OFHEO WORKING PAPERS

Working Paper 06-1

**Removing Appraisal Bias from a Repeat-Transactions
House Price Index: A Basic Approach**

Andrew Leventis

February 2006

Abstract

Responding to concerns that house price appraisals are biased measures of true home values, some researchers have suggested that repeat-transaction house price indices should be calculated using only price data from home purchases. This paper builds on previous literature and presents an alternative way of inoculating such indices from appraisal bias. Trends in and estimation precision for the "bias-neutral" index are compared against the properties of a purchase-only index, as well as the basic OFHEO house price index. Empirical estimates for three states and two metropolitan areas indicate that the "bias-neutral" index closely tracks the other two indices, while benefiting from relatively small standard errors.

Introduction

Over the last two decades, a significant literature has arisen on the subject of biases in house price appraisals. Researchers have analyzed operational and empirical evidence to determine whether home appraisers systematically over- or underestimate the ultimate selling prices of houses. Generally, their findings reveal some evidence that appraisals tend to be positively biased; that is, they tend to overpredict property selling prices.

Over approximately the same period, economists have developed and modified a number of methodologies for constructing house price indices. These various indexing approaches, which include repeat-sales models, hedonic models, and various hybrid approaches, have been compared and refined to optimize the use of available data.

Unfortunately, little literature has dealt with the intersection of the two topics: about how to construct an index if appraisal and transaction price data are both available. To some extent, the dearth of research is understandable. As researchers noted in the mid 1990s, indexing methodologies "tend to be custom-designed to the data"[1] and many researchers do not have access to both transaction prices and appraisal data. They thus are forced to construct their indices using only one type of valuation.[2]

In light of the potential bias associated with appraisals, this paper aims to provide an answer to the following question: If appraisal and sales price data are both available, how should a price index be constructed? The solution will attempt to avoid the simplistic approach of merely discarding appraisal data. Such an approach is highly problematic because limited house price information is available for small geographic regions and time periods. In such situations, the removal of appraisal information may dramatically increase estimation imprecision or, in some cases, may actually leave the estimation procedure with no observations to use in constructing an index.

I proceed by first providing some background on the relevant appraisal-bias and house price index literatures, paying particular attention to the repeat-sales methodology underlying OFHEO's House Price Index (HPI). I then discuss the theoretical and empirical properties of purchase-only repeat-sales indices. The mathematics associated with a very simple index that is neutral to appraisal-bias are then described. This index, which uses both appraisals and transaction prices data but "strips" the bias from the appraisal observations, is then computed for a number of geographies and is compared against a purchase-only index, as well as the basic HPI.

One important innovation in this paper relative to prior research is that the loan purpose behind appraisals is known and is accounted for in the empirical analysis. Appraisals performed as part of refinances may have different bias characteristics depending on whether the refinances adjust the mortgage interest rate or term ("Rate-Term" refinances)

---

[1] See Stephens *et al* (1995).
[2] Transaction prices typically would be the only valuation source available.

or if equity is being taken out of the property as part of the transaction ("Cash-Out" refinances).[3] Recognizing the distinction, in early 2004, Fannie Mae mandated that originators review *cash-out* appraisals because it was seeing "a trend toward inflated appraisal values on cash-out refinancings."[4] Recent OFHEO releases have made the distinction as well.[5]

One subject related to appraisal bias involves the issue of price "smoothing." In an Advanced Notice of Proposed Rulemaking (ANPR) in 1999, OFHEO in fact defined "appraisal bias" to include this phenomenon, stating that appraisal bias:

> "…occurs when the use of appraisals to value property at refinancing…smooth[s] the fluctuations in housing values because appraisals are derived from comparisons with properties that have either been sold or listed for sale within the past several months and may fail to indicate more recent changes in housing value."[6]

This is a very broad definition of bias, as it includes mismeasurement of house price *volatility*. In this paper, I define bias in the usual statistical sense in which it only pertains to differences between averages (i.e., the true average and the average of the "estimator," in this case the appraisal). To simplify the discussion, I do not pay particular attention to whether the appraisals artificially "smooth" indices. The presence of appraisal smoothing does not enjoy the same level of consensus that appraisal bias has.[7] The issue, although certainly important and related, should be resolved in future research.

In my analyses, I compare appraisals associated with mortgage refinances with what will be described as "purchase-related" or simply "purchase" valuations. For mortgages that are created to finance home purchases (as opposed to supplanting existing mortgages), the home price provided to OFHEO by the Enterprises is the lesser of the purchase price and the appraisal valuation performed in connection with the mortgage application. For the purchase-related transactions, OFHEO is not provided with the source of the valuation data.

As others have done, I maintain a level of abstraction to simplify the discussion.[8] When I construct a "purchase-only" index, it should be recognized that it will include both purchase prices and, in some cases, valuations from purchase-related appraisal valuations. In some situations, comparing refinance-related appraisals and the "purchase-only" index in effect entails comparing two different types of appraisals. Given that appraisals performed in connection with home purchases tend to exceed purchase prices,[9] however, this may not be a particularly common occurrence in the data.

---

[3] Note that these are not necessarily mutually exclusive categories. The data available from the Enterprises only allow one descriptive category, however.

[4] See "Fannie Gets Tough on Inflated Appraisals" *Dow Jones*, January 13, 2004.

[5] See HPI Release—Quarter 3, 2004 (available at www.ofheo.gov/media/pdf/3q04hpi.pdf).

[6] See OFHEO's comments in the ANPR, Federal Register April 13, 1999 (Volume 64, Number 70).

[7] See, for example, Calhoun *et al* (1998).

[8] See Calhoun *et al* (1998).

[9] See, for example, Cho and Megbolugbe (1996) and Loebs (2005).

It should be noted that the literature on house price indexing techniques remains active and, as will be discussed, the repeat sales approach used in the HPI is just one of several basic indexing methodologies whose properties are frequently compared.[10] This paper will remain focused on removing appraisal bias from repeat-sales indices and will not pay particular attention to strategies for removing such bias from other indices. Further review, however, may find that the approach employed here has relevance for other analytical approaches as well.

Background

Literature Review: Appraisal Bias and Volatility

Although much of the recent price indexing literature has tried to identify the flaws and benefits associated with the various methodologies, the issue of appraisal bias has been discussed for an extended period—since at least the late 1980s.

Dotzour (1988) studied the topic in some detail and found no strong empirical evidence for appraisal bias. Although its empirical support was not clear, by 1995, a consensus was reported to exist supporting the notion that bias was systematically present in appraisals. In that year, research by Stephen *et al* (1995) confidently asserted that "appraisals are widely believed to contain significant measurement error." Noting that some researchers believe that there may be an upward bias to appraisals, contemporaneous research by Myers and Pitkin (1995), by contrast, contended that a *downward* bias may exist. Their hypothesis was that, because they are hired by the lender, appraisers only need to validate threshold loan-to-value (LTV) ratios that support mortgage applications. Their presumption was that appraisers would not bother setting valuations above the required threshold level.

Cho and Megbolugbe (1996) performed their own empirical investigation and found that 80 percent of appraisals associated with house sales were at the sales price or between zero and 5 percent above the price. They suggested that "in many cases, property appraisals are discrete rather than continuous: appraisers assign different value estimates [vis-à-vis the selling price] only when differences between perceived values and transaction prices are large." The pressure to ensure that low appraisals do not harm transactions was cited as a reason for the upward bias.

Chinloy, Cho and Megbolugbe (1997) expanded on the earlier research, continuing to argue that appraisal bias was present.[11] Using a sample of 1993 purchase price data for which they also had the appraisal information, they compared purchase prices against

---

[10] Indeed, even within the repeat-sales indexing literature, a number of alternative implementation approaches have recently been analyzed. See, for example, Dreiman and Pennington-Cross (2002).
[11] U.S. Patent Number 6,401,070B1 (Inventors: Douglas A. McManus and Sol T. Mumey; June 4, 2002) extends the methodology employed in the Chinloy, Cho, and Megbolugbe (1997) paper.

appraisals to determine whether there were systematic differences. They estimated an upward bias of two percent and found that appraisals exceeded purchase price in approximately 60 percent of the cases. The authors discussed the potential causes for appraisal bias, citing the existence of "…a moral hazard incentive to complete the deal." They suggested that the "appraiser(s) face asymmetric costs from overstating versus understating."[12]

The authors also discussed the relative volatility of appraisals, cataloging the theoretical reasons that appraisals could over- or understate purchase price volatility. That appraisers "extrapolate" valuations from recent results and have a vested interest in ensuring that their valuations appear reasonable (and perhaps consistent) to the originators suggest that the volatility of appraised values may be lower. At the same time, the authors believe that the appraisals' reliance on a small number of comparables "almost surely" leads to "more volatility than marketwide prices." Their supposition is that appraisal valuations and sales prices can be thought of sample averages. Appraisals are effectively sample averages based on a small number of comparable properties. Transaction prices, by contrast, incorporate information from a much larger set of properties. Consequently, as a simple matter of statistics, with their reliance on smaller samples, appraisal valuations should have higher variability.

Empirically, the authors found no clear empirical evidence for pervasive appraisal smoothing. The data evidenced "moderate smoothing" nationwide, but substantial regional differences were found. For example, in the West, statistical tests revealed that appraisals were *more* volatile than purchase prices. The same held in the South, although this result did not pass statistical tests for significance.

After 1997, the literature on appraisal bias and artificial price smoothing became less active. Research by Lacour-Little and Malpezzi (2003) studied appraisal bias tangentially as they analyzed the relationship between appraisal quality and mortgage defaults. Edelstein and Quan (2004) found some evidence of smoothing in commercial real estate returns, estimating that "variances [we]re undervalued by 55 percent for…[their] overall real estate return index."[13]

A 2005 report by the Collateral Assessment and Technologies Committee, an industry trade group focused on "increasing the awareness of the merits of collateral assessment technologies," provides a recent empirical analysis of appraisal biases. The report quantified and compared biases for "full appraisals" against "non-full" appraisals, which included valuations done through automated valuation models (AVMs). Using a very large dataset extending back to 1977, the analysis found significant upward bias in full appraisals associated with purchase transactions.[14] The report also found that, in the context of house refinancings, AVM-based valuations were more conservative than full-

---

[12] The authors note that: "Undervaluation involves costs of additional verification and the risk of the deal failing if the seller refuses to make concessions."
[13] See Edelstein and Quan (2004).
[14] The results are quite similar to those in Cho and Megbolugbe (1996).

appraisals. The overall level of AVM bias (if any) is difficult to discern from the reported results, however.[15]

Literature Review: The Use of Appraisal Data in House Price Indices

The relevant literature has generally found that appraisal bias does impact house price indices, but has focused little on developing prescriptive methodologies for how indices ought to be calculated. A consensus seems to exist that: "excluding refinancings has noticeable effects, whose direction can vary over divisions and periods…"[16] but the empirical characteristics of approaches that remove the bias have drawn relatively little attention.

After noting that "price index techniques… tend to be custom-designed to the data and are seldom broadly applicable," Stephens *et al* (1995) discuss various indexing methodologies, including Freddie Mac's Conventional Mortgage House Price Index (CMHPI), and enumerated their various flaws. They noted that "many believe that the highest quality measure of market value comes from an arm's length purchase" and caution against the use of self-assessments and tax assessments, arguing that each has its own sources of biases and, in the case of self-valuations, can be quite noisy. They contend that the use of refinancing appraisals is problematic as well, not only because of appraisal bias, but also because of sample selection considerations. To the extent that low appraisals prohibit some from refinancing and thus preclude certain low-appraisal observations from appearing in any dataset, any house price indexing methodology that relies on appraisal data is said to be prone to sample selection bias.

Ultimately, however, the authors do not arrive at a conclusion as to what ought to be done with appraisals. They suggest that "…the potential bias from using refinancing transactions must be weighted against the potential gain in index accuracy from regional disaggregation…" but do not arrive at "an optimal policy regarding the use of refinancing transactions." All of their empirical calculations include refinances wherever possible.

Chinloy *et al* (1997) test whether appraisal bias affects repeat-sales price indices, separating the matched-pair sample by the loan purpose combinations and constructing separate price indices using the four different permutations: purchase-purchase, appraisal-appraisal, appraisal-purchase, and purchase-appraisal pairings. They find that F-tests universally reject the hypothesis that these indices are the same. Appraisal-based price indices are found to exhibit slightly higher growth rates than purchase-only indices, although no consistency is evident across geographic regions. A purchase-only index for the West, for example, reveals faster growth than for an appraisal-based index.

---

[15] Herein after, to simplify the discussion, no distinction will be made between full-appraisals and "non-full" appraisals. Accordingly, the theoretical and empirical assessments of "appraisal bias" will effectively refer to *average* bias (across all types of appraisals). OFHEO is not provided with valuation type information in the house price data provided by the Enterprises.
[16] See Stephens *et al* (1995).

After finding these differences, the authors conclude their discussion with the *caveat* that aggregation of appraisal and purchase data may be "unavoidable" in certain situations because of "data constraints."

As part of their analysis, Chinloy *et al* (1997) rely on a very simple adjustment to the basic repeat sales model that allows for the detection of appraisal bias. The basic modification seems to entail the inclusion of a series of time-specific appraisal dummy variables as regressors in the basic indexing model. These coefficients measure the extent to which appraisals exceed purchase prices for various time periods and effectively "strips" any appraisal-bias from the basic index coefficients. This allows the coefficients to mimic a purchase-only index while taking advantage additional observations. The authors do not, however, discuss the relative merits of using this approach over using a purchase-only repeat-sales index.

Calhoun *et al* (1998) proceed similarly to Chinloy *et al* (1997). Aiming to study growth rates and volatility measures for purchase-only and aggregated house price indices, they employ both a differing-coefficients and a separate-samples approach. They introduce a few basic modifications in the process,[17] but, like *Chinloy et al*, focus primarily on testing for the *existence of differences* across various indices.

The authors compare their version of a purchase-only index against appraisal-only indices for four large metropolitan areas. Although the empirical results strongly suggest that the "appraisal-based indices differ from purchase based indices," the authors struggled to find any consistency in their results in terms of appreciation rates or relative volatility levels. They found higher appreciation rates for three of the four cities with the appraisal-only index, but warned that the findings were "not necessarily conclusive of an upward bias." One-year volatility was generally found to be greater for the purchase-only index, but over five years, the result depended on the city.

The authors then use the "interaction approach," which estimated house price index values using a pooled (appraisal and purchase price) dataset but allowed index coefficients to differ for appraisals. Under this approach the authors found that appreciation was generally "highest for the sale-appraisal process, followed by appraisal-appraisal, sale-sale and appraisal-sale." As was the case with the separate-samples estimation, the relative volatility of the purchase-only vis-à-vis the appraisal-only indices depended on whether one was looking at one-year or five-year volatility measures.

In analyzing the interaction models, the authors did not provide statistics revealing the precision with which the various index parameters were estimated. Except in a very qualitative way, they also did not compare the results of the interaction models against the results of the separate-samples approach. Their focus was on using both techniques to determine whether differences existed between appraisal- and purchase-based indices.

---

[17] For example, when pairing transactions, the authors skip over ("leapfrog") transactions of different types and pair those transactions having homogeneous loan purposes. The approach in Chinloy *et al*, by contrast, involved pairing consecutive transactions and then selecting only those with two purchases.

The authors concluded that the use of appraisals generally resulted in the estimation of higher appreciation rates. They stressed, however, that one could not conclude that appraisal indices were biased upward because the results were sensitive to the year in question.

Reviewing the Use of Appraisal Data in Contemporary House Price Indices

Two of the five most commonly used housing price indices in use today employ appraisal information. The HPI and the Fannie Mae-Freddie Mac CMHPI both use appraised house prices in addition to purchase prices.[18] By contrast, price indices published by the National Association of Realtors (NAR)[19] and Fiserv CSW only use purchase price information.[20] In these cases, it is not clear whether the publishers would use appraisal data if it were readily available to them. Comments in recent published research by a representative of Fiserv CSW suggest that, at least in the latter case, it is unlikely.[21]

In its 1999 ANPR, after acknowledging that a perception of appraisal bias exists, OFHEO indicated that it did not believe it would be prudent to adjust its HPI model until more "conclusive research" becomes available on the topic. OFHEO noted however that:

> "Appraisal bias could, theoretically, affect the rates generated by the
> stress test if the method of computing the HPI were changed in some
> way to account for appraisal bias or if appraisal bias were found to be
> significantly different in more recent data than in the historical data
> used to estimate the models. OFHEO does not believe the change in
> the amount of appraisal bias in the HPI, if any, is significant."

While Freddie Mac seemed to agree with OFHEO's decision to use the existing pooled model,[22] the Mortgage Risk Assessment Corporation (MRAC), contended that the inclusion of appraisals biased appreciation rates upward by one percentage point, and supported the use of a purchase-only index ("if practical").

Recent comments by Fannie Mae's chief economist suggest that Fannie Mae now believes that the preferable way of calculating home price indices is to use only purchase-related valuations. Chief Economist David Berson, in a recent weekly commentary,

---

[18] Beginning with the third quarter 2004 release, OFHEO recently has provided additional information on how appraisals affect the index.
[19] NAR publishes median prices for a number of geographic regions and metropolitan areas.
[20] See www.cswv.com/products/redex/case/
[21] See www.cswv.com/pdfs/CATC_SystemicRisks1.1.pdf: "There is generally wide acceptance of the notion that arms-length sale transactions are generally the most reliable indicators of market value for residential properties…the systematic bias [evidence in accompanying tables presented]…suggest[s] that, more often than not, the subject property valuations within these full valuation appraisals are of dubious value…"
[22] OFHEO reported that Freddie Mac did not want to make any alterations to the HPI model to account for appraisal bias.

states that: "…looking only at the purchase figures (*which we feel provides a more accurate measure of underlying values than including the value of homes that were refinanced*)[23] home values jumped…"

The Purchase-Only "Solution"

The vast majority of the paired transactions used in the construction of the repeat-transactions indices include refinance-related appraisals. Removing refinance-related appraisals thus substantially reduces the number of observations available for model estimation. Stephens *et al* (1995) found that 84 percent of the paired observations in their sample included some type of refinancing-related valuation. Using more recent data from the Enterprises, Tables 1a and 1b reveal that, for pairings having at least one of their valuations in the last ten quarters, the proportion that were purchase-only pairings is extremely low. For Illinois, Washington State, and California, depending on whether the pairing approach selects only sequential purchases or merely pairs all available purchases, the proportion of observations that were "purchase-only" varied from about 4 percent to 27 percent.

In deciding whether to use a purchase-only index, this reduction in observations has been cited as an important consideration. Noting that "excluding refinancings has noticeable effects…," Stephens *et al* suggested that "...the potential bias from using refinancing transactions must be weighed against the potential gain in index accuracy from regional disaggregation."[24]

Another consideration lobbying against the use of a purchase-only index is that refinance-related appraisals may mitigate *transactions* bias. Research has suggested that appreciation rates for houses that sell may not be the same as appreciation rates for the rest of the housing stock.[25] Empirical analyses in the early 1990s presented conflicting evidence on the validity of these concerns, but the majority of later estimates suggest that

---

[23] Emphasis has been added.
[24] One notable issue is that, if OFHEO were to begin using a purchase-only index, it would have small, but significant ramifications for information released to the public. Under OFHEO's existing publication policies, the index value for a metropolitan area is published for a given period if: (a) at least 1,000 transactions have occurred in that area and (b) at least ten transactions occur during the period. Appreciation rate rankings are provided for metropolitan areas that have had 15,000 or more transactions over the prior ten years. (See HPI Frequently Asked Questions (FAQ) publication available at: www.ofheo.gov/hpifaq.asp). Analysis of OFHEO's latest HPI release indicates that index values for one metropolitan area--Hinesville-Fort Stewart, GA—would no longer be published if the purchase-only index were used. As of the second quarter of 2005, the Hinesville area failed to meet both the 1,000 transactions and ten quarterly transactions criteria. OFHEO would no longer provide rankings for more than 100 of the 265 metropolitan areas that it ranked in the second quarter 2005 release.
[25] See, for example, Stephens *et al* (1995).

the bias is real and significant.[26] The presence of refinance observations would, at least in theory, dilute such bias.

If, given the various considerations, one opts to use a purchase-only index, as suggested in the literature review, a "purchase-only" index can be constructed in two ways. The first approach, which might be described as the "pre-pairing," first constructs transaction pairs and then culls out those that have two purchases. As indicated in the literature review, the Chinloy papers seem to have used this approach.[27] The second approach, "leapfrogging," begins by removing the non-purchase transactions from the raw transactions data, and then pairing the remaining purchase observations. This latter approach was used by Calhoun *et al* (1998).

The datasets produced under the different approaches are significantly different, as is evident in Table 1a and Table 1b.[28] Because it requires that *successive* transactions involve purchases, as opposed to finding pairings for all purchase transactions, the pre-pairing approach is much more constraining and thus produces smaller sample sizes. This may present significant estimation problems for smaller geographic regions. Pre-pairing is, nevertheless, consistent with the basic matching methodology employed in the usual HPI construction.[29]

Pre-pairing may have certain advantages if controlling for quality changes is particularly important in index construction. The exclusive use of successive purchase transactions may better ensure that price changes reflected in the index omit the effects of quality improvements.

For example, if a cash-out refinance occurs between two purchase transactions, a possibility exists that the funds from the cash-out refinance were used to improve the home.[30] In this circumstance, because the two purchase transactions would be paired, the leapfrogging index would reflect the effects of any quality improvement. By contrast, the pre-pairing index would omit both observations and thus not embed the quality improvement. Figure 1, which shows relative appreciation rates for homes having these purchase- cash-out - purchase mortgage sequences, presents evidence that the quality improvement effect may be non-trivial.[31]

---

[26] See Calhoun (1991); Abraham and Schauman (1991); Case, Pollakowski, and Wachter (1991); Case, Pollakowski and Wachter (1997); Gatzlaff and Haurin, (1993); Gatzlaff and Haurin, (1997); Hwang and Quigley (1999); and Cheung, Yau, and Hui (2004).

[27] The authors state: "The purchase series is the observation of two *consecutive* sales" (Italics added).

[28] Table 3 will also reveal some of their differences.

[29] Beginning with its release of data from the third quarter of 2004, OFHEO began reporting a "purchase-only" index for informational purposes only (see www.ofheo.gov/media/pdf/3q04hpi.pdf). This index employs the leapfrogging approach when pairing purchase transactions.

[30] As opposed to merely repairing the home.

[31] An alternative explanation for the phenomenon evident in the graph is that homes that rapidly appreciate have a greater tendency to induce cash-out refinancings than other homes. This topic will be discussed in greater detail in later sections.

Another issue is that the leapfrogging approach, by construction, produces pairings that have much longer intervals between transactions. If houses exhibiting shorter periods between transactions somehow better reflect price appreciation, then again, the pre-pairing approach might be desirable.[32]

This general topic lies beyond the scope of this paper, but deserves further attention. For the purposes of this analysis, I will focus on the leapfrog approach because it has the known advantage of a relatively large sample size. Also, the possibility that the leapfrogging approach may overstate appreciation due to improving quality is not a significant problem given how OFHEO uses the data.[33]

Figures 2a-2e compare price appreciation trends for several geographic areas using both the HPI as well as a purchase-only index.[34] Results are shown for three states— California, Illinois, and Washington—and two metropolitan statistical areas (MSAs)— Los Angeles[35] and San Diego.[36] These geographies should be sufficiently diverse in terms of sample sizes and coverage, as to be broadly representative of the type of impact that different selection approaches will have on index values.

The five graphs reveal that appreciation rates derived from the purchase-only index do not consistently exceed or lie below the appreciation rates estimated using the HPI (the "pooled" sample). The differences in the series do suggest, however, that refinance-related appraisals may differ from purchase-related valuations.

As illustrated in Figures 3a to 3b, the direction of any appraisal-related bias is not immediately obvious. The first line in the figures, the solid line, plots the difference between appreciation rates reflected in the HPI and those reflected in a purchase-only index.[37] The second line plots the percentage of loans that were refinances. If refinances have predictable impact on appreciation rates, then the lines should show a consistent positive or negative relationship. They generally do not, although there seems to be some consistency within certain time periods. For periods prior to about 1998, the prevalence of refinances seems to be positively correlated with the difference between the HPI and the purchase-only index. After about 1998 and until very recently, the two series seem to be negatively correlated.

---

[32] OFHEO's implementation of the repeat-sales index does down-weight houses having long intervals between transactions. It is unclear, however, whether the effect of the leapfrogging approach would be fully offset by the downweighting methodology currently used.

[33] The HPI is used to estimate the value of the collateral underlying mortgages owned or guaranteed by the Enterprises. Regardless of whether the quality improvements or basic house price inflation are causing increases in collateral value, OFHEO's price index needs to capture both sources of valuation augmentation.

[34] Note that transactions for which the loan-purpose was "unknown" have been removed from the samples prior to estimation. For this reason, the "pooled" index results may differ from those reported in the OFHEO's usual HPI releases.

[35] The OMB-defined area known as the Los Angeles-Long Beach-Glendale, Metropolitan Statistical Area Division (MSAD).

[36] The OMB-defined area known as the "San Diego-Carlsbad-San Marcos, CA" MSA.

[37] The difference is calculated as: HPI Four-Quarter Growth Rate minus Purchase-Only Index Four-Quarter Growth Rate.

A number of possible factors could underly the phenomena observed in Figures 3a and 3b. One possibility is that the appraisal-bias associated with rate-term refinances has shrunk in recent years. Another possibility is that the ratio of rate-term refinances to total refinances may have changed significantly in recent years. Regardless of the cause, however, the figures highlight the fact that the difference between the HPI and the purchase-only index is a function of a number of time-specific factors and that it would be difficult to readily anticipate the impact of appraisals on the index without more empirical rigor.

Figures 4a and 4b illustrate the extent to which the use of the purchase-only index increases estimation imprecision. Both figures focus on Lima, Ohio, a small metropolitan area whose price series should be demonstrative of the problems associated with using the purchase-only index when sample sizes are already small.

Figure 4a simply compares four-quarter appreciation rates over time for the HPI and the purchase-only index. As is evident in the graph, the purchase-only series estimates large swings in four-quarter appreciation rates. For example, the purchase-only series estimated four-quarter rates of between roughly -1 and 12 percent during the period between 2003 and 2004. By contrast, appreciation rates for the pooled HPI ranged from 2 to 7 percent during that same period.

Figure 4b plots confidence intervals around the two indices between 1992 and 1994, a period in which the number of purchase and refinance transactions was relatively significant for the Lima area. The graph reveals that, even during this relatively active period, the confidence bands surrounding the purchase-only index are much wider than that surround the usual HPI. Further, the graph also shows that the point estimates for the purchase-only index (i.e., the center of the confidence bands) are significantly more volatile than the HPI estimates.

The Model

The following describes a very simple adjustment to the basic HPI methodology. With very minor alterations, the model is nearly identical to that presented in Chinloy *et al* (1997) and Calhoun *et al* (1998).[38] The approach acts as an efficient alternative to the purchase-one index, stripping from the basic index coefficients any effects of appraisal bias. The approach effectively constructs indices of cash-out and rate-term appraisal bias and uses those indices to remove the net effect of any biases on the basic index coefficients.

---

[38] To be precise, given the ambiguities in the prior papers, this model does not differ substantially from the models that *seem to have been* proffered previously.

It is useful to begin by reviewing the basic math underlying the HPI, much of which was set forth in Calhoun (1996) and prior work by Case and Shiller (1987, 1989).[39] The basic methodology begins by assuming that house prices are determined by the following equation:

$$\ln(P_{i\tau}) = \beta_\tau + H_{i\tau} + N_{i\tau} \tag{1}$$

The log of the transaction price for house i transacting in period τ is a function of three sets of parameters. The first set, β_τ, is the basic house price level for the relevant geographic region in the time period in which the transaction occurs. The second set, $H_{i\tau}$, is a house- and time-specific random-walk that describes the evolution of that homes' value over time. The third set, $N_{i\tau}$, was initially contemplated to be home- and time-specific white-noise terms. However, to ensure consistency between empirical results and the underlying theory, OFHEO and others sometimes make the assumption that the white noise does not vary over time. Hence, the $N_{i\tau}$ in the above equation is replaced by N_i, white noise that influences a given home's value equally in all time periods.

For any given pair of transactions, the change in value between the two transactions can be expressed as:

$$\Delta V_i = \ln(P_{it}) - \ln(P_{is}) \tag{2}$$

and

$$\Delta V_i = \sum_{\tau=0}^{T} (\beta_\tau + H_{i\tau} + N_\tau) D_{i\tau} \tag{3}$$

where

$D_{i\tau} = 1$ if first transaction occurred in the period τ
$D_{i\tau} = -1$ if second transaction occurred in the period τ
$D_{i\tau} = 0$ if i did not transact in period τ

One can estimate the values of the betas using the specification:

$$\Delta V_i = \sum_{\tau=0}^{T} (\beta_\tau D_{i\tau} + \varepsilon_{i\tau}) \tag{4}$$

As described in Calhoun *et al* (1996), equation (4) is estimated in three steps. The first involves estimating (4) with ordinary least squares. The second and third steps aim to improve the efficiency of the first-stage index coefficients, accounting for the possibility that the estimation error is positively related to the time between transactions. In the second stage, the squared errors from the first stage are regressed on the time interval between transactions[40] and the time interval squared. Subsequently, in the final step, the original first-stage regression is repeated except that the observations are weighted by the predicted errors from the second-stage.

---

[39] The fundamental methodology was originally described in Bailey, Muth and Nourse (1963).
[40] E.g., the number of quarters between the first and second transactions.

Model Adjustments to Account for Appraisal-Related Valuation Differences

The valuation equations can be adjusted to account for the possibility that refinance-related valuations systematically differ from purchase-related transactions. In reconstructing the equations, one allows for the possibility that the bias: (a) differs for cash-out and rate-term refinances and (b) varies by period. Under these modifications, equation (1) becomes:

$$\ln(P_{i\tau}) = \beta_{\tau} + CO_{\tau} + RT_{\tau} + H_{i\tau} + N_{i\tau} \tag{5}$$

The CO_{τ} and RT_{τ} coefficients track the magnitude of rate-term and cash-out appraisal bias over time.

Equation (3) can then be rewritten as:

$$\Delta V_i = \sum_{\tau=0}^{T}(\beta_{\tau} + CO_{\tau} + RT_{\tau} + H_{i\tau} + N_{i\tau})D_{i\tau} \tag{6}$$

where $D_{i\tau}$ is defined as it was before:

$D_{i\tau} = 1$ if first transaction occurred in the period τ
$D_{i\tau} = -1$ if second transaction occurred in the period τ
$D_{i\tau} = 0$ if i did not transact in period τ

The analogue to equation (4)—which was the primary estimation equation previously—now has a few more appraisal-related terms:

$$\Delta V_i = \sum_{\tau=0}^{T}\beta_{\tau}D_{i\tau} + CO_{\tau}Y_{i\tau} + RT_{\tau}Z_{i\tau} + \varepsilon_{i\tau} \tag{7}$$

where $D_{i\tau}$ is defined exactly as it was before, and Y and Z are defined similarly:

$Y_{i\tau} = 1$ if first transaction occurred in the period τ and was a cash-out refinance
$Y_{i\tau} = -1$ if second transaction occurred in the period τ and was a cash-out refinance
$Y_{i\tau} = 0$ if i did not transact in period τ

$Z_{i\tau} = 1$ if first transaction occurred in the period τ and was a rate-term refinance
$Z_{i\tau} = -1$ if second transaction occurred in the period τ and was a rate-term refinance
$Z_{i\tau} = 0$ if i did not transact in period τ

Results

Index Estimates

Figures 5a-5e and 6a-6e plot levels of and changes in three different prices indices: the HPI (which can be described as a "pooled" index), the purchase-only index, and the bias-neutral index. Estimates are shown for the five different geographic regions for the last fifteen years. The time period reflected in these figures is the longest period that OFHEO can study given the data provided by the Enterprises. The loan purpose field, which identifies mortgages originated for new home purchases and those that are refinancings, is consistently identified beginning in the early 1990s.[41]

The most striking aspect of the empirical results is the substantial similarity between the indices in all of the geographic regions. The magnitude of and changes in the appreciation rates reflected in the three different indices correspond very closely throughout the entire time frame.[42] The correlation coefficient in four-quarter appreciation rates for Washington State, for example, is approximately .96 when comparing the pooled and bias-neutral indices and .99 when comparing the bias-neutral and purchase-only indices.[43] Correlation coefficients for other states are also quite high, with the lowest correlation being .86 (comparing the pooled and bias-neutral indices in Illinois).

None of the three indices tends to demonstrate "smoother" appreciation patterns as measured using the coefficient of variation on the four-quarter price changes. The bias-neutral index has the smoothest appreciation patterns in Illinois, while the pooled index is least volatile in Washington State. The purchase-only index demonstrates the least volatility in California as a whole, as well as in the two metropolitan areas studied.[44]

In Washington State, California, and the two metropolitan areas in California, the purchase-only series tends to show slightly higher appreciation rates in recent years than the other two indices. Between the second quarter of 1997 and the second quarter of 2005, total estimated price appreciation in Washington State was 79.1 percent for the purchase-only series and 74.8 percent for the bias-neutral index. In California, total estimated appreciation was 199.7 percent for the purchase-only series and 184.3 percent for the bias-neutral index.

---

[41] For estimation purposes, transactions occurring in prior periods are, of course, included in the estimation. The filter applied to those records (as well as all other records) is that the "loan purpose" field must not be designated as "Unknown."

[42] It should be noted that the time period analyzed span several macroeconomic cycles and thus the results may not be idiosyncratic to the period studied.

[43] The correlation coefficient is calculated using appreciation rates between the first quarter of 1990 and the second quarter of 2005.

[44] San Diego-Carlsbad-San Marcos, CA Metropolitan Statistical Area (MSA).

Given that the bias-neutral index aims to remove bias from non-purchase-related valuations, the fact that there is any material difference between its estimates and those of a purchase-only index may be surprising. One reason for the difference is the fact that the purchase-only index omits information that the bias-neutral index takes into account. By construction, homes that have exactly one purchase transaction in the Enterprise data are not used in the purchase-only index.[45] As long as these homes have refinance-related valuations, the bias-neutral approach will include these in its estimation.

Figure 7 reveals that, for Washington State, an index constructed using only those observations that have exactly one purchase rises at a slower slow pace vis-à-vis an index constructed from houses that sold at least twice. This phenomenon, which is broadly consistent with previous results presented by in Case *et al* (1997),[46] thus explains some of the difference in the appreciation rates.

Other than the higher recent appreciation rates for the purchase-only index, Figures 5a-5e and 6a-6e do not reveal any additional obvious empirical regularities. The bias-neutral index exceeds the purchase-only index in some geographies and periods, but these periods do not share any notable commonalities in terms of macroeconomic or housing market conditions.

Impact on the Explanatory Power of the Model

For all of the geographic regions analyzed, the inclusion of the cash-out and rate-term bias indexing variables had a statistically significant impact on the explanatory ability of the models. F-Tests confirm that the two sets of variables, both separately and jointly, lead to positive contributions to the regression R-Squared values for the first stage.[47]

Tables 2a and 2b present the average squared error from the first and second regression stages for three different models. Results for the basic pooled and bias-neutral models are presented, as are results for a model run on only purchase observations.

For all of the geographic areas studied it is clear that the bias-neutral index leads to a nontrivial reduction in the average squared regression error relative to the basic house price index specification. For the first-stage, the bias-neutral model produces squared

---

[45] The leapfrogging methodology "skips over" refinance observations and seeks out purchase observations to complete pairs. The pre-pairing approach simply omits pairs that are "Purchase-Cash-Out", "Cash-Out, Purchase," "Purchase-Rate-Term" and "Purchase-Cash-Out."

[46] In that paper, the authors found that higher appreciation rates were associated with houses selling multiple times and that homes with exactly one purchase transaction during their 10-15 year sample period had relatively small growth rates.

[47] Separate testing entails testing the cash-out variables as a group and then testing the rate-term variables as a group.

When tested jointly, the calculated F-values for Washington State, Illinois, California, San Diego and Los Angeles always exceed one hundred and, on several occasions, exceeded one thousand. The critical value at the 95 percent confidence level is approximately 1.22.

errors that are approximately 2.6 to 6.8 percent lower than the basic specification. The reduction in squared errors for the second stage ranges from approximately 1.7 to 10.5 percent. It should be noted that, because the error from the first stage has been shown to be a function of the time between transactions, the findings from the second-stage, which shows the error reduction after accounting for time between transactions, are perhaps more relevant.

The tables also reveal the first and second stage errors for the bias-neutral approach are even more impressive relative to the errors from the purchase-only estimation. In comparing the results for the bias-neutral and purchase-only approaches, caution must be exercised, however. The regressions employ different datasets and, to the extent that the samples have fundamentally different properties, the distribution of the regression errors may naturally differ.

One obvious difference in the two samples is that the purchase-only dataset, because of the way it is constructed, exhibits much greater durations between transactions than the data used for the pooled and bias-neutral regressions. As shown in Table 3, the intervals are between 56 and 78 percent longer for the purchase-only sample than the usual sample. If the second-stage regression is not as accurate at predicting price changes for pairs having longer intervals,[48] the fact that the purchase-only dataset exhibits longer intervals will mean that, by construction, the purchase-only estimates will tend to have greater second-stage squared errors. Indeed, using the basic pooled data, Table 4 indicates that squared errors do indeed tend to be much greater for pairings having longer intervals.[49]

I produce a very crude comparison of the purchase-only errors against the bias-neutral errors by accounting for this difference. I modify the average squared error for the bias-neutral results to account for the differing time intervals. After calculating average squared errors for the bias-neutral approach for each time interval (e.g., one quarter between transactions, two quarters between transactions, etc.), the average errors are multiplied by the fraction of the *purchase-only* observations that have that exact interval. The adjusted average error for the bias-neutral approach is the sum of these products.

Mathematically, the time-adjusted error for the bias-neutral model can be expressed as:

$$\text{Time - Adjusted Average Error (Bias - Neutral)} = \sum_{t=1}^{T} \text{POWeight}_t \bullet \text{AvgSqError(ABNeutral)}_t \qquad (8)$$

where

$$\text{POWeight}_t = \frac{\text{Number of Purchase - Only Pairings Having Interval of t periods}}{\text{Total Number of Purchase - Only Pairings}} \qquad (9)$$

---

[48] That is, the second stage's errors are themselves heteroskedastic.

[49] Table 5, which presents regression coefficients for a much less constrained model, suggests that part of the problem may be second-stage misspecification. Note that estimated coefficients for the interval[3] and interval[4] variables are statistically significant at conventional error levels.

and

$$\text{AvgSqError(ABNeutral)}_t = \frac{\text{Sum of Squared Errors for pairings having interval of t}}{\text{Number of Pairings with interval of t}} \qquad (10)$$

As shown in Table 6, using this adjustment procedure, I find that the adjusted average errors for the bias-neutral methodology are lower than the purchase-only methodology in two of the three states analyzed.

In all of the geographic regions studied, the bias-neutral approach leads to a significant reduction in index standard errors relative to standard errors for the purchase-only approach. Figures 8a-8e illustrate the tightening of the estimation range for a recent representative quarter for the five geographic regions studied. As would be expected given results presented in the last section, the figures also reveal that the bias-neutral index also leads to increased precision relative to the pooled HPI.

Bias Indices

As discussed, the coefficients on the cash-out and rate-term dummy coefficients (CO_τ and RT_τ) in the basic index regression measure the magnitude of bias associated with each type of appraisal. Figures 9a-9e show time trends in these coefficients for each of the five geographic regions analyzed.

Unlike Figures 5a-5e and 6a-6e, which show levels of and changes in the house price *index*, Figures 9a-9e report the regression *coefficients* from the basic price change model. The house price index values were constructed by exponentiating the regression coefficients and then multiplying by 100. Here, the bias coefficients themselves are reported because they have a direct and convenient interpretation: they are the percentage estimated bias for cash-out and rate-term appraisals (relative to valuations obtained in purchase transactions).[50]

As is evident in the Figures 9a-9e, the coefficients differ significantly across the two appraisal types. For appraisals associated with rate-term refinances, depending on the geography and time period, the estimated bias tends to range between negative seven and seven percent, with a clustering of values in the zero to five percent range. For appraisals associated with cash-out refinances, the estimated bias tends to range between zero and eight percent.

---

[50] To be precise, the coefficients are an *approximation* of an estimated percentage bias. To arrive at the actual percentage estimate, one can exponentiate the coefficient and subtract one. For small coefficients, this will be close to the original coefficients.

Both series suggest that the magnitude of any bias may have shrunk in recent years. For cash-out refinances, which still seem to demonstrate some positive biases, estimated bias has dropped noticeably in the last six quarters. For example, in Washington State and California, the estimated bias averaged 4.6 and 5.4 percent respectively during the period between the first quarter of 1991 and the fourth quarter of 1999. By contrast, the estimated cash-out appraisal bias has averaged 3.9 and 1.2 percent over the last six quarters. For Illinois, the estimated bias for cash-out appraisals actually increased slightly from 3.2 percent in the 1990s to 4.4 percent in recent quarters.

The estimated bias associated with rate-term refinances has fallen to such an extent that the estimated bias has become negative[51] in recent quarters. Data from the 1990s suggest that rate-term appraisals in Washington State, California, and Illinois tended to overpredict purchase valuations by 2.1, 4.5 and 1.6 percent respectively. In recent quarters, by contrast, the rate-term valuations were systemically lower than purchase-valuations by between .7 and 7.2 percent.

The underlying cause of the declining bias is not entirely clear, although one *a priori* possibility is that outdated appraisals are being used. When mortgagors refinance their loans multiple times within the same year (as is sometimes done when interest rates are falling rapidly), the same house appraisal may be used for multiple loan applications.[52] The use of such "recycled"[53] appraisals provides the price index regression with information implying that either no price appreciation has occurred between two dates or that refinance-related appraisals tend to be biased downward. Because the valuations associated with purchases will have grown over the subject period,[54] the regression attributes the lack of appreciation to a downward bias for refinance-related appraisals.

Unfortunately, this hypothesis is not supported by the data. Figure 10 shows the results of removing from estimation all refinance pairings for which: (a) the time interval between transactions was less than one year and (b) the two valuations were identical. As is evident in the figure, which plots the results for Washington State, the trends in the rate-term and cash-out bias coefficients do not change materially upon excluding these observations.[55] The use of a more general filter, one that simply removes all observations for which the time between transactions is less than one year, also fails to support the hypothesis.

Bias Indices: Complications and Confounding Factors

The discussion thus far has ignored some more complex issues related to the construction and interpretation of the refinance-related appraisal coefficients, what have heretofore

---

[51] That is, rate-term appraisals tend to lie below valuations associated with home purchases.

[52] Appraisals can generally be used for up to one year, assuming that certain conditions are met.

[53] See Loebs, T. (2005) "Systemic Risks in Residential Property Valuations: Perceptions and Reality."

[54] In the last several years, appreciation rates over very short time periods such as these have been quite substantial.

[55] Approximately 2,000 pairs (out of approximately 900,000 total pairs) were removed by this filter.

been described as "bias" coefficients. Ultimately, the nature of the methodology described in this paper is such that the refinance coefficients will be a function of several effects, not all of which are bias-related. A brief example of the estimation mechanics illuminates these deeper issues.

The coefficient for a given refinance variable (say, $CO_{Q2,\ 1998}$—the indicator for cash-out refinances in the second quarter of 1998) is estimated using a number of different types of transaction pairs. Transaction pairs whose *first* valuation is from a cash-out refinance in the second quarter of 1998 are inputs,[56] as are pairs whose *second* valuation is a cash-out refinance in that quarter.[57] To estimate a coefficient for cash-out refinancings for that quarter, least squares compares the relative appreciation rates for these two types of pairings against the appreciation rates for non-cashout pairings. If cash-out appraisal valuations are unusually high, perhaps because of appraisal bias, appreciation rates for those pairings whose first transaction was a cash-out should have relatively weak appreciation rates. Similarly, those pairings having a cash-out refinance as their second transaction should have relatively large appreciation rates. In this situation, least squares would assign a positive value to the cash-out refinance coefficient for the second quarter of 1998.

Note that the different pairings that contribute to the cash-out refinance coefficients, in theory, could estimate significantly different cash-out effects. If the coefficient were estimated using only those pairings having a cash-out valuation as the *first* valuation, the cash-out coefficient might be significantly different from one calculated using those pairings that have a cash-out refinance as their *second* valuation. The approach employed in this paper effectively constrains the estimation so that a single, aggregated cash-out effect is estimated.[58]

The question then arises: Why would the cash-out coefficients differ across the two pairing types? The answer entails some considerations that have not yet been addressed and generically involves various sample selection and cohort effects.

It is important to note that, even with the model augmentation suggested in this paper, the repeat-transactions model is quite parsimonous. Appreciation rates are solely ascribed to a geography, the timing of the transactions, and—in this augmented case—the type of valuations. A variety of other factors may be correlated with appreciation rates but are not fed to the regression model. For example, homes that transact frequently may have higher or lower appreciation rates than others.[59] Also—homes in certain neighborhoods[60]

---

[56] These pairings would include homes whose second valuation either involved a purchase or a rate-term appraisal.

[57] These pairings would include homes whose first valuation involved either a purchase or a rate-term appraisal.

[58] A cursory analysis of refinance valuations in Seattle-Bellevue-Everett MSAD suggests that the cash-out coefficient may indeed differ significantly when estimated using different pairing types. When calculated using "Purchase-Cash-out" pairings, the cash-out coefficients tended to be in the range of approximately 4 to 9 percent. Estimated with "Cash-out - Purchase" pairings, the coefficients tended to be between 0 and -4 percent.

[59] See, for example, Case e*t al* (1997) and Dreiman and Pennington-Cross (2002).

may have much different appreciation rates than others in the same general geographic area.

Houses from a particular cohort (e.g., those properties whose first valuation was a cash-out refinance in the second quarter of 1998) may have "unobserved" attributes that lead to faster or slower growth rates. To the extent that cash-out and rate-term refinances are correlated with these "unobserved factors," the "bias" coefficients estimated in the model may capture some of these effects in addition to appraisal bias.

In an extreme case, the valuation type itself may be *direct evidence* that its appreciation rate was relatively high or low. For example, rapidly-appreciating homes may be more likely to be involved in cash-out refinances because the owners have more equity to access and loan approval would be relatively easy.[61] In this case, the cash-out "bias" coefficients would include some of this sample selection effect. That is, the cash-out "bias" coefficients will be different than they would be if appraisal bias were the sole cause of differences in measured appreciation rates. The "bias-neutral" index coefficients will reflect lower growth rates than they would if such a sample-selection issue were not present. The presence of the "bias" coefficients, in short, may dilute the impact of fast-appreciating homes on the base index coefficients.[62]

A cursory review of appreciation patterns in Washington State fortunately suggests that, under specific assumptions, this type of selection bias may not be particularly significant. The results of the analysis, which focused on the Seattle-Bellevue-Everett and Tacoma MSADs, are shown in Figure 11.

The graphs plot two different house price indices. One is constructed using houses for which a cash-out refinance occurred at some point during the data period. Importantly, this "cash-out" index is constructed using only those pairings that occur prior to the first cash-out refinance. The other index reported in Figures 11 is simply constructed using all homes that had no cash-out refinancings.

If cash-out refinances are strongly associated with high appreciation rates, assuming that relative appreciation rates for a given home are correlated over time, one would expect the cash-out index to grow at a faster rate. As is evident in Figure 11, the appreciation patterns for homes that had ultimately had cash-out refinances are nearly identical to appreciation rates for other homes. This is imperfect evidence that sample-selection effects are not present.[63]

---

[60] For example, houses in a particular part of a metropolitan area may appreciate at a faster rate than for the city as a whole.

[61] If the fixed-costs associated with rate-term refinances are nontrivial, it is possible that rate-term refinancings may also be marker for rapidly-appreciating homes.

[62] I note that the dilution may not be entirely pernicious. To the extent that they are outliers, some of the rapidly appreciating properties may have an undue impact on index estimates.

[63] It is also possible, of course, that the underlying assumption of the analysis is invalid. That is, for multiple transaction pairings of the same property, the pairings' relative appreciation rates may not be correlated. In this context, the empirical exercise described here is not meaningful.

It should be noted that, even if such effects are present, the use of a purchase-only index would not be a solution to the problem. Because the purchase-only index excludes data from many homes having cash-out refinances, it is also susceptible to underestimating appreciation rates in an environment where cash-out refinancings are positively correlated with appreciation rates.

Conclusions

Despite the possibility that various confounding effects may be spuriously absorbed into the "bias" coefficients, on balance, the bias-neutral indexing approach seems to be a promising methodology that warrants further review. It alleviates some concerns about appraisal bias while obviating the need to discard refinance-related appraisal data. As discussed in this paper, discarding refinance appraisals from the estimation process significantly increases the imprecision of house price index estimates.

New Research Questions

Although the results of this preliminary investigation suggest that the bias-neutral index may not differ significantly from the HPI that is currently used, it is possible that the results for the specific geographies used here may be idiosyncratic. As such, it will be useful to extend the empirical work to include additional states and all of the U.S. census divisions.

Because small changes in the HPI values can have significant effects on the capital requirements determined by OFHEO's risk-based capital (RBC) model, it would be useful to calculate the sensitivity of the RBC requirements to the use of the bias-neutral index. Such an analysis would be most enlightening if it were performed retrospectively, so that the HPI and the bias-neutral index results are compared across several historical periods. Given that the bias-neutral index does not seem to consistently estimate higher or lower price appreciation rates than the HPI, it is unlikely that the RBC requirements will *systematically* differ between the two approaches. However, the retrospective analysis would allow for the identification of any periods in which the two approaches would have yielded different Enterprise capital classifications under OFHEO's capital classification system.

Two additional promising topics arise from some of the empirical findings discussed in this paper. The first involves the fact that appraisal bias on rate-term refinance appraisals seems to have shrunk over the last several years. The house price modeling process would benefit from a determination of whether this phenomenon is widespread (i.e., is the bias shrinking across all geographic regions?) and whether it will likely continue.

The second subject relates to the timing of and purpose for successive house valuations and the information they provide in index estimation. For example, one of this paper's empirical findings is that appreciation rates are relatively high for houses that have a cash-out refinance sandwiched between two purchase transactions. While that result may be more germane to the debate concerning the degree to which the HPI reflects quality improvements,[64] the use of information found in financing patterns might be quite useful in estimating appraisal bias. For example, if a subject home were sold several times immediately prior to a refinancing, the refinance-related appraisal might be significantly more accurate than other appraisals *ceteris paribus*.

One final related subject that has not been explored in this analysis entails the relationship between appraisal bias and LTV ratios. The bias-neutral index discussed in this paper provides estimates for the *average* level of refinance-related appraisal bias, but does not measure how those levels differ for loans having different LTV ratios. A simple modification to the bias-neutral index, the introduction of interactive variables that are the product of the LTV and the refinance dummy variables, would facilitate such a measurement.[65] A review of the coefficient estimates would then provide evidence for whether appraisal bias is most severe for high LTV mortgages.

---

[64] See, for example, McCarthy and Peach (2004). I note that the finding would obviously not be relevant to that debate if it merely reflects a positive correlation between appreciation rates and cash-out refinance activity.

[65] To the extent that LTV ratios are correlated with the type of appraisal valuation performed, the coefficient on the interactive variable will, to some extent, capture the different biases associated with different appraisal types.

References

Abraham J. M. and W. S. Schauman. (1991). "New Evidence on Home Prices from Freddie Mac Repeat Sales." *AREUEA Journal*, 19(3): 333-52.

"Advance Notice of Proposed Rulemaking." (1999, April 13). *Federal Register*, 64(70): 18083-18300.

Bailey, M., R. Muth and H. Nourse (1963). "A Regression Method for Real Estate Price Index Construction." *Journal of the American Statistical Association*, 58, 933-942.

Berson, D. W. (2005, August 8). "Continued Strong Home Price Gains." Available at http://www.fanniemae.com/media/pdf/berson/weekly/080805.pdf.

Calhoun, C. A., P. Chinloy, and I. F. Megbolugbe. (1995). "Temporal Aggregation and House Price Index Construction." *Journal of Housing Research*, 6(3): 419-439.

Calhoun, C. A. (1991, June/July). "Estimating Changes in Housing Values from Repeat Transactions." Paper presented at the Western Economic Association International Meetings, Seattle, WA.

Calhoun, C. A. (1996, March). "OFHEO House Price Indexes: HPI Technical Description." *HPI Methodology Report*, Available at www.ofheo.gov/Media/Archive/house/hpi_tech.pdf.

Calhoun, C. A., M. C. Harter, and M. R. VanderGoot. (1998, July). "Appraisals, Repeat Mortgage Transactions, and House Price Indices." Paper presented at the ERES/AREUEA International Conference, Maastricht, The Netherlands.

Case, B., H. O. Pollakowski, and S. M. Wachter. (1991). "On Choosing among House Price Index Methodologies." *AREUEA Journal* 19 (3): 286-307.

Case, B., H. O. Pollakowski, and S. M. Wachter. (1997). "Frequency of Transactions and House Price Modeling." *Journal of Real Estate Finance and Economics*, 14(1-2): 173-87.

Case, K. and R. Shiller. (1987, Sept/Oct). "Prices of Single-Family Homes since 1970: New Indexes for Four Cities." *New England Economic Review*, 87: 45-56.

Case, K. and R. Shiller. (1989). "The Efficiency of the Market for Single-Family Homes." *American Economic Review*, 79(1): 125-137.

Cheung, U.S., K. K. W Yau, and Y.V. Hui (2004). "The Effects of Attributes on the Repeat Sales Pattern of Residential Property in Hong Kong." *Journal of Real Estate Finance and Economics*, 29(3): 321-339.

Chinloy, P., M. Cho and I. F. Megbolugbe. (1997). "Appraisals, Transaction Incentives, and Smoothing." *Journal of Real Estate Finance and Economics*, 14(1): 89-111.

Cho, M. and I. F. Megbolugbe. (1996). "An Empirical Analysis of Property Appraisal and Mortgage Redlining." *Journal of Real Estate Finance and Economics*, 13(1): 45-55.

Dotzour, M. G. (1988). "Quantifying Estimation Bias in Residential Appraisal." *Journal of Real Estate Research*, 3(3): 1-11.

Dreiman, M. H. and A. Pennington-Cross. "Alternative Methods of Increasing the Precision of Weighted Repeat Sales House Prices Indices." *Office of Federal Housing Enterprise Oversight Working Paper 02-02*, available at www.ofheo.gov.

Edelstein, R. H. and D. C. Quan. (2004). "How Does Appraisal Smoothing Bias Real Estate Returns Measurement?" *The Journal of Real Estate Finance and Economics*, 32(1).

"Fannie Gets Tough on Inflated Appraisals." (2004, January 13). *Dow Jones*.

Gao, A. and G. H. K. Wang. (2005, April 29). "Multi Transactions Model for Constructing House Price Indices." Working paper presented to Washington Area Finance Association Meeting.

Gatzlaff, D. H. and D. R. Haurin. (1993, January). "Sample Selection and Biases in Local House Value Indices." Draft paper presented at the American Real Estate and Urban Economics Association Meetings.

Gatzlaff, D. H. and D. R. Haurin. (1997). "Sample Selection Bias and Repeat-Sales Index Estimates." *Journal of Real Estate Finance and Economics* 14(1-2): 33-50.

Hwang, Min and John M. Quigley. "Selectivity, Quality Adjustment and Mean Reversion in the Measurement of House Values*." Journal of Real Estate Finance and Economics* 28(2-3): 161-178.

Lacour-Little, M. and S. Malpezzi, "Appraisal Quality and Residential Mortgage Default: Evidence from Alaska." *Journal of Real Estate Finance and Economics* 27(2): 211-233.

Loebs, T. (2005, June). "Systemic Risks in Residential Property Valuations: Perceptions and Reality." *Collateral Assessment and Technologies Committee*, Available at http://www.cswv.com/pdfs/CATC_SystemicRisks1.1.pdf.

McCarthy, J. and R.W. Peach, "Are Home Prices the Next 'Bubble'?" *FRBNY Economic Policy Review* 10(3): 1-17.

McManus, Douglas A. and Sol T. Mumey. June 2002. "System and Method for Providing House Price Forecasts Based on Repeat Sales Model." U.S. Patent 6,401,070B1.

"Metropolitan Area Existing-Home Prices and State Existing-Home Sales." National Association of Realtors. Available at www.realtor.org/research.nsf/pages/MetroPrice?OpenDocument.

Myers, D. and J. R. Pitkin. (1995). "Evaluation of Price Indices by a Cohort Method." *Journal of Housing Research*, 6(3): 497-518.

Stephens, W., Y. Li, V. Lekkas, J. Abraham, C. Calhoun, and T. Kimner. (1995). "Conventional Mortgage Home Price Index." *Journal of Housing Research* 6(3): 389-418.

Wang, F. T. and P. M. Zorn. (1997). "Estimating House Price Growth with Repeat Sales Data: What's the Aim of the Game?" *Journal of Housing Economics*, 6(2): 93-118.

Figure 1: Appreciation in Homes Having
Pu-Co-Pu Financing Compared with Other Financing Patterns

Illinois

Pu-CO-Pu

A "Pu-Co-Pu" sequence of mortgages is one in which a purchase transaction is followed by a cash-out refinance and then another purchase.

Others

House Price Index (1980Q1=100)

370.00

320.00

270.00

220.00

170.00

199603 199701 199703 199801 199803 199901 199903 200001 200003 200101 200103 200201 200203 200301 200303 200401 200403 200501
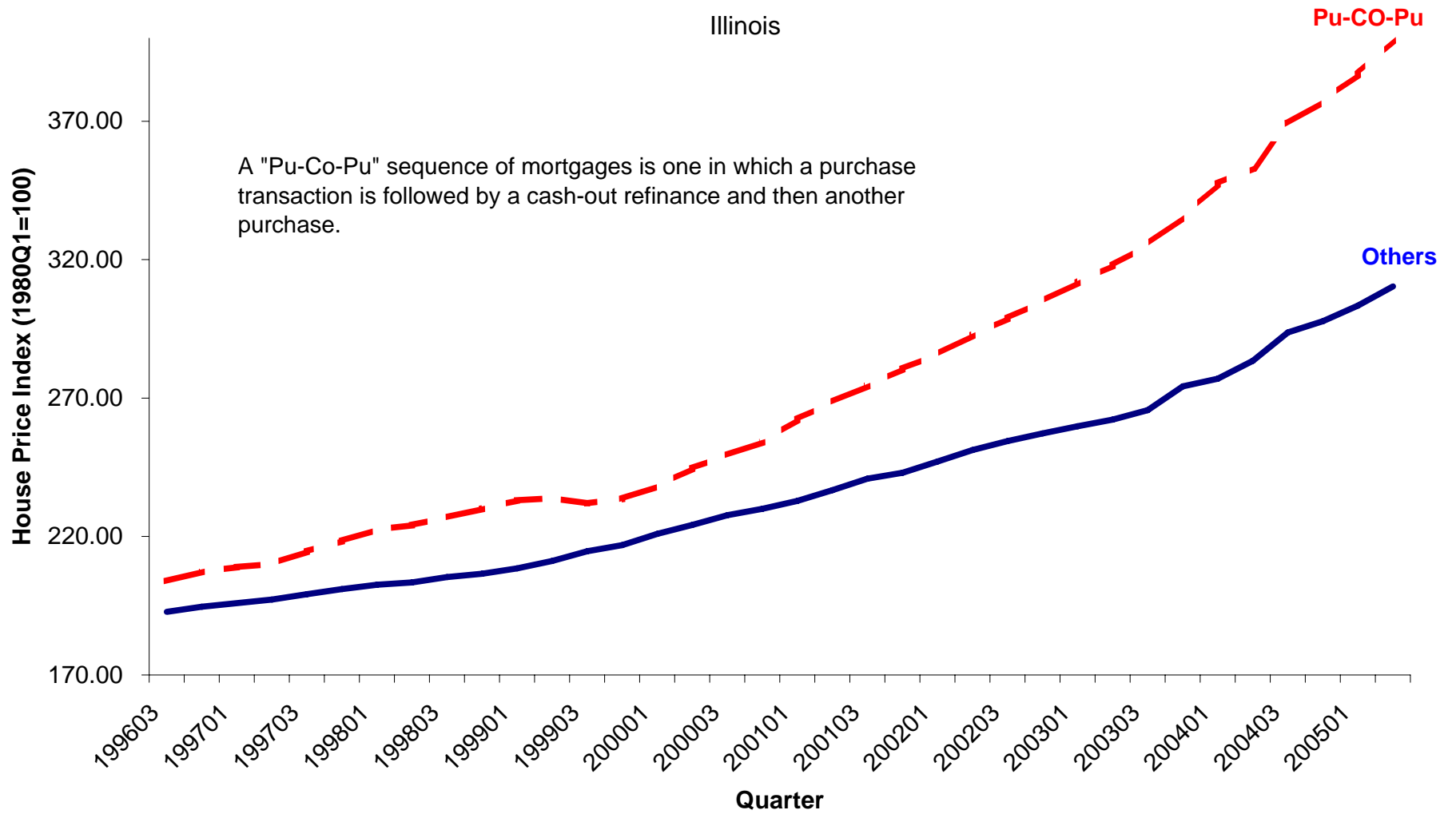
Quarter

Figure 2a: Four-Quarter Percentage Change in House Prices: By Methodology

Washington State
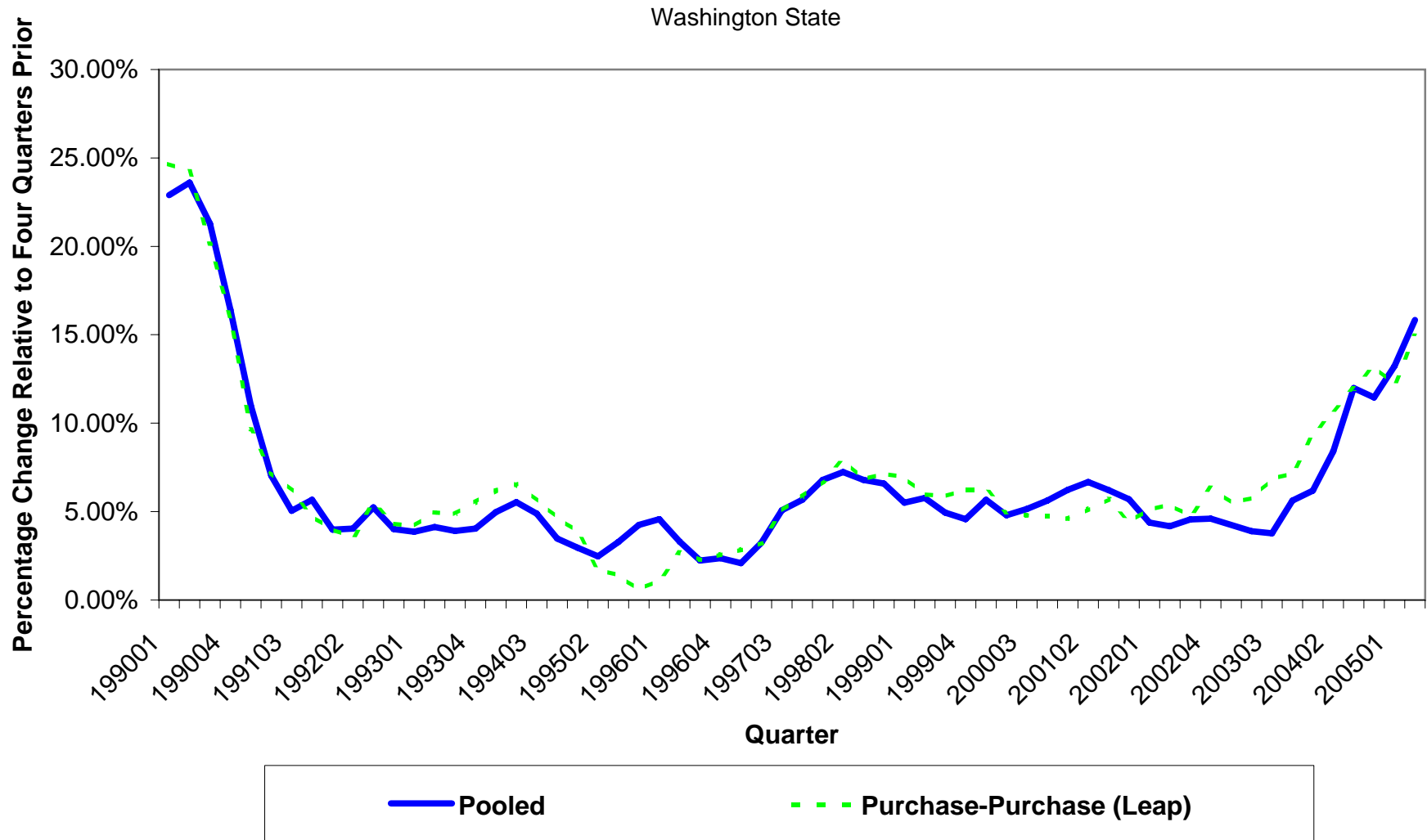
Figure 2b: Four-Quarter Percentage Change in House Prices: By Methodology

California

Percentage Change Relative to Four Quarters Prior

Quarter

— Pooled   - - - Purchase-Purchase (Leap)
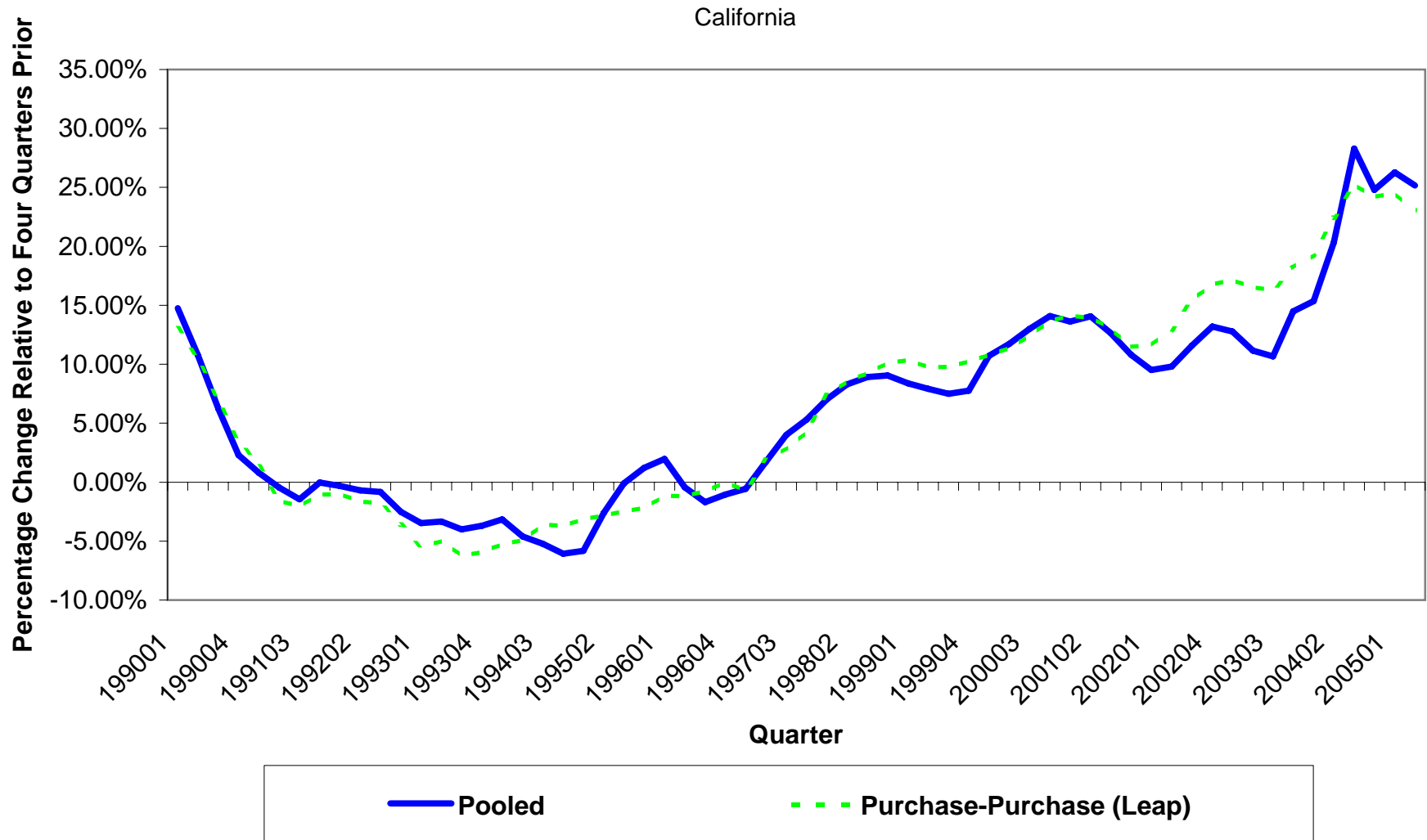
Figure 2c: Four-Quarter Percentage Change in House Prices: By Methodology

Illinois
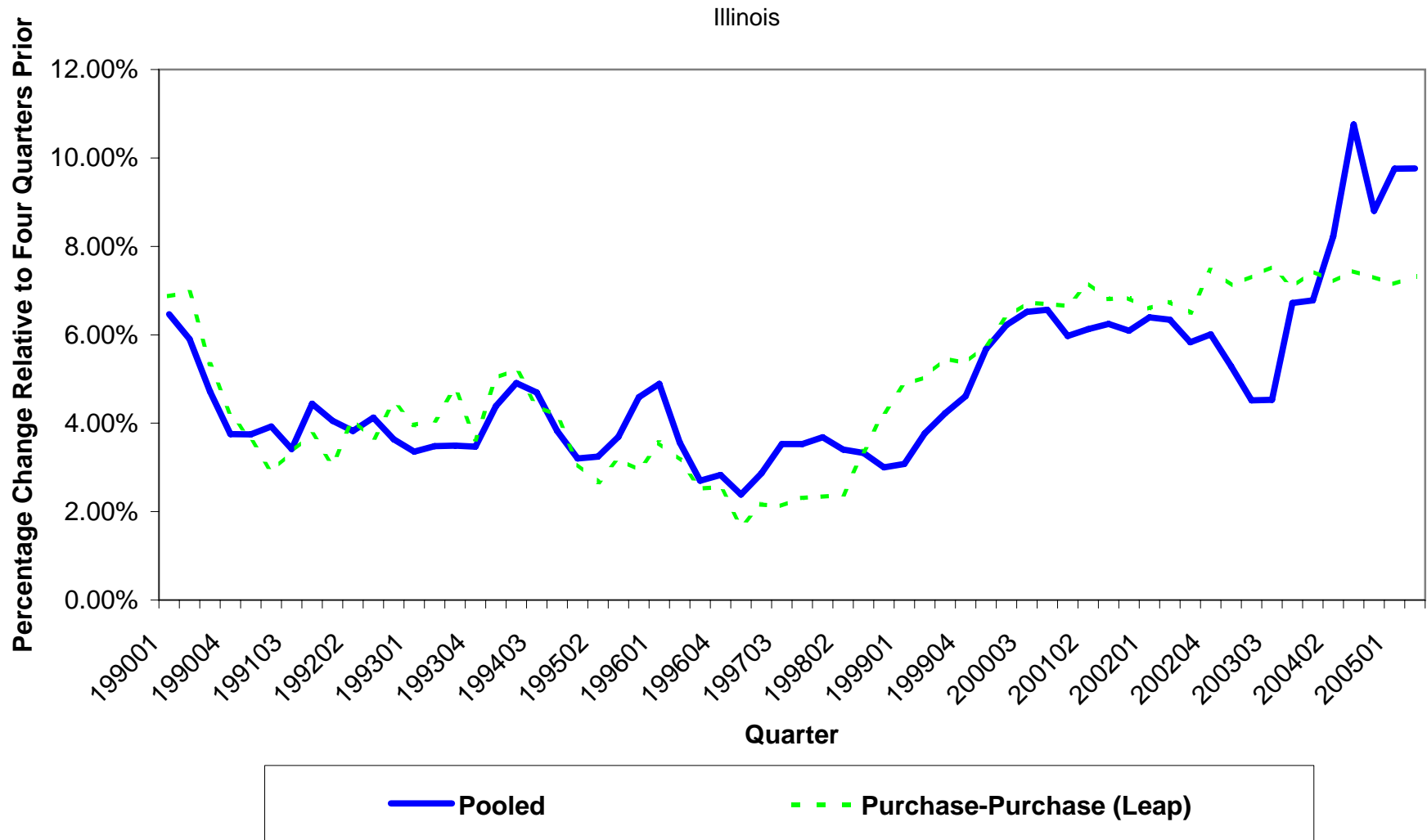
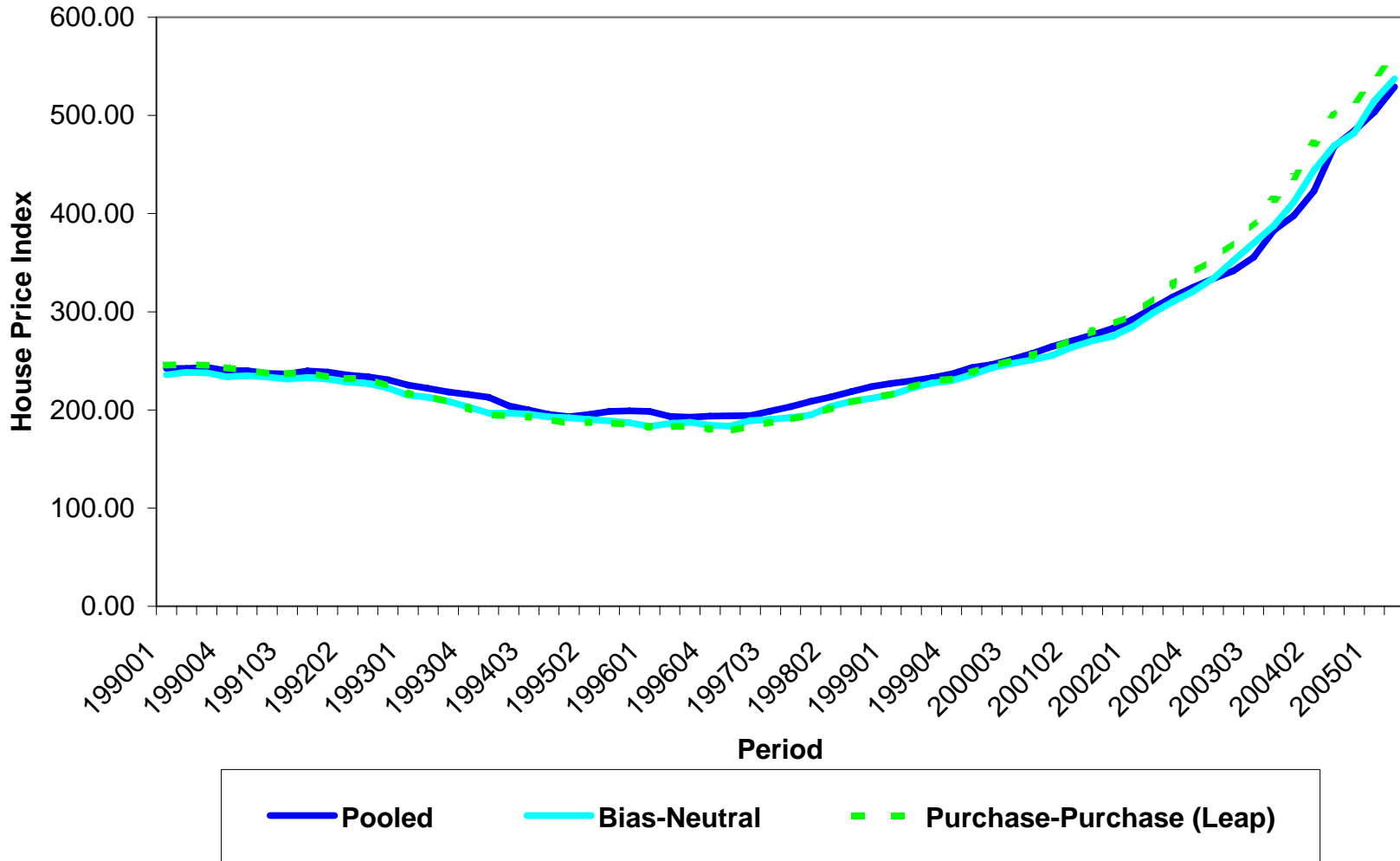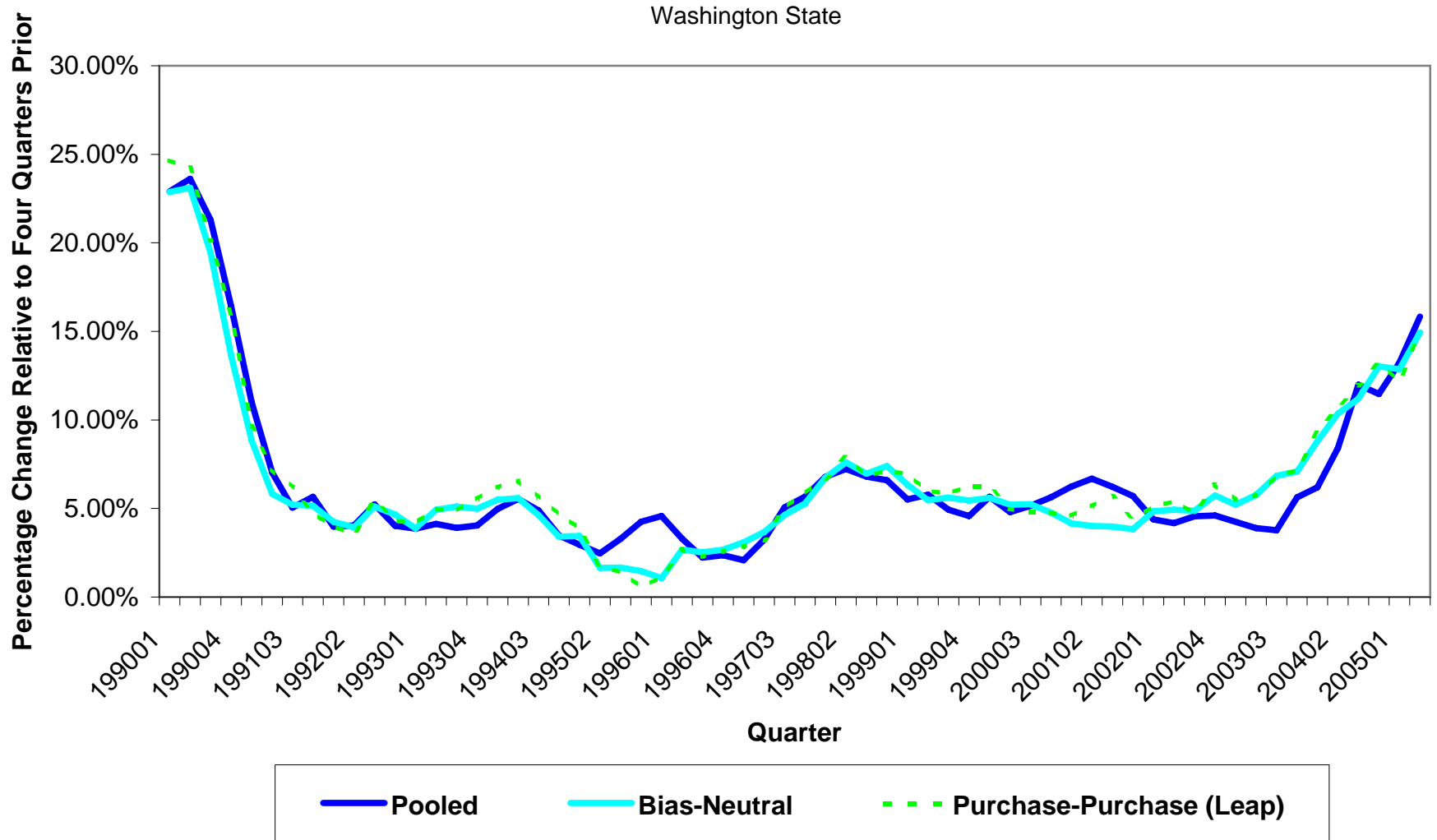**Figure 2d: Four-Quarter Percentage Change in House Prices: By Methodology**

San Diego MSA
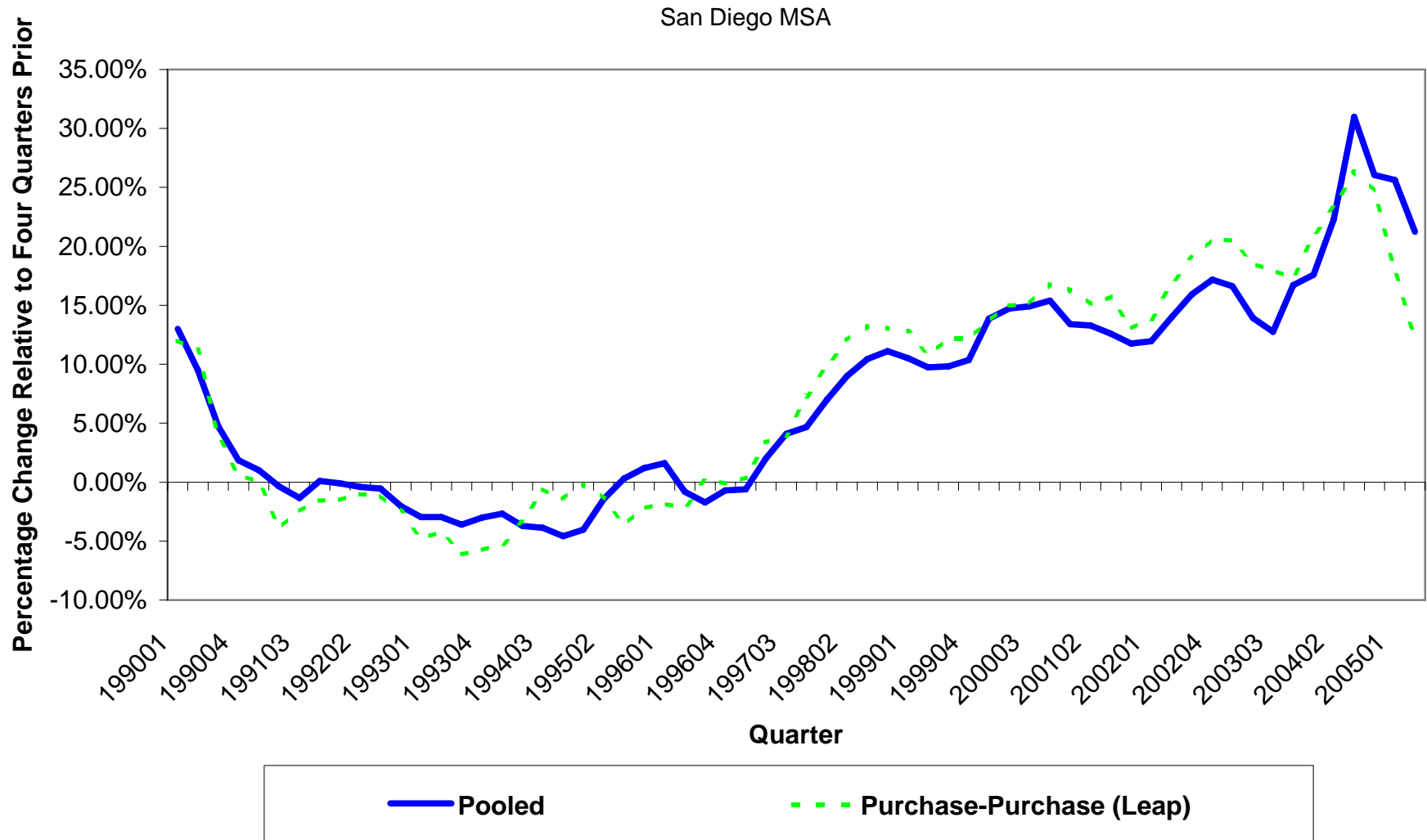
Figure 2e: Four-Quarter Percentage Change in House Prices: By Methodology

Los Angeles-Long Beach-Glendale, CA MSAD
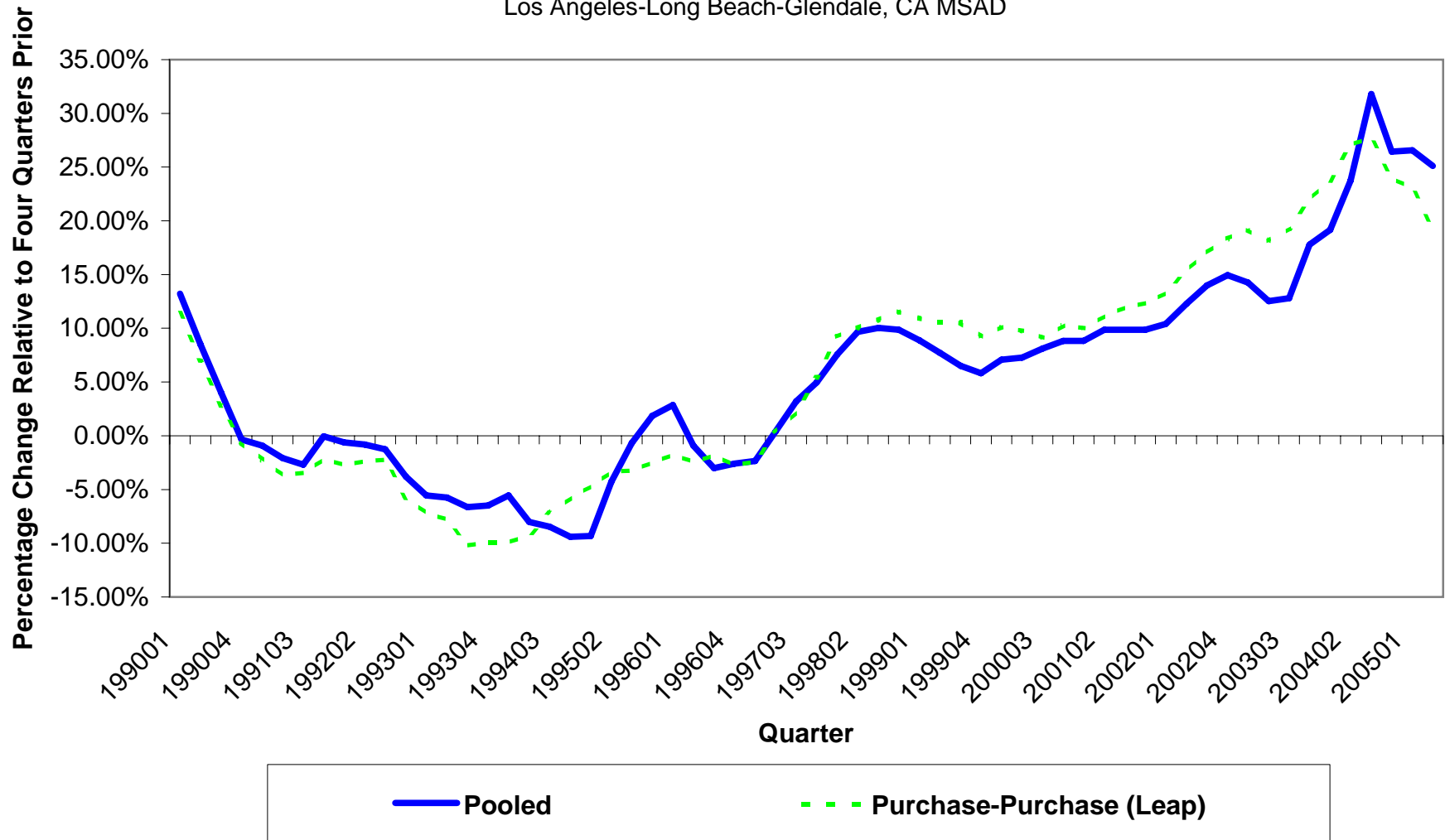
Figure 3a: Index Growth Rate Differences versus Proportion of Loans that are Refinances

Washington State

Difference between Two Indices' Four-Quarter Growth Rates

Proportion of Refinance Mortgages

Quarter

— Difference between Pooled and Purchase-Only Indices Four-Quarter Growth Rates

⋯ Proportion of Loans that were Refinacings (Second Y-Axis)
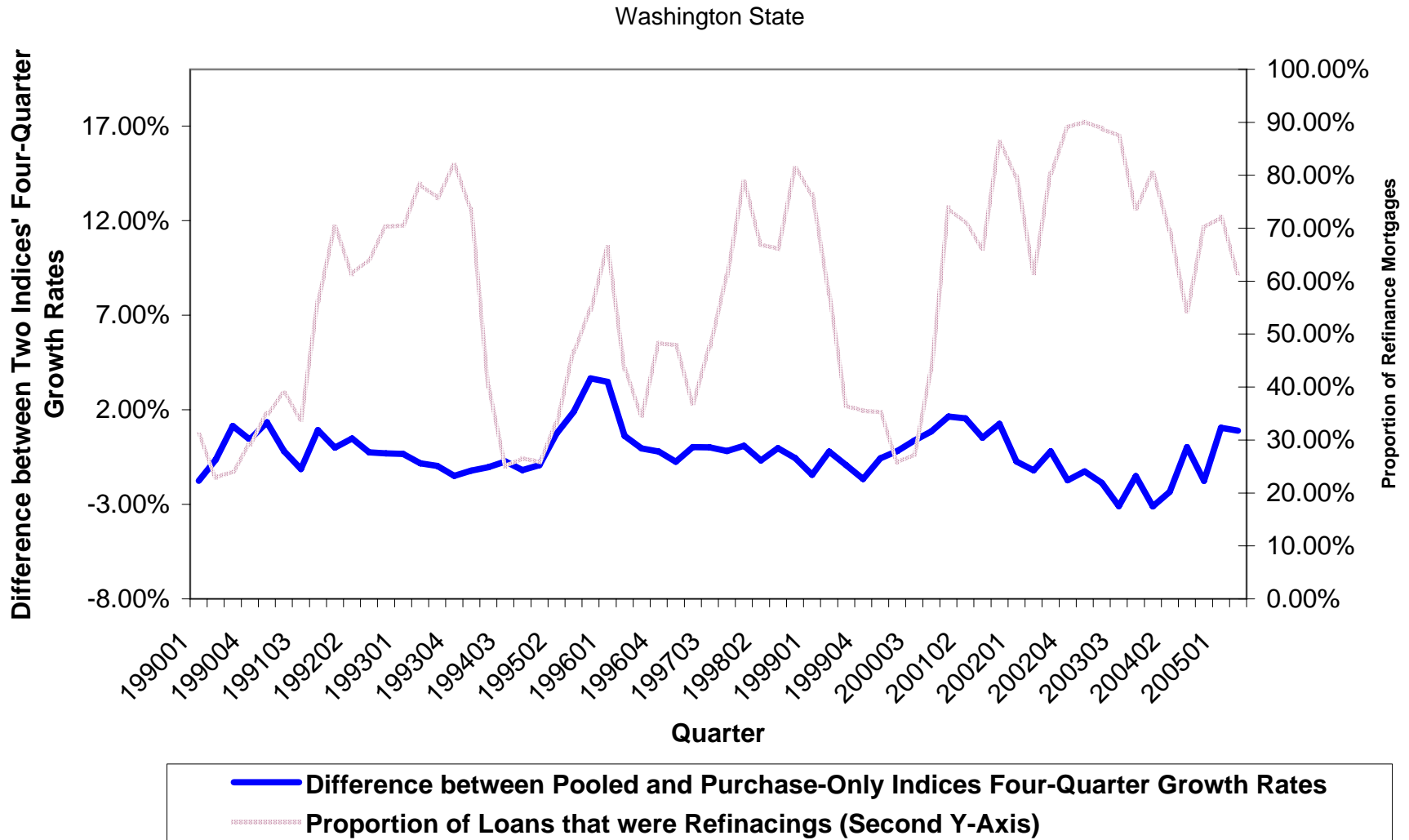
Figure 3b: Index Growth Rate Differences versus Proportion of Rate-Term Refinances
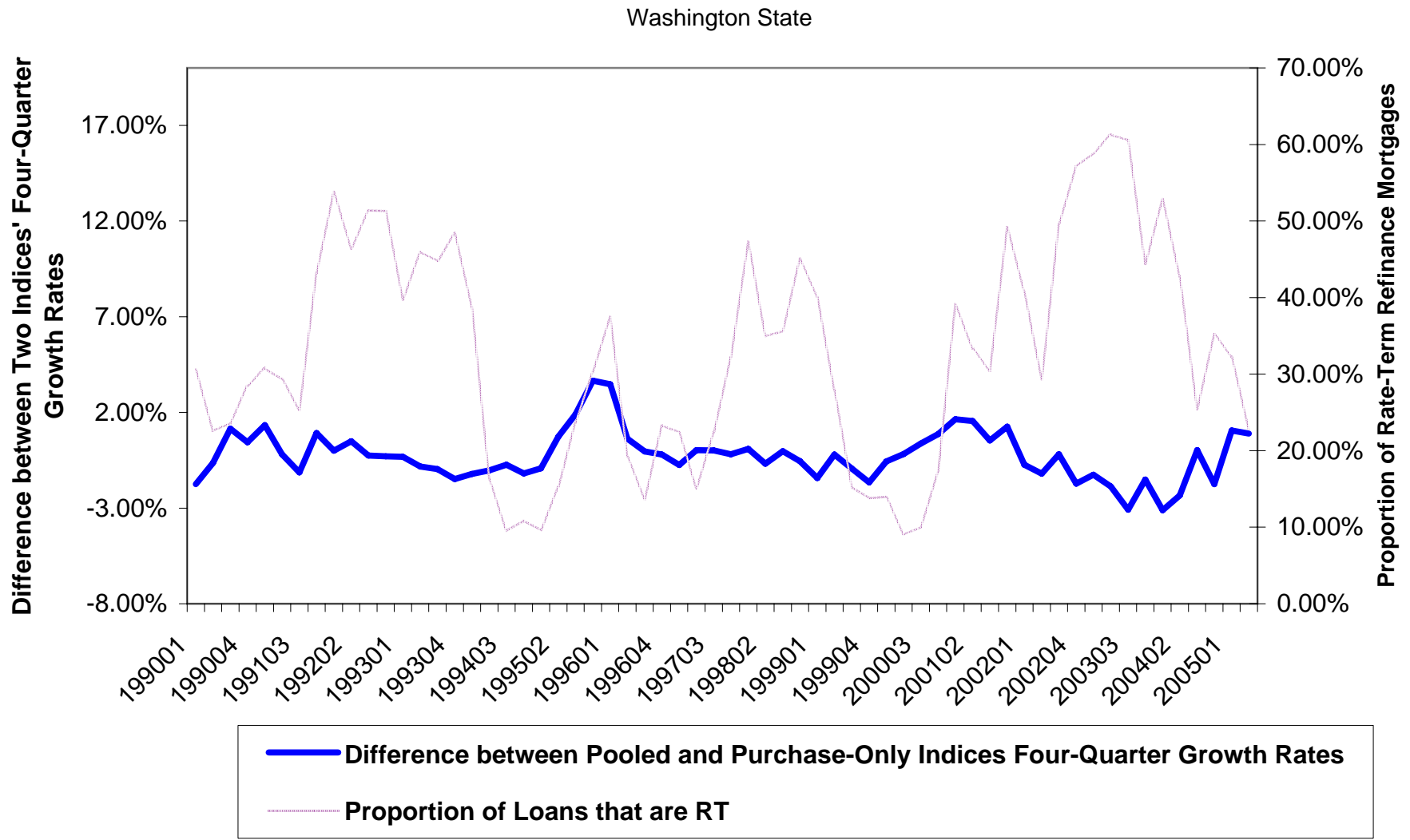
Washington State

Figure 4a: Relative Imprecision of Purchase-Only Index under Small Samples

HPI Changes for Lima, Ohio

Percentage Change Relative to Four Quarters Prior

Quarter

Legend: Pooled — — — Purchase-Purchase (Leap)
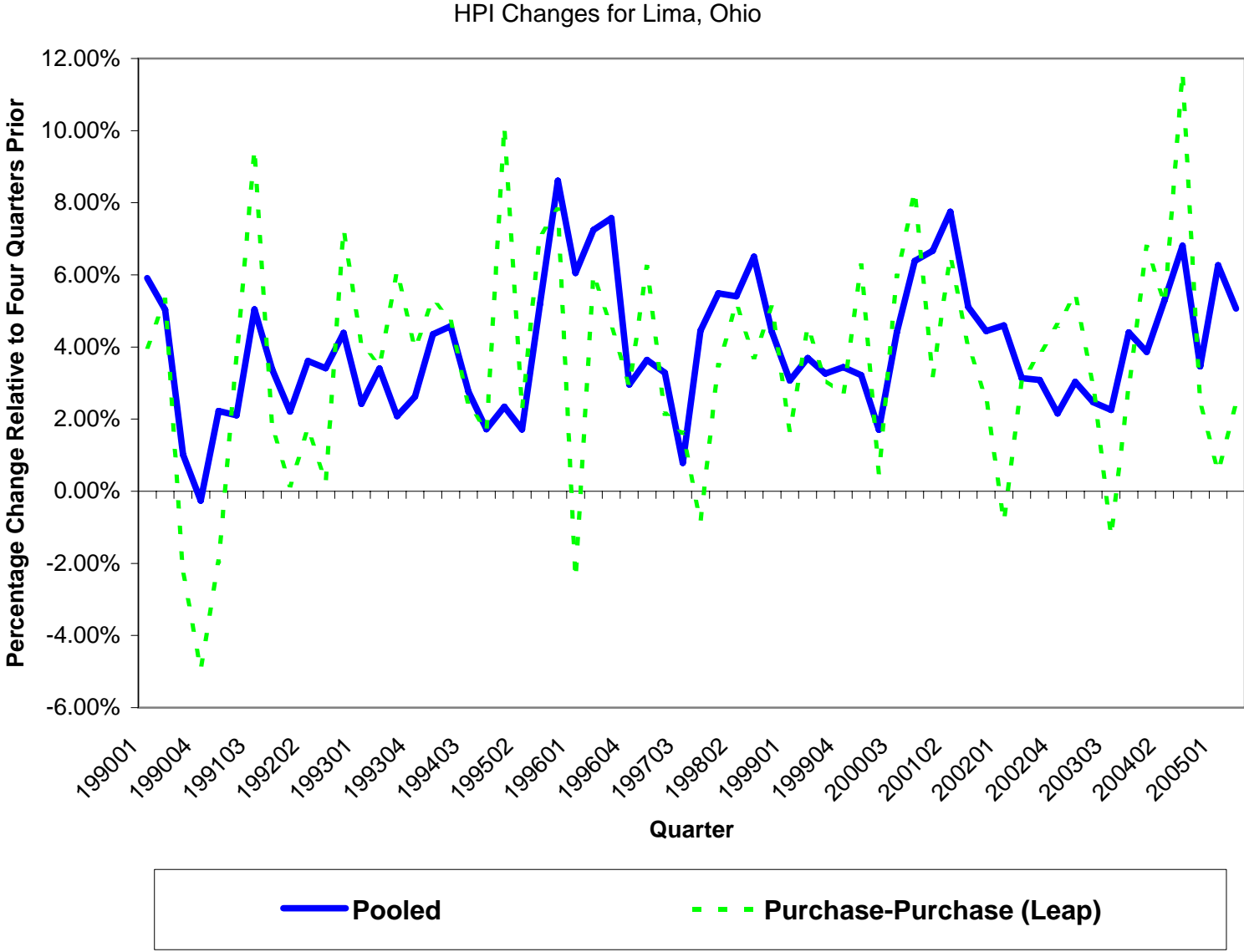
Figure 4b: Confidence Bands for Lima, Ohio by Indexing Method

1992-1994
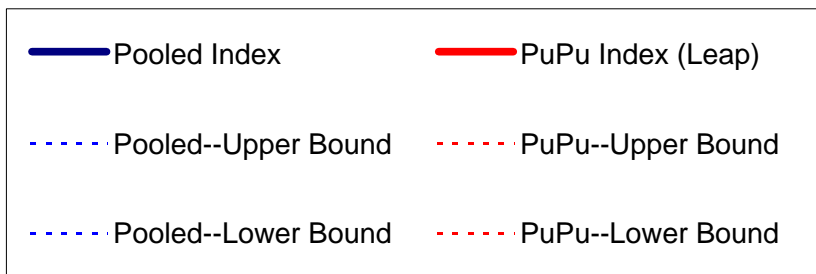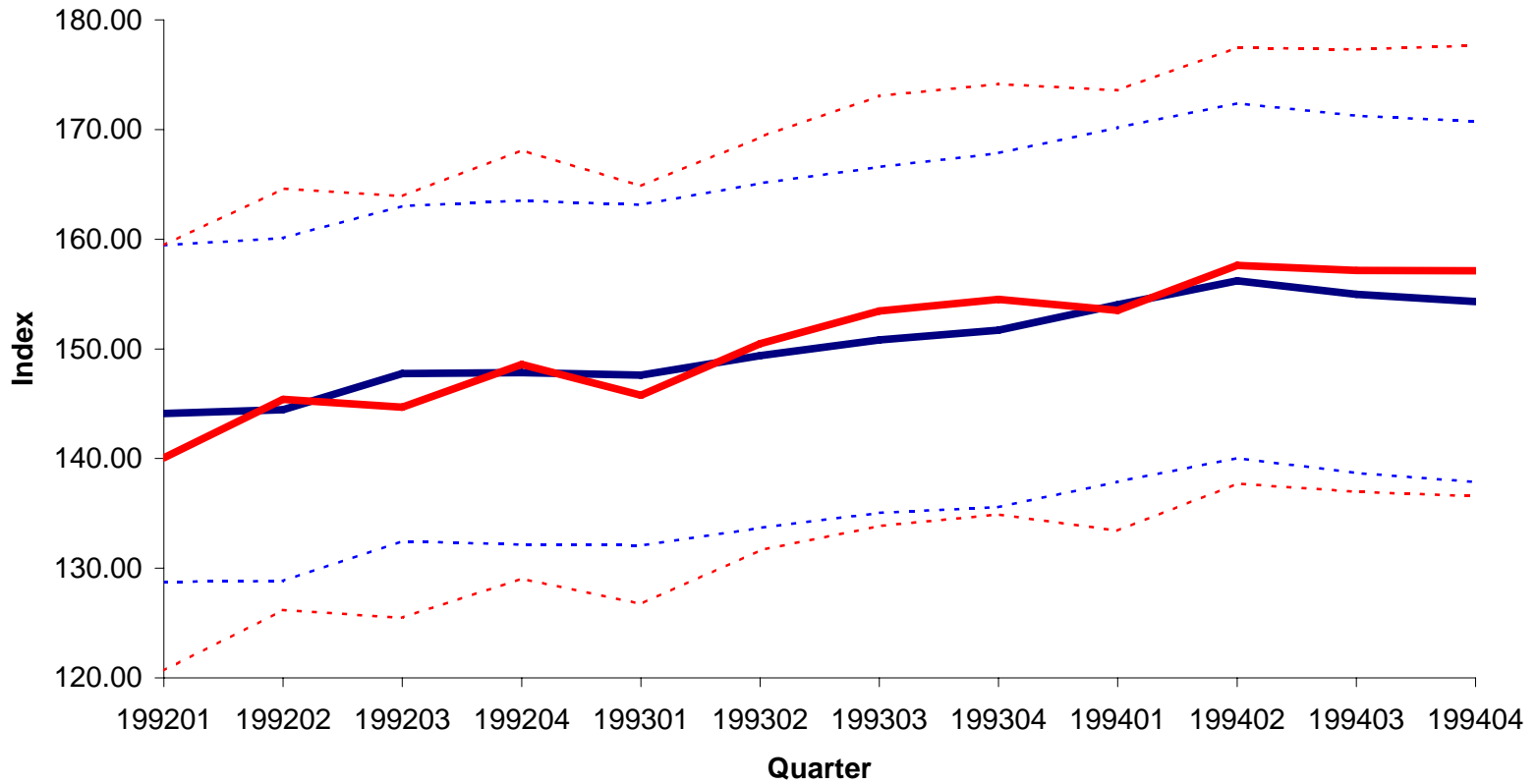
Figure 5a: House Price Indices: By Methodology

Washington State
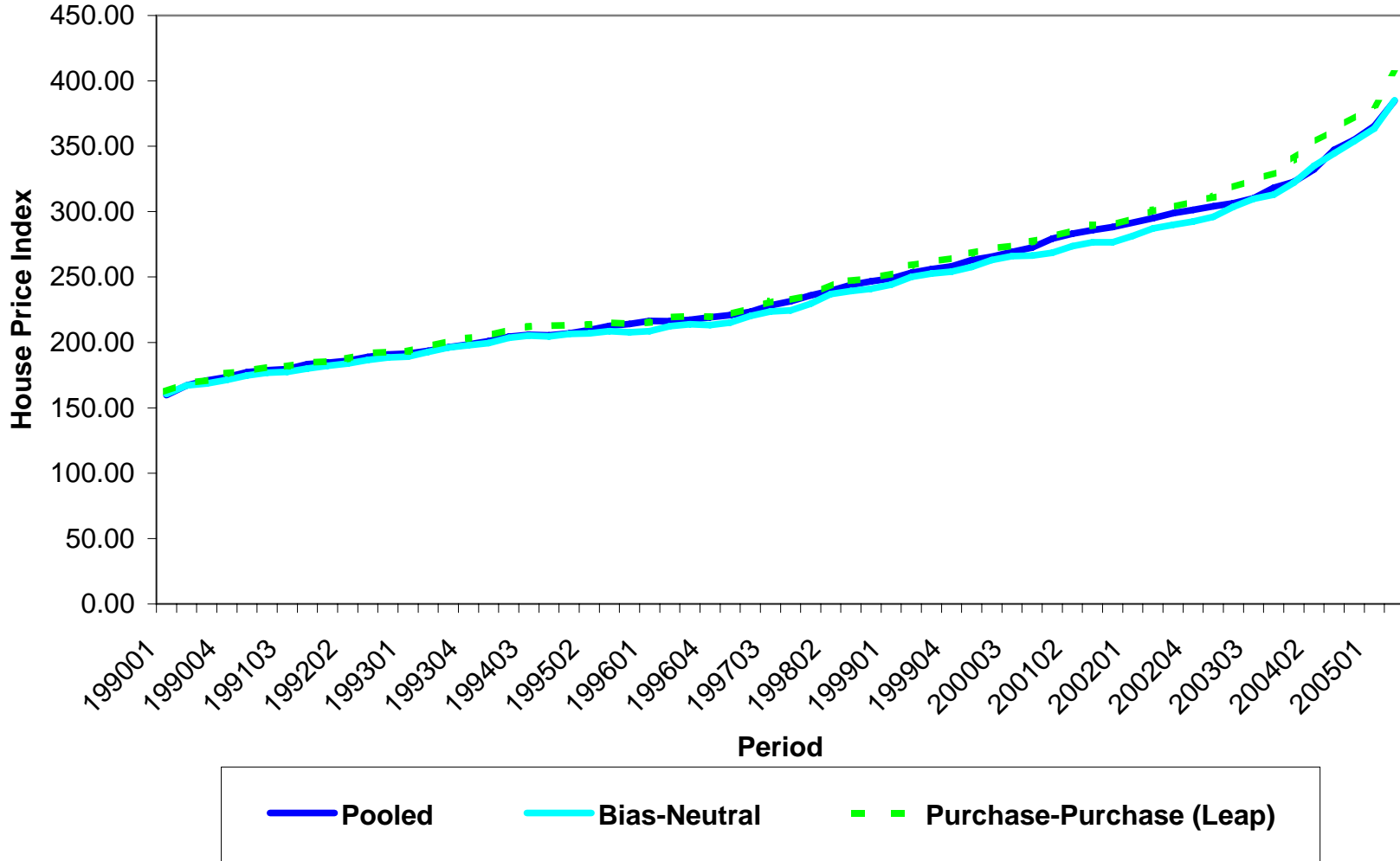
Figure 5b: House Price Indices: By Methodology

California

Legend:
- **Pooled** (dark blue solid line)
- **Bias-Neutral** (cyan solid line)
- **Purchase-Purchase (Leap)** (green dashed line)

Y-axis: House Price Index (0.00 to 700.00)

X-axis: Period (199001 to 200501)
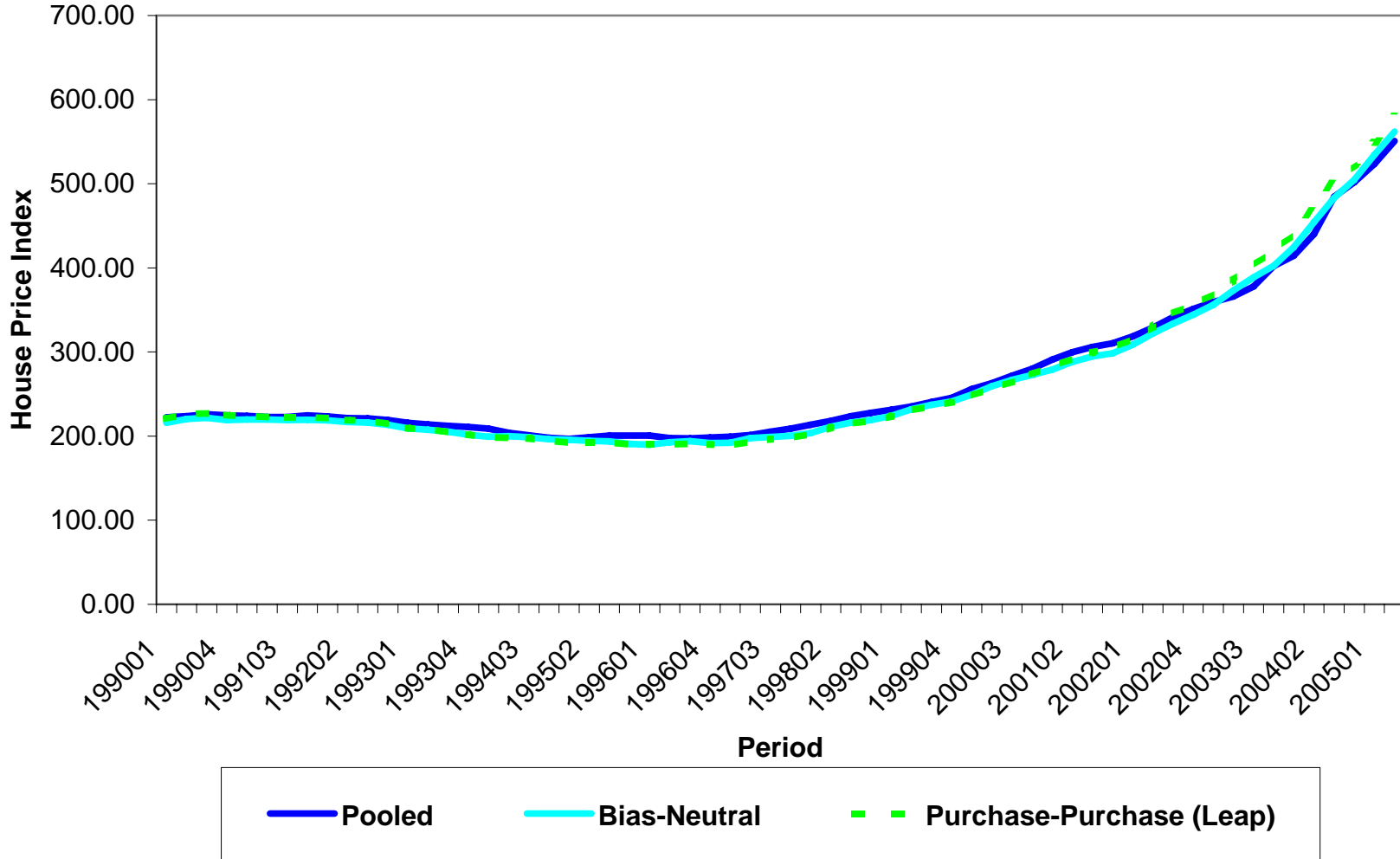
Figure 5c: House Price Indices: By Methodology

Illinois

Legend: Pooled — Bias-Neutral — Purchase-Purchase (Leap)
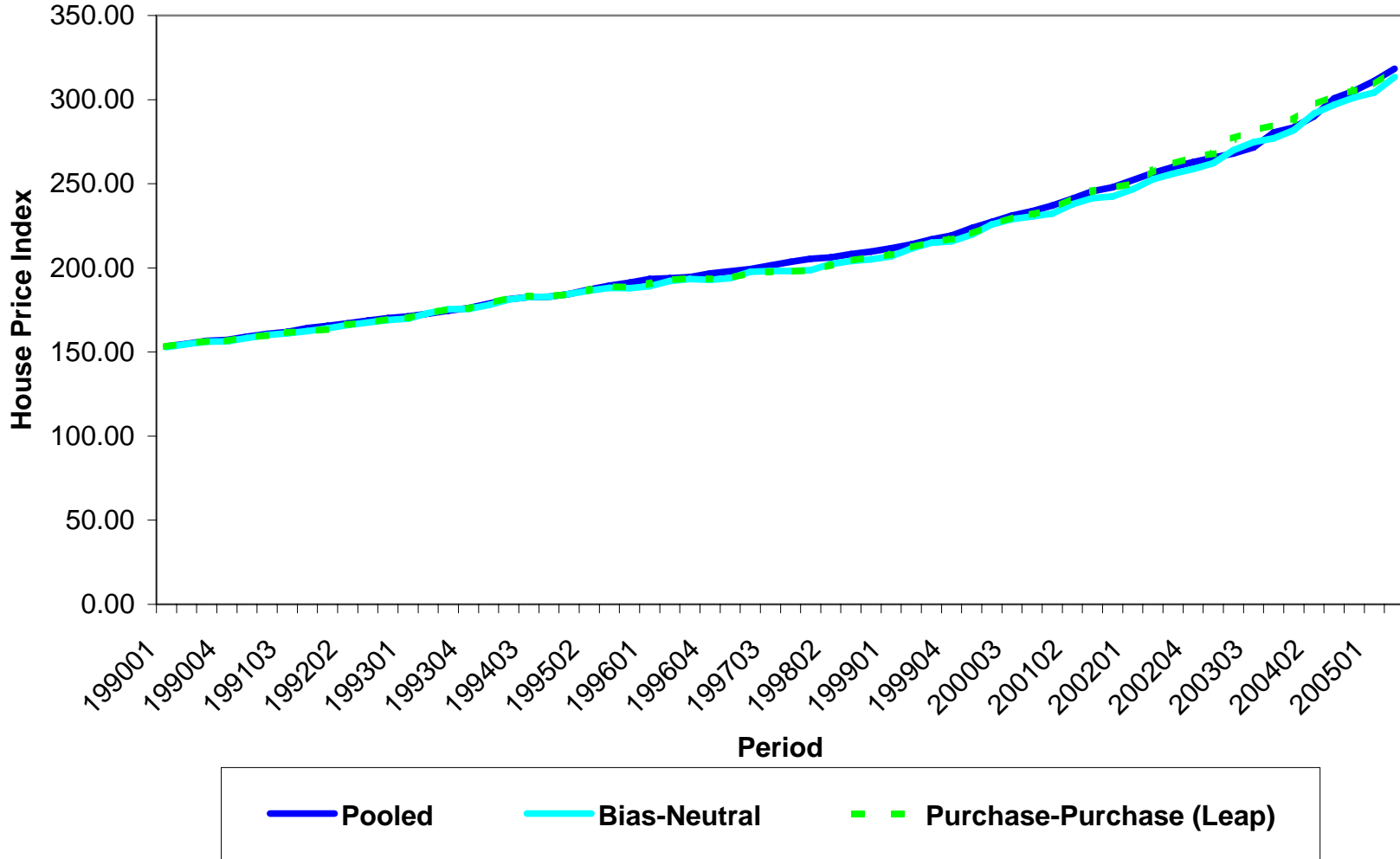
Figure 5d: House Price Indices: By Methodology

San Diego MSA

| | Pooled | | Bias-Neutral | | Purchase-Purchase (Leap) |
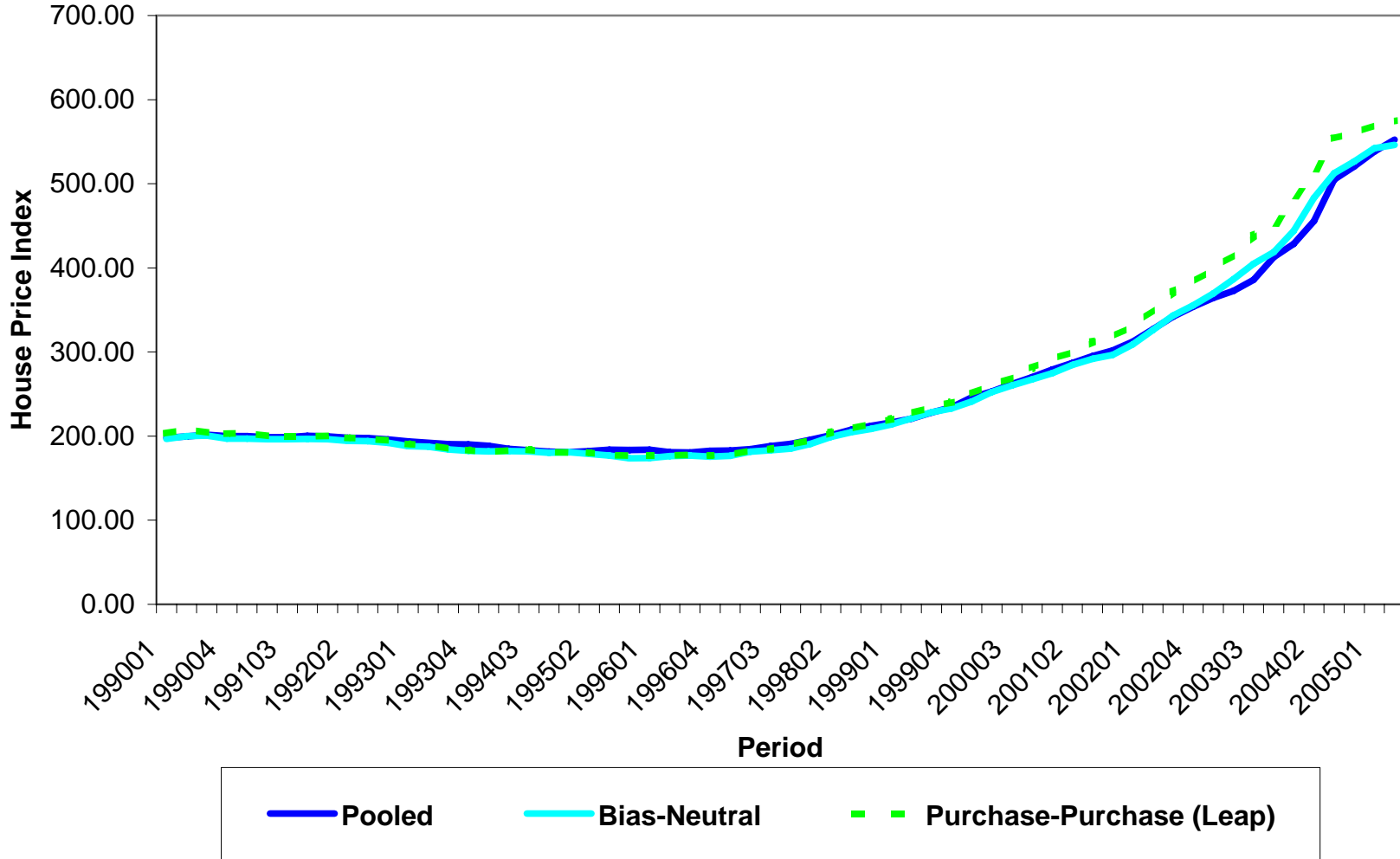
Figure 5e: House Price Indices: By Methodology

Los Angeles-Long Beach-Glendale, CA

Figure 6a: Four-Quarter Percentage Change in House Prices: By Methodology

Washington State

Pooled — Bias-Neutral — Purchase-Purchase (Leap)

Figure 6b: Four-Quarter Percentage Change in House Prices: By Methodology

California
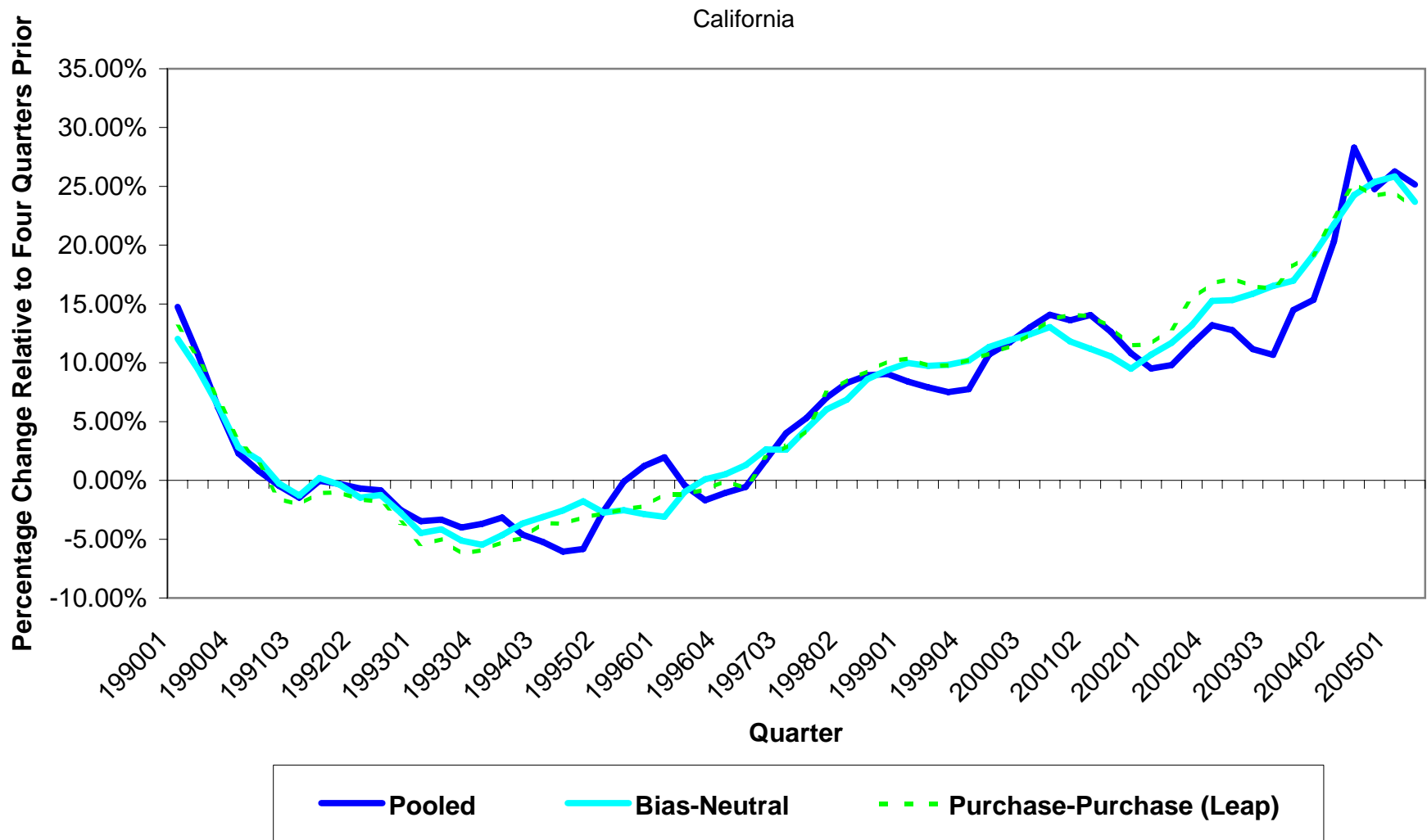
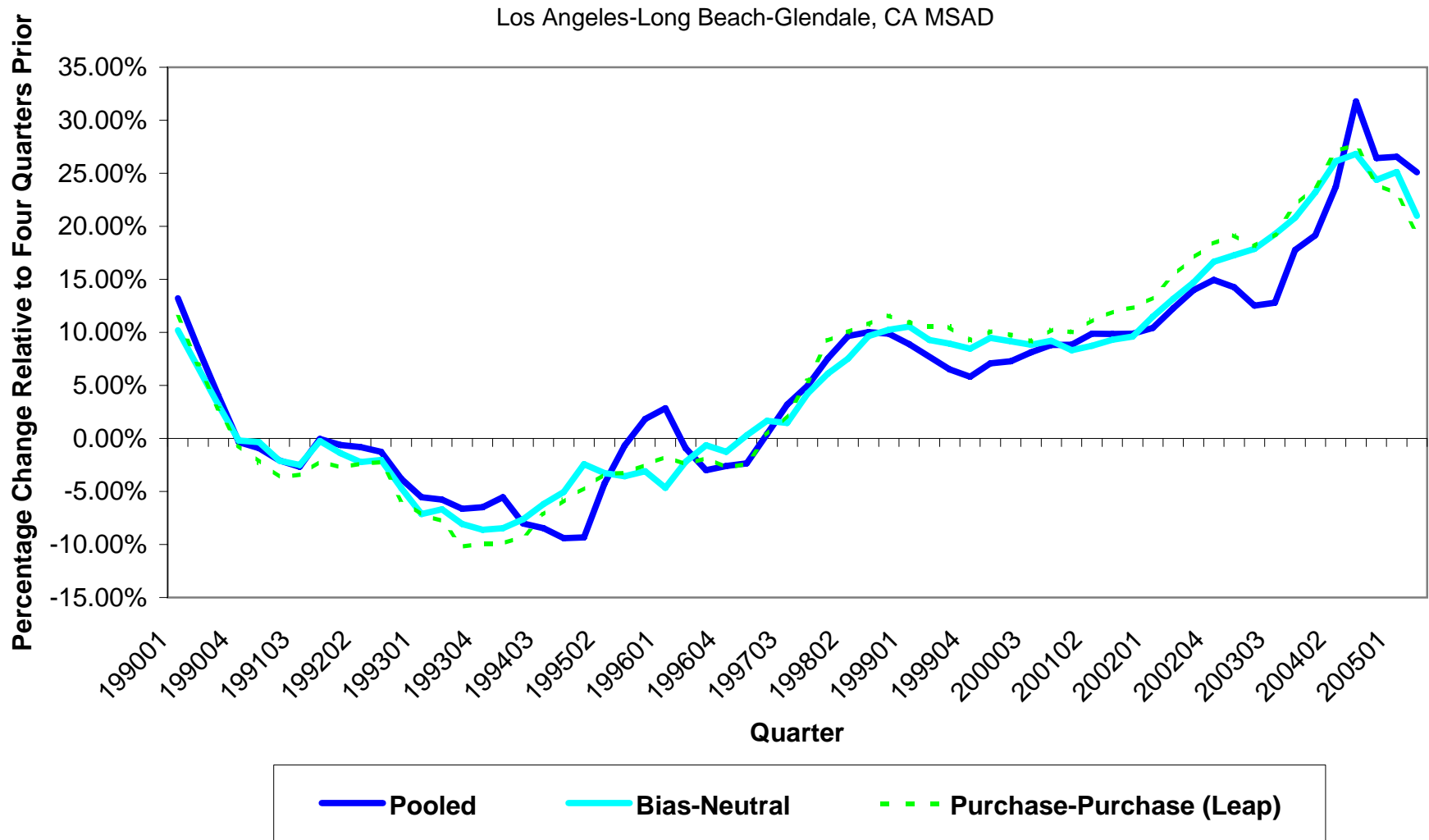**Figure 6c: Four-Quarter Percentage Change in House Prices: By Methodology**

Illinois
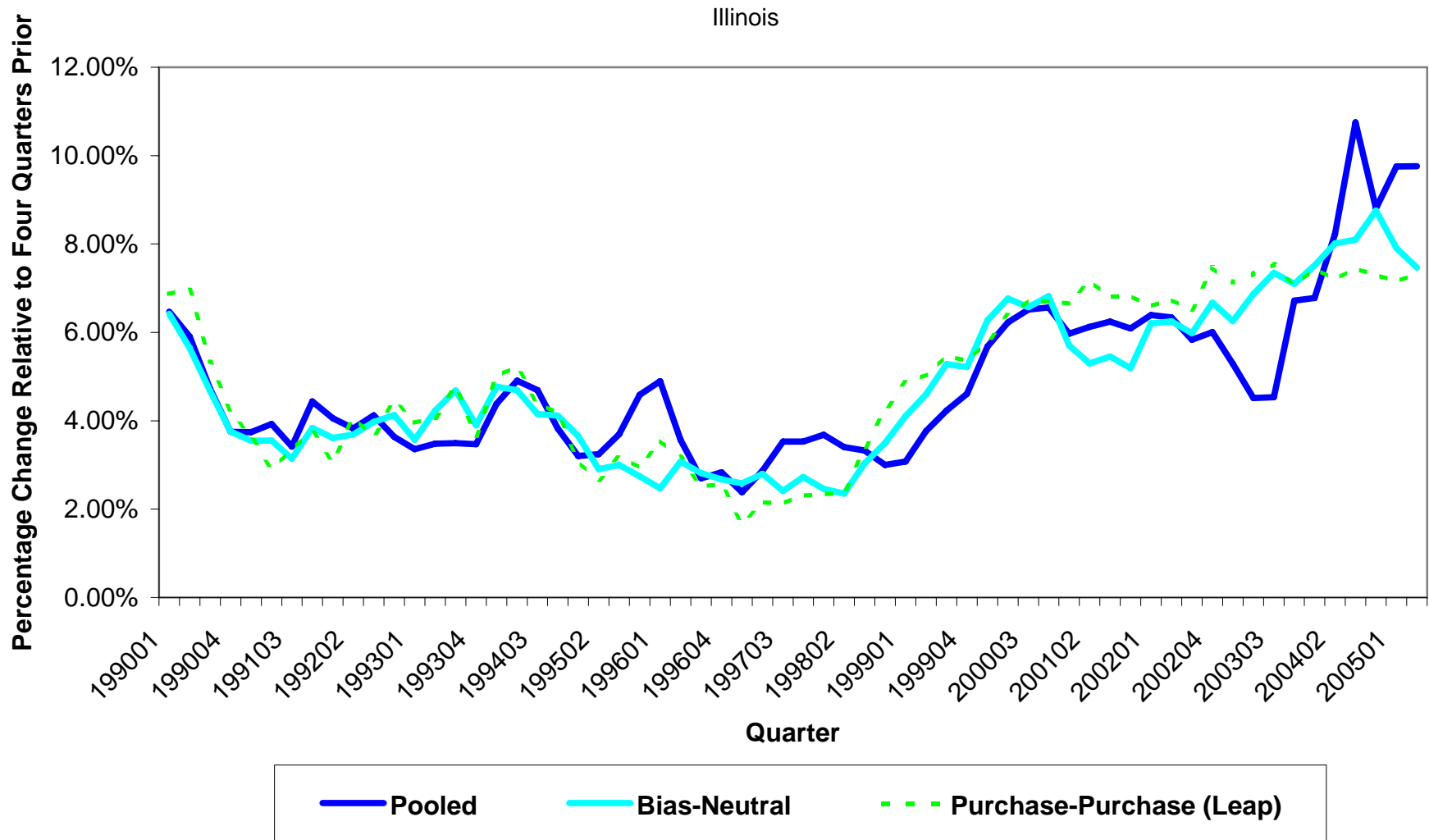
Figure 6d: Four-Quarter Percentage Change in House Prices: By Methodology

San Diego MSA
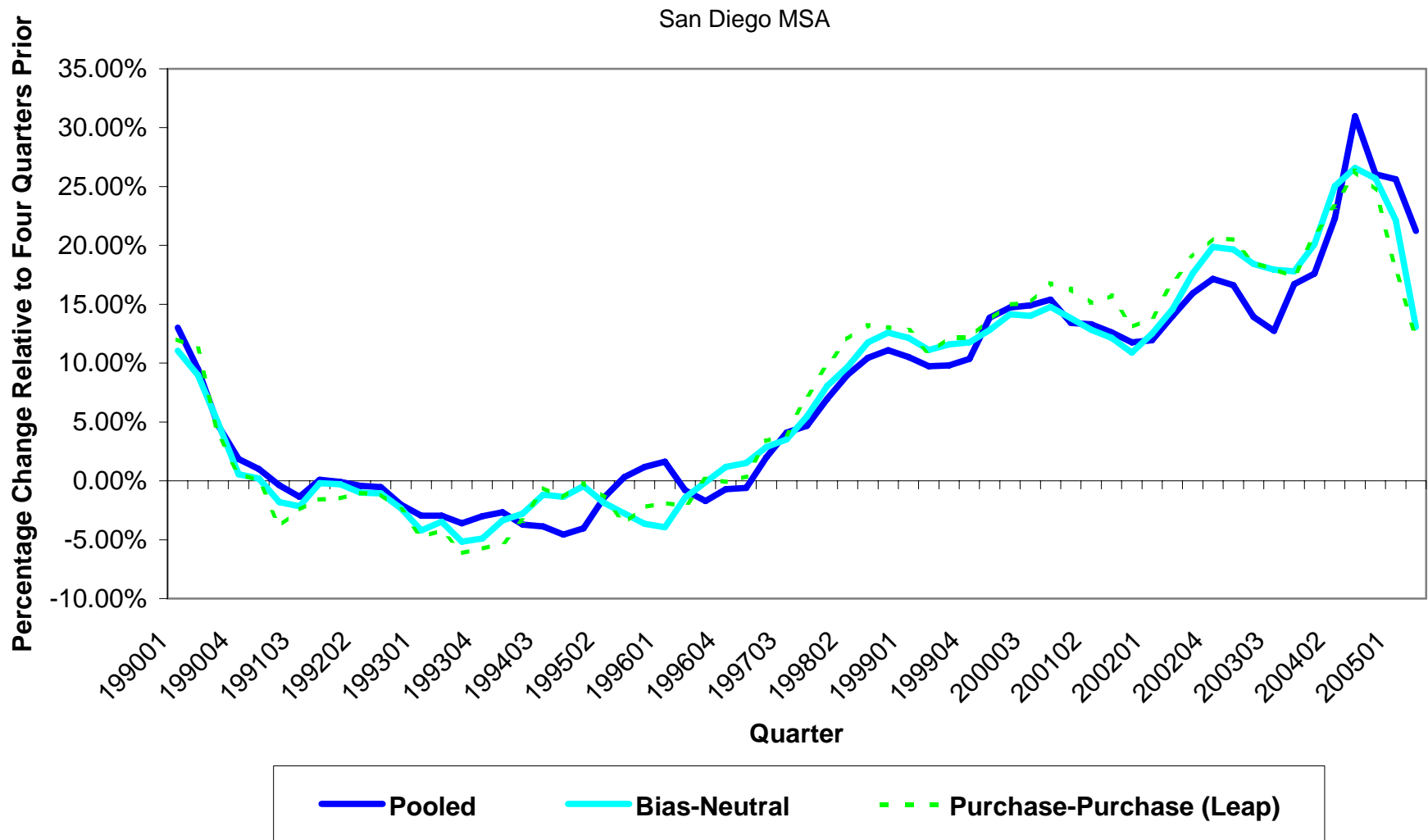
Figure 6e: Four-Quarter Percentage Change in House Prices: By Methodology

Los Angeles-Long Beach-Glendale, CA MSAD

Figure 7: Estimated Appreciation for Sampled Homes Having One and More than One Purchase

Washington State
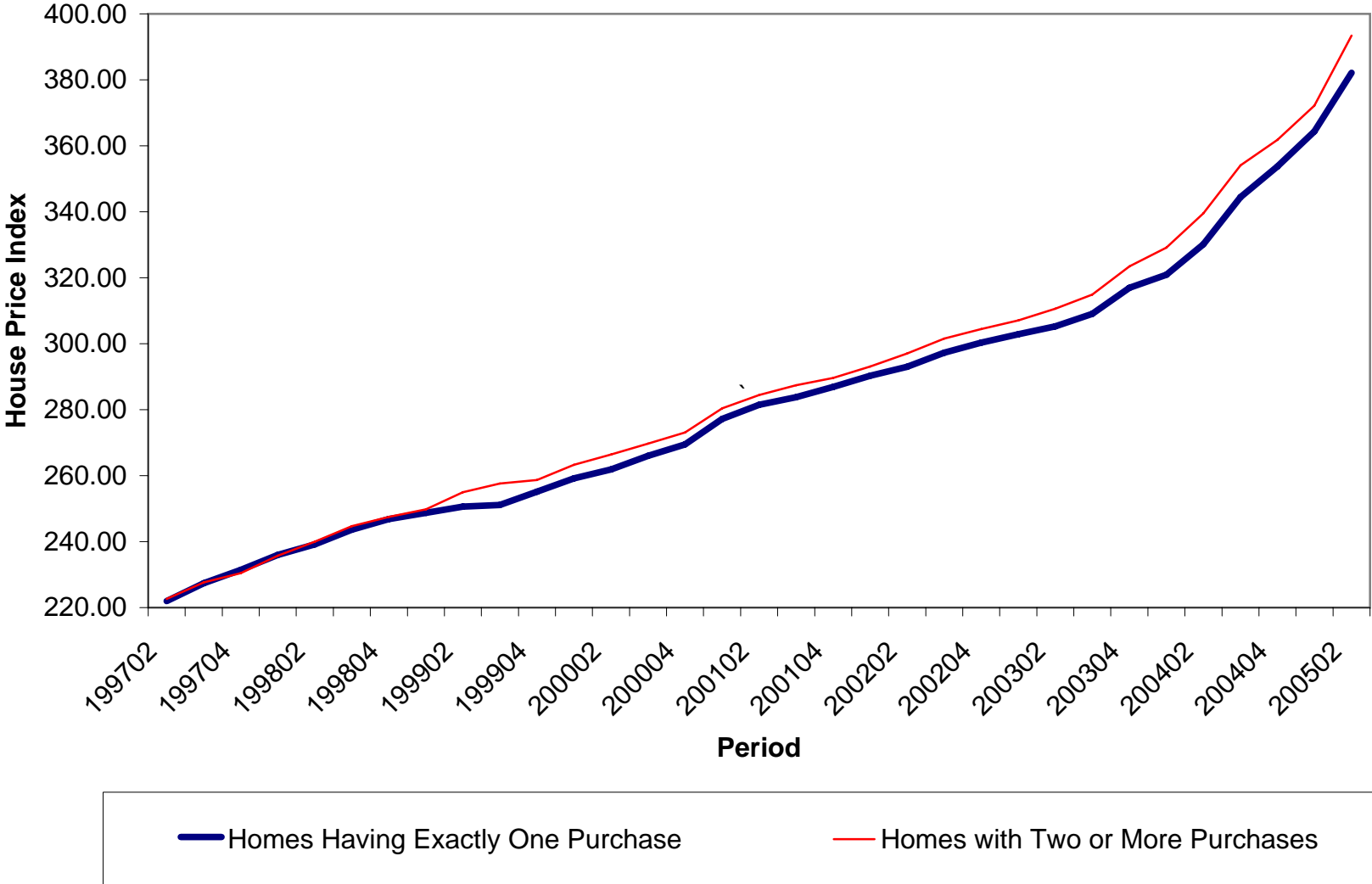
Figure 8a: Sampling Distributions of HPI Estimates by Methodology
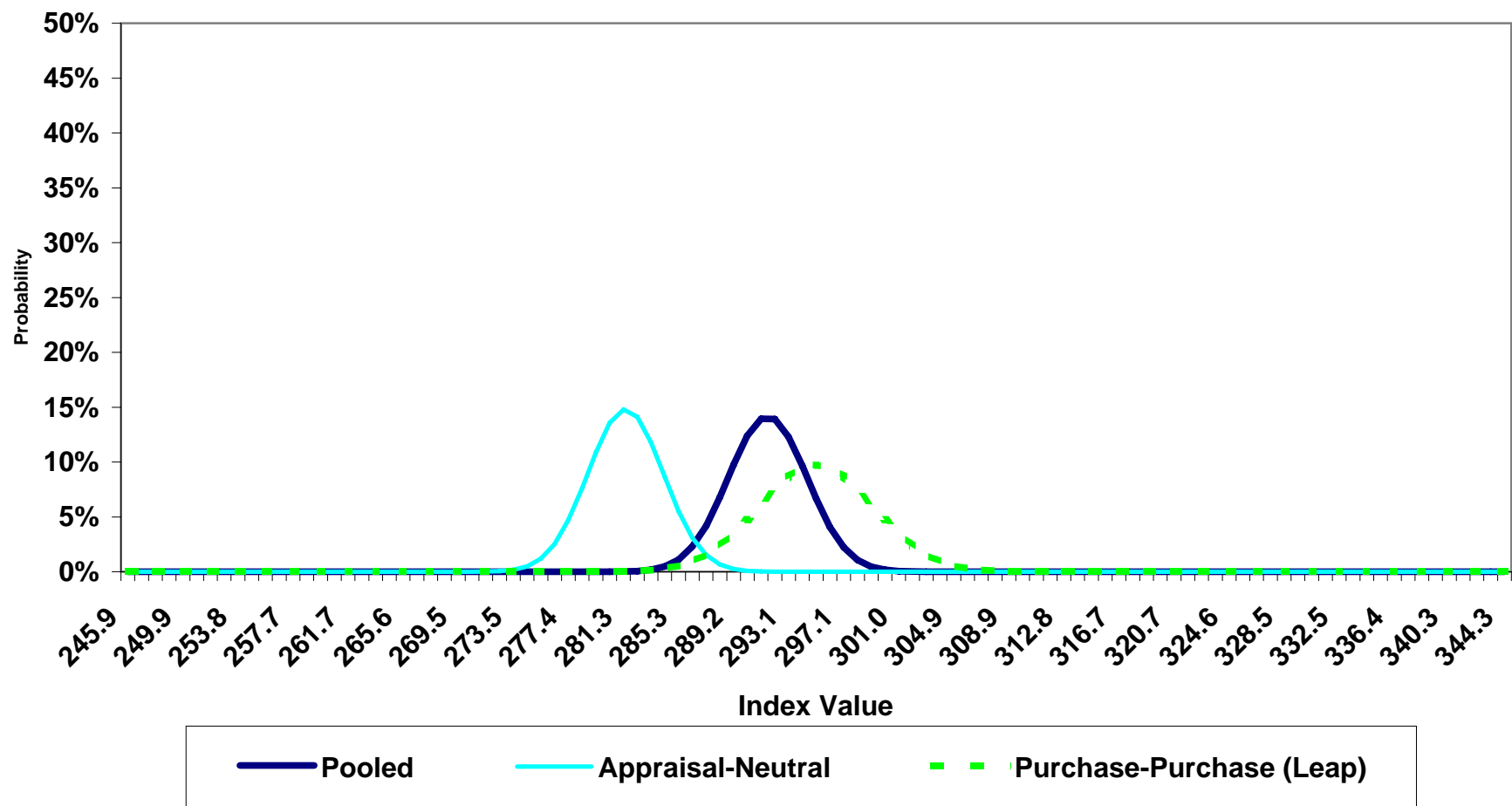
Washington State -- 2002 Quarter 1

Probability

Index Value

Pooled — Appraisal-Neutral — — Purchase-Purchase (Leap)

Figure 8b: Sampling Distributions of HPI Estimates by Methodology
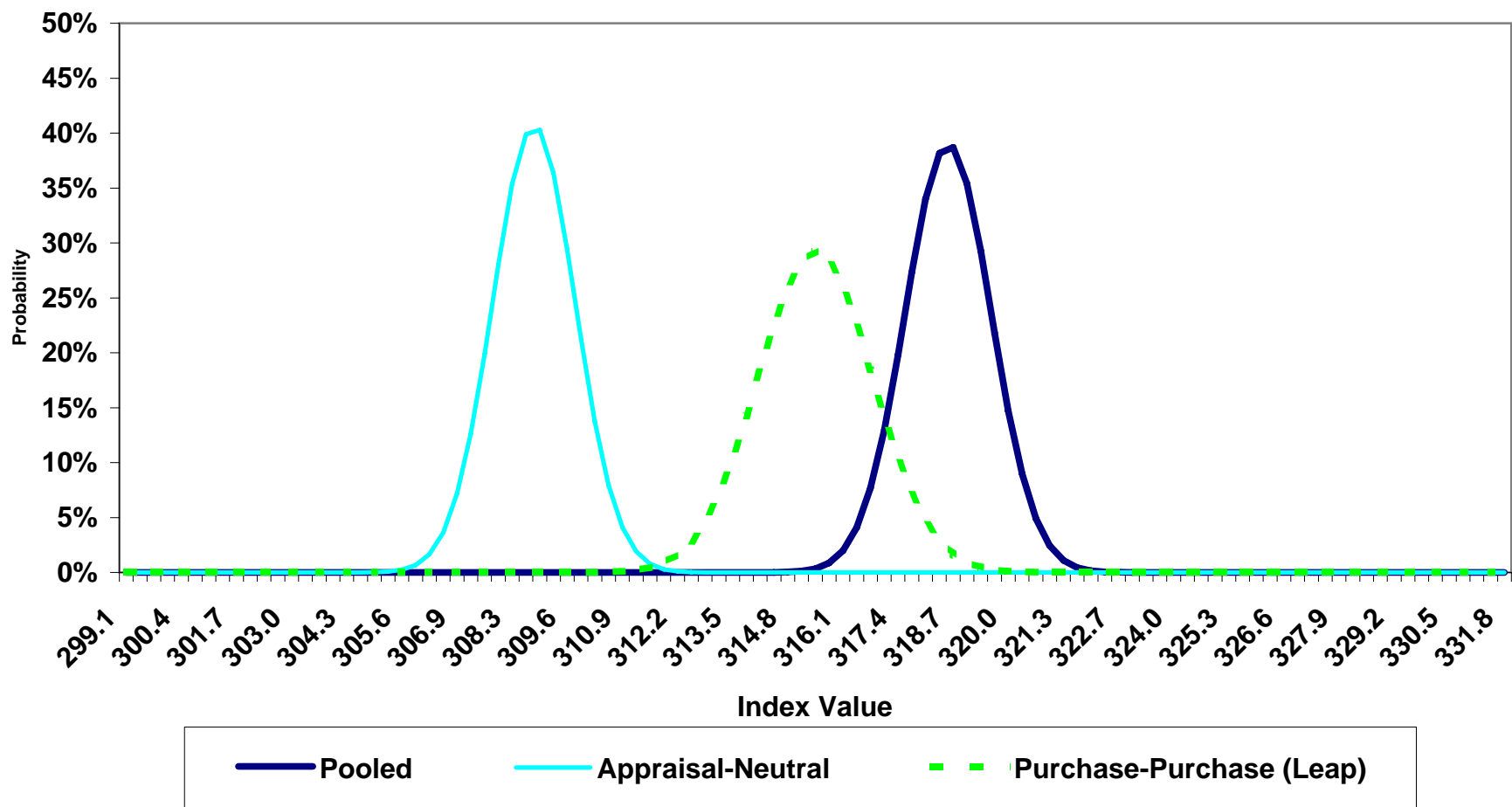
California -- 2002 Quarter 1

Figure 8c: Sampling Distributions of HPI Estimates by Methodology
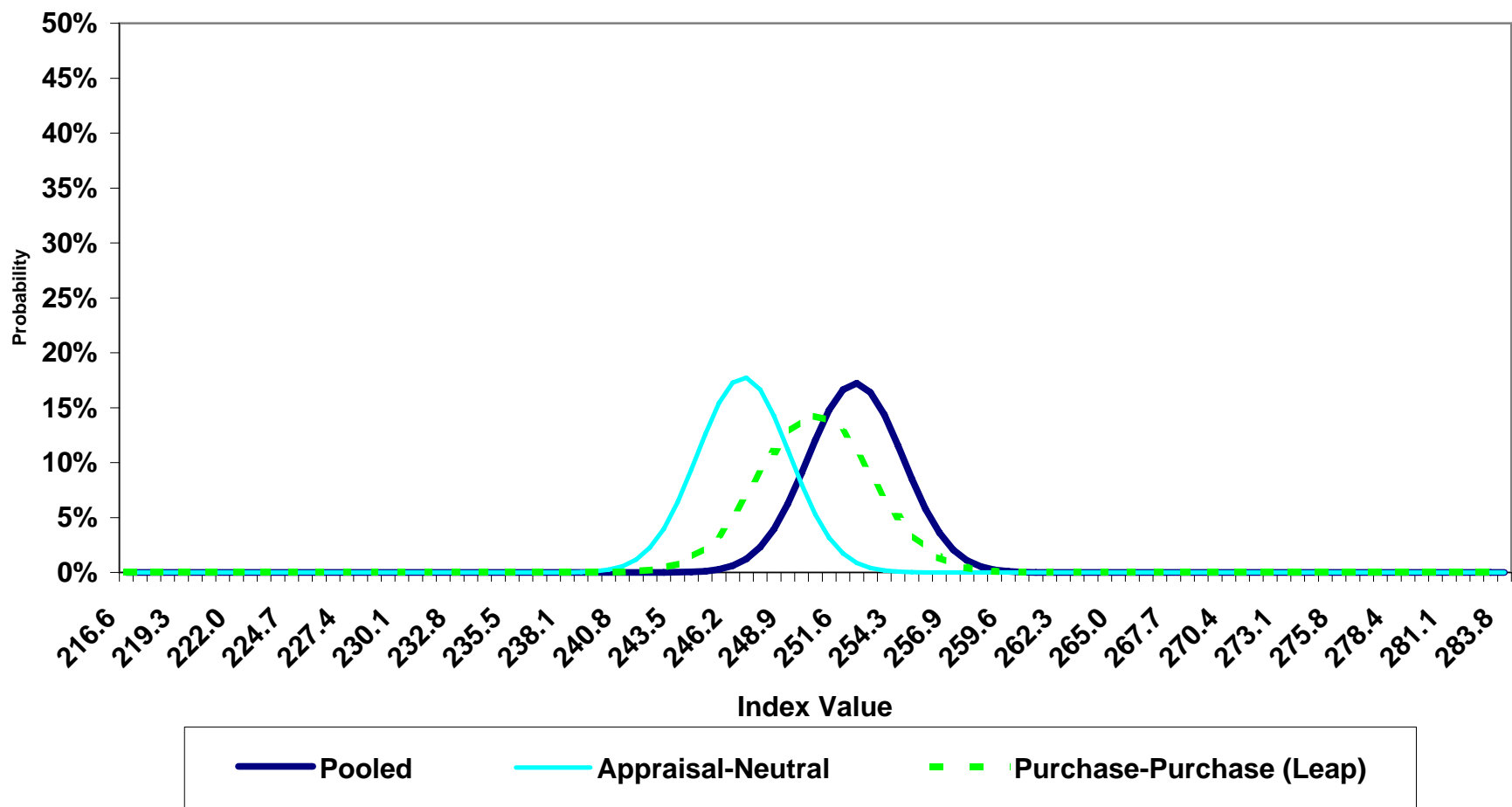
Illinois -- 2002 Quarter 1

Figure 8d: Sampling Distributions of HPI Estimates by Methodology
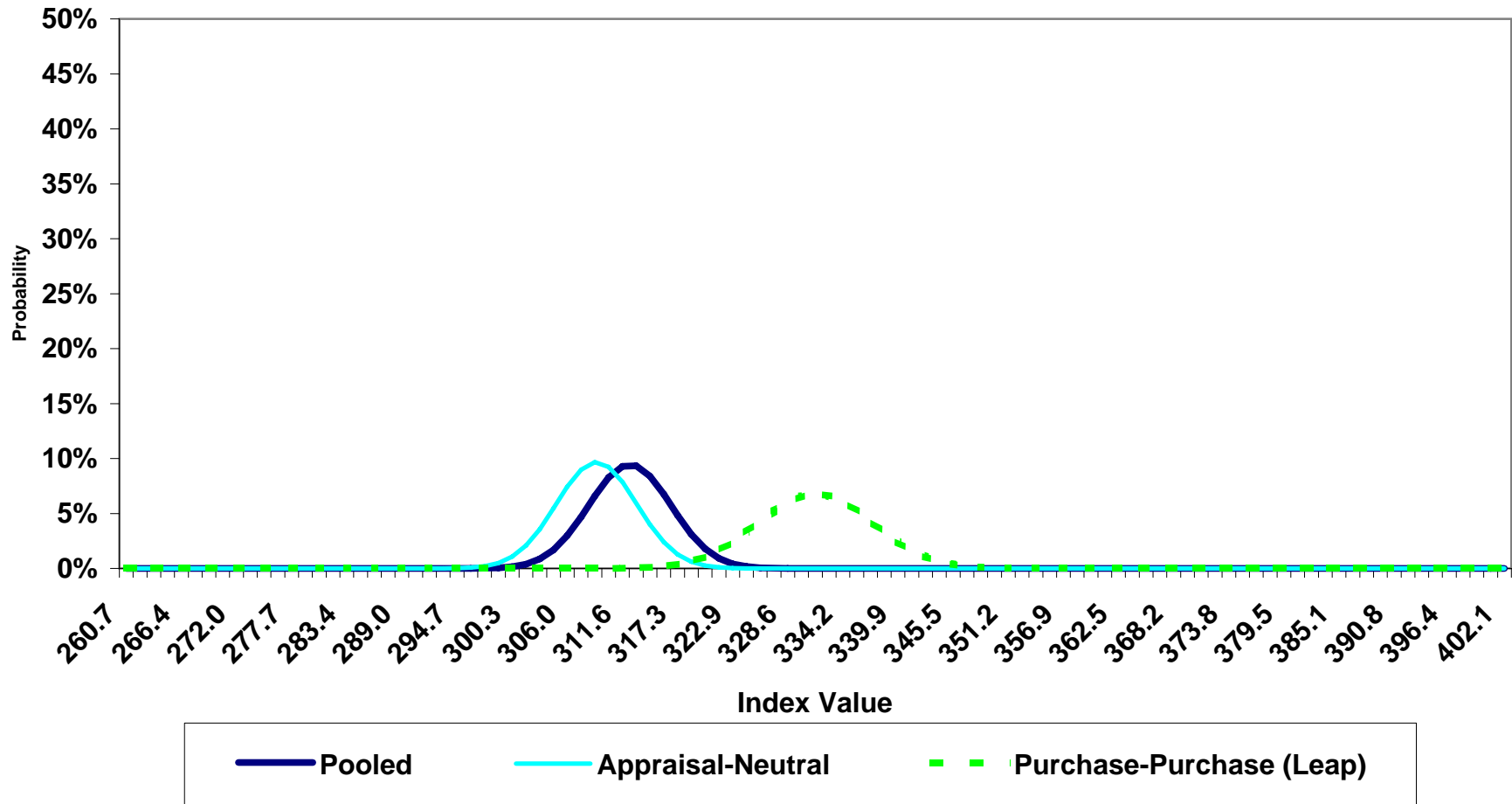
San Diego MSA -- 2002 Quarter 1

Figure 8e: Sampling Distributions of HPI Estimates by Methodology
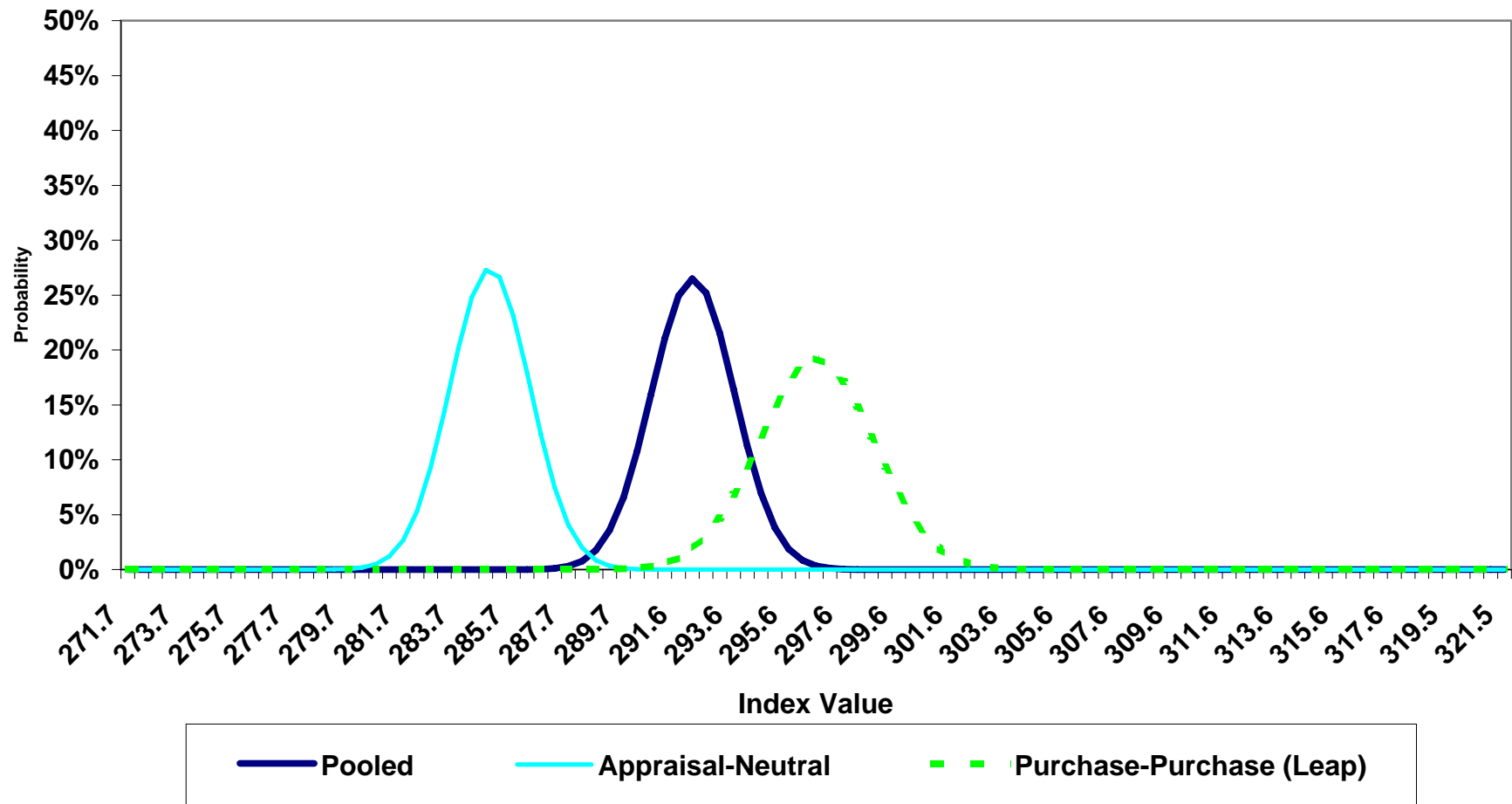
Los Angeles-Long Beach-Glendale, CA MSAD -- 2002 Quarter 1

Legend: Pooled — Appraisal-Neutral — Purchase-Purchase (Leap)

Figure 9a: Rate-Term and Cash-Out Bias Coefficients

Washington State

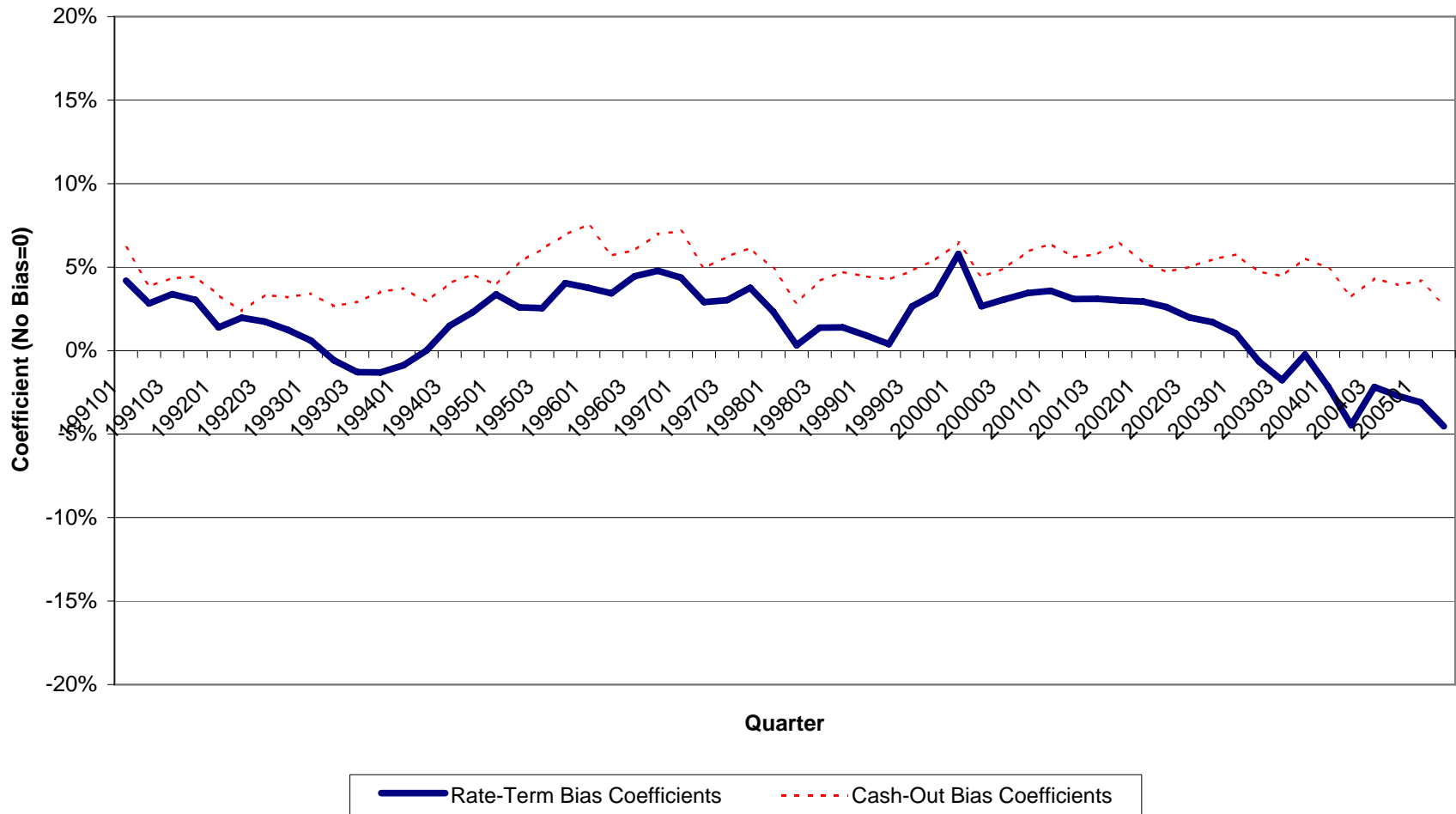**Rate-Term Bias Coefficients**   **Cash-Out Bias Coefficients**

Figure 9b: Rate-Term and Cash-Out Bias Coefficients
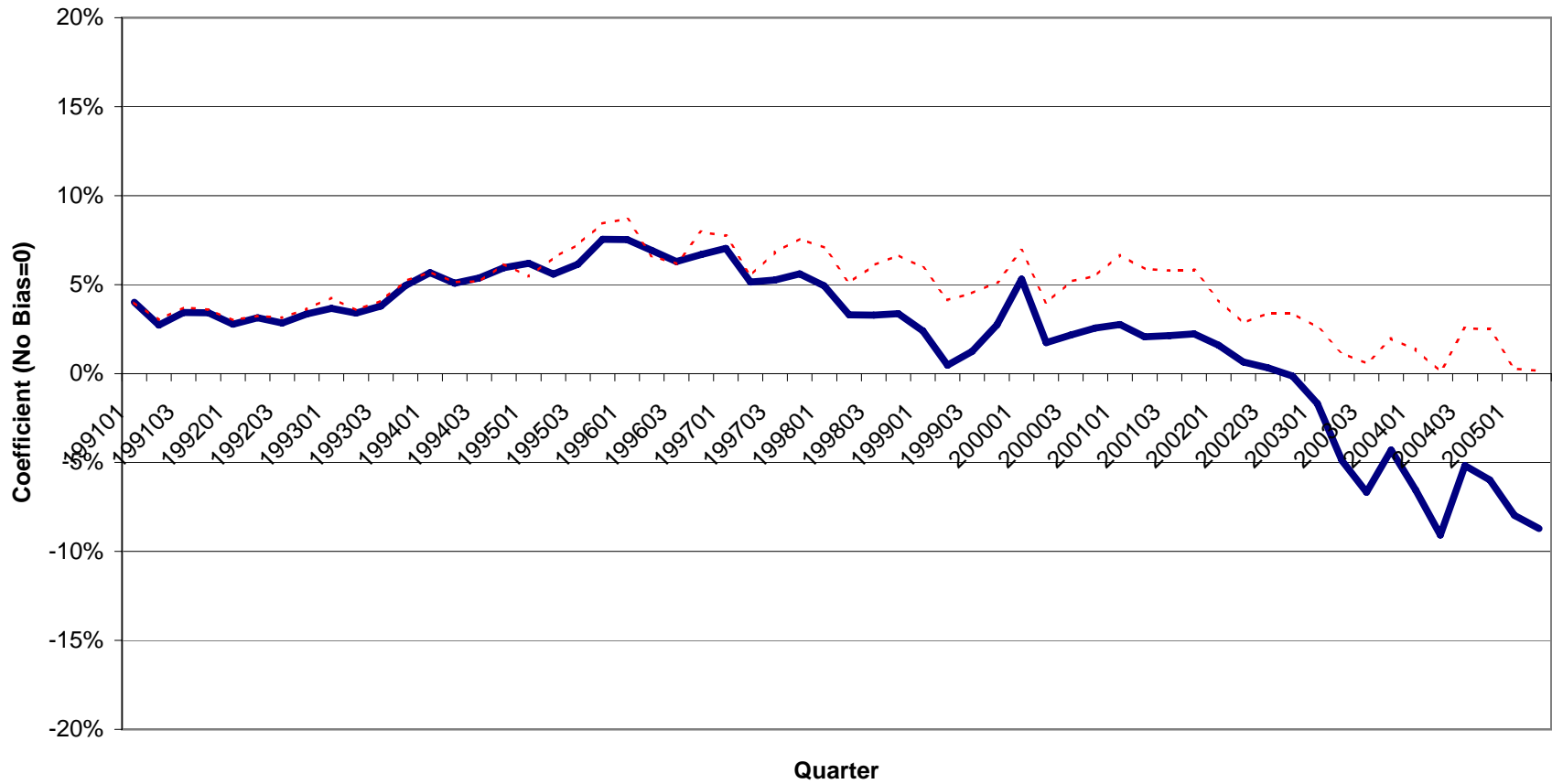
California

Figure 9c: Rate-Term and Cash-Out Bias Coefficients

Illinois

Quarter
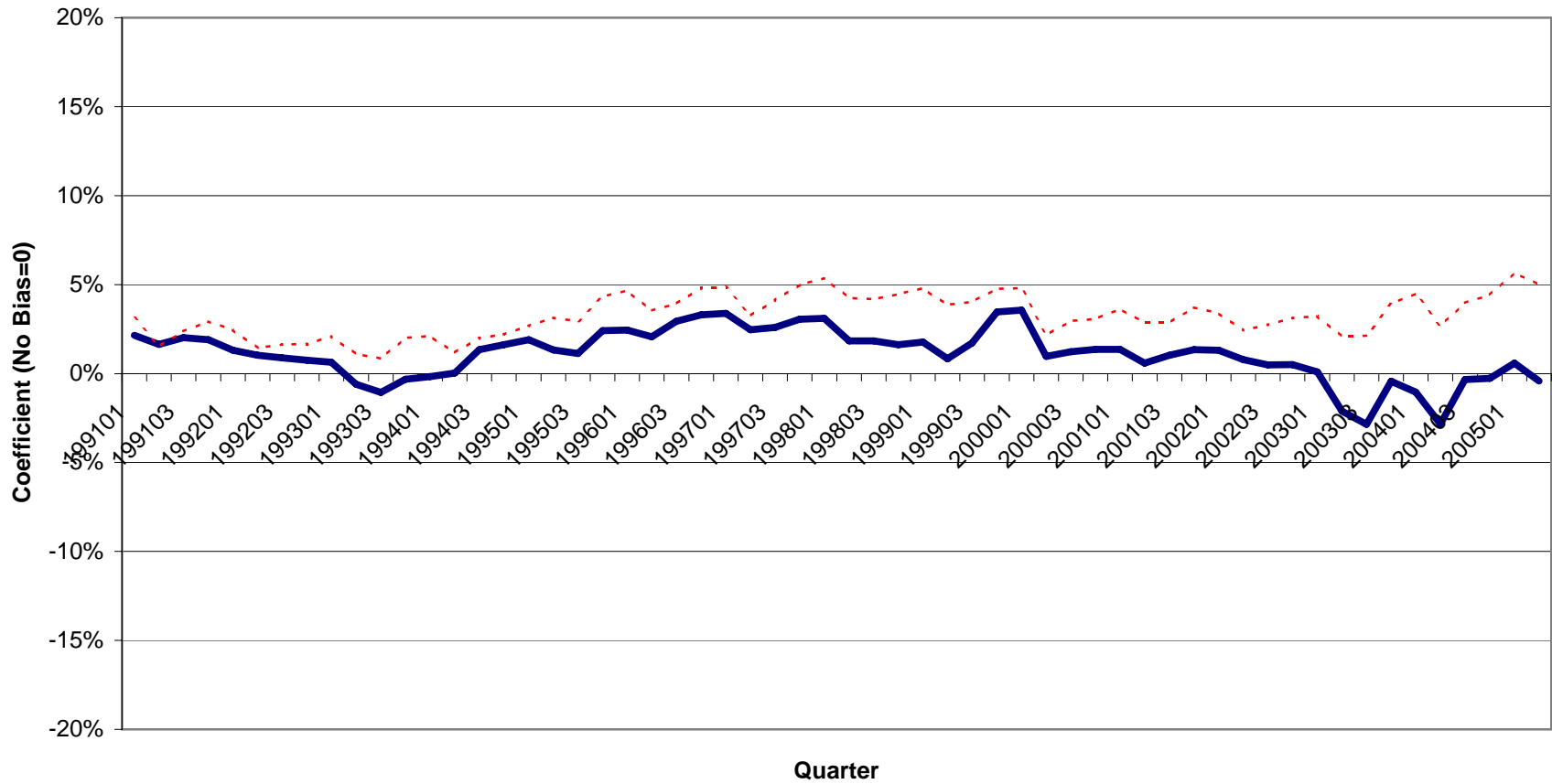
Rate-Term Bias Coefficients .... Cash-Out Bias Coefficients

Figure 9d: Rate-Term and Cash-Out Bias Coefficients

San Diego MSA
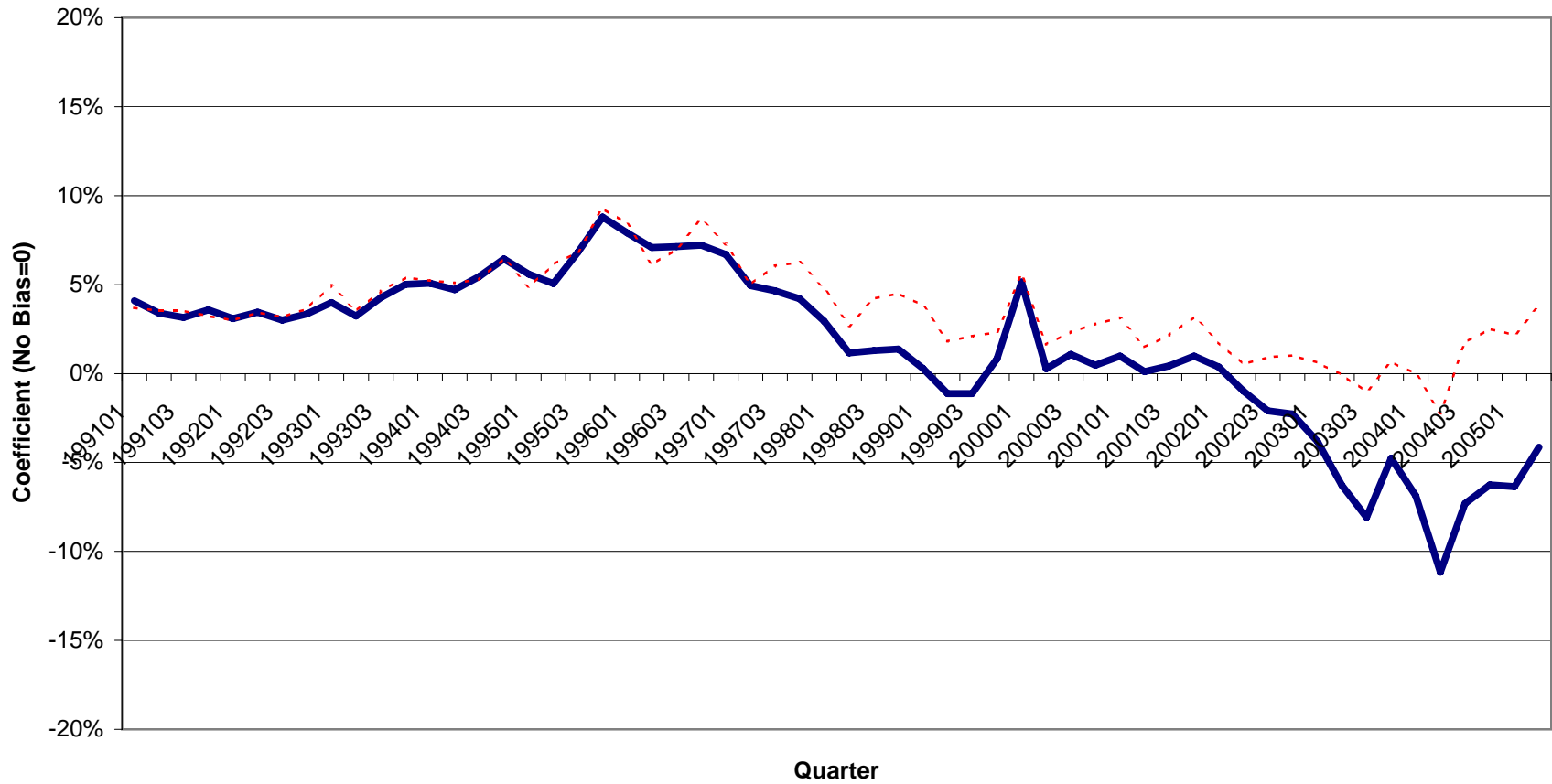
Figure 9e: Rate-Term and Cash-Out Bias Coefficients

Los Angeles-Long Beach-Glendale, CA MSAD

Figure 10: The Effect of Removing Quick-Turnaround, Zero Price Change Observations from Bias-Neutral Regressions

Coefficients from Washington State

Figure 11: House Price Appreciation Patterns for Cash-Out Houses versus Other Houses

Seattle-Bellevue-Everett MSAD

Houses that have at least One Cash-Out (Pairings prior to First Cash-Out)

Other Houses

Tacoma, WA MSAD

Table 1a: Number of Paired Samples Having at least One Observation in Appointed Quarter:

Pooled versus Purchase-Only (Leapfrog) Approach

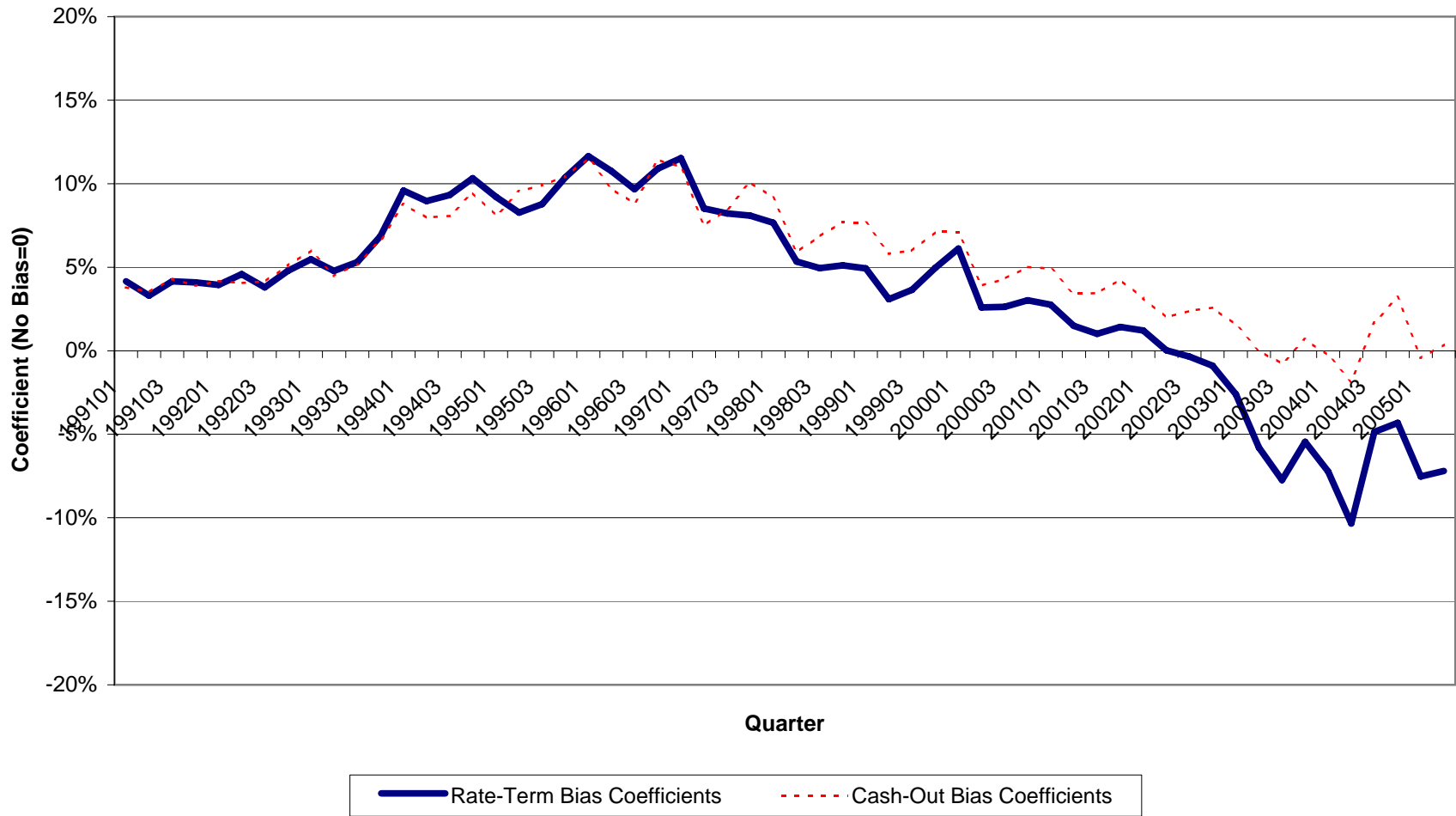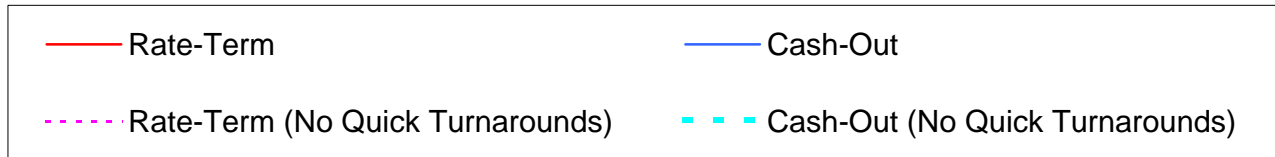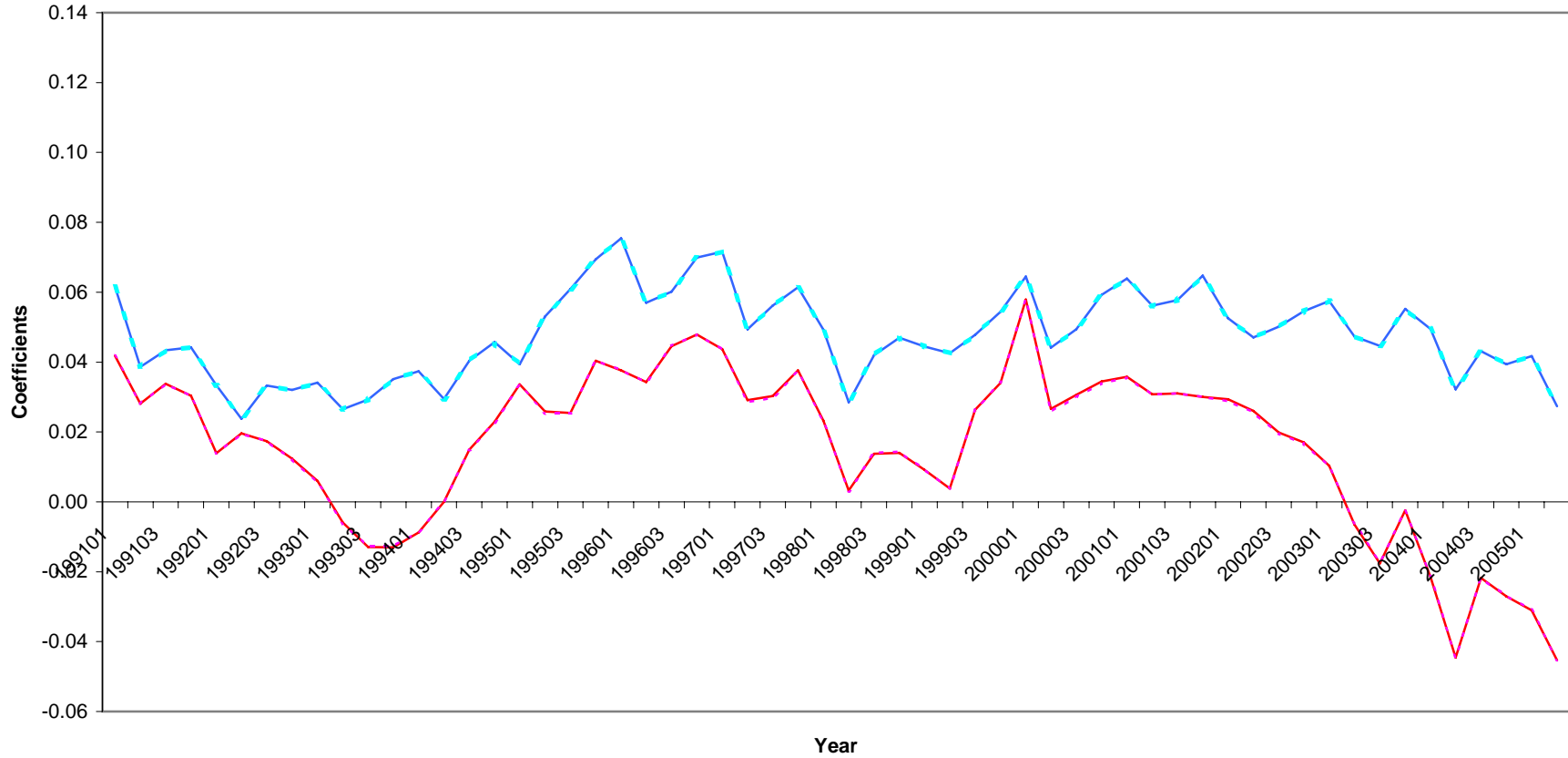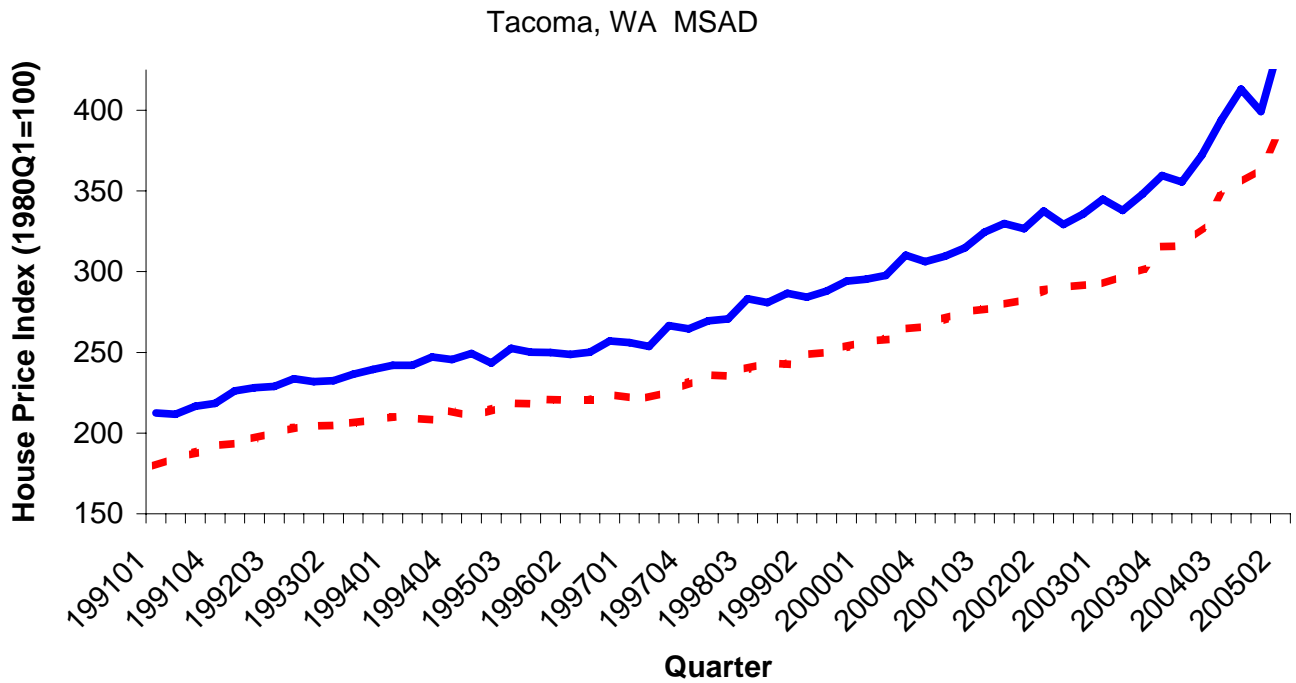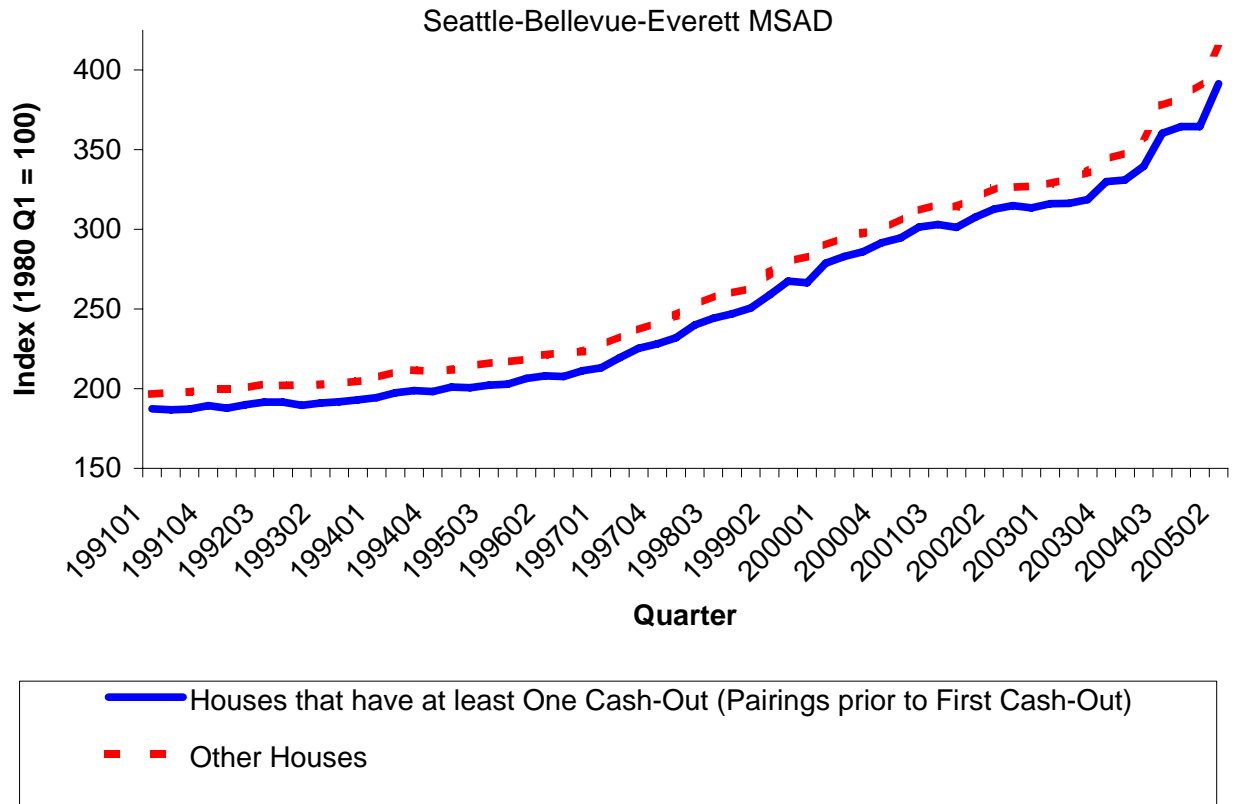|	2003Q1	2003Q2	2003Q3	2003Q4	2004Q1	2004Q2	2004Q3	2004Q4	2005Q1	2005Q2
California										
Total	347,592	393,182	314,888	119,866	131,400	113,882	61,119	68,616	59,871	47,297
Purchase-Only (Leapfrog)	12,002	14,983	14,067	9,189	7,241	7,429	6,386	4,926	3,739	3,955
Purchase-Only as % of Total	3.5%	3.8%	4.5%	7.7%	5.5%	6.5%	10.4%	7.2%	6.2%	8.4%
Illinois										
Total	126,396	141,928	122,991	36,280	38,708	42,621	24,248	26,688	22,627	20,510
Purchase-Only (Leapfrog)	4,588	8,295	8,298	4,848	3,631	6,807	6,557	4,529	3,347	5,542
Purchase-Only as % of Total	3.6%	5.8%	6.7%	13.4%	9.4%	16.0%	27.0%	17.0%	14.8%	27.0%
Washington										
Total	62,288	71,889	64,759	19,985	21,515	22,551	13,753	15,127	13,678	12,816
Purchase-Only (Leapfrog)	3,297	4,663	4,674	2,979	2,507	4,093	3,732	2,742	2,348	3,088
Purchase-Only as % of Total	5.3%	6.5%	7.2%	14.9%	11.7%	18.1%	27.1%	18.1%	17.2%	24.1%

Source: Enterprise data submissions for HPI Release Q2 2005

Table 1b: Number of Paired Samples Having at least One Observation in Appointed Quarter:

Pooled versus Pre-Pairing Approach

	2003Q1	2003Q2	2003Q3	2003Q4	2004Q1	2004Q2	2004Q3	2004Q4	2005Q1	2005Q2
California										
Total	347,592	393,182	314,888	119,866	131,400	113,882	61,119	68,616	59,871	47,297
Purchase-Only (Pre-Pair)	7,069	8,238	7,433	4,708	3,695	3,484	2,993	2,320	1,773	1,827
Purchase-Only as % of Total	2.0%	2.1%	2.4%	3.9%	2.8%	3.1%	4.9%	3.4%	3.0%	3.9%
Illinois										
Total	126,396	141,928	122,991	36,280	38,708	42,621	24,248	26,688	22,627	20,510
Purchase-Only (Pre-Pair)	2,680	4,277	4,078	2,360	1,718	2,864	2,707	2,020	1,517	2,348
Purchase-Only as % of Total	2.1%	3.0%	3.3%	6.5%	4.4%	6.7%	11.2%	7.6%	6.7%	11.4%
Washington										
Total	62,288	71,889	64,759	19,985	21,515	22,551	13,753	15,127	13,678	12,816
Purchase-Only (Pre-Pair)	2,042	2,720	2,594	1,631	1,271	1,929	1,808	1,366	1,109	1,483
Purchase-Only as % of Total	3.3%	3.8%	4.0%	8.2%	5.9%	8.6%	13.1%	9.0%	8.1%	11.6%

Source: Enterprise data submissions for HPI Release Q2 2005

Table 2a: Average Squared Errors--First Stage

	Washington State	California	Illinois	San Diego	Los Angeles-Long Beach-Glendale, CA
Pooled	0.031	0.031	0.023	0.022	0.026
Purchase-Only (Leap)	0.044	0.040	0.039	0.026	0.034
Bias-Neutral Approach	0.029	0.029	0.022	0.020	0.024

Table 2b: Average Squared Errors--Second Stage

	Washington State	California	Illinois	San Diego	Los Angeles-Long Beach-Glendale, CA
Pooled	0.0067	0.0067	0.0088	0.0058	0.0064
Purchase-Only (Leap)	0.0138	0.0127	0.0326	0.0100	0.0115
Bias-Neutral Approach	0.0065	0.0061	0.0087	0.0052	0.0059

Source: Enterprise data submissions for HPI Release Q2 2005

Table 3: Average Number of Quarters Between Transactions

	Washington State	California	Illinois	San Diego	Los Angeles-Long Beach-Glendale, CA
Pooled / Bias-Neutral	18.4	18.8	15.6	17.6	20.6
Purchase-Only (Leap)	28.8	32.0	26.4	31.4	35.2

Source: Enterprise data submissions for HPI Release Q2 2005

Table 4: Average Squared Second-Stage Errors by Time Between Transactions
(Pooled Model)

	Washington State	California	Illinois	San Diego	Los Angeles-Long Beach-Glendale, CA
0-20 Quarters	0.002	0.002	0.001	0.002	0.002
21-40 Quarters	0.010	0.010	0.012	0.010	0.009
41-60 Quarters	0.026	0.019	0.051	0.016	0.015
61-80 Quarters	0.062	0.051	0.155	0.077	0.044
81 -100 Quarters	0.051	0.057	0.362	0.043	0.055
101-122 Quarters	0.061	0.073	0.318	0.013	0.030

Source: Enterprise data submissions for HPI Release Q2 2005

Table 5: Second-Stage Regression Coefficients (and Standard Errors)

	Washington State	California	Illinois	San Diego	Los Angeles-Long Beach-Glendale, CA
Qtrs Between Transactions	0.00215	0.00284	0.00195	0.0025	0.00304
	(92.09)	(271.45)	(93.06)	(73.62)	(139.06)
Qtrs Between Transactions^2	-0.00002558	-0.00006616	-0.0000465	-0.00008121	-0.00010593
	(-16.2)	(-95.43)	(-31.23)	(-35.19)	(-74.91)
Qtrs Between Transactions^3	2.669E-07	8.434E-07	0.00000104	0.0000013	0.00000164
	(8.58)	(61.68)	(33.65)	(28.13)	(59.59)
Qtrs Between Transactions^4	-8E-10	-3.4E-09	-5.7E-09	-6.8E-09	-0.000000008
	(-4.43)	(-42.66)	(-30.72)	(-24.82)	(-50.63)

Source: Enterprise data submissions for HPI Release Q2 2005
Note: T-Statistics in Parentheses

Table 6: Average Second-Stage Error--Correcting for Time Between Transactions

	Washington State	California	Illinois	San Diego	Los Angeles-Long Beach-Glendale, CA
Pooled (Adjusted Average)	0.0137	0.0144	0.0271	0.0135	0.0138
Purchase-Only (Leap)	0.0138	0.0127	0.0326	0.0100	0.0115
Bias-Neutral (Adjusted Average)	0.0135	0.0137	0.0268	0.0124	0.0132

Source: Enterprise data submissions for HPI Release Q2 2005