dependent variables in our regressions are the frequencies with which mortgages default, refinance, or are prepaid because of a (non-defaulted) home sale. Accordingly, we normalize our measures of the volume of default, refinancing, and mobility-related terminations by estimates of the stock of loans in each geographic area. The 1990 Census provides estimates of the fraction of owner-occupied units in each county for which a mortgage is outstanding. We estimate the stock of loans outstanding by multiplying this mortgage share by the California Department of Finance estimates of the stock of one- to four-unit homes in each county.