Burnout from pools to loans: Modeling refinancing prepayments as a self-selection process

Gan, Jumwu

Fannie Mae

31 May 2009

Burnout from pools to loans: Modeling refinancing prepayments as a self-selection process

Junwu Gan*

Applied Financial Technology, 160 Sansome Street, Suite 1200, San Francisco, CA 94104-3718

(May 31, 2009)

Abstract

In this paper we present compelling evidence from a detailed analysis of historical prepayment data to demonstrate that a mortgage cohort remembers the level of the previous mortgage rate troughs experienced by the cohort. This is a general property, observed ubiquitously, that inescapably leads to refinancing models with a continuous distribution of refinancing incentive thresholds (elbows). We present such a new refinancing model, derived from the first principle, based on a single assumption that each loan has an incentive threshold above which its borrower will refinance. In this model, the refinancing prepayment of a cohort is a dynamic self-selection process that evolves by itself according to the encountered mortgage rate environment with the cohort concurrently acquiring its memory along the way.

---

*Current address: Model Review & Oversight, Fannie Mae, 3900 Wisconsin Ave. NW, Washington, DC 20016-2892. Email: junwu_gan@fanniemae.com

I. INTRODUCTION

A mortgage prepayment model is supposed to project prepayment speeds under any possible future scenario of market conditions, of which the dominant factors affecting refinancing prepayments are future mortgage rates. Apart from some sporadic market implied measurements, the major source of data for model calibration comes from the historical prepayment behavior. A major challenge, and possibly the ultimate one, is to ensure that a model calibrated to a single historical scenario remains qualitatively correct for any possible future scenario.

It is likely that a thorough understanding of the sufficient conditions for a qualitatively correct prepayment model is still beyond our reach. Nevertheless, there are certain necessary conditions that any qualitatively correct prepayment model must not violate. Beautiful fitting of historical prepayment speeds is important. It is one important measure of the model caliber. It is the first test a model must pass. But it is not the last test. It is far more important to extract correct qualitative features from historical data and have them built into the model. The difference lies in the predictive power of the model.

For a prepayment model to have the potential of possessing predictive power, it must capture the fundamental econometric relationship [5]. The function of a prepayment model is to predict how efficiently borrowers exercise their prepayment option when presented with financial incentives. It is hopeful that the fundamental econometric relationship is relatively stable in time after the effects of changing economic environment have been identified and treated separately. For instance, changes to loan program guidelines or underwriting standards can all affect financial incentives in the eyes of borrowers making refinancing decisions. The impact of the changing economic environment is better taken into account by computing an incentive measure that is comparable over time, leaving the core prepayment model to capture the fundamental and hopefully stable econometric relationship.

Because future mortgage rates are unknown, the core prepayment model capturing the econometric relationship should not be tainted by the specific shape of the historical mortgage rate curves against which the model is calibrated. In fact, the econometric relationship should be independent of the shape of the historical mortgage rate curves! (If one wonders whether the existence of the so-called media effect would invalidate the above argument, please read on.) One should always check whether one would end up with a different model if sequences, relative heights, and distances between mortgage rate peaks and troughs were different. A blind statistical regression of the historical prepayment data against the single scenario of the historical mortgage rates and housing prices etc. risks failing this test.

In general, newly originated mortgages should not depend on the historical mortgage rates before their origination. Superficially, the existence of the media effect may present a possible exception. The media effect represents the fact that when mortgage rates reach historical lows, the prepayment speeds are likely to see an extra boost because borrowers are bombarded by the news media about their refinancing opportunities. At any given time, past mortgage rates are used to determine whether the current mortgage rate qualifies as a historical low. Even in this case, the past mortgage rates only enter into the calculation of a numeric measure for quantifying how low historically is the current mortgage rate. This numerical measure only describes the market condition. It is independent of mortgages themselves and their characteristics. For instance, this numerical measure is the same for all mortgage product types. The econometric relationship for the media effect measures the extra prepayment boost for a given numerical measure of how low historically is the mortgage rate under consideration and generally varies with mortgage product types. For instance, this extra boost is proportionally stronger for conforming than jumbo fixed rate mortgages historically. After separating out the numeric measure representing the market condition, the econometric relationship representing the media effect is independent of the historical mortgage rates before mortgage origination. It is true that the model parameters for the media effect would be different if a different numerical measure for the market condition were chosen. Nevertheless, the logical separation between the econometric relation capturing the borrowers' response to the market condition and the numerical measure quantifying the market condition is unambiguous.

For seasoned mortgages, their future behavior does depend on the mortgage rates before the current pricing date due to the well known burnout effect [4]. This dependence arises from the loans' exposure to those mortgage rates since origination. The model at the time of origination of these mortgages does not depend on the historical mortgage rates before origination.

The constraint that a prepayment model be independent of the shape of the historical mortgage rate curves before mortgage origination is not difficult to satisfy by itself. For example, a simple model with a monotonic dependence of the refinancing prepayment speed on the incentive would satisfy the constraint. However, the richness of the burnout effect makes satisfying the constraint more challenging.

The usual description of the burnout effect is that a seasoned cohort that has experienced refinancing prepayments before will prepay slower in the future than a similar cohort of newly originated mortgages. This effect was correctly attributed to the inhomogeneity of the mortgages in the cohort [4]. Various approaches have been developed to capture the burnout effect, from a simple depression of future refinancing prepayments by a cumulative measure of past refinancing exposure to a continuous distribution of populations with different refinancing intensities [1,3,5–8]. These approaches share a common understanding of the burnout effect that the inhomogeneity results in populations with different refinancing amplitudes. Conceptually, these approaches are qualitatively equivalent to a two-population mechanism of the burnout effect [4]. In this simple mechanism, a cohort is considered virtually made up of one fast prepaying population and another slow or non-prepaying population. As the fast population is depleted, the refinancing prepayments of the cohort will be suppressed. The difference between the various approaches [1,3,5–8] lies in the details in the model flexibility to fit historical data.

One common deficiency of the above mentioned traditional approaches to the burnout effect is that the depression of the future refinancing prepayments is across the entire spectrum of refinancing incentives. For example, one can imagine a scenario in which a cohort experiences a long period of small refinancing incentives. Though the refinancing prepayments will be low due to the small incentives, the cohort will nevertheless be burned out after a long period. Consequently, this burned-out cohort will not participate in future refinancing prepayments no matter how high future incentives will be. Experienced practitioners know that this is not true. For instance, it is shown in Figures 10 and 11 of the reference [2] that the effectiveness of the burnout effect decreases for lower mortgage rates: "Even a seasoned borrower would be enticed to refinance if presented with a refinancing incentive that is greater than any that he has ever seen". To alleviate the obvious deficiency of complete burnout by an extended period of marginal refinancing exposure, one may introduce a curing (also referred to as reconstitution) effect. The idea is that whatever challenging conditions that caused slow or non-prepaying populations not participating in previous refinancing waves are cured gradually over the time. While the economic basis for the curing effect is sound, it is practically delicate to determine the curing rate. In a practical situation somewhat similar to the above discussed simplistic example where a cohort exposed to a long period of small refinancing incentives runs into a big refinancing wave, the calibrated curing effect would likely depend on the specific shape of the historical mortgage rate curves if the curing rate were simply chosen to produce desired refinancing peaks for the late refinancing wave. One has to make sure that the calibrated curing rate would stay the same if the late big refinancing wave occurred sooner or later in time, with different depth for the mortgage rate troughs, and if the period of small refinancing incentives were shorter or longer.

When the burnout effect is treated by depressing the future refinancing amplitude across the entire incentive spectrum, the shape of the S-curve (the refinancing response function in terms of incentives) is suppressed vertically. Though the shape of the total S-curve for a cohort can evolve in time in multi-population models, each population still has its own S-curve and the depression of the future refinancing prepayments of each population is uniform across the entire spectrum of the refinancing incentives. The presence of more populations allows better historical fitting. But the positioning of the elbows of the fictitious intermediary refinancing populations is treacherous, with the undesired side effect of oscillating convexities. The risk of over-parameterization also undermines the predictive power of the model by including in the model untested sensitivities with respect to the changing shape of the mortgage rate curves.

The difficulties of the traditional treatments of the burnout effect can be traced back to the conceptual misunderstanding of the inhomogeneity resulting in different refinancing amplitudes. In the next section and Appendix A, we shall present a detailed analysis of historical prepayment data to demonstrate unambiguously that the consequence of the inhomogeneity is a distribution of refinancing incentive thresholds (elbows), not the refinancing amplitudes. The misunderstanding of the role of the inhomogeneity not only caused practical difficulties but also lead to qualitatively incorrect answers about when higher incentives can correspond to lower refinancing prepayments.

The specific fact we shall establish through the analysis of the next section and Appendix A is that a mortgage cohort with past refinancing exposure remembers the level of previous mortgage rate troughs. The future behavior of the cohorts will be different, depending on whether or not future mortgage rates will fall below the previous lows. Since future mortgage rates are unknown, a cohort of newly originated mortgages can experience multiple periods of different refinancing incentives. The first trough of the future mortgage rates can be at any level and the cohort can remember that level. Therefore, a refinancing prepayment model mus be able to remember any mortgage rate value. A refinancing model that can remember any mortgage rate value must have a continuous distribution of refinancing incentive thresholds. It is true that mortgages with different refinancing incentive thresholds generally have different refinancing amplitudes. But the different amplitudes are the manifested consequence, not the cause.

When prepayment speeds are differentiated by more and more loan-level attributes and when a pool-level cohort is split into loan-level cohorts, how does the burnout effect evolve? For example, a cohort of fixed rate mortgages with pool-level characteristics in the traditional sense of agency pools has only two indicatives, i.e. the average coupon and origination date. The loan-level attributes, such as loan size, loan-to-value, credit score, geological location and documentation level etc., are all considered part of the mortgage inhomogeneity of the pool-level cohort. When these attributes are pulled out of the inhomogeneity, it is hard to imagine that the distribution of the refinancing elbows due to the inhomogeneity will stay unchanged. Indeed, historical evidence presented in Section VI will show that some loan-level attributes such as loan size and loan-to-value do interact with the burnout effect. The new refinancing model with a continuous distribution of elbows provides a natural framework to treat the evolution of the burnout effect from pool-level to loan-level cohorts.

It is clear from the above discussions that a prepayment model must satisfy at least two constraints. First, the validity of the econometric relationship of the model should not be constrained to the historical scenario of market conditions against which the model is calibrated. Especially, the econometric relationship should not be tainted by the particular shape of the historical mortgage rate curves. A mortgage cohort simply evolves from origination according to whatever scenario of the market conditions encountered by the cohort. The model has no preferential scenario. It just happens that for the historical scenario of the market conditions, the model generates projections close to the actual prepayments. Secondly, as the cohort becomes seasoned, it remembers the level of the previous mortgage rate troughs it has experienced.

It is easy to fall into the trap of reducing prepayment modeling to fitting a bunch of historical prepayment curves, sometimes even unconsciously. For instance, one may start with a vague but not incorrect notion that a prepayment model is a multi-variate function of time, market condition variables and loan attributes, then make a quantum leap to construct such a multi-variate function numerically with essentially a multi-dimensional look-up table. Various statistical methods can help the scarcity of loan data in certain regions of the multi-dimensional variable space. But a look-up table, or anything of similar nature, violates both constraints just discussed above and therefore is fundamentally wrong.

The knowledge about mortgages is never likely to be complete, if one remembers that there is an indescribable human being behind each mortgage. There are always a lot of unknowns. Consequently, mortgages in a cohort will always exhibit inhomogeneity. It is a complex problem to construct a correct effective description of unknowns. It is even more challenging to understand the interplay between the knowns and unknowns when some unknowns become known. It is our hope to stimulate the study of this complex problem by demonstrating that the collective effect of unknowns for the refinancing prepayments is observable from the actual prepayment speeds in terms of an effective distribution of the refinancing elbows.

The organization of the paper is as follows. In Section II, we present an analysis of the historical prepayment data for jumbo 30-year fixed rate mortgages. In Section III, we derive the new refinancing model. In Section IV, we present calibration results. In Section V, we describe the new burnout mechanism. In Section VI, we discuss the evolving burnout effect in the pool-level to loan-level model transition. In Section VII, we discuss the effect of burnout induced S-curve steepening on duration and convexity. In Section VIII, we discuss the interplay between the knowns and unknowns. An analysis similar to Section II for the 30-year fixed rate FNCL pools is included in Appendix A to establish the generality of the property exhibited by the jumbo mortgages that a cohort remembers the level of the previous mortgage rate troughs.

II. REMEMBERING THE LEVEL OF THE PREVIOUS MORTGAGE RATE TROUGHS

In this section, we present a detailed analysis of the historical prepayment speeds of the prime jumbo 30-year fixed rate mortgages. In Appendix A, we shall confirm our findings from the 30-year fixed rate FNCL pools.

Generally, a borrower exercises his refinancing option only if there is a financial gain. While some borrowers exercise their options more efficiently than others, it is natural to expect higher refinancing prepayments for higher incentives, everything else being equal. However, it was discovered almost from day one in the study of agency pools that refinancing prepayments of a seasoned pool are depressed by past refinancing exposure [4]. This is the so-called burnout effect. An example of the burnout effect is shown in Figure 1. Due to the burnout effect, refinancing prepayments become interesting and challenging both intellectually and practically because modeling refinancing prepayments is no longer a simple proportional fitting of the historical prepayment speeds.

Current treatments [1,3,5–8] of the burnout effect are essentially based on an understanding [4] that can be derived from Figure 1. It has been recognized that some mortgages within the cohort prepay faster than the others due to inhomogeneity. As the fast prepaying mortgages are burned out, the remaining ones will prepay slower when future opportunities arise. Now we show in the rest of this section that this understanding is at least incomplete.

Figure 2 shows four aggregated cohorts originated in 1998 with coupon bin width of 25 basis points. For the refinancing peak around December 2001, a cohort with a higher coupon indeed prepaid faster than lower coupon cohorts. When mortgage rates fell to a new low in late 2002 after an initial rise in early 2002, the order of refinancing peak height is reversed for the three high coupon cohorts (green, yellow and pink curves in Figure 2). In late 2002, the cohort represented by the pink curve has the highest peak, followed by the yellow one. The peak of the green cohort is lower than the yellow one. This reversal of the order of the refinancing peak heights took place even though the difference between the cumulative prepayments of the three high coupon cohorts from the first refinancing peak is relatively small, in comparison to Figure 1. Even the absolute magnitude of the largest burnout measured by the area under the green curve over the period of the first refinancing wave is mediocre. Yet the prepayment curves of the three high coupon cohorts crossed each other before rising to the second refinancing peak around November 2002.

Figures 3 and 4 show the behavior of crossing prepayment curves similar to Figure 2 over the same time period for 1999 and 2001 vintage cohorts. Figure 5 also shows similar behavior to Figure 2 for 2000 vintage cohorts but over a different time period, because mortgages originated in 2000 tend to have high coupons. Thus, one can safely conclude that the behavior of crossing prepayment curves exhibited in Figure 2 is not an exception.

One should note that the refinancing peaks around December 2001 in Figures 2–4 are not very high by the standard of traditional treatments of the burnout effect for the crossing of the prepayment curves to take place. Figure 6 shows the actual prepayments of five 2002 vintage cohorts aggregated by origination date and coupon. The aggregation bin widths are twelve months for the origination date and 25 basis points for the coupon. The center of the coupon aggregation bins is indicated by the legends in Figure 6. One can see that even though the 2003 refinancing peaks are giant, the prepayment curves do not cross each other in Figure 6. Figures 7 and 8 show the same behavior of non-crossing prepayment curves for 2003 and 2004 vintage cohorts.

So what causes prepayment curves to cross each other in Figures 2–5 but not in Figures 6–8? One difference is that in Figures 2–5 the mortgage rates fell to a new low at the time of the second refinancing peak, as one can see in Figure 2. In contrast, the mortgage rates after the first refinancing peak in Figures 6–8 did not fall below the previous trough. More concurring evidence from FNCL pools will be presented in Appendix A.

Figure 9 shows the rate of prepayment increase per unit of mortgage rate decrease for several cohorts from Figures 2–4, as well as the historical mortgage rate curve on the right axis. The historical mortgage rates decrease almost linearly from May 2002 to November 2002. The rate of prepayment increase rises sharply around August 2002 when the mortgage rate falls below the previous low, indicating that the cohorts remember the level of the previous mortgage rate trough, or equivalently, the largest experienced refinancing incentive.

When a prepayment model is applied to cohorts of newly originated mortgages, an unknown future mortgage rate scenario can have multiple troughs. The level of the first mortgage rate trough may or may not be surpassed by the second trough. The prepayment model should be able to distinguish these different scenarios. Furthermore, the level of the first mortgage rate trough can be at any value of which the prepayment model should be able to remember. If a prepayment model can remember any possible future mortgage rate, the model must contain a continuous distribution of refinancing populations with different refinancing incentive thresholds.

If a refinancing model has a continuous distribution of refinancing incentive thresholds, then fresh refinancing populations suffered no previous burnout will participate in the refinancing prepayments when the mortgage rates fall below the previous low. The sharp rise of the rate of prepayment increase corresponds to the activation of the fresh populations suffered no previous burnout. This burnout mechanism is depicted in Figures 13a and 13b.

III. SELF-SELECTION REFINANCING MODEL

Let us consider a cohort of fixed rate mortgages of the same loan type, originated at time t_0 with average coupon c_0. Each loan belonging to the cohort is described by a set of known and unknown attributes. Let us assume that a loan with known attributes k and unknown attributes u has a refinancing incentive threshold $E_{k,u}$. When mortgage rates fall so that the incentive is above the threshold $E_{k,u}$, the financial gain is large enough that there is a finite probability for the loan's borrower to go through the hurdles to refinance. The refinancing prepayment speed in the unit of percentage SMM (Single Month Mortality) for the cohort at time t_i measured in month is

$$P(t_i) = \frac{1}{B(t_{i-1})} \sum_{k,u} b_{k,u}(t_{i-1}) p_{k,u}(t_i, t_0, c_0)\, \theta\left[I(t_i, c_0) - E_{k,u}\right], \tag{1}$$

$$B(t_{i-1}) = \sum_{k,u} b_{k,u}(t_{i-1}), \tag{2}$$

where t_{i-1} is the previous month of t_i, $b_{k,u}(t_{i-1})$ is the balance of the loans with known and unknown attributes k and u at time t_{i-1}, $p_{k,u}(t_i, t_0, c_0)$ is the refinancing probability for the month t_i, and $I(t_i, c_0)$ is the refinancing incentive for the loans with coupon c_0 at time t_i. $I(t_i, c_0)$ is essentially a measure of market conditions. Obviously, $B(t_{i-1})$ is the total balance of the cohort at time t_{i-1}. Finally, $\theta(x)$ is the step function

$$\theta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0. \end{cases} \tag{3}$$

Conceptually, the refinancing incentive $I(t_i, c_0)$ can be considered as the ratio between the gross coupon c_0 of the loan and the prevailing mortgage rate at time t_i with proper look-back. In practice, the calculation of the incentive is more involved. But the details are beyond the scope of this paper.

It is conceivable that there are borrowers who will never refinance no matter how high is the incentive. This feature has to be built into the model. In other words, we do not make the assumption that the behavior of every borrower is financially rational. In the discussions of the previous sections, we pointed out that there are always unknowns about mortgages. The functional dependence of the incentive threshold $E_{k,u}$ on the loan attributes is unknown. As we have shown in Section II, what is important and observable is an effective measure of unknown attributes manifested as a distribution of refinancing incentive thresholds. For this purpose, we introduce a density function in terms of the refinancing incentive thresholds,

$$D_k(t_i, \omega) = \frac{1}{B_k(t_{i-1})} \sum_u b_{k,u}(t_{i-1}) \delta(\omega - E_{k,u}), \tag{4}$$

$$B_k(t_{i-1}) = \sum_u b_{k,u}(t_{i-1}), \tag{5}$$

where $\delta(x)$ is the Dirac function. Intuitively, it can be viewed as

$$\delta(x) = \lim_{\sigma \to 0} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right). \tag{6}$$

For borrowers with different refinancing thresholds, the refinancing intensity can be different too. Let us define a cohort-level refinancing probability function also in terms of the refinancing incentive thresholds,

$$p_k(t_i, \omega, t_0, c_0) = \frac{1}{D_k(t_i, \omega) B_k(t_{i-1})} \sum_u b_{k,u}(t_{i-1}) p_{k,u}(t_i, t_0, c_0) \delta(\omega - E_{k,u}). \tag{7}$$

We emphasize that the summation in equations (4) and (7) is only over the unknown loan attributes u, reflecting the fact that only the unknowns are transformed into a distribution of refinancing incentive thresholds.

Inserting the identity

$$\int_{-\infty}^{\infty} d\omega \, \delta(\omega - E_{k,u}) = 1, \tag{8}$$

into equation (1), the refinancing prepayment of equation (1) can be written as,

$$P(t_i) = \frac{1}{B(t_{i-1})} \int_{-\infty}^{\infty} d\omega \, \theta \left[I(t_i, c_0) - \omega \right] \sum_{k,u} b_{k,u}(t_{i-1}) p_{k,u}(t_i, t_0, c_0) \delta(\omega - E_{k,u}). \tag{9}$$

With the help of equation (7), the last equation (9) can be recast as,

$$P(t_i) = \frac{1}{B(t_{i-1})} \sum_k B_k(t_{i-1}) P_k(t_i), \tag{10}$$

$$P_k(t_i) = \int_{E_l(k)}^{E_h(k)} d\omega \, p_k(t_i, \omega, t_0, c_0) D_k(t_i, \omega) \, \theta \left[I(t_i, c_0) - \omega \right], \tag{11}$$

where the actual integration range is introduced in equation (11). $P_k(t_i)$ represents the prepayment speeds of a loan-level cohort with known loan attributes k. The lower limit $E_l(k)$ is the lowest incentive threshold of all loans in the cohort with the same known attributes k,

$$E_l(k) = \min_u (E_{k,u}). \tag{12}$$

The upper limit $E_h(k)$ is not so straightforward. As we discussed earlier, there may be borrowers who will never refinance. Consequently, the actual upper limit could be infinite. In practice, we can impose a finite upper limit $E_h(k)$ and use it to approximate all loans with incentive thresholds higher than $E_h(k)$. To allow for the fact that there are borrowers who will never refinance, we require the refinancing probability function $p_k(t_i, \omega, t_0, c_0)$ to fall to zero for ω at a value below $E_h(k)$ and remain zero thereafter.

The burnout is a natural consequence of the inhomogeneity manifested by the existence of a distribution of thresholds. In equation (1), the evolution of the refinancing populations after participating in refinancing prepayments simply results from the weighting balance $b_{k,u}(t_{i-1})$. Consequently,

$$D_k(t_{i+1}, \omega) = D_k(t_i, \omega) \left[1 - \frac{\bar{p}_k(t_i, \omega, t_0, c_0)}{100} \right], \quad \text{for } \omega \leq I(t_i, c_0), \tag{13}$$

where $\bar{p}_k(t_i, \omega, t_0, c_0)$ is a seasonally adjusted refinancing probability function,

$$p_k(t_i, \omega, t_0, c_0) = \bar{p}_k(t_i, \omega, t_0, c_0) \, \mathrm{Seas}(t_i). \tag{14}$$

The seasonal adjustment coefficients $\mathrm{Seas}(t_i)$ are just twelve numerical values, with one for each month. By separating out the seasonal coefficients explicitly, the density function

$D_k(t_{i+1}, \omega)$ can be largely immune to the seasonal variation. The population density at any time t_{i+1} is automatically normalized according to

$$D_k(t_{i+1}, \omega) \longrightarrow \frac{D_k(t_{i+1}, \omega)}{\int_{E_l(k)}^{E_h(k)} d\omega \, D_k(t_{i+1}, \omega)}. \tag{15}$$

Equations (11) and (13) constitute our new self-selection refinancing model. It is derived based on a single assumption of the existence of a refinancing incentive threshold for each loan.

There are two important points we want to emphasize. First, equations (4) and (7) express mathematically the notion discussed in Introduction that the unknown mortgage attributes represented by u are transformed into the observable refinancing population density and probability functions in terms of the incentive threshold ω. The functional dependence of the incentive threshold $E_{k,u}$ itself on the mortgage attributes u probably will never be known. In the practical sense, that dependence is not essential. What is important is to determine the fairly unique population density at origination $D_k(t_i = 0, \omega)$ and the ω-dependence in the probability function $\bar{p}_k(t_i = 0, \omega, t_0, c_0)$, both of which can be inferred from the actual prepayment speeds. Only the population density at origination is determined by model parameters, because the population densities at later times are evolved from the one at origination according to equation (13).

Secondly, the content of the known loan attributes k in equations (4) and (7) can be changed by the way of cohort formation according to the practical need. By assigning some mortgage attributes to the category of the known attributes and aggregating accordingly, one can form cohorts with various degrees of loan-level characteristics. The self-selection model derived in this section provides a unified and consistent approach to study the interplay between the known and unknown mortgage attributes.

IV. APPLICATION TO PRIME JUMBO 30-YEAR FIXED RATE MORTGAGES

In this section, we present some calibration results of the application of the new refinancing model to prime jumbo 30-year fixed rate mortgages. For cohorts mimicking agency pools, loans are aggregated by origination date and coupon only. All loan-level attributes are averaged out. Thus, the label k for the known loan-level attributes in Equation (11) represents an empty set for pool-level cohorts and will be dropped in this section. For simplicity, we assume a factorized form for the refinancing probability function in equation (11),

$$p(t_i, \omega, t_0, c_0) = q(\omega)\, m(t_i, t_0, c_0)\, \mathrm{Seas}(t_i). \tag{16}$$

Equation (11) becomes

$$P(t_i) = m(t_i, t_0, c_0)\, \mathrm{Seas}(t_i) \int_{E_l}^{E_h} d\omega\, q(\omega) D(t_i, \omega)\, \theta\left[I(t_i, c_0) - \omega\right]. \tag{17}$$

The burnout equation (13) becomes

$$D(t_{i+1}, \omega) = D(t_i, \omega)\left[1 - \frac{q(\omega) m(t_i, t_0, c_0)}{100}\right], \quad \text{for } \omega \leq I(t_i, c_0). \tag{18}$$

This pool-level refinancing model given by equations (17) and (18) is described by parametric functions $D(\omega) = D(t_0, \omega)$, $q(\omega)$, and $m(t_i, t_0, c_0)$, apart from the seasonality coefficients. The meanings are: $D(\omega)$ represents the population density distribution in terms of the refinancing incentive thresholds at origination; $q(\omega)$ is the core refinancing probability function for populations with different refinancing incentive thresholds; and $m(t_i, t_0, c_0)$ can be viewed as an overall multiplier.

The probability multiplying function $m(t_i, t_0, c_0)$ is not the focus of this study. Its main ingredients are widely known in the industry, including at least an aging ramp, a premium origination effect (SATO effect) and a media effect (or publicity effect). For the model results presented in this section, the commercially available multiplying function for the 30-year fixed rate jumbo mortgages from Applied Financial Technology [1] is used. The media effect for the 30-year fixed rate jumbo mortgages is weaker proportion-wise than for 30-year fixed rate agency pools. Perhaps due to similar reasons, the effect of the positive housing price appreciation is also weaker for jumbo mortgages than for agency pools. This is partly the reason why we choose to present the results for the 30-year fixed rate jumbo mortgages.

For simplicity, the probability multiplying function $m(t_i, t_0, c_0)$ used to generate the model results throughout this paper has its housing price effect turned off. (This was the decision of late 2007 and early 2008. The situation has been very different since late 2008. But the subject of an updated prepayment model in a severe bear housing market and for a future mortgage market emerging from credit crisis is beyond the scope of this paper.) Figure 10 shows the refinancing probability function $q(\omega)$ and the initial density function $D(\omega)$ for the model calibrated to all 30-year fixed rate jumbo pool-level cohorts. Also shown is the refinancing S-curve at origination defined as

$$S_0(I) = \int_{E_l}^{E_h} d\omega\, q(\omega) D(\omega) \theta\,(I - \omega)\,. \tag{19}$$

Both the probability function $q(\omega)$ and the initial density function $D(\omega)$ are independent of the origination time t_0.

Figures 11a to 11c show the model performance for three 2001 vintage cohorts aggregated by origination year and gross coupon shown in Figure 4. The coupon bins are centered at 7.25, 7.5 and 7.75 respectively and the bin width is 25 basis points. Figures 11a to 11c correspond to the case where the historical prepayment curves of different coupons cross each other between the first and the second refinancing peaks in Figure 4.

Figures 12a to 12c show the model performance for three 2002 vintage cohorts shown in Figure 6 with coupon bins centered at 6.0, 6.25 and 6.5 respectively. The coupon bin width is also 25 basis points. This is the case where the historical prepayment curves of different coupons do not cross each other between the first and the second refinancing peaks.

It may be helpful to point out that the pool-level model was calibrated against all vintage-coupon cohorts created from all available 30-year fixed rate jumbo mortgages. The calibration is not limited to the prepayment history shown in the graphs of this section. Furthermore, the height of 2003 peaks in Figures 12a to 12c is affected by the refinancing aging ramp which varies sporadically at the pool-level without detailed knowledge about the pool composition. Nevertheless, better fitting to the historical prepayment speeds than Figures 11a to 12c can be achieved by turning on the housing price effect and improving the treatment of the SATO effect. The latter point will be discussed in Appendix A where the actual data indicates that the depression of refinancing prepayments by SATO effect is uneven for refinancing populations with different incentive thresholds. But these subjects are beyond the scope of this paper.

V. BURNOUT MECHANISM

It was pointed out in Section II that there is a crucial difference between the behavior of the cohorts in Figures 6–8 and Figures 2–4. In the former case, a cohort with a higher coupon always prepays faster than a lower coupon one during the successive refinancing waves. In the latter case, some cohorts with lower coupons actually prepay faster than others with higher coupons in the second refinancing wave. This is an intriguing phenomenon of the burnout effect: higher incentives not necessarily lead to higher prepayments for seasoned cohorts exposed to refinancing prepayments before. The prepayment curves of two cohorts with different coupons may cross each other between successive refinancing peaks.

It is important to distinguish the mechanisms responsible for the crossing of the prepayment curves of cohorts with different coupons between the traditional burnout models and the new self-selection model. There are various treatments of the burnout effect reported in literature [1,3,5–8], ranging from a simple depression of future refinancing prepayments by a cumulative measure of past refinancing exposure (or past incentives) to complex distributions of refinancing populations with different refinancing amplitudes. All these treatments manipulate the refinancing amplitudes of the populations introduced to capture the mortgage inhomogeneity and are qualitatively the same. When this qualitatively same mechanism of the existing treatments is applied to two seasoned cohorts with different coupons, the condition for the crossing of the two prepayment curves between successive refinancing waves only depends on the amount of cumulative burnout during the first refinancing wave. As long as the cohort with a higher coupon is sufficiently burned out so that the remaining fast prepaying populations are nearly depleted, its response during the second refinancing wave will be lower than the other lower coupon cohort.

The burnout mechanism in the self-selection model is completely different. What matters most in the self-selection model is not the cumulative burnout in the first refinancing wave. Instead, it is whether or not mortgage rates fall to new lows so that fresh populations with higher refinancing incentive thresholds can participate in the second refinancing wave.

Figure 13a illustrates the situation when the mortgage rates do not fall below the trough reached during the first refinancing wave. From equation (17), the refinancing prepayments in the second refinancing wave for two cohorts with the same origination but different coupons are,

$$P(t_2) = \mathrm{Seas}(t_2) m(t_2, t_0, c) \int_{Blue} d\omega \, q(\omega) D(t_2, \omega), \tag{20}$$

$$P'(t_2) = \mathrm{Seas}(t_2) m(t_2, t_0, c') \left[\int_{Blue} + \int_{Yellow} \right] d\omega \, q(\omega) D'(t_2, \omega), \tag{21}$$

where c and c' are the coupons of the cohorts A and B respectively, with $c < c'$. In the self-selection refinancing model, the population densities at origination for the two cohorts are the same. At time t_2 after the first refinancing wave, their population densities are no longer the same and they are denoted by $D(t_2, \omega)$ and $D'(t_2, \omega)$ respectively. The refinancing prepayment $P(t_2)$ of cohort A has a contribution only from the blue area in the upper graph of Figure 13a, while $P'(t_2)$ of cohort B has contributions from the blue and yellow areas in the lower graph. If neither cohort was originated at significant premium, both $m(t_2, t_0, c)$ and $m(t_2, t_0, c')$ should not have much dependence on the coupons c and c', leading to $m(t_2, t_0, c) \simeq m(t_2, t_0, c')$. From Figure 13a, we observe that $D(t_2, \omega) \simeq D'(t_2, \omega)$ in the blue area. In other words, the burnout in the blue areas is about equal for both cohorts even though cohort B produced higher prepayments during the first refinancing wave. In fact, $D(t_2, \omega)$ is slightly smaller than $D'(t_2, \omega)$ if the normalization effect given by equation (15) is taken into account. Consequently, the contribution to $P(t_2)$ for cohort A from the blue area is not higher than the corresponding contribution to $P'(t_2)$ for cohort B. But $P'(t_2)$ has an extra contribution from the yellow area, therefore $P'(t_2) > P(t_2)$ is always true. In other words, the cohort with a higher coupon will prepay faster than a lower coupon cohort in the second refinancing wave as long as the mortgage rates in the second refinancing wave do not fall below the previous low, no matter how much burnout has been suffered by the higher coupon cohort from the first refinancing wave.

Figure 13b illustrates the situation when the mortgage rates reach new lows in the second refinancing wave. First of all, we observe that the contributions to the second wave refinancing prepayments from the blue and pink areas are about equal for the two cohorts. Cohort A with a lower coupon will prepay faster than cohort B with a higher coupon in the second refinancing wave if

$$\int_{Green} d\omega\, q(\omega) D(t_0, \omega) > \int_{Green} d\omega\, q(\omega) D'(t_2, \omega) + \int_{Orange} d\omega\, q(\omega) D(t_0, \omega). \tag{22}$$

The left side of the last equation corresponds to the contribution to the refinancing prepayments of cohort A in the second wave from the green area in the upper graph of Figure 13b, while the two terms on the right side of equation (22) correspond to the green and orange areas in the lower graph for cohort B. We recall that $D(t_0, \omega)$ represents the density at origination. Because the green area for cohort A and the orange area for cohort B in Figure 13b have not participated in the first refinancing wave, the population densities in these areas remain roughly the same as at origination, up to the normalization effect. $D'(t_2, \omega)$ represents the density of cohort B after suffering burnout from the first refinancing wave. Generally speaking, the condition in equation (22) can be satisfied only if the probability function $q(\omega)$ is much lower in the orange region $[I(t_2), I'(t_2)]$ than in the green region $[I(t_1), I'(t_1)]$. This requires that the orange region $[I(t_2), I'(t_2)]$ be in the saturation area of the initial S-curve (19). The crossing of the prepayment curves of the two cohorts with the same origination but different coupons between successive refinancing peaks can occur only if the mortgage rates in the second refinancing wave fall below the previous low. But having the mortgage rates reaching a new low is not a sufficient condition. The incentives for the high coupon cohort must be high enough in the second refinancing wave so that equation (22) is satisfied.

The shape of the density functions in Figures 13a and 13b is for illustration purpose only. The area under the density function should always be normalized to one. The normalization will raise the solid curves on the right side of $I(t_1)$ and $I'(t_1)$ above the corresponding dotted curves. For the clarity of the illustration, the normalization effect is intentionally omitted to emphasize the fact that the burnout from the first refinancing wave only depletes the population densities below $I(t_1)$ and $I'(t_1)$ respectively.

VI. BURNOUT IN POOL TO LOAN MODELS: A UNIFIED FORMULATION

A pool-level cohort of fixed rate mortgages from a given loan program is specified by its origination date and average coupon. Because all other loan-level attributes are averaged out, the symbol k in equation (4) denoting known loan-level attributes represents an empty set, as in Section IV. When prepayment speeds are differentiated by more and more loan-level attributes after being included in the aggregation criteria, a pool-level cohort is split into numerous loan-level cohorts. For non-agency mortgages, the list of the loan-level attributes represented by symbol k can contain, not exhaustively, coupon dispersion, average loan size, loan to value, credit score, geographical region, documentation level, and debt to income ratio etc. The self-selection model allows a uniform treatment across-the-board from pool-level to loan-level models.

From equation (4), the relation between the pool-level population density function $D(\omega)$ and the loan-level $D_k(\omega)$ is

$$D(t_i, \omega) = \frac{1}{B(t_{i-1})} \sum_{k,u} b_{k,u}(t_{i-1}) \delta(\omega - E_{k,u}) = \sum_k \frac{B_k(t_{i-1})}{B(t_{i-1})} D_k(t_i, \omega), \tag{23}$$

$$E_l = \min_k (E_l(k)), \tag{24}$$

where E_l is the refinancing threshold for the pool-level cohort and $E_l(k)$ for the loan-level cohorts defined in equation (12). Similarly, the relation between the pool-level and loan-level refinancing probability functions follows from equation (7),

$$p(t_i, \omega, t_0, c_0) = \sum_k \frac{B_k(t_{i-1}) D_k(t_i)}{B(t_{i-1}) D(t_i)} p_k(t_i, \omega, t_0, c_0). \tag{25}$$

Equations (23) and (25) simply state that the pool-level population density and probability functions result from averaging the loan-level ones. The practical need is actually opposite.

One usually obtains a pool-level model first. So the practical problem is how to obtain loan-level models from a pool-level model. Equations (23) and (25) provide the foundation to arrive at loan-level models. One just needs to find loan-level modifications to the pool-level population density and probability functions that are consistent with equations (23) and (25). In principle, the pool-level population density and probability functions can be modified in an infinite number of ways by the loan-level attributes. In practice, some simple and straightforward modifications are likely to be enough for most purposes.

For the probability function, the simplest modification is an overall multiplier,

$$p_k(t_i, \omega, t_0, c_0) = \lambda_k \, p(t_i, \omega, t_0, c_0). \tag{26}$$

Sometimes, an incentive dependent (ω-dependent) multiplier to the probability function may be needed, such as in the case of Figures 19a and 19b discussed in Appendix A. But our scope here is limited to the simplest way of modifying the probability function.

For the population density function, an overall multiplier has no effect since the area under the density function is always normalized to one. One simple way to modify the density function is the well known "elbow shift",

$$E_l(k) = E_l + \Delta_k. \tag{27}$$

Another simple way to modify the density function specific to the self-selection model is to tilt the density function at origination,

$$D_k(\omega) = \left[1 + (\gamma_k - 1)\frac{E_h - \omega}{E_h - E_l}\right] D(\omega). \tag{28}$$

The tilt coefficient boosts the density function at E_l by a factor γ_k while holding the other end at E_h unchanged. Because it is assumed that the density function $D_k(\omega)$ will always be normalized first before being used in calculation, no attention was paid to the normalization in equation (28).

Equations (26), (27), and (28) represent three simple ways to adjust a pool-level model into loan-level models. They are the overall refinancing multiplier λ_k to the probability function, the elbow shift Δ_k and tilt coefficient γ_k to the density function. It is helpful to point out the difference between the elbow shift Δ_k in equation (27) and the commonly known elbow shift applied to the S-curve. In the former case of equation (27), the elbow shift itself only adds or subtracts an extra contribution to the S-curve as one can see from equation (19) by replacing E_l with $E_l + \Delta_k$. It does not shift the steeply rising part of the S-curve horizontally. In the latter case of the usual elbow shift applied to the S-curve,

$$S_0(I) \rightarrow S_0(I + \Delta_k), \tag{29}$$

the entire S-curve is shifted horizontally.

What distinguishes the loan-level models constructed from equations (26)–(28) from the traditional ones in the literature is that the burnout effect is consistently built into the construction of the self-selection model. Specifically, the differentiation of refinancing prepayments by loan-level attributes is constructed in such a way that not only the differentiation evolves with the mortgage age but also the dependence of the evolution on the mortgage rate environment is consistently accounted for.

Due to correlations among the loan-level attributes represented by k, it is by no means a simple task to calibrate loan-level models. Furthermore, when a pool-level cohort is split into numerous loan-level cohorts, the quality of statistics is reduced. Consequently, more sophisticated statistical methods may be needed. It is beyond the scope of this paper to present a full-fledged loan-level model for jumbo mortgages. Instead, we would like to conclude this section by presenting an example illustrating the dynamic nature of the loan-level differentiation.

Figure 14 shows the actual prepayment speeds for three cohorts of different loan size split from a 2002 vintage pool-level cohort with coupon centered at 6.25 previously shown in Figures 6 and 12b. The prepayment model speeds from the pool-level model are also shown for comparison. Figures 15a and 15b show the loan-level adjustments to the pool-level model for capturing the loan size differentiation.

It is worth pointing out that Figure 14 corresponds to the case when the mortgage rates in the second refinancing wave in 2004 did not fall below the previous low in 2003. For the other case of successive refinancing waves with mortgage rates reaching new lows, the prepayment curves of the same vintage cohorts with the same average coupon but different loan size can cross each other. This is the case for some 1998 and 1999 vintage cohorts. The pattern is similar to Figures 2 and 3. The key point is that not only the loan size effect evolves with the age but also the evolution depends on the mortgage rate environment. Fortunately, the evolution of the differentiation of the refinancing prepayments by some attributes like FICO score and geographical location are generally simpler than loan size and less dependent on the encountered mortgage rate environment.

VII. EFFECT OF BURNOUT INDUCED S-CURVE STEEPENING ON DURATION AND CONVEXITY

The self-selection refinancing model captures the inhomogeneity in borrowers' refinancing behavior through an effective distribution of refinancing incentive thresholds. This effective distribution of incentive thresholds has very different consequences on duration and convexity, in comparison to an effective distribution of refinancing amplitudes. The shape of the S-curve evolves in a specific way as the consequence of the burnout effect, as opposed to an across-the-board suppression of the S-curve when the inhomogeneity is represented by an effective distribution of refinancing amplitudes.

To illustrate the effect of the evolving S-curve shape on duration and convexity, we have chosen one TBA and two passthroughs from the July 2008 reports on OAS, duration and convexity by Wall Street dealers to compare their durations and convexities. The three TBA or passthroughs have the same net coupon but different age, being originated in the last three years. Their historical prepayment speeds are obtained by aggregating the corresponding FNCL pools and shown in Figure 16. It is clear from Figure 16 that both 2006 and 2007 vintage passthroughs have suffered burnout in early 2008. Figure 17 shows the evolving S-curves as of July 2008 for the three aggregated cohorts in our calibrated Fannie 30-year fixed rate pool model. The 2008 vintage has not suffered much burnout and the S-curve depicts its shape at origination. The steepening of the S-curves due to burnout is prominent and unambiguous, since historical model fitting is very sensitive to the S-curve. As the S-curve steepens, the behavior of seasoned premium cohorts will be shifted toward at-the-money ones and the at-the-money cohorts will be shifted toward the discounts. This shifting behavior is qualitatively different from the one expected from an across-the-board suppression of the S-curve. If only the S-curve amplitude were suppressed by burnout, durations would generally be lengthened due to a longer mortgage life span resulting from lower future prepayment speeds and a reduced sensitivity to the mortgage rates in the mortgage spread duration. For convexities, the effect of an across-the-board suppression of the S-curve would reduce the magnitude of the convexity which is usually negative. In contrast, if the shape of the S-curve is changed by burnout, the magnitude of the convexity is not always suppressed. Figure 18 compares the convexities of three TBA or passthroughs as of July 2008 after experiencing prepayments shown in Figure 16. The steepening of the S-curve leads to more negative convexities for the 2006 and 2007 vintage passthroughs. The full consequences of the burnout effect on seasoned cohorts need to be further investigated.

VIII. UNDERSTANDING THE INTERPLAY BETWEEN THE KNOWNS AND UNKNOWNS

When a prepayment model is developed, it is usually calibrated against the history of all available loans of the same type for better statistics. However, when a security is priced or its risk matrix is calculated, only the loans in the collateral of the security are relevant. It is natural to ask how one can ensure that a model calibrated against all available loans is appropriate for the security? The question becomes even more challenging for loan-level models since not only the prepayments of the security but also the evolution of the composition of its known loan attributes can be different from the large set of all available loans. The loan-level cohorts constructed from the large set of all available loans still exhibit strong burnout effect after the usually known loan-level attributes such as loan size, loan-to-value, FICO, geography etc. have been treated separately. The effect of inhomogeneity from unknown variables on the refinancing prepayments is still important and can vary from security to security. The historical prepayment speeds of a large set of all available loans of the same type only capture the average effect of unknown variables for the large set of loans. For a given security, the effect of unknown variables for the underlying collateral of the security may not fall around the average point on the spectrum of varying effects of unknown variables. If so, the security will not be priced correctly without appropriate security specific adjustment.

For a security whose collateral consists of newly originated loans, one has no way to know whether or not the effects of unknown variables of the loans in the security are typical of the large set of loans. In fact, it is not guaranteed that the behavior of these newly originated loans will resemble the old vintages of the same type, even after the effect of changing market conditions has been isolated and properly treated. One can only assume that the loans in the collateral of the security will behave like the large set of older loans with similar average loan-level characteristics.

For a security whose collateral consists of seasoned loans, the effect of unknown variables of the security may be observable from their historical prepayment speeds, depending on the amount of available history. Therefore, it may be possible to deduce the relative level of the effects of unknown variables of the loans in the security on refinancing prepayments as compared to the large set of loans. Intuitively, one can think of the large set of loans as consisting of a fictitious series of securities with a distribution of the effects of unknown variables. What is important is to determine the relative level of the effects of unknown variables for the security among the fictitious series of securities. A simple extrapolation to the future of the deviation of the past prepayments of the security from the large set of loans generally will not produce correct security specific adjustment. The prepayment history of the security is a manifestation of both the known and unknown loan variables.

A security is usually backed up by only a few thousand loans. The evolution of the composition of the known loan variables varies from security to security. For loan-level prepayment models calibrated to the large set of loans, the differentiation of the prepayment speeds by the known loan variables have already been determined. For the security whose collateral consists of seasoned loans, the security specific adjustment should only contain the discrepancy in the effects of unknown variables between the loans in the security and the large set of loans. Thus, the correct security specific adjustment should be amount to the would-be historical deviation of the prepayment speeds between the loans in the security and the large set of loans, after excluding the effects of discrepant historical evolution of the composition of the known loan variables.

Let us illustrate the interplay between the known and unknown variables with a simple example of 30-year fixed rate jumbo mortgages. For simplicity, let us assume that the only known loan-level variable is the FICO score and a loan-level model differentiating FICO score has been calibrated against all 30-year fixed rate jumbo mortgages. Intuitively, one can imagine that the loan-level model captures the behavior of cohorts aggregated by origination date, coupon and FICO. Furthermore, one can conceptually assume that cohorts of the same origination month are aggregated with very narrow bin widths for the coupon and FICO score. The aggregated cohorts have very good statistics, since all available 30-year fixed rate jumbo loans are used. For simplicity, let us further assume that the constituent loans of the concerned security have the same origination month and the same narrow bin width for the coupon as the aggregated cohorts from all 30-year fixed rate jumbo loans. But the dispersion of the FICO scores for the constituent loans of the security is wide. If the loans in the security historically prepaid faster than the corresponding large sample cohort with the same origination month, coupon and the same average FICO score at origination, it is not guaranteed that the same security will prepay faster in the future. Generally, loans with higher FICO scores prepay faster than loans with lower FICO scores when everything else is equal. If the security prepaid faster in the past by depleting more loans with higher FICO scores, then the same security may actually prepay slower in the future. Both the discrepancy in the inhomogeneity from unknown variables and the changing composition of the FICO scores in the collateral of the security contribute to the different historical prepayment speeds between the security and the large set of loans. One has to separate out the effect of the changing composition of the FICO scores in the historical prepayment speeds of the security, with the help of the FICO score effect of the loan-level model, to produce correct security specific adjustment.

The above example only highlights the problem conceptually. In reality, it is much more complex. First, there are more than one loan-level variables. So one should think of the FICO score in the above example as representing a collection of loan-level variables. Second, the loan-level variables are correlated. Third, the loan-level models are effectively calibrated against cohorts with finite aggregation widths for the coupon and other known loan-level variables, and sometimes even with finite aggregation widths for the origination date. Thus, both the discrepancy in the time evolution of the known loan-level variables and the different distribution of the inhomogeneity from unknown variables contribute to the difference in the actual prepayment speeds between the security and the large set of loans. The task of finding a security specific adjustment is to exclude the effects of changing compositions of the known variables from the deviation of the historical prepayment speeds between the security and the large set of loans. This is a complex but practically very important problem. With the birth of the self-selection refinancing model, it is possible to attack this problem consistently because the dynamic burnout effect has been built into the model construction, as illustrated by the treatments of the pool-level to loan-level transition in Section VI. The effects of unknown variables are transformed to an effective distribution of refinancing incentive thresholds. The task of finding a security specific adjustment is to find appropriate adjustments, similar to equations (26)–(28), to the refinancing population density and probability functions of the generic loan-level model to capture solely the difference in the effects of unknown variables, after excluding the discrepancy in the effects on refinancing prepayments of changing compositions of the known loan variables between the security and the large set of loans.

IX. CONCLUSIONS

In this paper, we have presented compelling evidence from historical prepayment data that unavoidably leads us to a refinancing prepayment model with a continuous distribution of refinancing incentive thresholds. We have derived the self-selection refinancing prepayment model from the first principle, instead of preferentially choosing one from many possible formulations. We have presented selected results of the calibration for the 30-year fixed rate jumbo mortgages to illustrate the new burnout mechanism of the self-selection model. We have shown that the self-selection refinancing model with an effective distribution of refinancing incentive thresholds is qualitatively different from models with effective distributions of refinancing amplitudes. Furthermore, the self-selection model provides a unified and consistent formulation to treat the evolving burnout effect from pool-level to loan-level models, as well as the interplay between the known and unknown loan-level variables. It is our hope that this paper will stimulate further rigorous researches on the challenging subject of the interplay between the knowns and unknowns.

ACKNOWLEDGMENTS

It is my pleasure to take this opportunity to thank my former colleagues at the once existed AFT (Applied Financial Technology) who helped me develop my understandings on mortgage prepayments. I am especially grateful to Michael Bykhovsky who introduced me to the prepayment modeling and whose extensive knowledge I had benefited from during the years of building up AFT business. I also wish to thank Kyle Lundstedt for his support in the late days of LPS/Applied Analytics when final touches to this paper were put in place.

APPENDIX A: REMEMBERING THE LEVEL OF THE PREVIOUS MORTGAGE RATE TROUGHS: EVIDENCE FROM FNCL POOLS

When analyzing historical mortgage data, it is imperative to establish the generality of any observed feature because there are so many unknown factors that can influence the prepayments of certain mortgages. In this Appendix, we present our analysis of FNCL pools to establish the generality of our findings of Section II, that mortgage cohorts remember the level of the previous mortgage rate troughs and the cohorts' behavior is qualitatively different depending on whether or not the mortgage rates fall below the previous lows. All cohorts presented in this Appendix are aggregated from FNCL pools with annual aggregation for the origination date and 50 basis points for the net coupon.

First of all, we have confirmed that FNCL pools exhibit prepayment behavior similar to Figures 2–8 over the same time period. We shall not repeat similar discussions here. But FNCL pools have history extending back to 1980s. Starting from 1992, the decreasing mortgage rates reached successive new lows, enabling FNCL pools to provide rich history for analyzing the burnout effect.

Figures 19a to 19c show the actual prepayment speeds of 1988 vintage cohorts. The monthly averages of the weekly Mortgage Banker Association survey rates for the 30-year fixed rate mortgages are displayed in each figure with two month delay for an approximate alignment between the prepayment speeds and the mortgage rates. Each figure is limited to three prepayment curves for clarity. Various mortgage rate troughs are marked in Figure 19a for references in later discussions. The prepayment curves cross each other throughout Figures 19a to 19c and cohorts with lower coupons can prepay faster than higher coupon cohorts after crossing, but only when mortgage rates fall below the previous lows.

Because of the wide range of coupons covered by the cohorts in Figures 19a to 19c, the burnout effect is mingled with the so-called SATO effect. The fact that borrowers were taking out mortgages with coupons significantly higher than the average mortgage rate available at origination signals their challenging financial and credit conditions. SATO effect depresses refinancing prepayments for mortgages originated at significant premium. This effect is clearly seen for the yellow curve in Figure 19a where the area under the yellow curve from 1991 to 1994 is significantly smaller than the other two curves. The depression of the refinancing prepayments by SATO effect may be more severe at higher incentives, as the yellow curve in Figure 19a seems to be depressed proportionally most for the lowest mortgage rate troughs from November 1992 to December 1993. However, the difference between the prepayment curves after 1995 becomes insignificant. As mortgages become seasoned, SATO effect diminishes since the borrowers' credit conditions improve and their equity in the house builds up over time. The reduced refinancing response due to SATO effect also leads to reduced burnout. The higher than expected difference between the 1996 and 1998 refinancing peaks may signal the curing effect for populations which did not participate in earlier refinancing activities. The treatment of the interaction between SATO effect and the burnout effect represented by a simple factorization assumption of equation (16) needs to be examined thoroughly, especially for the loan-level models in which the contributions of the known loan-level attributes to SATO effect must be identified and properly separated out to avoid double counting.

Besides the crossing of prepayment curves, another goal of this Appendix is to establish the generality of the property shown in Figure 9, which indicates that the prepayment speeds increase much faster for the same amount of mortgage rate decrease when the mortgage rates fall below the previous troughs. The steep rise of the rate of prepayment speed increase occurs simultaneously for cohorts of various coupons so the effect cannot be attributed to the shape of a particular portion of the refinancing S-curve. Another possible explanation may be the coincidental onset of the media effect when the mortgage rates fall below the previous lows. By analyzing FNCL pools from early 1990s when the historical mortgage rates reached successive new lows with widely varying media effects, we argue that it requires too much coincidence to attribute the universally observed steep rise, of the rate of prepayment increase after the mortgage rates fall below the previous troughs, to the media effect.

Figure 20a shows the rate of prepayment increase per unit of mortgage rate decrease for the actual prepayments of the cohorts in Figure 19a. On their way to the trough in February 1992, the mortgage rates fell below the previous low of April 1991 around October 1991 marked by the small circle on the mortgage rate curve in Figure 20a. The sharp rise of the rate of prepayment increase is marked by the big circle in Figure 20a. From the historical mortgage rates displayed in Figure 19a, one can see that the media effect is nearly nonexistent around October 1991.

The calculation of the rate of prepayment increase requires aligning the prepayment speed time series with the mortgage rate time series. In Figure 20a, the alignment is two month delay for the mortgage rates. Since the alignment is a delicate issue, Figure 20b shows the results for the alignment of three month delay for the mortgage rates. The qualitative conclusion is independent of the alignment. The alignment was not a concern for Figure 9 where the mortgage rate decrease is almost linear.

Figures 21 and 22 show the steep rise of the rate of prepayment increase for the cohorts in Figures 19a to 19c for the next two times when the historical mortgage rates fell below the previous trough after the time period of Figure 20a. Figure 21 shows the steep rise of the rate of prepayment increase for the actual prepayment speeds when the historical mortgage rates fell below the previous trough marked by the arrow in Figure 19b. Figure 22 shows the same feature at the time marked by the arrow in Figure 19c.

Figure 23 shows four aggregated cohorts of 1994 vintage FNCL pools, in order to further establish the generality of the findings of Section II. The multiple crossings of the prepayment curves and the richness of the burnout effect are vividly seen. Figure 24 shows the steep rise of the rate of prepayment increase around the time marked by the arrow in Figure 23.

In summary, the rate of prepayment increase rises sharply when the mortgage rates fall below the previous troughs. This is a general feature. It is observed in different products, in cohorts of different vintages and different coupons, at different times. It cannot be attributed to the shape of a particular portion of the refinancing S-curve. It is also unlike to be attributable to the media effect which varies widely at different times in the examples presented in this paper. Our interpretation relies on the existence of refinancing populations with different incentive thresholds. The mechanism is illustrated by the diagrams in Figures 13a and 13b. The inhomogeneity of mortgage cohorts results in the existence of refinancing populations which are activated at different levels of refinancing incentives. When the mortgage rates reach new lows, the previously sleeping populations suffered no burnout start to participate in the refinancing prepayments, leading to the steep rise of the rate of the prepayment increase.

FIGURES

FIG. 1. An example illustrating the burnout effect as it is commonly known. The figure shows the actual prepayment speeds of two cohorts of prime jumbo 30-year fixed rate mortgages aggregated by the origination date and gross coupon. The aggregation bin widths are 12 months and 50 basis points respectively. Both cohorts were originated nearly at money. The 2001 vintage cohort has smaller refinancing peaks in 2003 and 2004 due to burnout even though its coupon is higher than the other cohort.

FIG. 2. Crossing of the prepayment curves of 1998 vintage cohorts with different coupons when mortgage rates reach a new low. The figure shows the actual prepayment speeds of four cohorts of prime jumbo 30-year fixed rate mortgages originated in 1998 aggregated by the origination date and gross coupon, with the aggregation bin widths of twelve months and 25 basis points respectively. For the prepayment peaks around December 2001, higher incentives correspond to higher prepayments. But three prepayment curves (green, yellow and pink) crossed each other before reaching November 2002 peak where higher incentives do *not* correspond to higher prepayments anymore. Also shown in the figure on the right axis is the monthly average of the survey rates for 30-year fixed rate mortgages from Mortgage Bankers Association. The mortgage rate curve is shifted by two months to approximately align with the prepayment curves. For example, the all time low of mortgage rates in June 2003 is shown at August 2003 in the figure.

FIG. 3. Crossing of the prepayment curves of 1999 vintage cohorts, similar to Figure 2.

FIG. 4. Crossing of the prepayment curves of 2001 vintage cohorts, similar to Figure 2.

FIG. 5. Crossing of the prepayment curves of 2000 vintage cohorts, similar to Figure 2. Due to the high mortgage rates in year 2000, the 2000 vintages have higher coupons. The crossing of the prepayment curves are seen between the peaks around May 2001 and December 2001, thus at a different time from Figures 2–4.

FIG. 6. Historical prepayment speeds of five aggregated cohorts of prime jumbo 30-year fixed rate mortgages originated in 2002. The prepayment curves do not cross each other when mortgage rates in 2004 did not fall below the previous low reached in 2003.

FIG. 7. No crossing of the prepayment curves of 2003 vintage cohorts, similar to Figure 6.

FIG. 8. No crossing of the prepayment curves of 2004 vintage cohorts, similar to Figure 6. The minor crossing between the green and yellow curves near the end of year 2004 is deemed an insignificant aberration.

FIG. 9. Remembering the level of the previous mortgage rate trough. This figure shows the rate of prepayment increase per unit of mortgage rate decrease, i.e. $-[P(t_i) - P(t_{i-1})]/[r(t'_i) - r(t'_{i-1})]$ with $t' = t - 2$, for aggregated cohorts shown in Figures 2–4. The thin dark blue curve marked by the bottom legend represents the mortgage rates shifted by two months and uses the right vertical axis. From May 2002 to November 2002, the mortgage rates decreased almost linearly and fell below the previous trough of December 2001 around August 2002 when the curves representing the rate of prepayment increase of the cohorts rise steeply.

FIG. 10. The refinancing probability function (yellow curve on the left axis), the initial population density at origination (blue curve on the right axis), and the refinancing S-curve at origination (red curve on the left axis) in terms of the refinancing incentives are shown for the model calibrated to the pool-level cohorts of the 30-year fixed rate prime jumbo mortgages. The shape of the S-curve is naturally generated by the integral in equation (19). The new model parameters are two parametric functions representing population density and refinancing probability. The undesirable hyper-sensitivity of the model speeds with respect to the S-curve change in traditional models is mitigated. The incentive can be conceptually viewed as the ratio between the coupon and the 30-year fixed mortgage rate. The area under the blue density curve is normalized to one.

FIG. 11a. Comparison between model (red curve) and historical (blue curve) prepayment speeds for a cohort in Figure 4, which is aggregated from prime jumbo 30-year fixed rate mortgages originated in 2001 with a coupon bucket centered at 7.25. The coupon bin width is 25 basis points. The green curve is the model housing turnover component in the total model prepayment speeds.

FIG. 11b. Model vs. actual prepayment speeds for another 2001 vintage cohort in Figure 4 with a coupon bucket centered at 7.5, similar to Figure 11a.

FIG. 11c. Model vs. actual prepayment speeds for another 2001 vintage cohort in Figure 4 with a coupon bucket centered at 7.75, similar to Figure 11a.

FIG. 12a. Comparison between model (red curve) and historical (blue curve) prepayment speeds for a cohort in Figure 6, which is aggregated from jumbo 30-year fixed rate mortgages originated in 2002 with a coupon bucket centered at 6.0. The coupon bin width is 25 basis points. The green curve is the model housing turnover component in the total model prepayment speeds.

FIG. 12b. Model vs. actual prepayment speeds for another 2002 vintage cohort in Figure 6 with a coupon bucket centered at 6.25, similar to Figure 12a.

FIG. 12c. Model vs. actual prepayment speeds for another 2002 vintage cohort in Figure 6 with a coupon bucket centered at 6.5, similar to Figure 12a.

FIG. 13a. Burnout mechanism in the self-selection refinancing model for the case when mortgage rates do not fall below the previous low. There are two graphs with each describing the evolution of the refinancing population density of one cohort. The horizontal axis of the two graphs are aligned. The cohorts encounter two refinancing waves at times t_1 and t_2 respectively. The incentives for the two refinancing waves are $I(t_2) < I(t_1)$ for cohort A and $I'(t_2) < I'(t_1)$ for cohort B. Since cohort B has a higher coupon, $I'(t_1) > I(t_1)$ and $I'(t_2) > I(t_2)$. The refinancing population density at origination represented by the dotted curve in each graph is the same. After the first refinancing wave, the population densities are reduced to the solid curves in both graphs due to the burnout effect. For cohort A in the upper graph, only the populations represented by the blue area under the density curve contribute to the second wave refinancing prepayments. For cohort B in the lower graph, the contributing populations are from the blue and yellow areas. Even though cohort B had higher refinancing prepayments during the first refinancing wave and therefore suffered more burnout, the blue area is about the same as for cohort A. But cohort B has an additional contribution to the second wave from the yellow area. Therefore, cohort B with a higher coupon prepays faster in the second refinancing wave than cohort A with a lower coupon.

FIG. 13b. Burnout mechanism for the case when mortgage rates fall below the previous low. The notations are the same as in Figure 13a. But in this case, $I(t_2) > I(t_1)$ and $I'(t_2) > I'(t_1)$. The contributions to the second wave refinancing prepayments for cohort A come from the blue, green and pink areas. For cohort B, the contributions come from the blue, green, pink and orange areas. The contributions from the blue and pink areas are comparable for the two cohorts. The contribution from the green area is higher for cohort A because it has not suffered burnout from the first refinancing wave. But cohort B has an additional contribution from the orange area. If the refinancing probability function $\bar{p}_0(\omega)$ in equation (17) is very small in the orange area, which is the case when the incentive $I'(t_2)$ is in the saturation area of the S-curve, the additional contribution from the orange area may not be enough to compensate the reduced contribution from the green area for cohort B. In this case, cohort A with a lower coupon prepays faster than cohort B with a higher coupon in the second wave.

FIG. 14. The actual prepayment speeds of three cohorts of different average loan size split from a 2002 vintage pool-level cohort in Figure 12b with a coupon bucket centered at 6.25. The dashed green line represents the pool-level model projections. Cohorts with higher average loan size have noticeably higher prepayments for the 2003 peak. But majority of the differentiation disappears in the 2004 peak.

FIG. 15a. Comparison of the loan size adjusted model against the actual prepayment speeds for the cohort of the lowest loan size bucket in Figure 14. The pool model is also shown. The 2003 peak of the actual prepayment curve is lower than the pool model. But most differentiation disappears in the 2004 peak.

FIG. 15b. Comparison of the loan size adjusted model against the actual prepayment speeds for the cohort of the middle loan size bucket in Figure 14. The pool model is also shown. The 2003 peak of the actual prepayment curve is higher than the pool model. But its 2004 peak is lower than the pool model.

FIG. 16. Historical prepayment speeds of three aggregated cohorts of 30-year fixed rate FNCL pools corresponding to one TBA and two passthroughs of three different vintage years but with the same 6.0% net coupon (slightly at premium as of July 2008) from dealers' duration reports of July 2008.

FIG. 17. Evolving S-curves as of July 2008 in the FNCL pool model for three aggregated cohorts in Figure 16. Though most discussions of this paper are focused on the 30-year fixed rate jumbo mortgages, the self-selection model has also been calibrated to the 30-year fixed rate FNCL pools. The steepening of the S-curves due to burnout is easily spotted.

FIG. 18. Convexities of one TBA and two passthroughs of three different vintages but with the same 6.0% net coupon (slightly at premium as of July 2008) from dealers' duration reports of July 2008. The convexities are normalized by the 2008 vintage TBA for clarity. The curves labeled "Street Model 1 and 2" (blue and pink) represent the convexities from the dealers' reports. The self-selection model convexities (yellow curve) are calculated using AFT (Applied Financial Technology) production two-factor short-rate interest rate model. Due to the steepening of the S-curves, the convexities for the passthroughs of the 2006 and 2007 vintages become more negative. Because the convexity for the 2008 vintage TBA is also negative, the yellow curve for the normalized convexities rises to above one at WALA values 11 and 23 months corresponding to the two passthroughs of the 2007 and 2006 vintages respectively. The tendency depicted by the yellow curve is robust and not much affected by which interest rate model is used.

FIG. 19a. Historical prepayment speeds of three yearly aggregated cohorts of FNCL pools originated in 1988 with net coupon bin width of 50 basis points. The mortgage rates (MR), i.e. the monthly averages of Mortgage Bankers Association survey rates for 30-year fixed rate mortgages (MB30), are shown on the right axis. The various troughs of the mortgage rates relevant for discussions in Appendix A are marked. The mortgage rate curve is shifted horizontally by two months. For example, the first trough marked as April 1991 actually represents the mortgage rate of February 1991. The yellow and red prepayment curves cross each other between the 1991 and the first 1992 peaks, though the obvious SATO effect displayed by the yellow curve mingles with the burnout effect. The 1991 peak for the blue curve is insignificant. Consequently, the blue curve only crosses the red one between the second and the third peaks.

FIG. 19b. Historical prepayment speeds of another three yearly aggregated cohorts of FNCL pools originated in 1988, similar to Figure 19a. The 1991 peaks are too small for the curve crossing between the 1991 and the first 1992 peaks to take place. While the crossing between the red and blue curves takes place between the first and second peaks in 1992, the crossing between the blue and green curves only takes place between the second 1992 and the first 1993 peaks. The arrow marks the location where mortgage rates fall below the previous low before the second 1992 peak, to be further examined in Figure 21. The SATO effect displayed by the red curve is easily noticed since the area under the red curve from 1991 to early 1995 is obviously smaller than other curves. Furthermore, the depression of the red curve by SATO effect seems more severe around the mortgage rate troughs from November 1992 to December 1993.

FIG. 19c. Historical prepayment speeds of three more yearly aggregated cohorts of FNCL pools originated in 1988, similar to Figure 19a but with lower coupons. Because of lower coupons, the crossing between the yellow and green curves only takes place between the two peaks in 1993 and no obvious SATO effect is observed. The arrow marks the location where mortgage rates fall below the previous low, to be further examined in Figure 22.

FIG. 20a. Evidence of cohorts of FNCL pools remembering the level of the previous mortgage rate trough. This figure shows the rate of prepayment increase for 1988 vintage cohorts of FNCL pools in Figure 19a, for establishing the generality of the property shown in Figure 9 for jumbo mortgages. The legends denote the vintage year and net coupon. The rate of prepayment increase per unit of mortgage rate decrease, i.e. $-[P(t_i) - P(t_{i-1})]/[r(t'_i) - r(t'_{i-1})]$ with $t' = t - 2$, rises sharply after the mortgage rates fall below the previous low of April 1991.

FIG. 20b. Demonstration of the robustness of the property shown in Figure 20a, where the alignment between the prepayment speeds and the mortgage rates is $P(t) \leftrightarrow MR(t-2)$. Since the alignment is a delicate issue practically, this Figure shows that the conclusion of Figure 20a stands unaffected if the alignment is changed to $P(t) \leftrightarrow MR(t-3)$ so that the rate of prepayment increase is calculated as $-[P(t_i) - P(t_{i-1})]/[r(t'_i) - r(t'_{i-1})]$ with $t' = t - 3$.

FIG. 21. More evidence about remembering the level of the previous mortgage rate trough, similar to Figures 9 and 20a. The legends denote the vintage year and net coupon. When mortgage rates fall below the previous low marked by the arrow in Figure 19b, the rate of prepayment increase per unit of mortgage rate decrease rises sharply.

FIG. 22. More evidence about remembering the level of the previous mortgage rate trough, similar to Figures 9 and 20a. The legends denote the vintage year and net coupon. When mortgage rates fall below the previous low marked by the arrow in Figure 19c, the rates of prepayment increase per unit of mortgage rate decrease rise sharply.

FIG. 23. Historical prepayment speeds of four 1994 vintage cohorts of FNCL pools. The legends denote the vintage year and net coupon. The richness of the burnout effect is vividly demonstrated in this figure where numerous crossings of the prepayment curves take place at various times when mortgage rates fall below previous lows. By the time when the mortgage rates reached the lowest point of 1998, three high coupon cohorts (green, yellow and red curves) had reversed their order in the height of refinancing peaks. But the refinancing peaks corresponding to the shallow mortgage rate troughs of 1999 and 2000 had their order of refinancing peak heights restored to the normal situation as in 1996, before starting to reverse the order again when the mortgage rates fell to low troughs since late 2001. By 2003, the order of all four prepayment curves is completely reversed as compared to the normal situation around March 1996. This behavior of reversing back and forth the order of refinancing peak heights is consistent with the burnout mechanism depicted in Figures 13a and 13b, but not explainable if the mortgage inhomogeneity results in a distribution of refinancing amplitudes. The arrow marks the location where the mortgage rates fall below the previous low, to be further examined in Figure 24.

FIG. 24. More evidence about remembering the level of the previous mortgage rate trough. The legends denote the vintage year and net coupon. When mortgage rates fall below the previous low marked by the arrow in Figure 23, the rate of prepayment increase per unit of mortgage rate decrease rises sharply.

REFERENCES

[1] Bykhovsky, Michael, Overview of Recent Prepayment Behavior and Advances in Its Modeling and Valuation, in *The Handbook of Mortgage-Backed Securities*, 6th Edition, Ed. Frank J. Fabozzi, McGraw Hill Professional (2005).

[2] Chen, Derek; Strand, Nicholas; Fan, Min; Ibanez-Meier, Rodrigo; and Maoui, Idriss, Prepayment Outlook, *Barclays Capital*, US Securitization Research, January 30, 2009.

[3] Hall, Arden, Controlling for Burnout in Estimating Mortgage Prepayment Models. *Journal of Housing Economics* **9**, pp. 215-232 (2000).

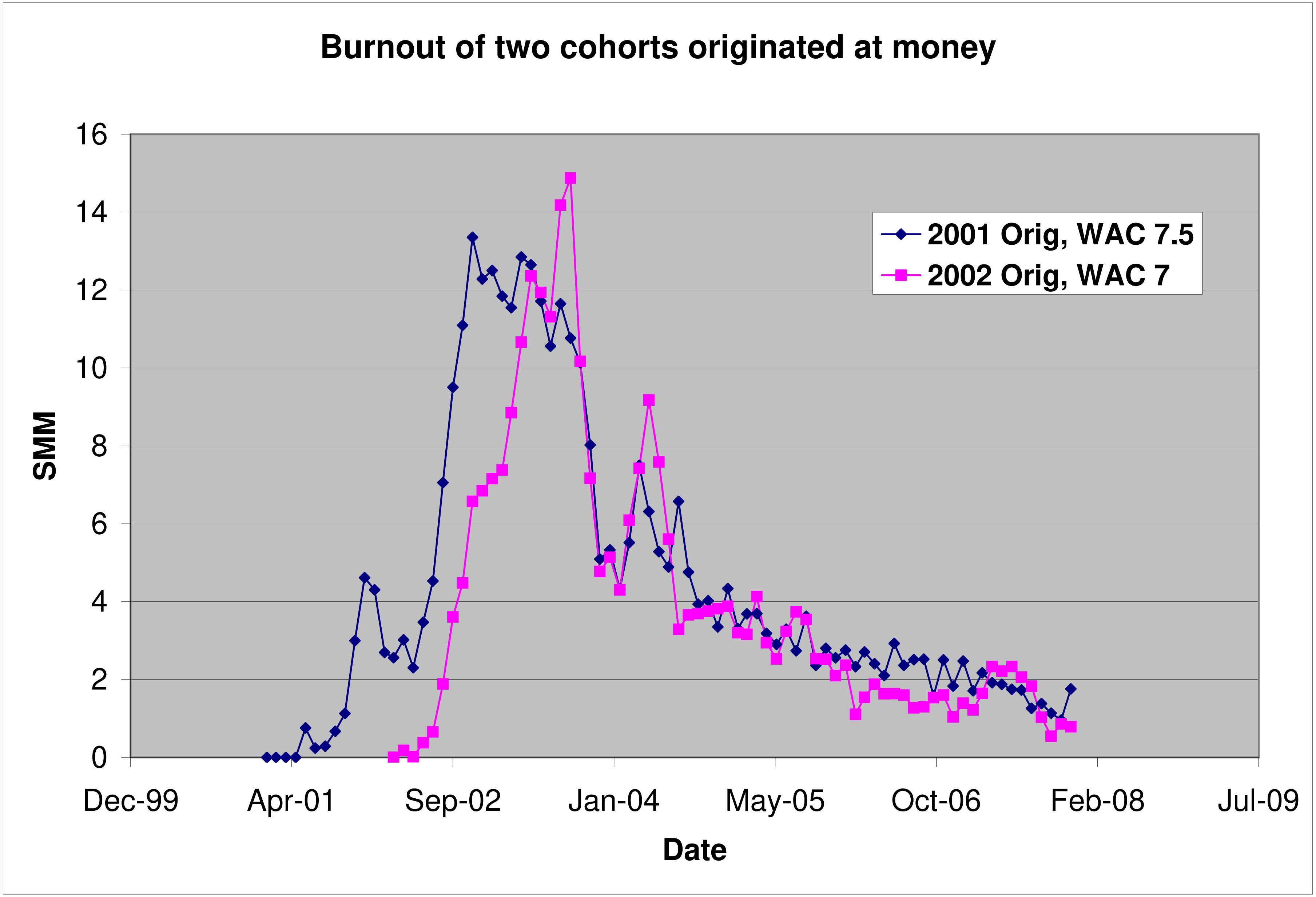
[4] Hayre, Lakhbir, A Simple Statistical Framework for Modeling Burnout and Refinancing Behavior. *Journal of Fixed Income* Volume **4** Number 3 (1994).

[5] Hayre, Lakhbir S.; Young, Robert; Teytel, Mikhail; and Cheng, Kevin, Anatomy of prepayments – The Citigroup Prepayment Model, March 2004.

[6] Levin, Alexander, Active-Passive Decomposition in Burnout Modeling. *Journal of Fixed Income* Volume 10, Number 4, pp. 27-40 (2001).

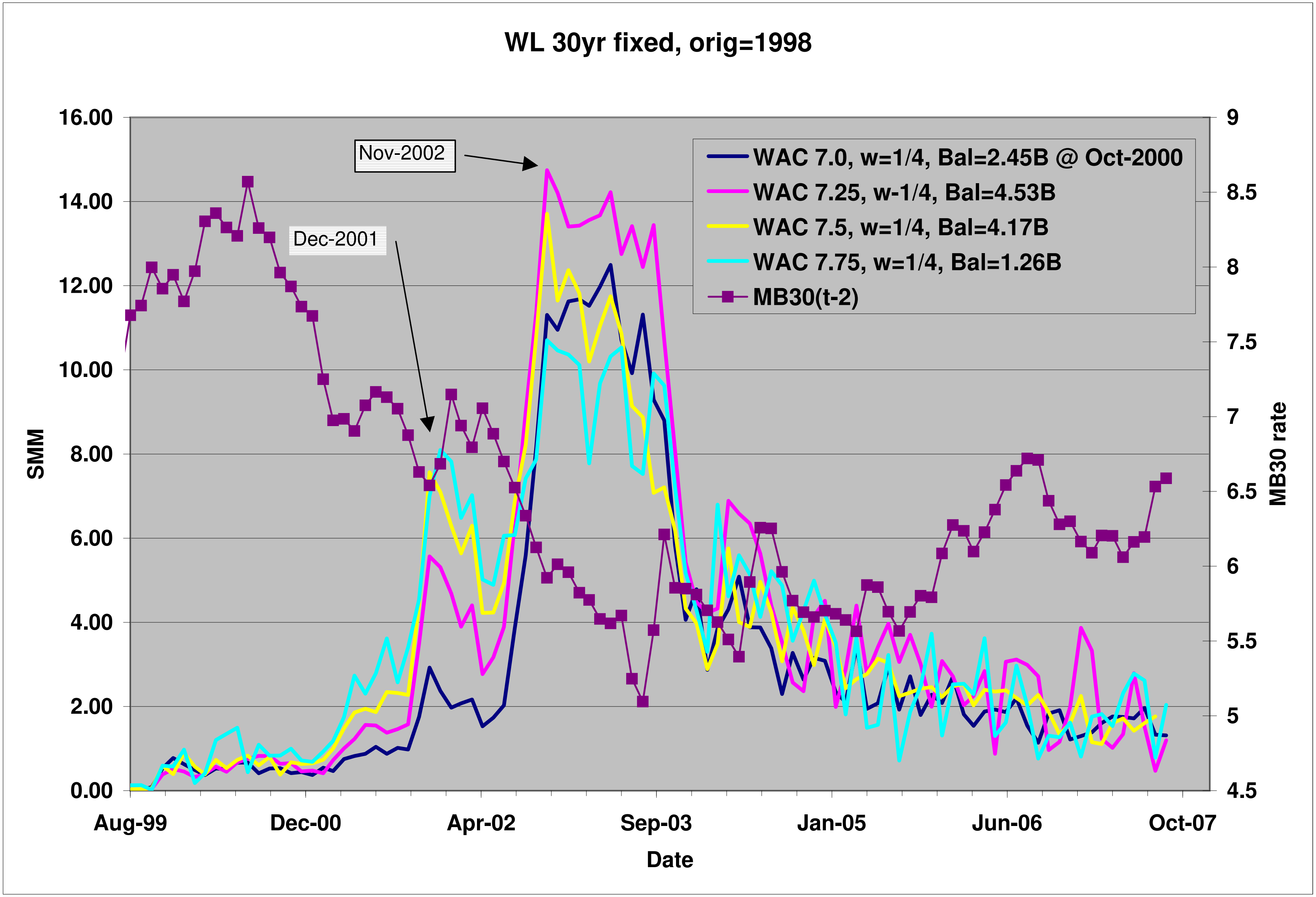
[7] Montano, David; Dabholkar, Rajan; Jha, Pankaj; Hay, Eliza; Sim, John; Wong, Nicole; and Ye, Brian, The JPMorgan Prepayment Model: It's All About Economics. *Mortgage Research*, JP Morgan Securities Inc. April 2003.

[8] Risa, Stefano; Khananusapkul, Phanwadee; Fan, Min; Modukuri, Srinivas; Subramanian, Prasanth; and Shilpiekandula, Vikas, Prime Fixed-Rate Prepayment Model. *Lehman Brothers*, February 9, 2007.



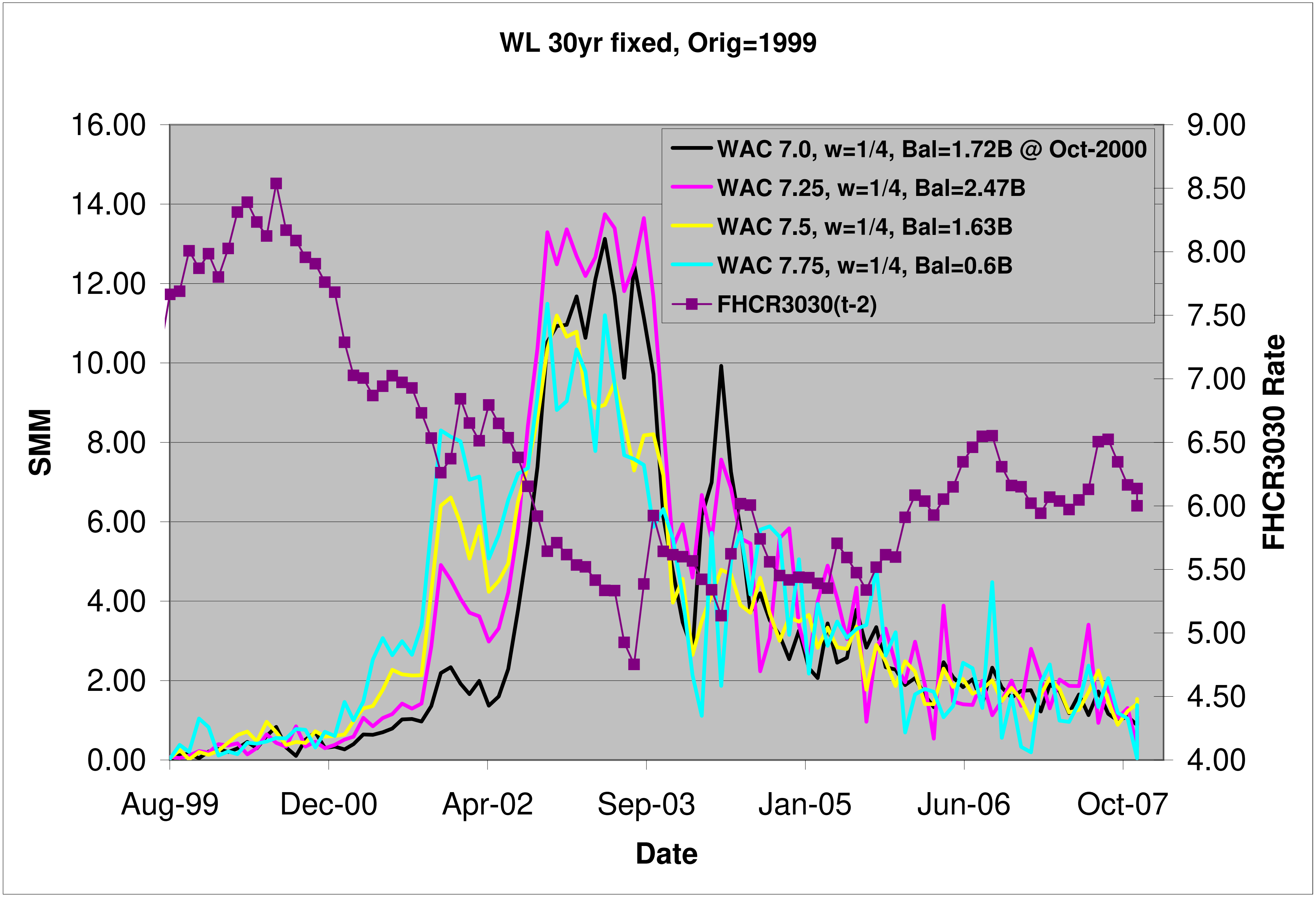
Burnout of two cohorts originated at money
2001 Orig, WAC 7.5
2002 Orig, WAC 7
SMM
16
14
12
10
8
6
4
2
0
Dec-99
Apr-01
Sep-02
Jan-04
May-05
Oct-06
Feb-08
Jul-09
Date

Fig 1



WL 30yr fixed, orig=1998
SMM
MB30 rate
Date
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
9
8.5
8
7.5
7
6.5
6
5.5
5
4.5
Aug-99
Dec-00
Apr-02
Sep-03
Jan-05
Jun-06
Oct-07
Nov-2002
Dec-2001
WAC 7.0, w=1/4, Bal=2.45B @ Oct-2000
WAC 7.25, w-1/4, Bal=4.53B
WAC 7.5, w=1/4, Bal=4.17B
WAC 7.75, w=1/4, Bal=1.26B
MB30(t-2)

Fig 2



WL 30yr fixed, Orig=1999
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
SMM
9.00
8.50
8.00
7.50
7.00
6.50
6.00
5.50
5.00
4.50
4.00
FHCR3030 Rate
WAC 7.0, w=1/4, Bal=1.72B @ Oct-2000
WAC 7.25, w=1/4, Bal=2.47B
WAC 7.5, w=1/4, Bal=1.63B
WAC 7.75, w=1/4, Bal=0.6B
FHCR3030(t-2)
Aug-99
Dec-00
Apr-02
Sep-03
Jan-05
Jun-06
Oct-07
Date

Fig 3

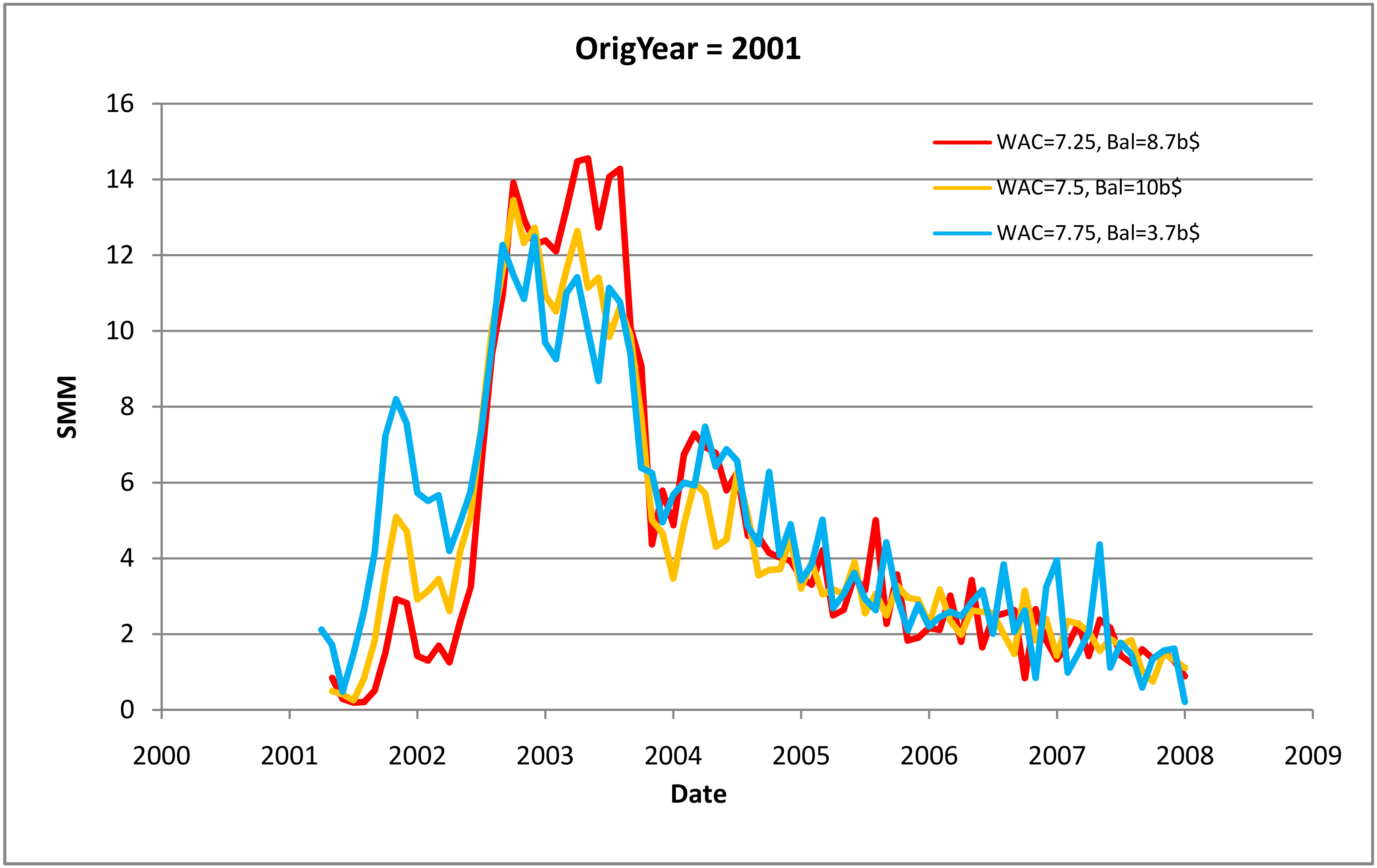
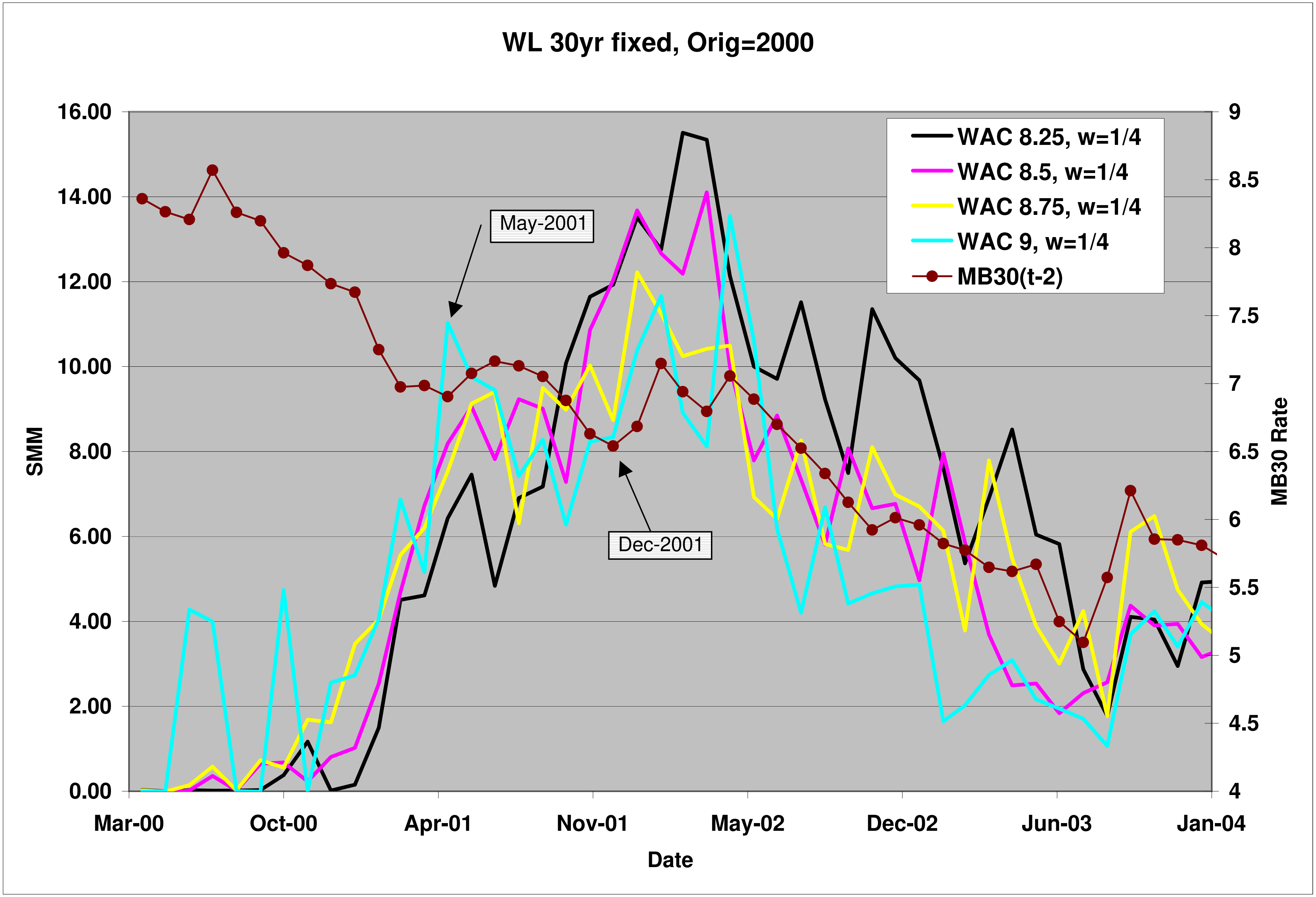
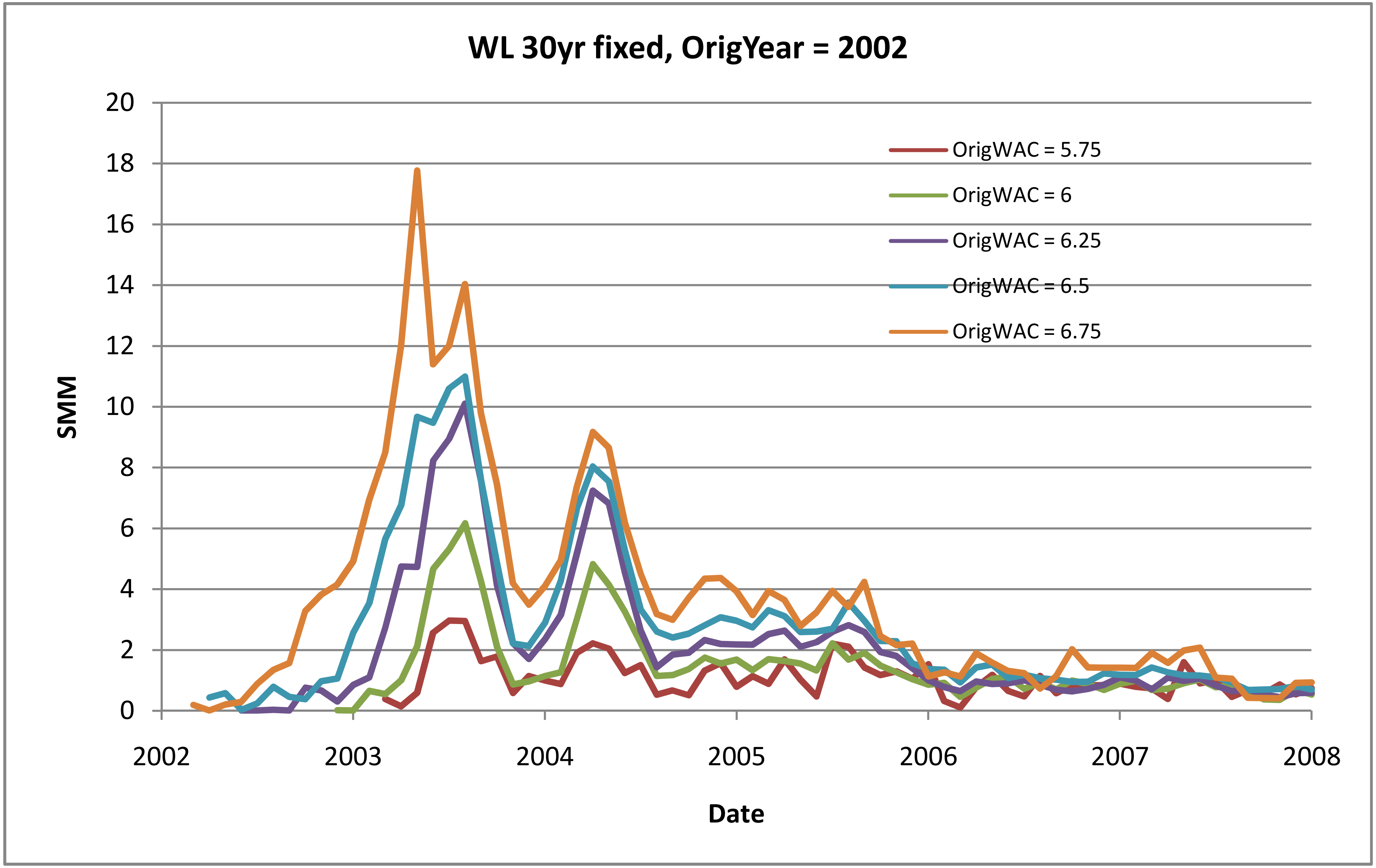

OrigYear = 2001
WAC=7.25, Bal=8.7b$
WAC=7.5, Bal=10b$
WAC=7.75, Bal=3.7b$
SMM
0
2
4
6
8
10
12
14
16
Date
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009

Fig 4



WL 30yr fixed, Orig=2000
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
SMM
9
8.5
8
7.5
7
6.5
6
5.5
5
4.5
4
MB30 Rate
WAC 8.25, w=1/4
WAC 8.5, w=1/4
WAC 8.75, w=1/4
WAC 9, w=1/4
MB30(t-2)
May-2001
Dec-2001
Mar-00
Oct-00
Apr-01
Nov-01
May-02
Dec-02
Jun-03
Jan-04
Date

Fig 5



WL 30yr fixed, OrigYear = 2002
SMM
20
18
16
14
12
10
8
6
4
2
0
2002
2003
2004
2005
2006
2007
2008
Date
OrigWAC = 5.75
OrigWAC = 6
OrigWAC = 6.25
OrigWAC = 6.5
OrigWAC = 6.75

Fig 6

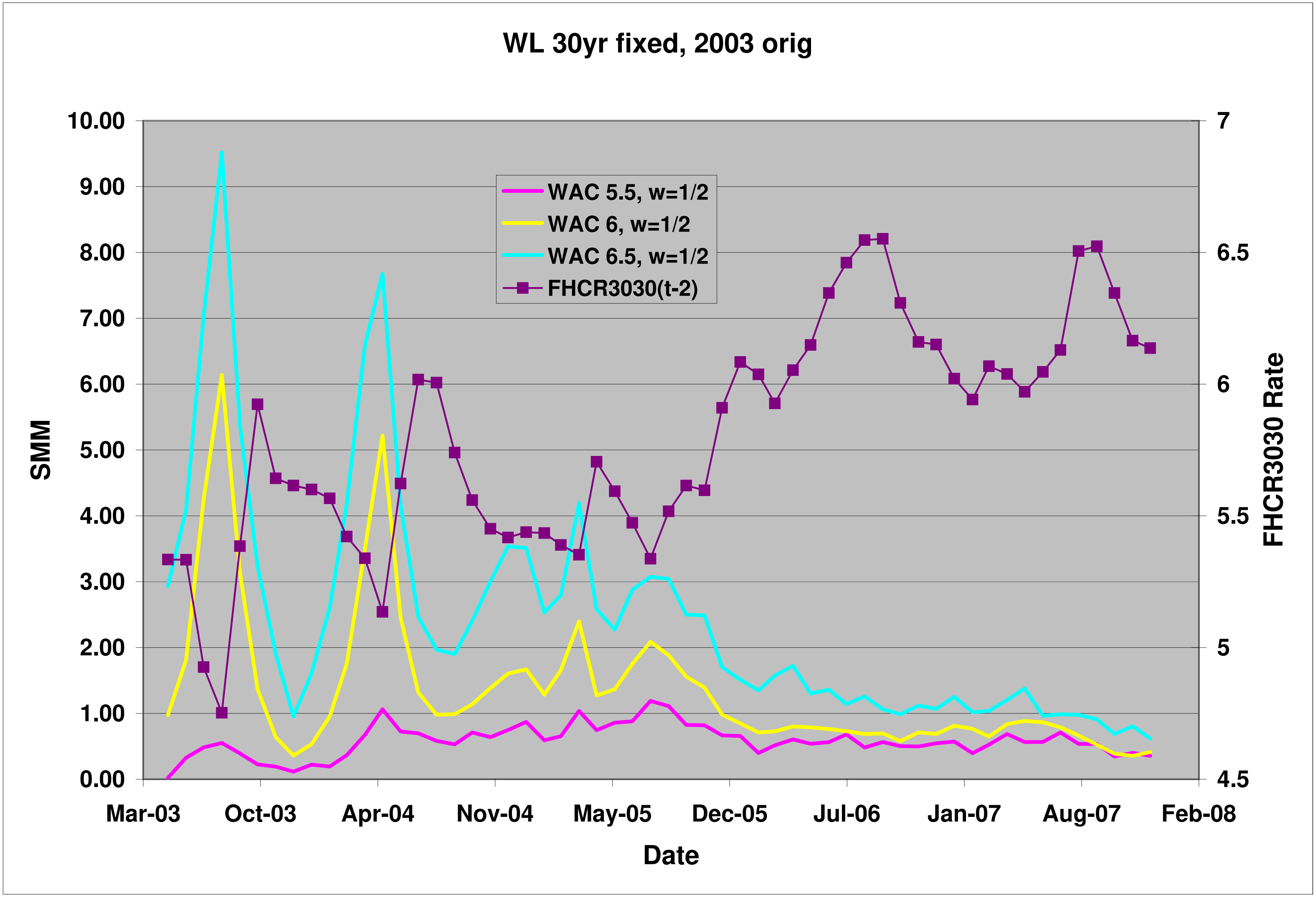
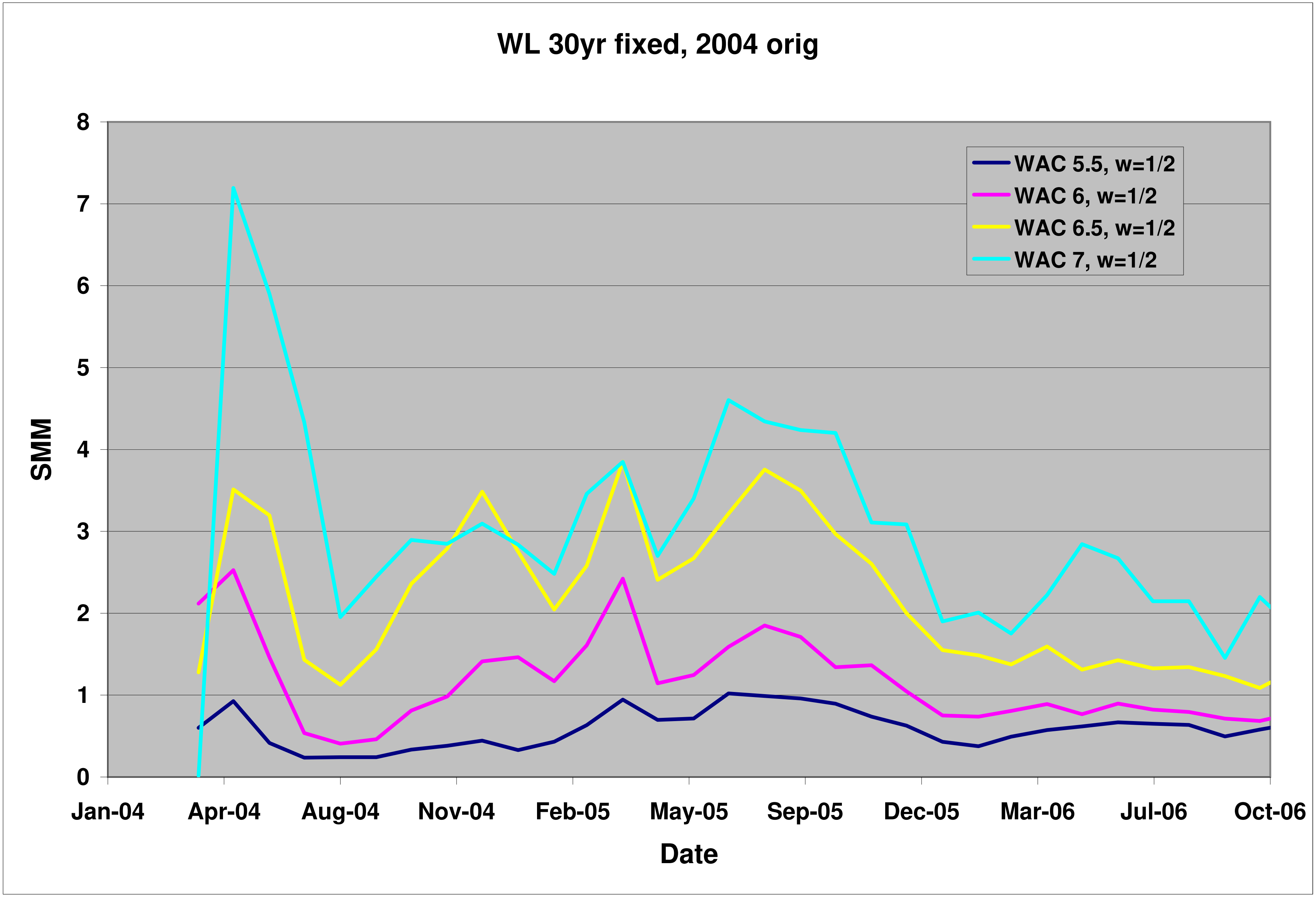

WL 30yr fixed, 2003 orig
WAC 5.5, w=1/2
WAC 6, w=1/2
WAC 6.5, w=1/2
FHCR3030(t-2)
SMM
10.00
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
FHCR3030 Rate
7
6.5
6
5.5
5
4.5
Mar-03
Oct-03
Apr-04
Nov-04
May-05
Dec-05
Jul-06
Jan-07
Aug-07
Feb-08
Date

Fig 7



WL 30yr fixed, 2004 orig
8
7
6
5
4
3
2
1
0
SMM
Jan-04
Apr-04
Aug-04
Nov-04
Feb-05
May-05
Sep-05
Dec-05
Mar-06
Jul-06
Oct-06
Date
WAC 5.5, w=1/2
WAC 6, w=1/2
WAC 6.5, w=1/2
WAC 7, w=1/2

Fig 8

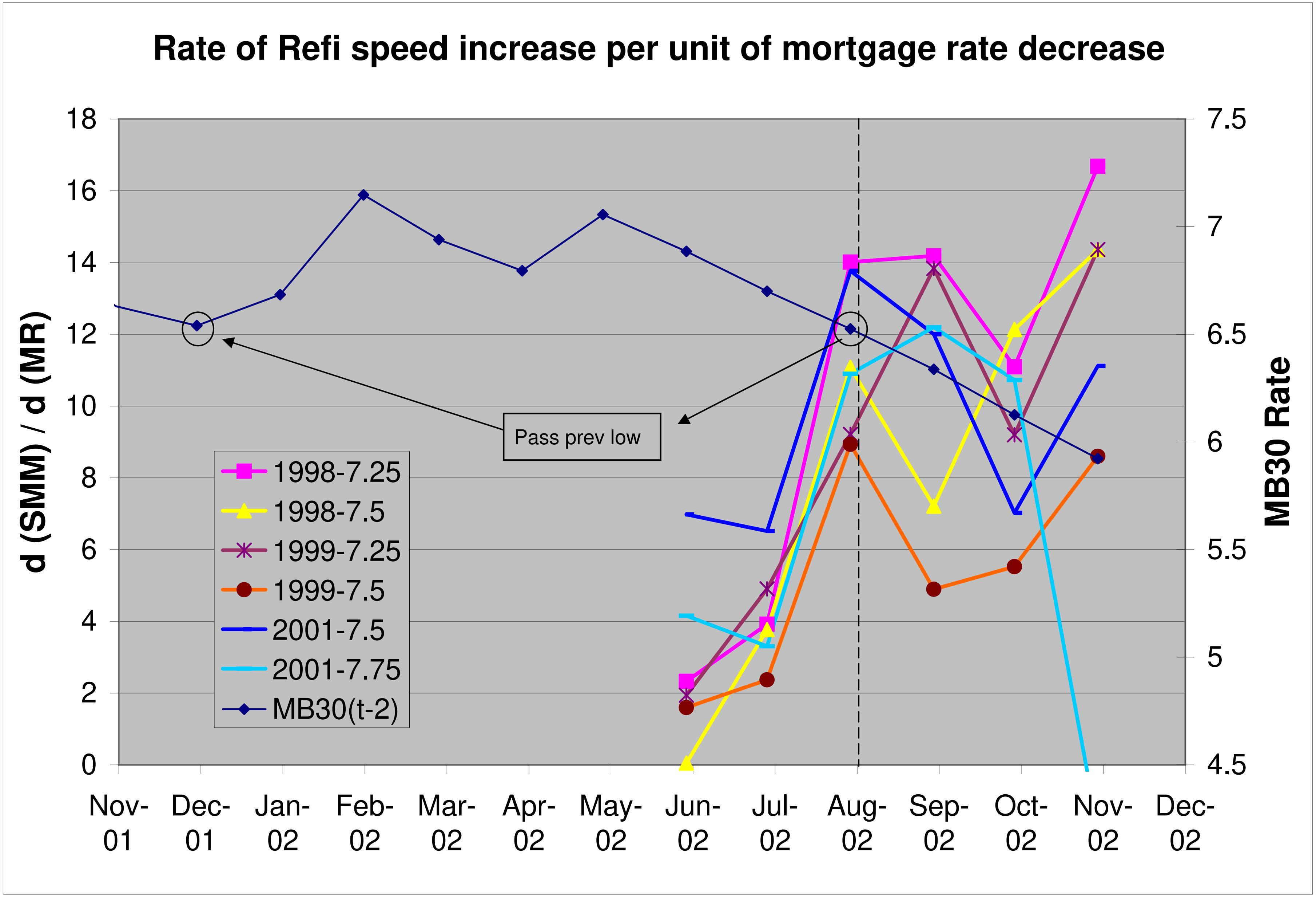
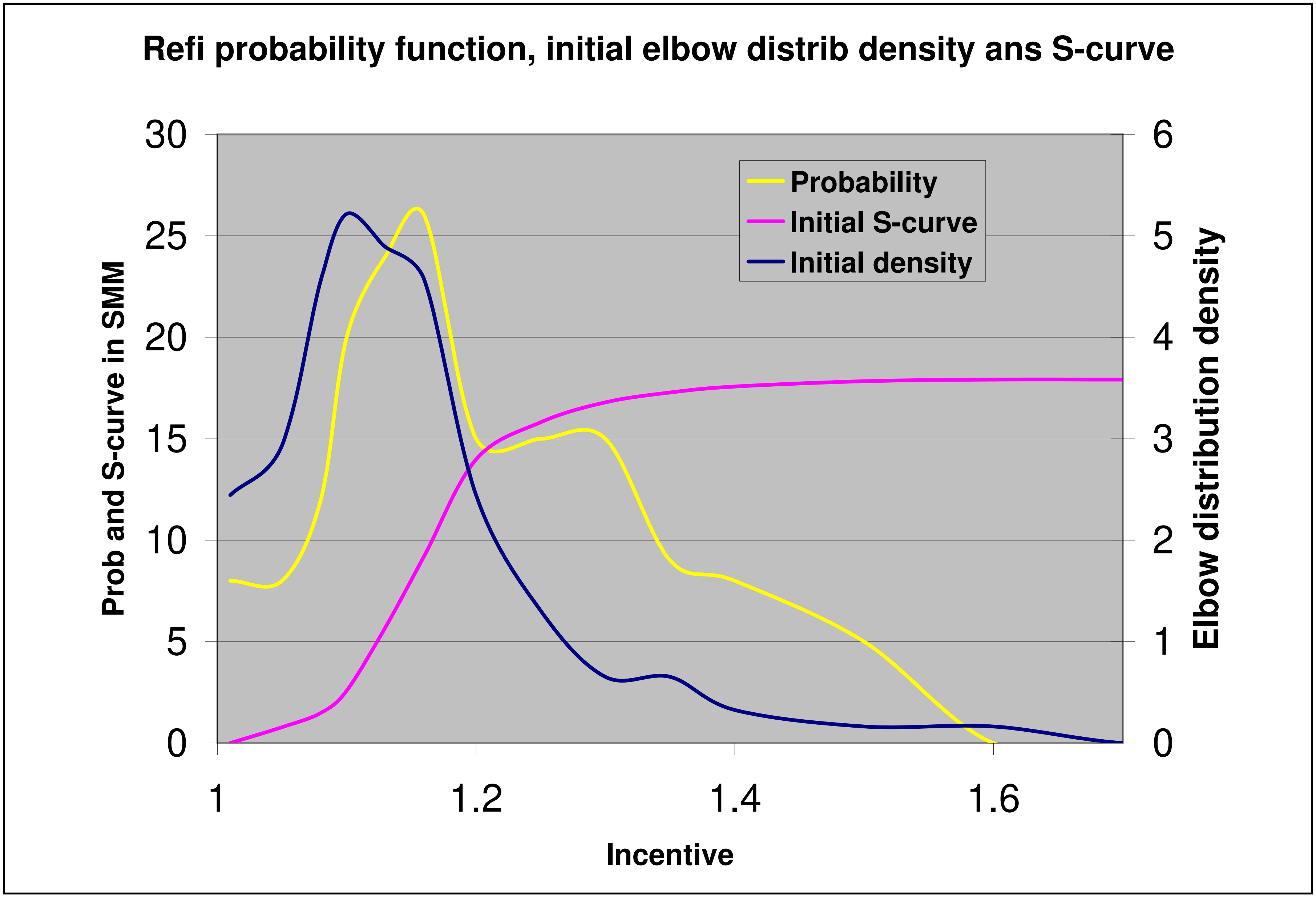

Rate of Refi speed increase per unit of mortgage rate decrease
d (SMM) / d (MR)
MB30 Rate
Pass prev low
1998-7.25
1998-7.5
1999-7.25
1999-7.5
2001-7.5
2001-7.75
MB30(t-2)
0
2
4
6
8
10
12
14
16
18
4.5
5
5.5
6
6.5
7
7.5
Nov-01
Dec-01
Jan-02
Feb-02
Mar-02
Apr-02
May-02
Jun-02
Jul-02
Aug-02
Sep-02
Oct-02
Nov-02
Dec-02

Fig. 9



Refi probability function, initial elbow distrib density ans S-curve
Probability
Initial S-curve
Initial density
Prob and S-curve in SMM
Elbow distribution density
Incentive
30
25
20
15
10
5
0
6
5
4
3
2
1
0
1
1.2
1.4
1.6

Fig 10

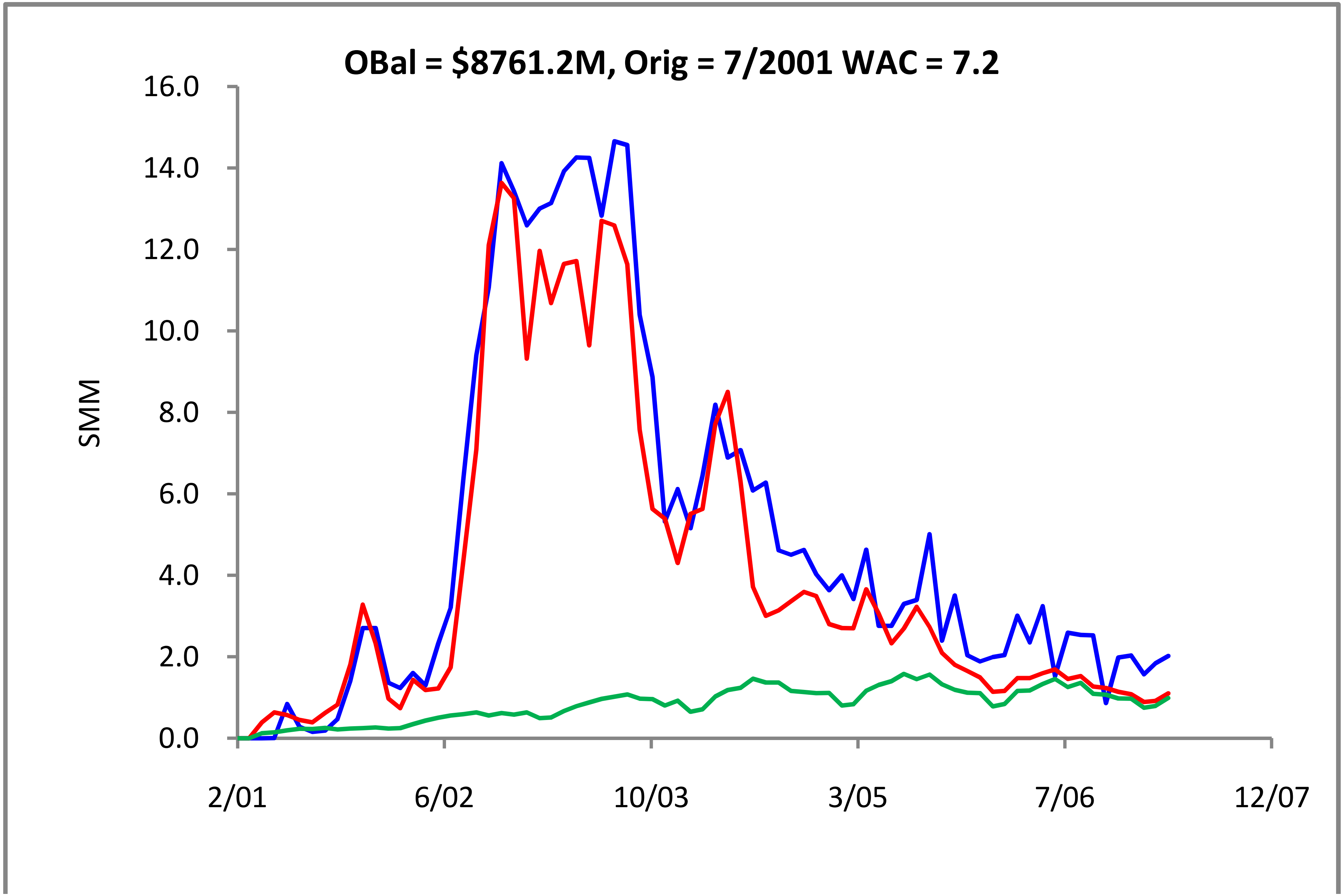

OBal = $8761.2M, Orig = 7/2001 WAC = 7.2
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
SMM
2/01
6/02
10/03
3/05
7/06
12/07

Fig 11a

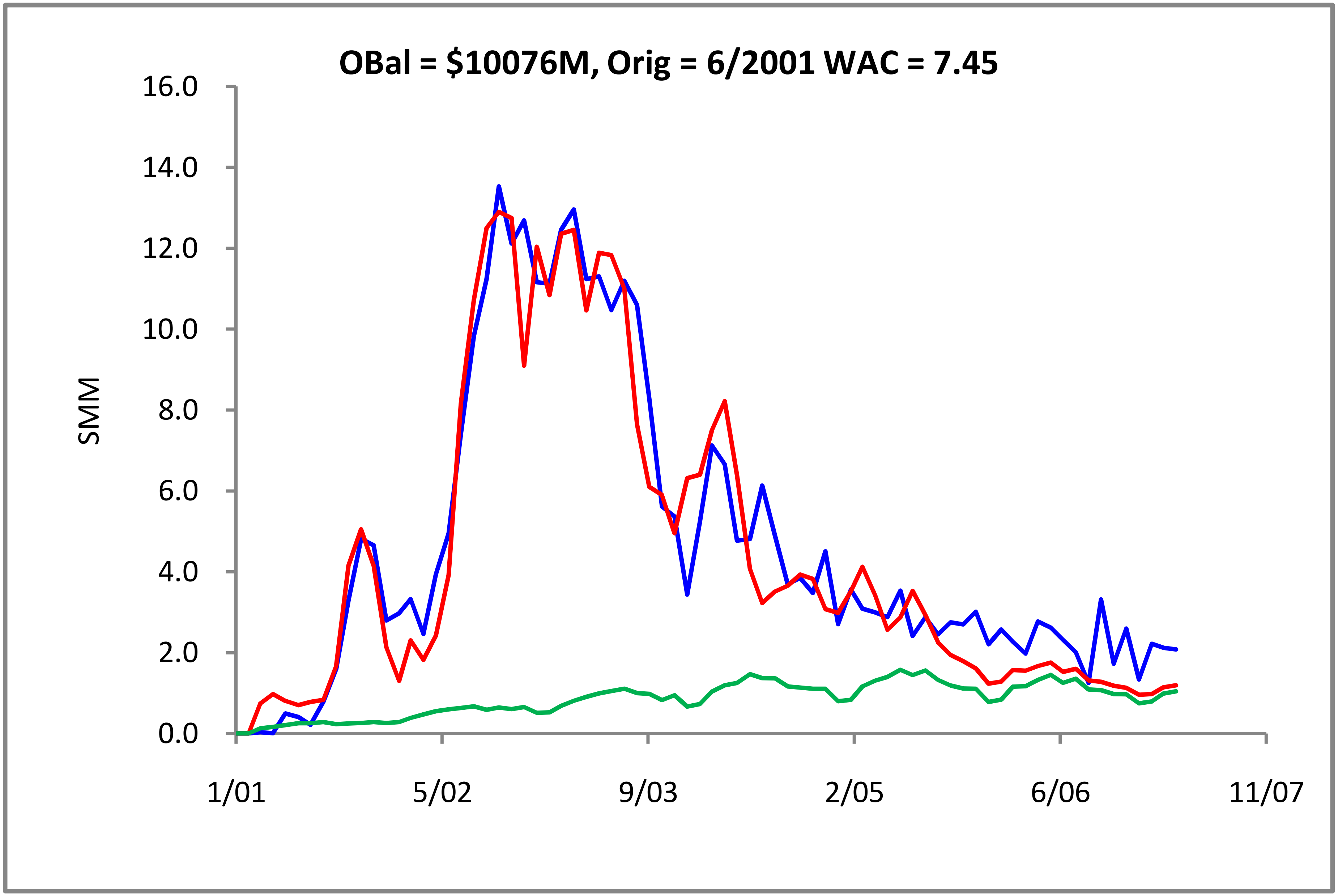

OBal = $10076M, Orig = 6/2001 WAC = 7.45
SMM
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
1/01
5/02
9/03
2/05
6/06
11/07

Fig 11b

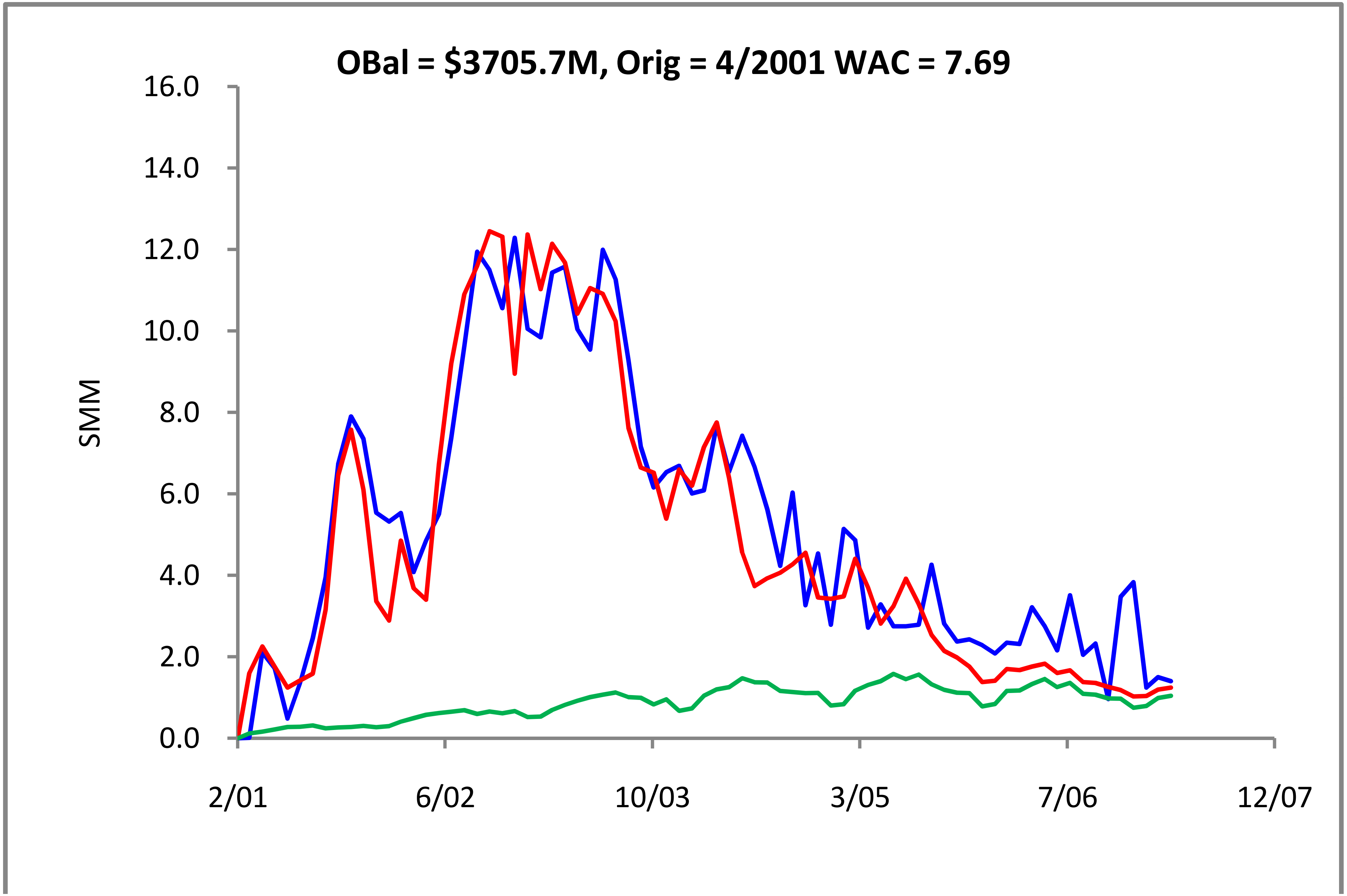

OBal = $3705.7M, Orig = 4/2001 WAC = 7.69
SMM
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
2/01
6/02
10/03
3/05
7/06
12/07

Fig 11c

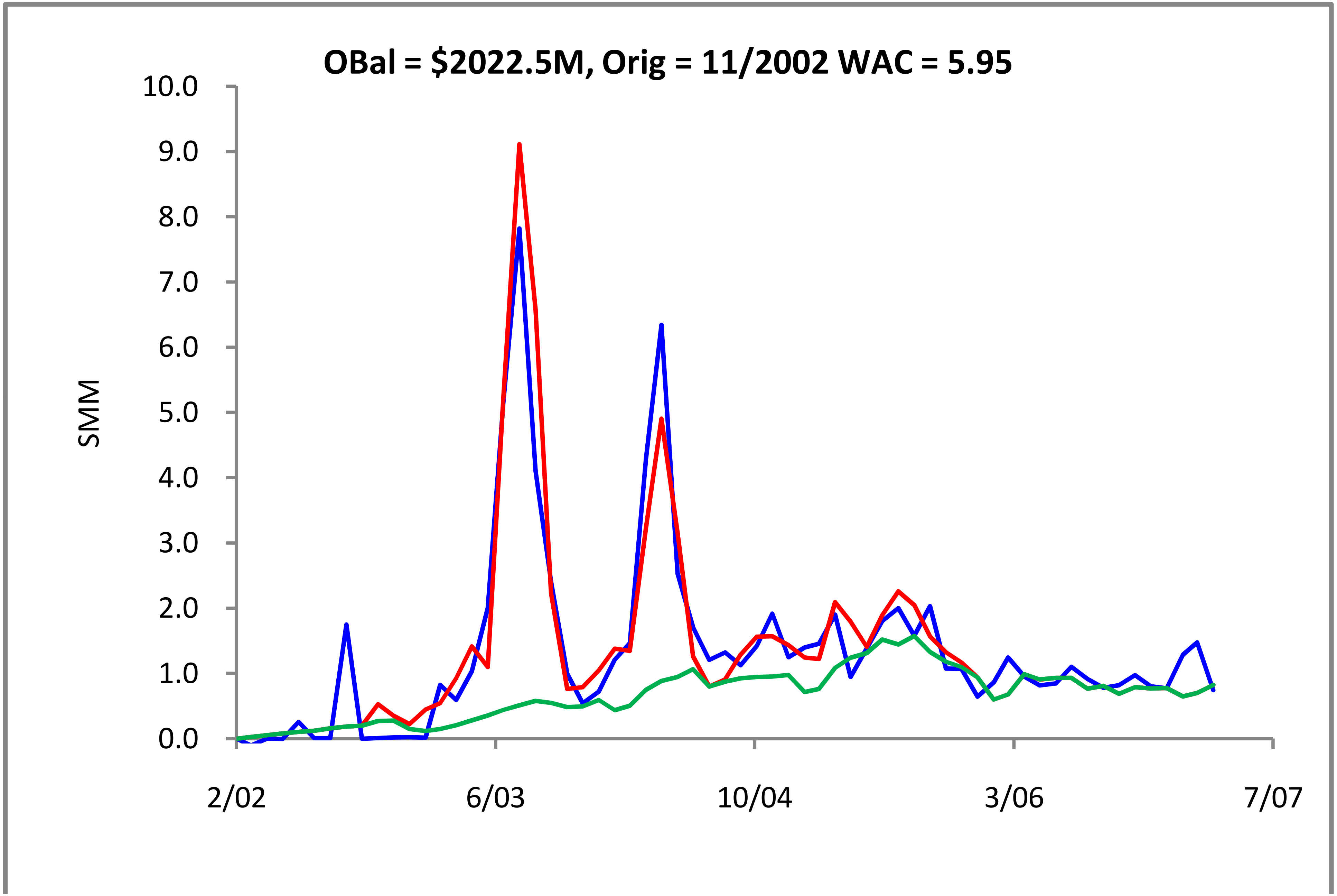
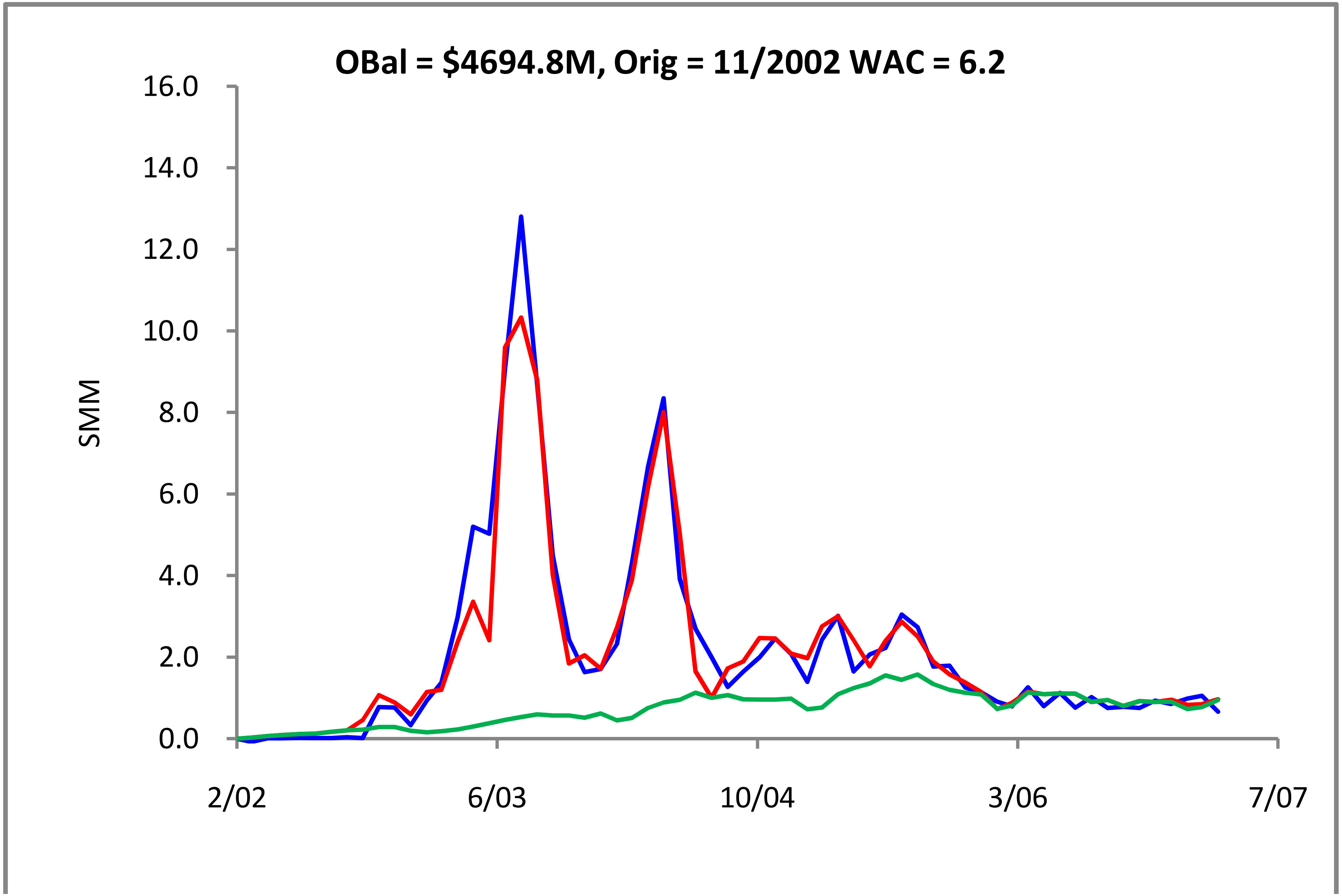

OBal = $2022.5M, Orig = 11/2002 WAC = 5.95
SMM
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
2/02
6/03
10/04
3/06
7/07

Fig 12a



OBal = $4694.8M, Orig = 11/2002 WAC = 6.2
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
SMM
2/02
6/03
10/04
3/06
7/07

Fig 12b

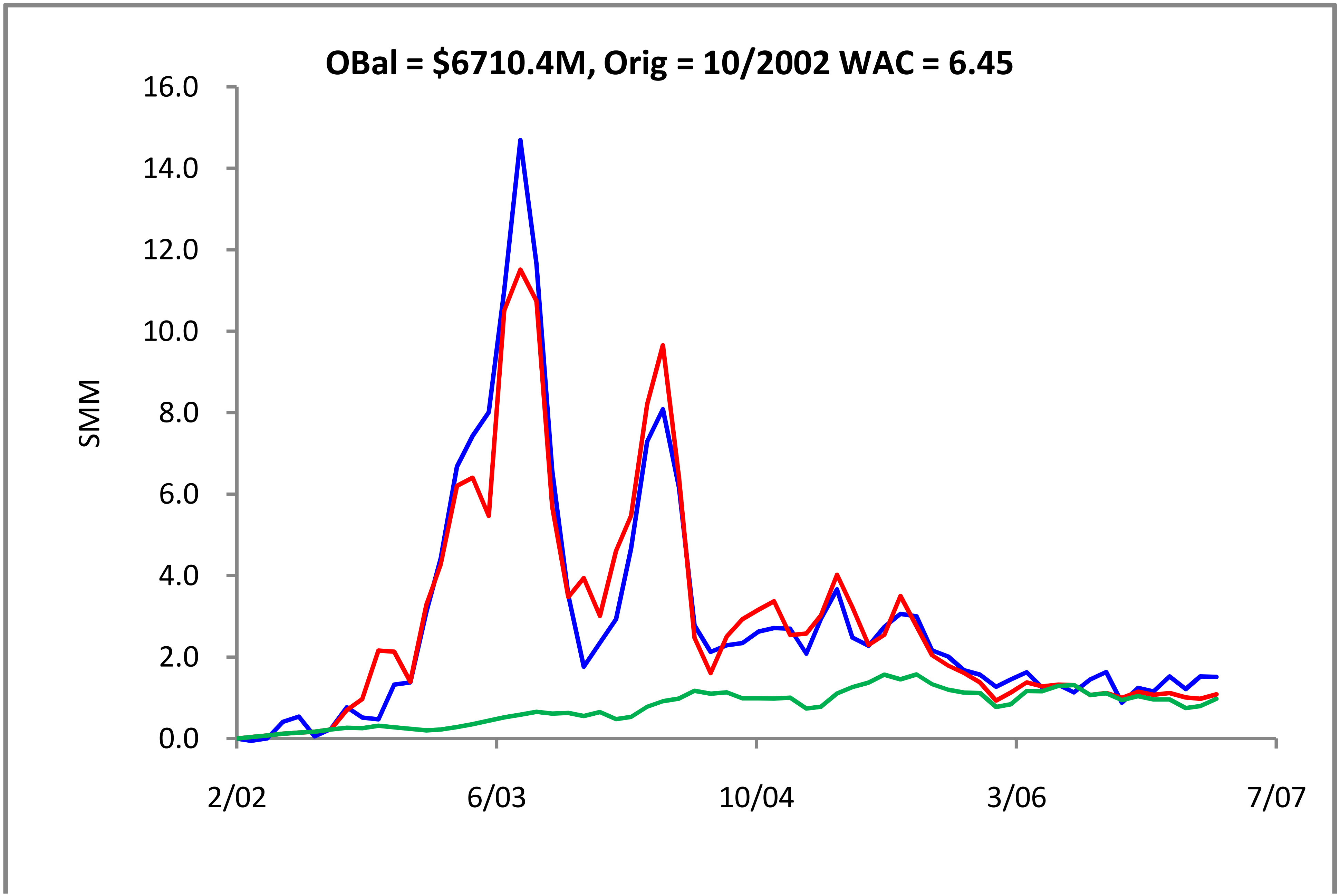
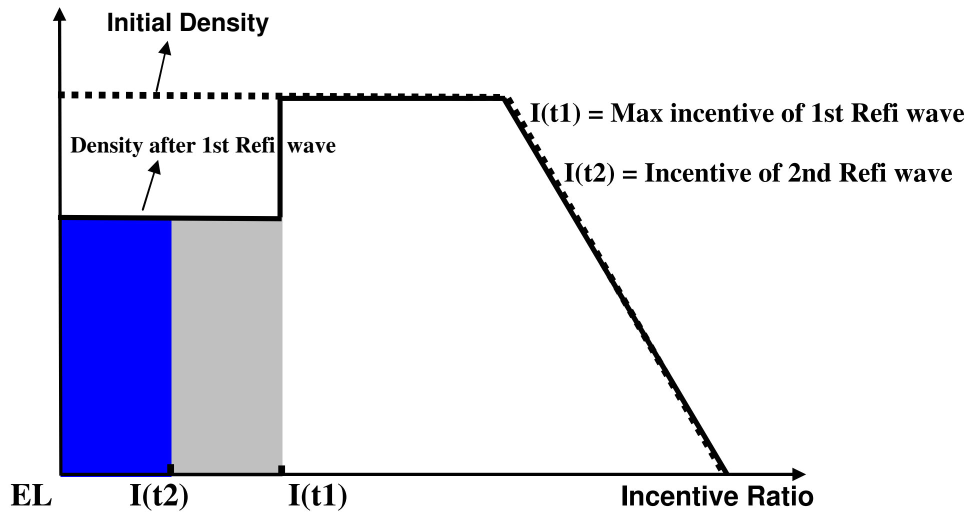

OBal = $6710.4M, Orig = 10/2002 WAC = 6.45
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
SMM
2/02
6/03
10/04
3/06
7/07

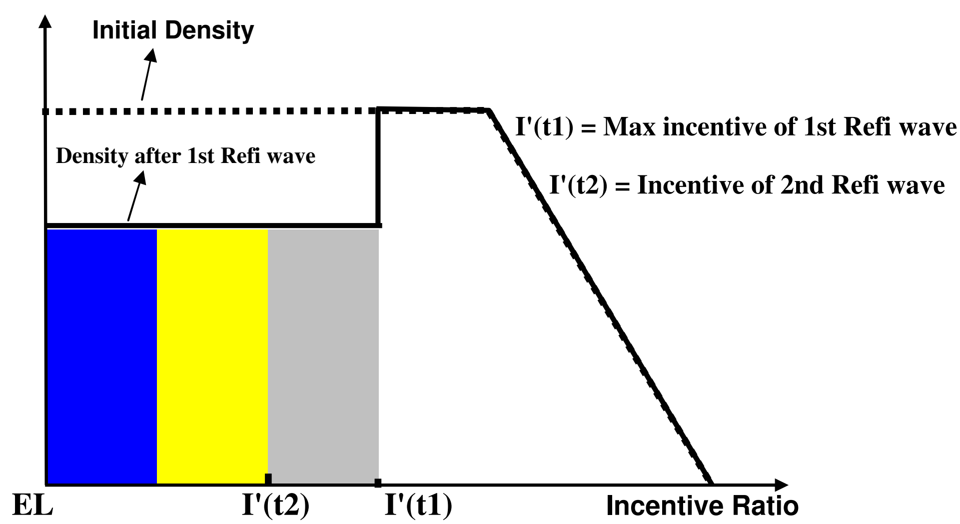
Fig 12c

Mortgage rates not surpassing prev low: High Coupon --> high prepay

Population Densities: Cohort A with lower coupon

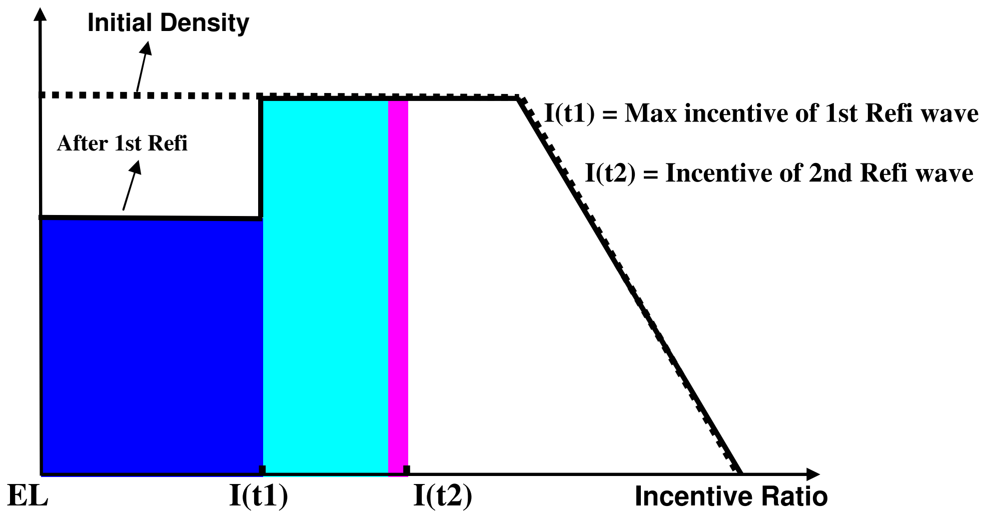


Population Densities: Cohort B with higher coupon



Mortgage rates surpass prev low: High Coupon - MAY NOT-> high prepay

Population Densities: Cohort A with lower coupon



Population Densities: Cohort B with higher coupon

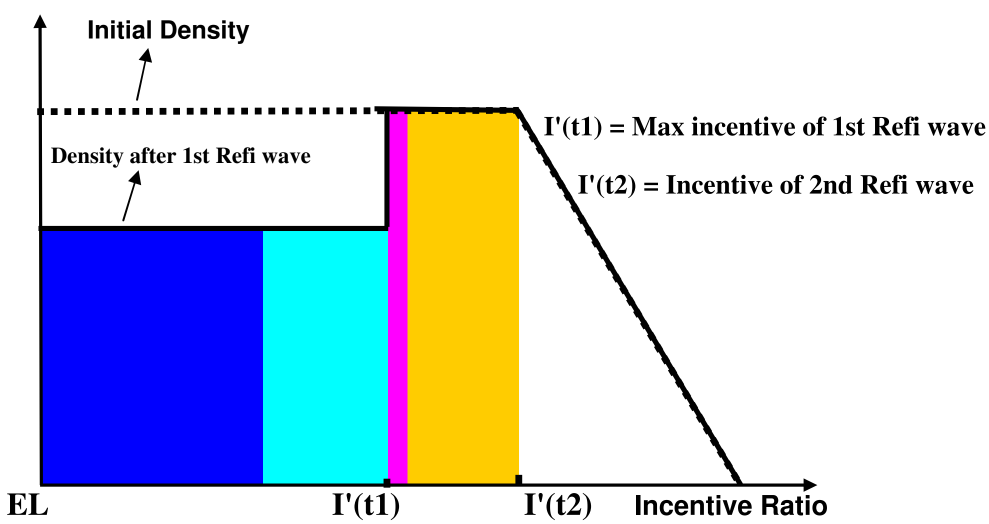


Fig. 13b

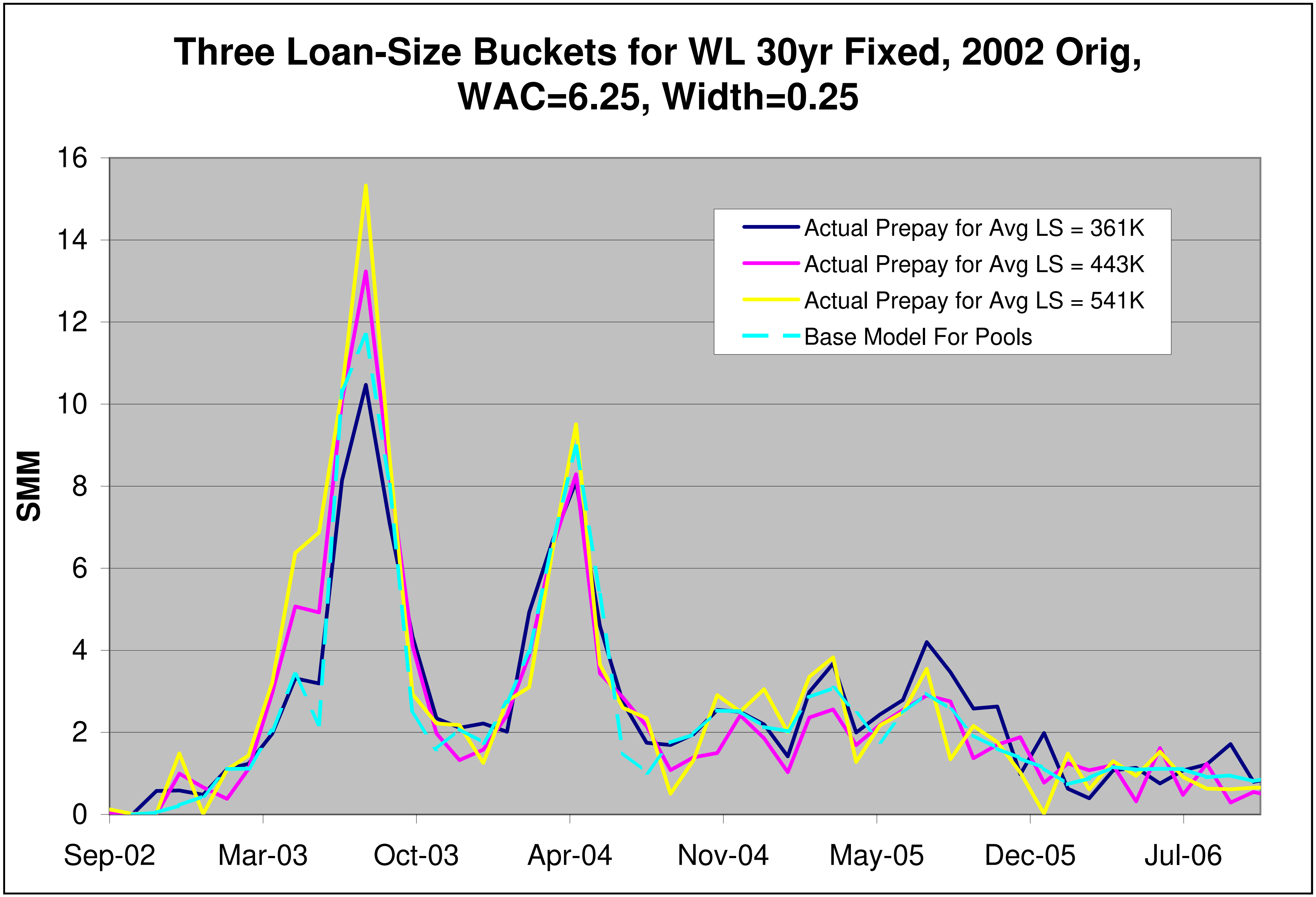

Three Loan-Size Buckets for WL 30yr Fixed, 2002 Orig, WAC=6.25, Width=0.25
SMM
0
2
4
6
8
10
12
14
16
Sep-02
Mar-03
Oct-03
Apr-04
Nov-04
May-05
Dec-05
Jul-06
Actual Prepay for Avg LS = 361K
Actual Prepay for Avg LS = 443K
Actual Prepay for Avg LS = 541K
Base Model For Pools

Fig 14

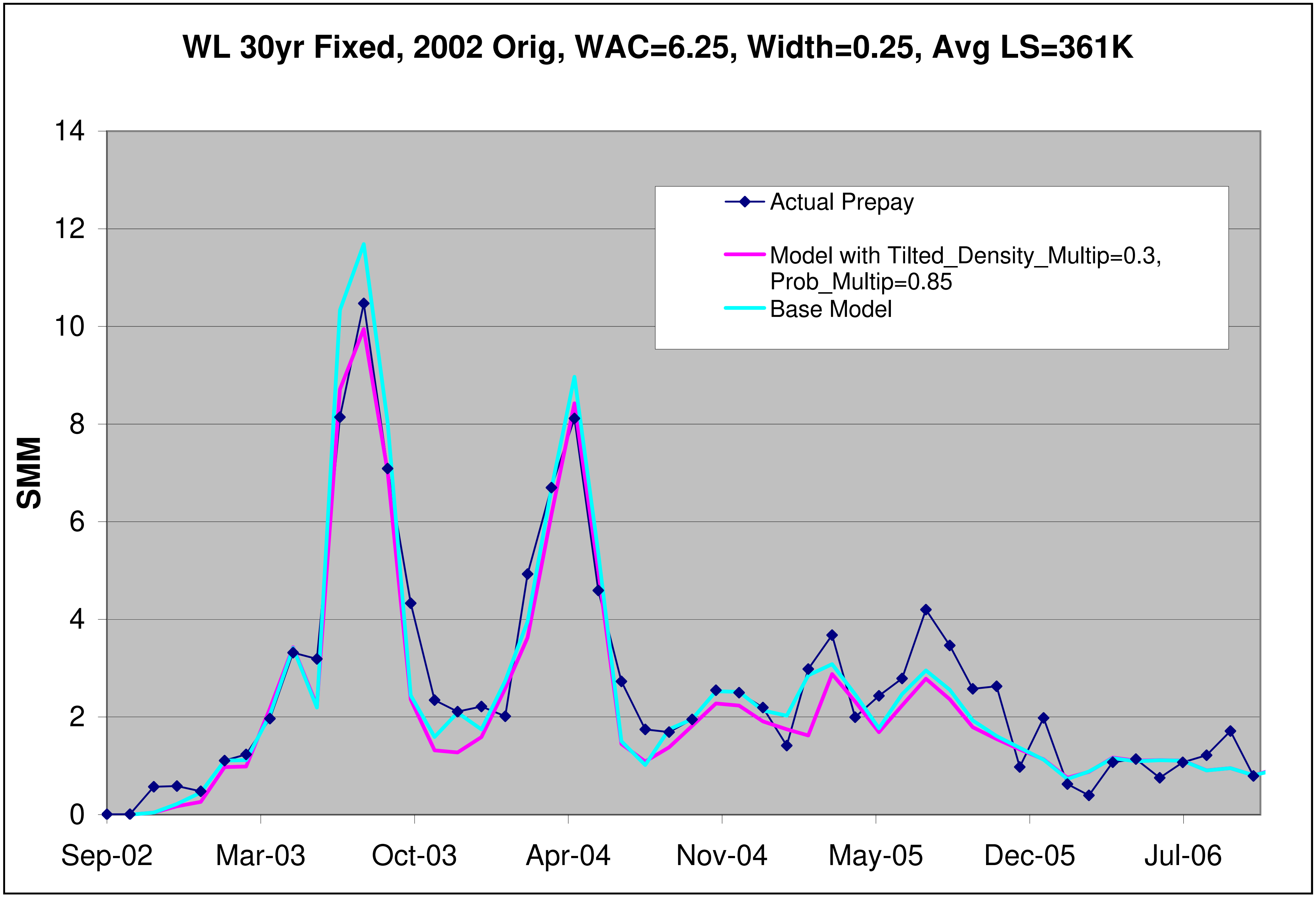

WL 30yr Fixed, 2002 Orig, WAC=6.25, Width=0.25, Avg LS=361K
SMM
0
2
4
6
8
10
12
14
Sep-02
Mar-03
Oct-03
Apr-04
Nov-04
May-05
Dec-05
Jul-06
Actual Prepay
Model with Tilted_Density_Multip=0.3, Prob_Multip=0.85
Base Model

Fig 15a

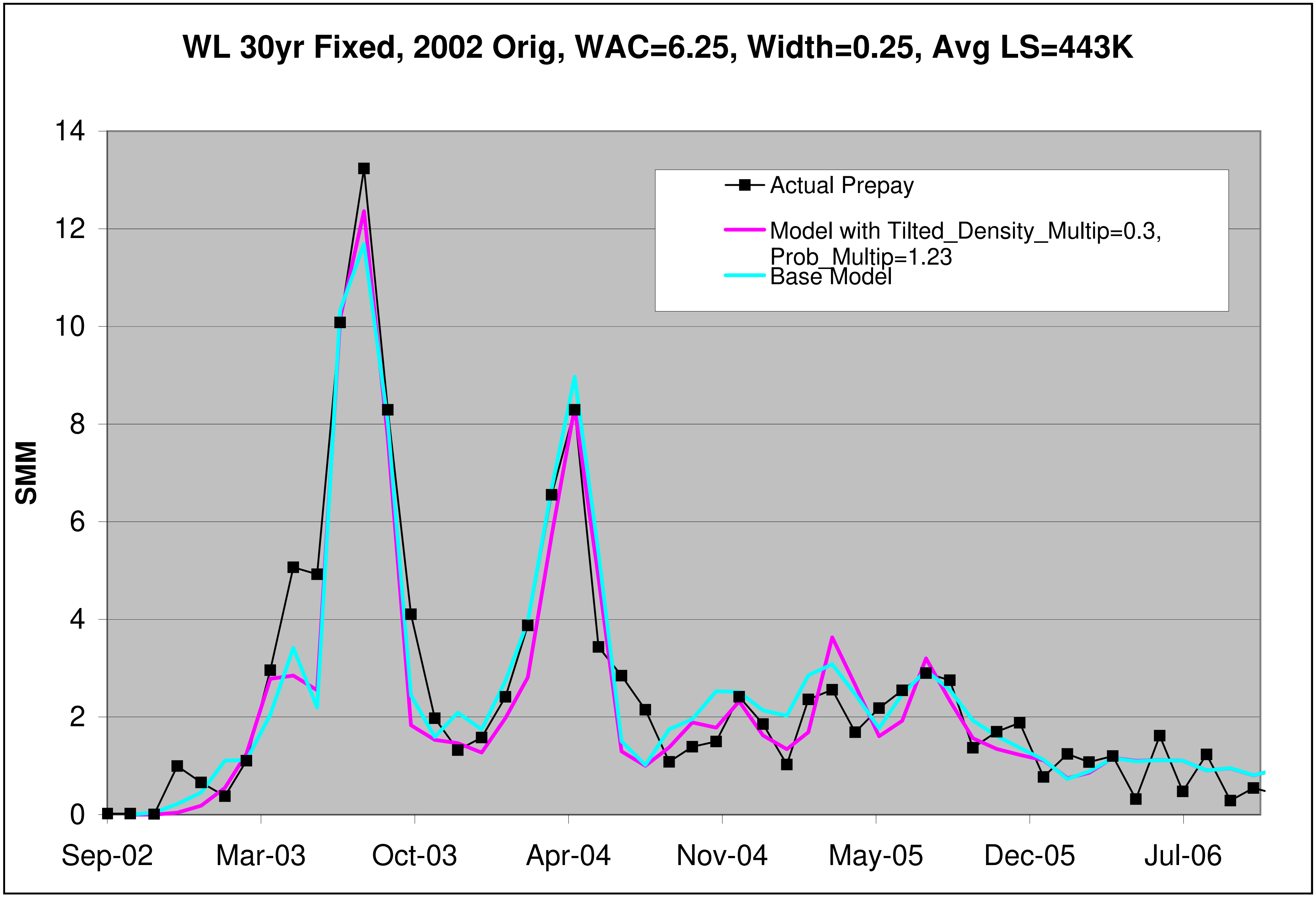

WL 30yr Fixed, 2002 Orig, WAC=6.25, Width=0.25, Avg LS=443K
Actual Prepay
Model with Tilted_Density_Multip=0.3, Prob_Multip=1.23
Base Model
SMM
14
12
10
8
6
4
2
0
Sep-02
Mar-03
Oct-03
Apr-04
Nov-04
May-05
Dec-05
Jul-06

Fig 15b

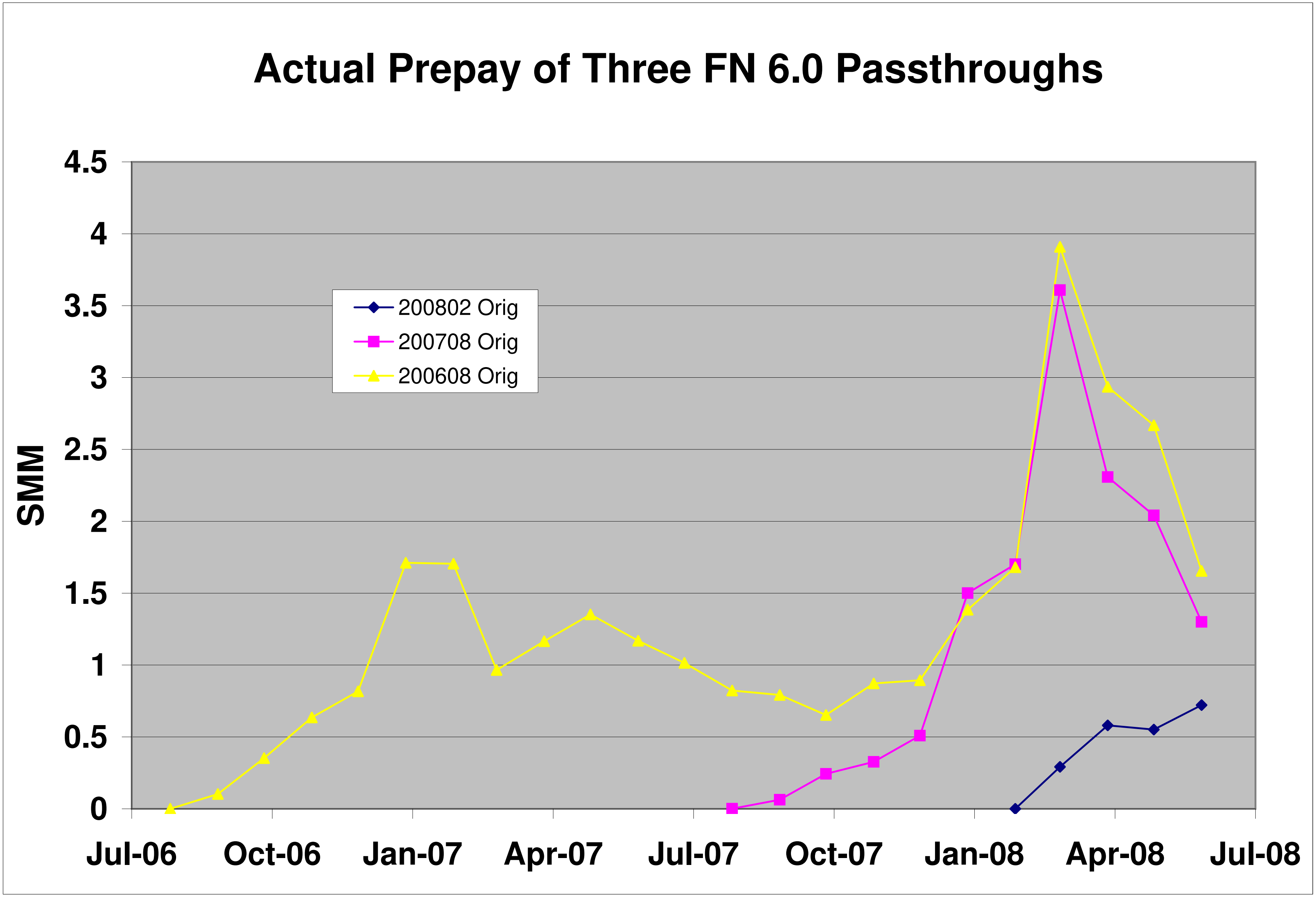

Actual Prepay of Three FN 6.0 Passthroughs
SMM
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Jul-06
Oct-06
Jan-07
Apr-07
Jul-07
Oct-07
Jan-08
Apr-08
Jul-08
200802 Orig
200708 Orig
200608 Orig

Fig 16

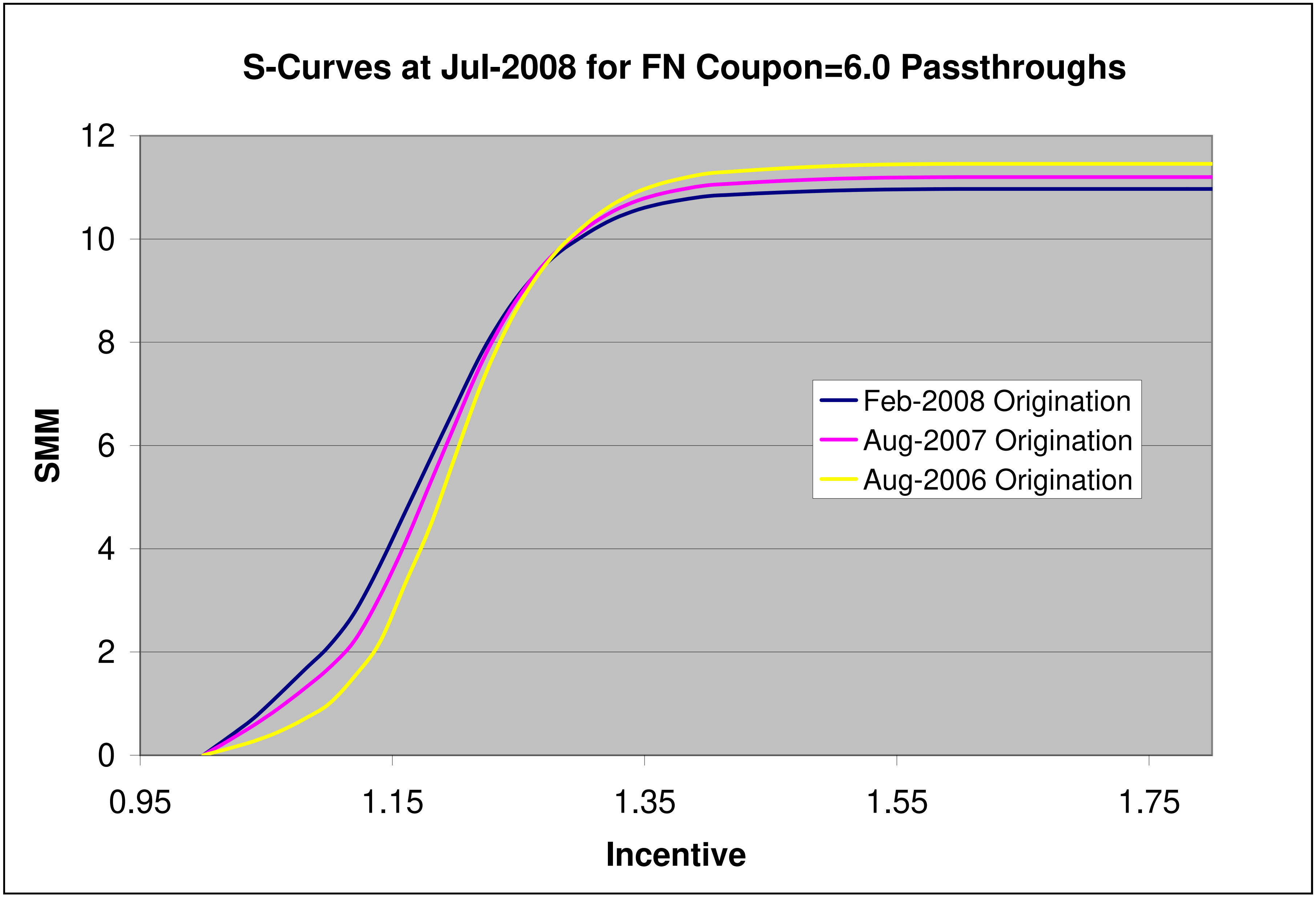

S-Curves at Jul-2008 for FN Coupon=6.0 Passthroughs
SMM
12
10
8
6
4
2
0
0.95
1.15
1.35
1.55
1.75
Incentive
Feb-2008 Origination
Aug-2007 Origination
Aug-2006 Origination

Fig 17

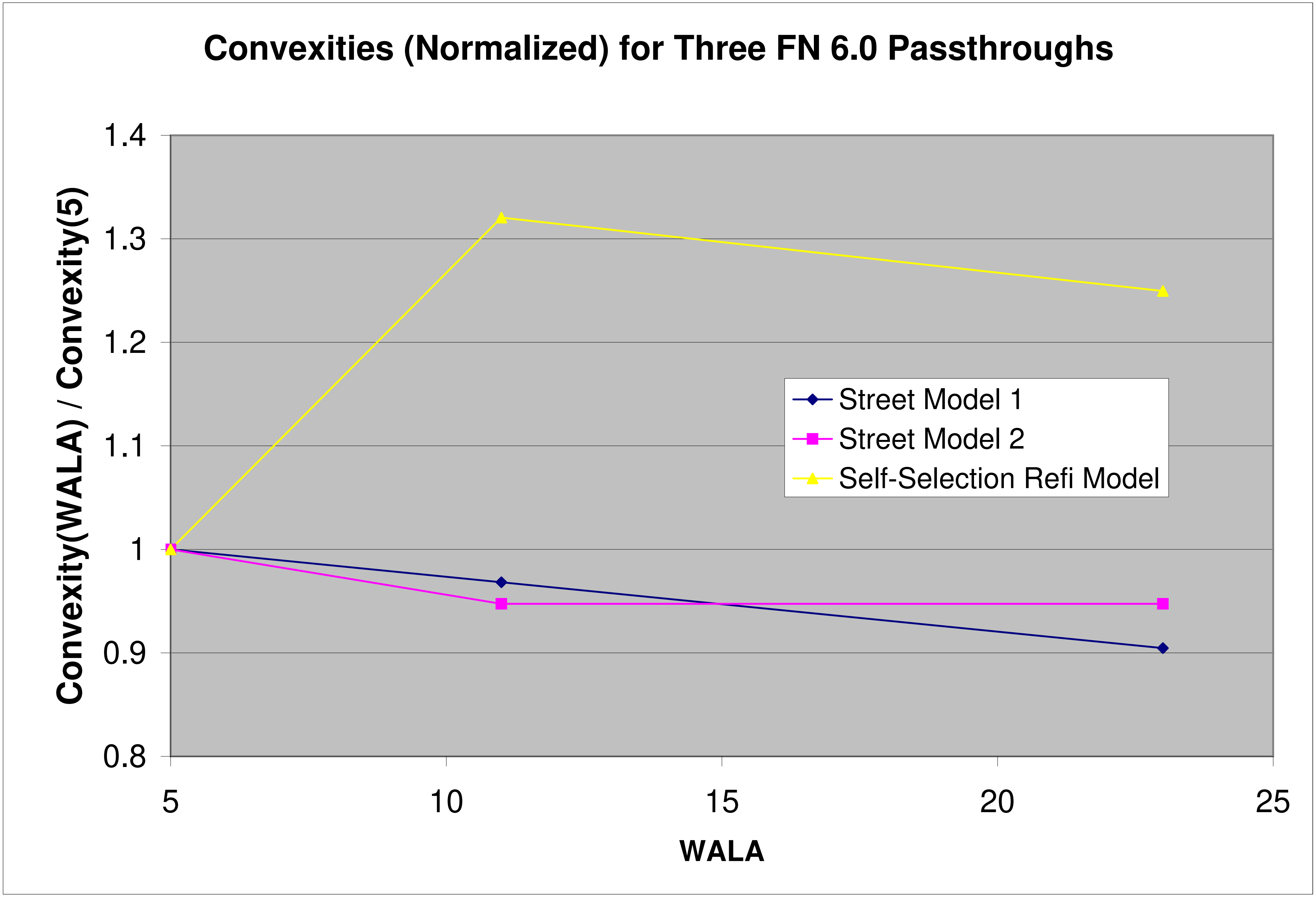
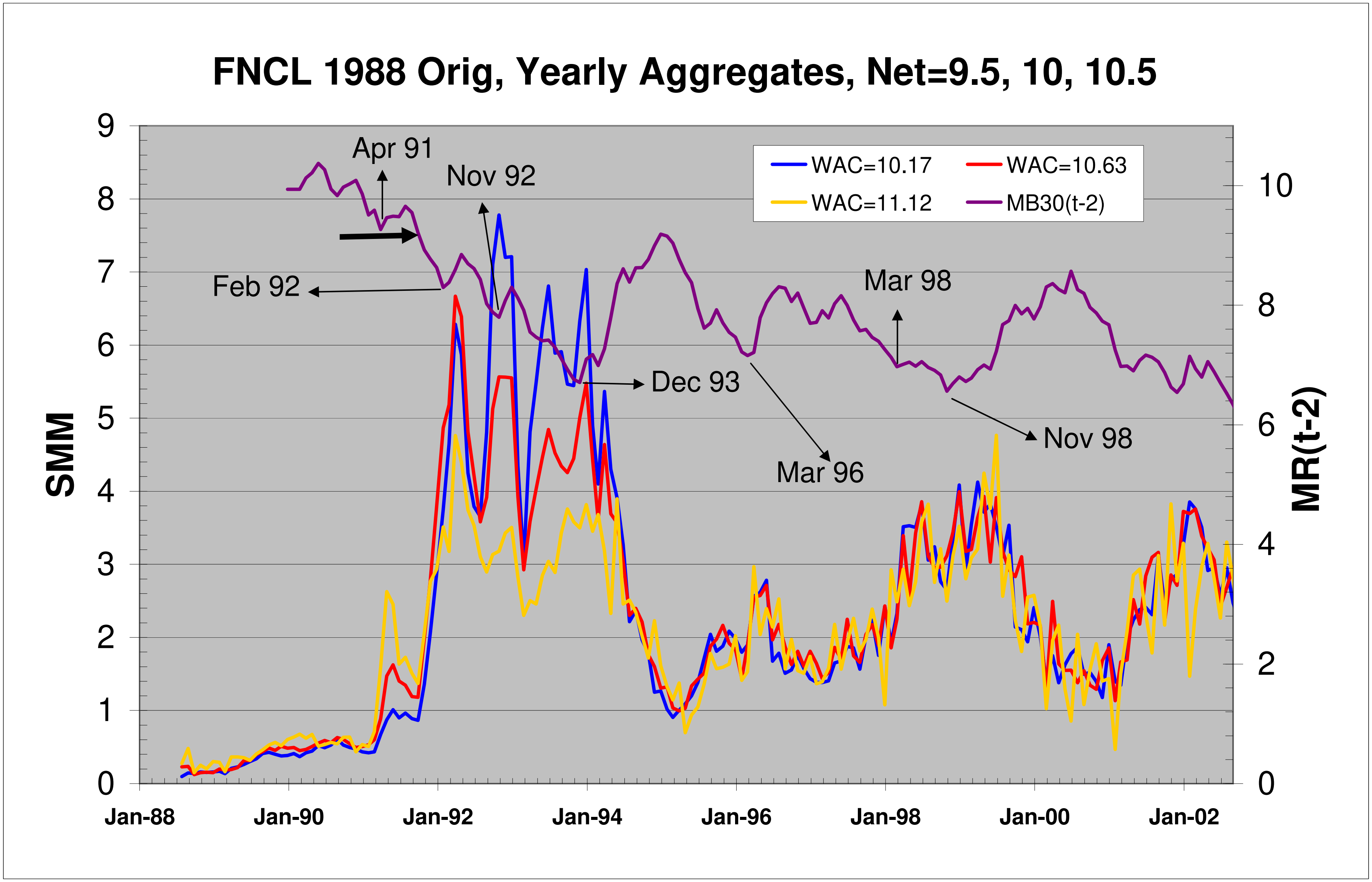
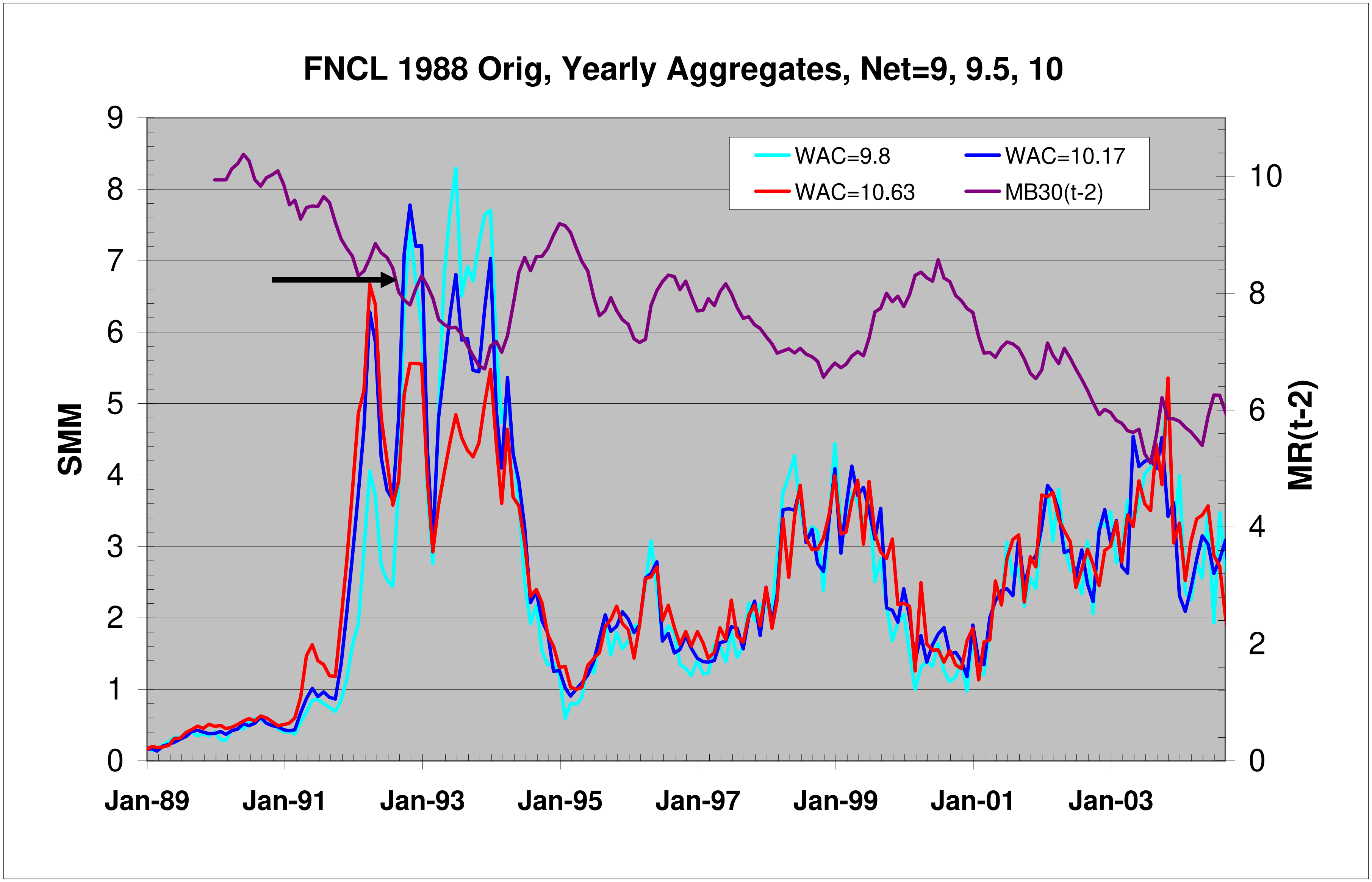

Convexities (Normalized) for Three FN 6.0 Passthroughs
Convexity(WALA) / Convexity(5)
WALA
1.4
1.3
1.2
1.1
1
0.9
0.8
5
10
15
20
25
Street Model 1
Street Model 2
Self-Selection Refi Model

Fig 18



FNCL 1988 Orig, Yearly Aggregates, Net=9.5, 10, 10.5
SMM
MR(t-2)
WAC=10.17
WAC=10.63
WAC=11.12
MB30(t-2)
Apr 91
Nov 92
Feb 92
Dec 93
Mar 96
Mar 98
Nov 98
9
8
7
6
5
4
3
2
1
0
10
8
6
4
2
0
Jan-88
Jan-90
Jan-92
Jan-94
Jan-96
Jan-98
Jan-00
Jan-02

Fig 19a



FNCL 1988 Orig, Yearly Aggregates, Net=9, 9.5, 10
SMM
MR(t-2)
WAC=9.8
WAC=10.17
WAC=10.63
MB30(t-2)
0
1
2
3
4
5
6
7
8
9
10
Jan-89
Jan-91
Jan-93
Jan-95
Jan-97
Jan-99
Jan-01
Jan-03

Fig 19b

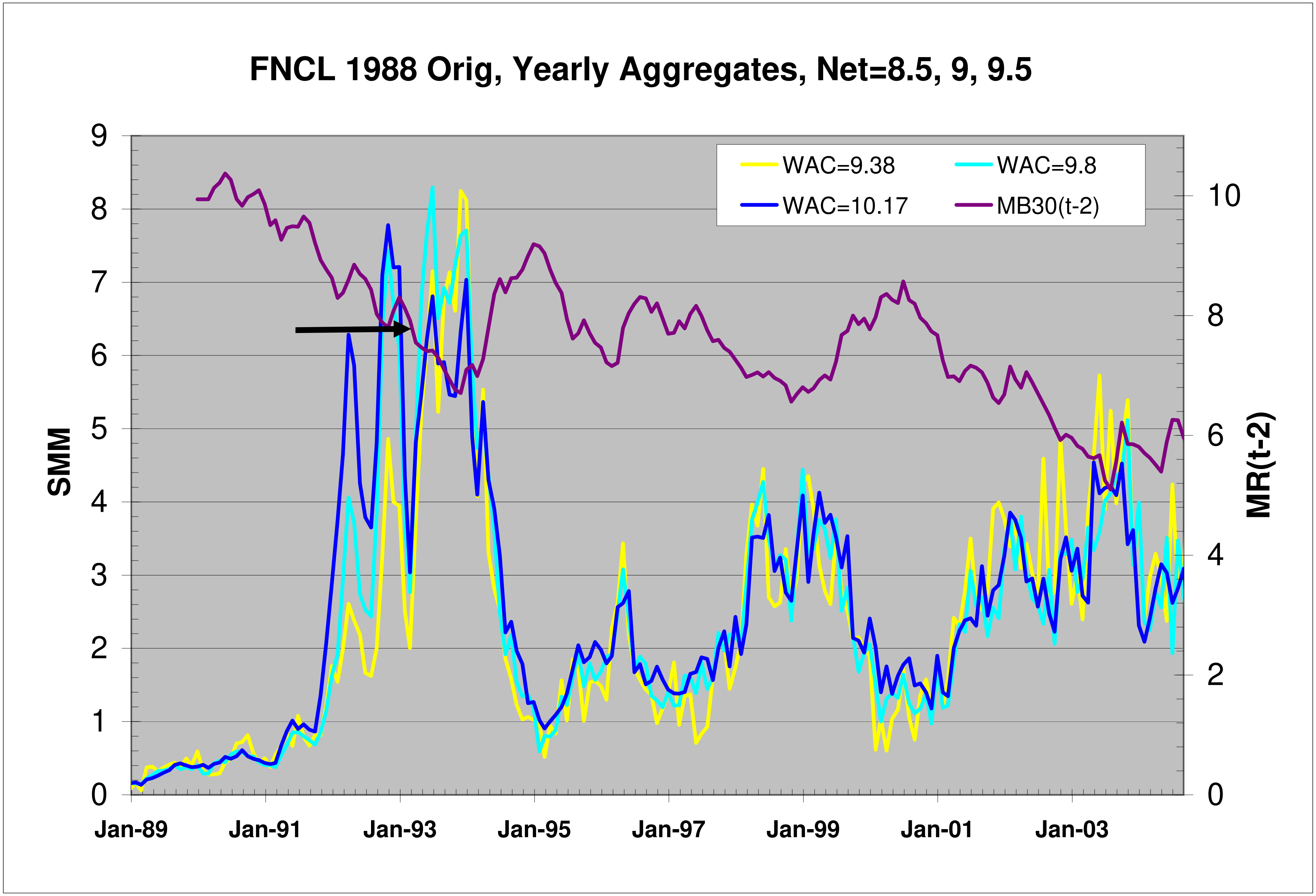

FNCL 1988 Orig, Yearly Aggregates, Net=8.5, 9, 9.5
SMM
MR(t-2)
WAC=9.38
WAC=9.8
WAC=10.17
MB30(t-2)
0
1
2
3
4
5
6
7
8
9
10
Jan-89
Jan-91
Jan-93
Jan-95
Jan-97
Jan-99
Jan-01
Jan-03

Fig 19c

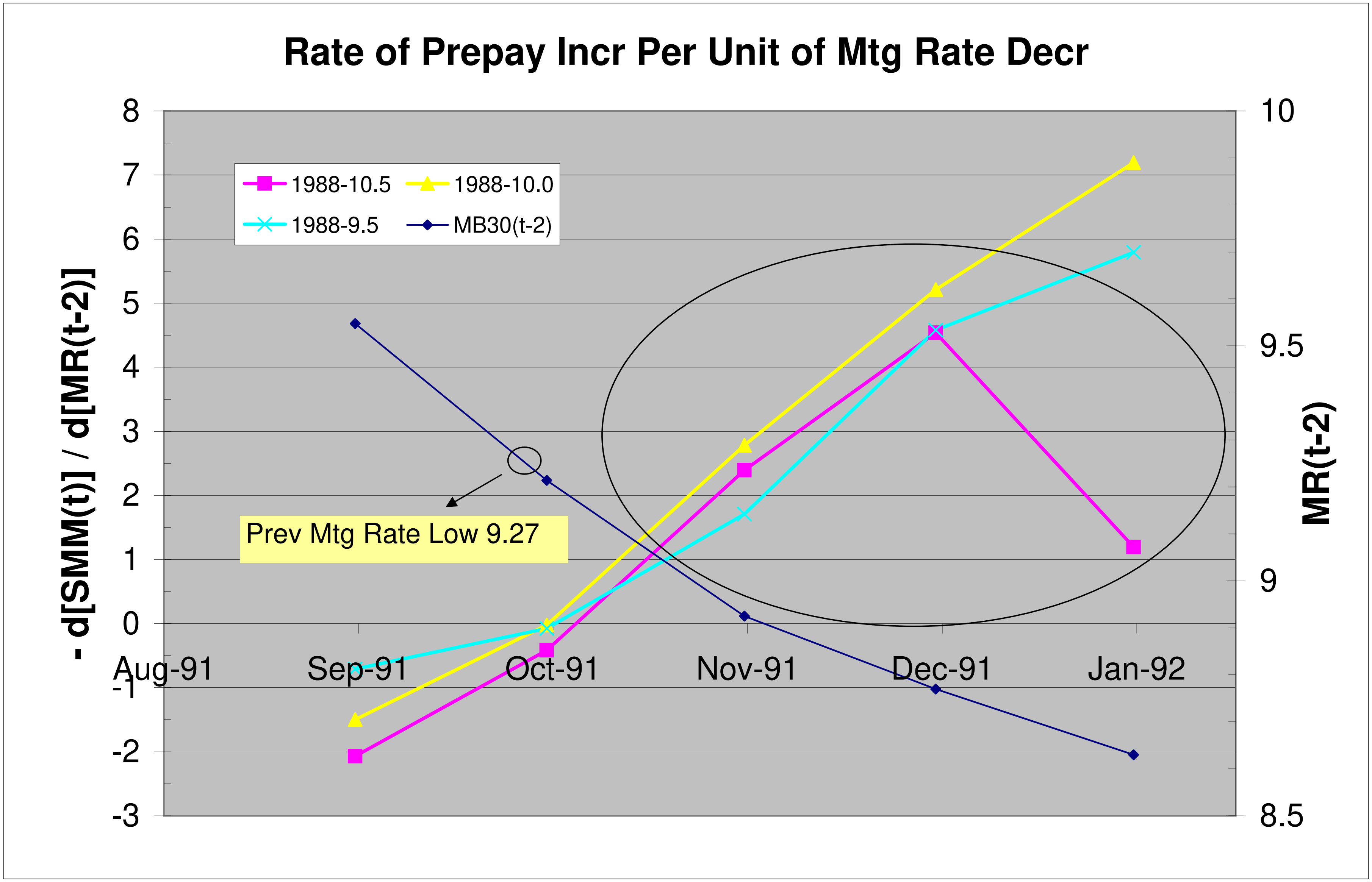

Rate of Prepay Incr Per Unit of Mtg Rate Decr
- d[SMM(t)] / d[MR(t-2)]
MR(t-2)
1988-10.5
1988-10.0
1988-9.5
MB30(t-2)
Prev Mtg Rate Low 9.27
8
7
6
5
4
3
2
1
0
-1
-2
-3
10
9.5
9
8.5
Aug-91
Sep-91
Oct-91
Nov-91
Dec-91
Jan-92

Fig. 20a

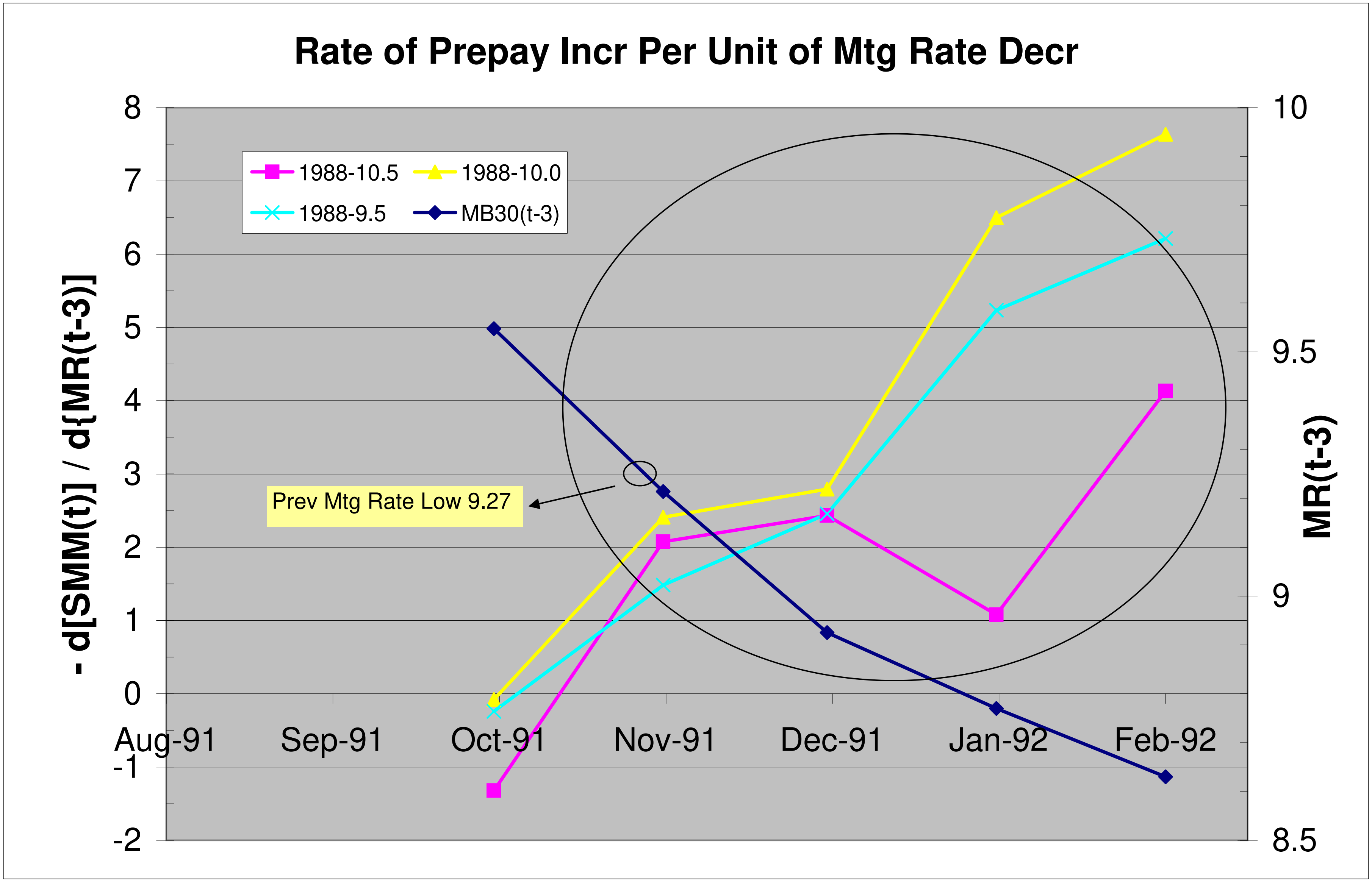

Rate of Prepay Incr Per Unit of Mtg Rate Decr
1988-10.5
1988-10.0
1988-9.5
MB30(t-3)
Prev Mtg Rate Low 9.27
- d[SMM(t)] / d{MR(t-3)]
MR(t-3)
8
7
6
5
4
3
2
1
0
-1
-2
10
9.5
9
8.5
Aug-91
Sep-91
Oct-91
Nov-91
Dec-91
Jan-92
Feb-92

Fig. 20b

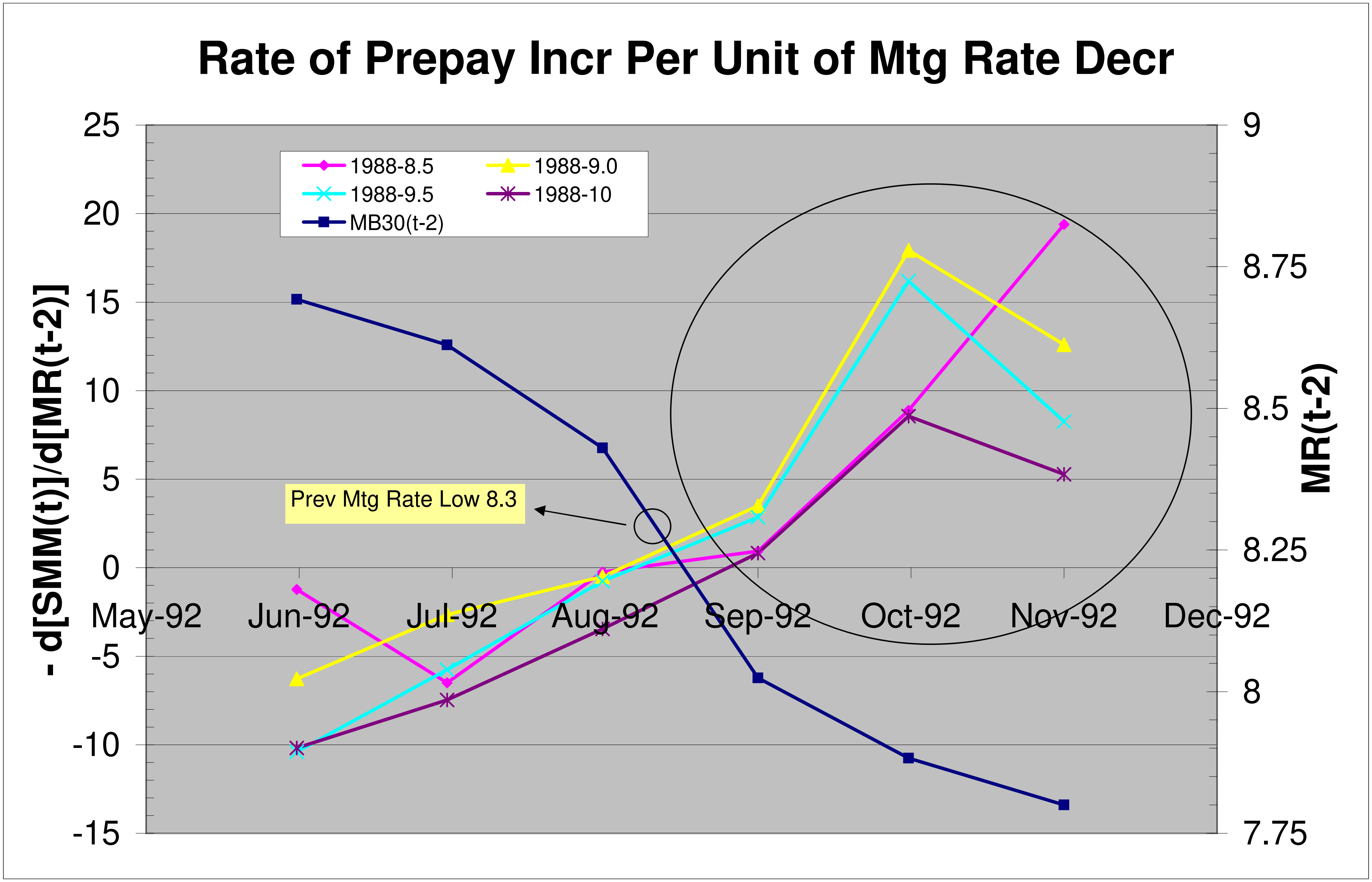
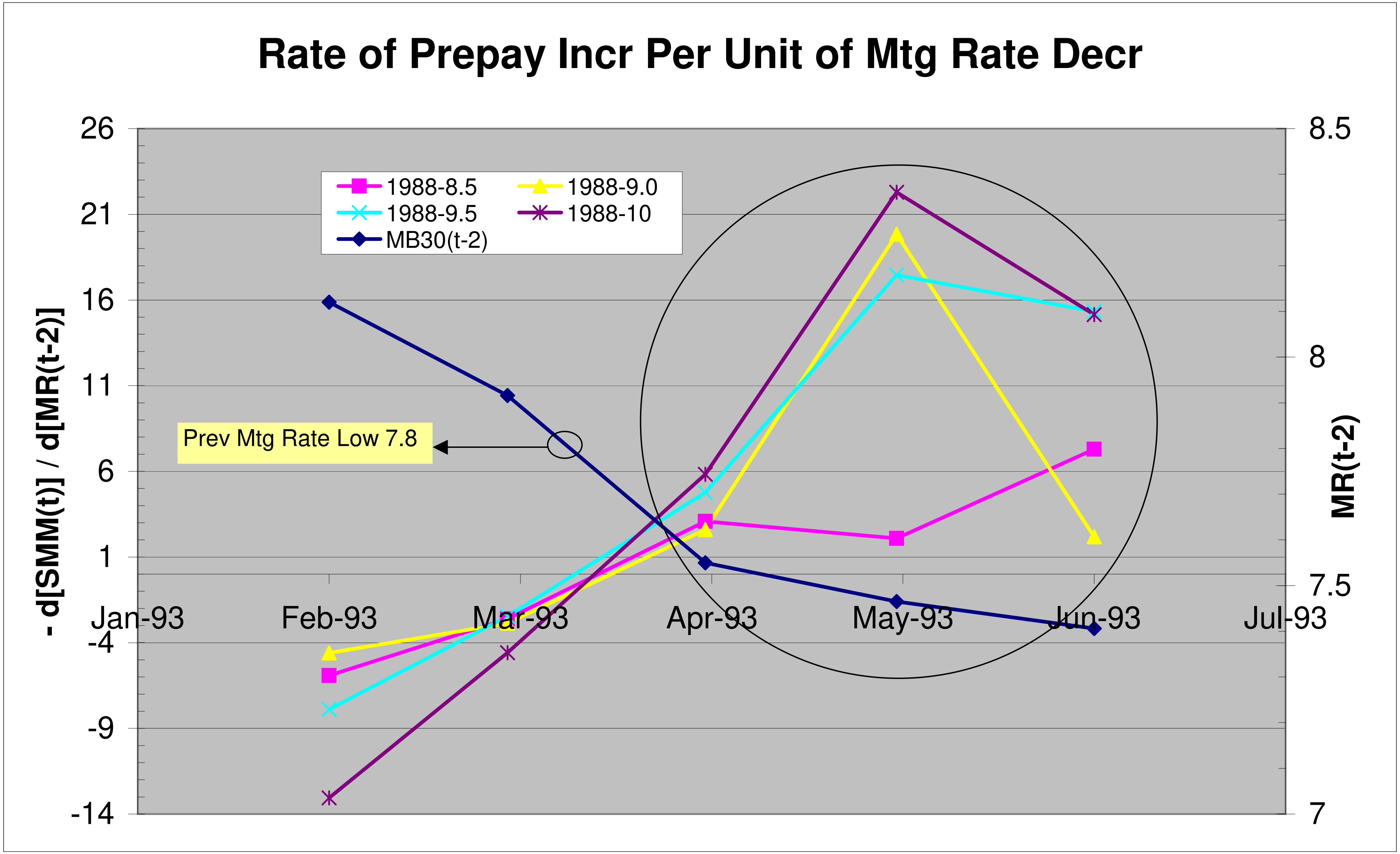

Rate of Prepay Incr Per Unit of Mtg Rate Decr
1988-8.5
1988-9.0
1988-9.5
1988-10
MB30(t-2)
Prev Mtg Rate Low 8.3
- d[SMM(t)]/d[MR(t-2)]
MR(t-2)
25
20
15
10
5
0
-5
-10
-15
9
8.75
8.5
8.25
8
7.75
May-92
Jun-92
Jul-92
Aug-92
Sep-92
Oct-92
Nov-92
Dec-92

Fig. 21



Rate of Prepay Incr Per Unit of Mtg Rate Decr
1988-8.5
1988-9.0
1988-9.5
1988-10
MB30(t-2)
Prev Mtg Rate Low 7.8
- d[SMM(t)] / d[MR(t-2)]
MR(t-2)
26
21
16
11
6
1
-4
-9
-14
8.5
8
7.5
7
Jan-93
Feb-93
Mar-93
Apr-93
May-93
Jun-93
Jul-93

Fig 22

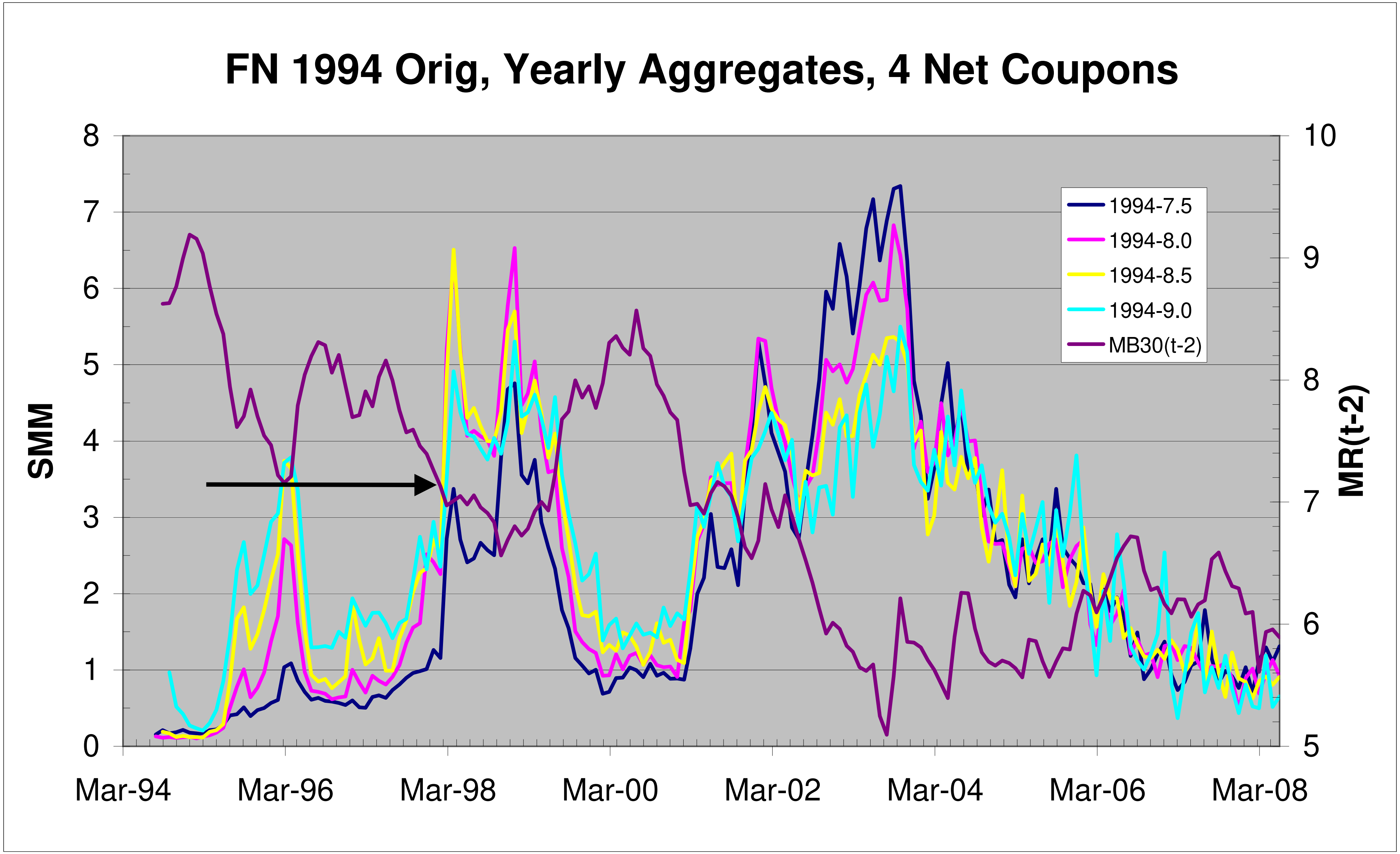
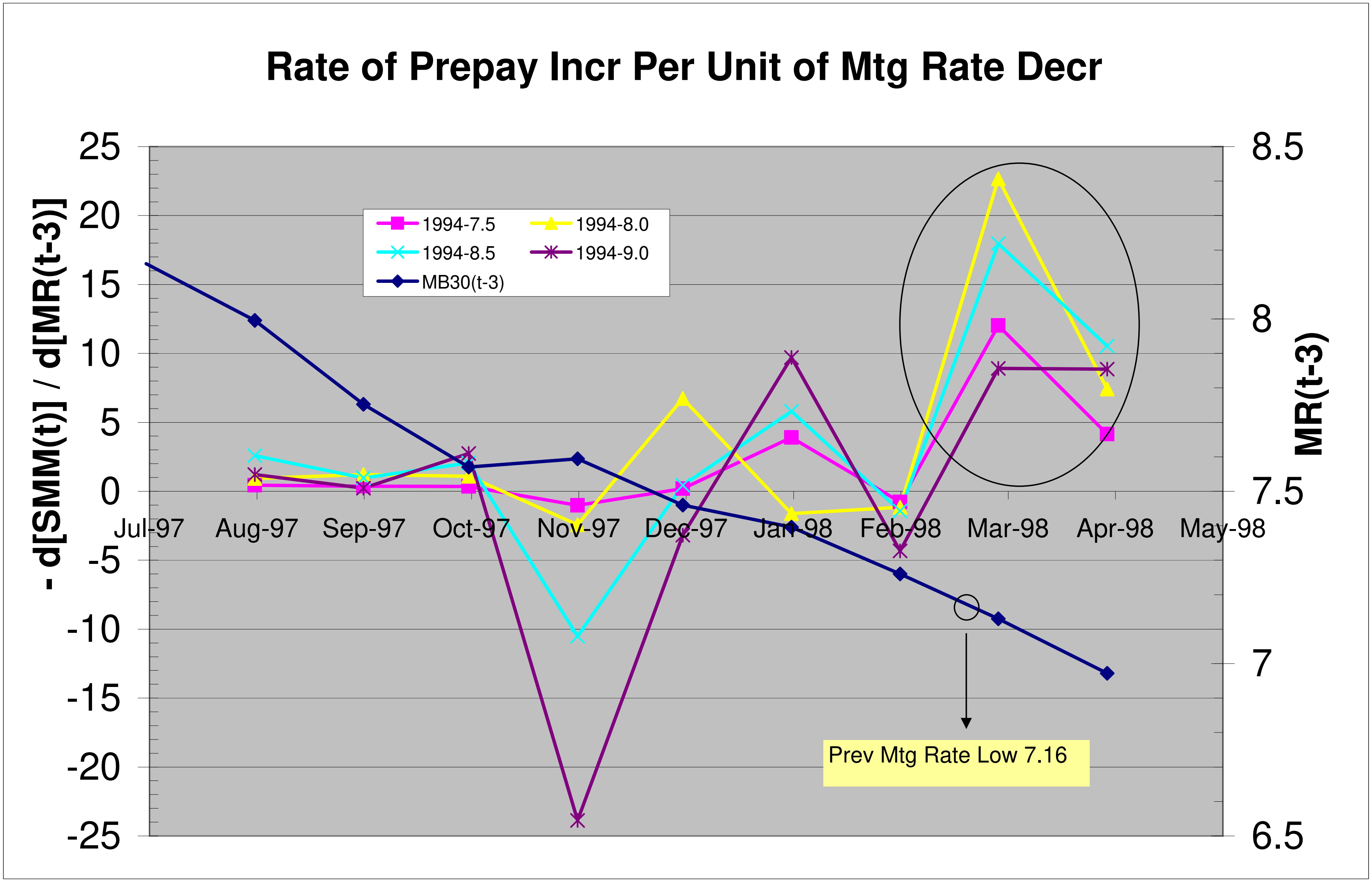

FN 1994 Orig, Yearly Aggregates, 4 Net Coupons
SMM
MR(t-2)
0
1
2
3
4
5
6
7
8
5
6
7
8
9
10
Mar-94
Mar-96
Mar-98
Mar-00
Mar-02
Mar-04
Mar-06
Mar-08
1994-7.5
1994-8.0
1994-8.5
1994-9.0
MB30(t-2)

Fig 23



Rate of Prepay Incr Per Unit of Mtg Rate Decr
- d[SMM(t)] / d[MR(t-3)]
MR(t-3)
1994-7.5
1994-8.0
1994-8.5
1994-9.0
MB30(t-3)
25
20
15
10
5
0
-5
-10
-15
-20
-25
8.5
8
7.5
7
6.5
Jul-97
Aug-97
Sep-97
Oct-97
Nov-97
Dec-97
Jan-98
Feb-98
Mar-98
Apr-98
May-98
Prev Mtg Rate Low 7.16

Fig. 24