The Repeat Rent Index*

Brent W. Ambrose, Pennsylvania State University†

N. Edward Coulson, Pennsylvania State University‡

and

Jiro Yoshida, Pennsylvania State University§

Current version: February 21, 2013

*We thank Walter D'Lima and Moussa Diop for their able research assistance and the Penn State Institute for Real Estate Studies for providing access to the RentBureau database.

†Institute for Real Estate Studies and the Department of Risk Management, The Pennsylvania State University, University Park, PA 16802-3306, bwa10@psu.edu

‡Department of Risk Management and Department of Economics, The Pennsylvania State University, University Park, PA 16802-3306, fyj@psu.edu

§Department of Risk Management, The Pennsylvania State University, University Park, PA 16802-3306, juy18@psu.edu

The Repeat Rent Index

Abstract

Studies of real estate markets have long been hamstrung by the lack of reliable information on the flow price of housing. In contrast to the voluminous information on constant-quality real estate sale prices (from e.g. the Federal Housing Finance Administration) comparable quarterly indexes for rents have not been available. The only widely available data comes from the Bureau of Labor Statistics (BLS), who compile survey data and construct rental indexes for the nation, the census regions, and for a limited number of metropolitan areas. This research improves upon these BLS indexes in three importnat ways. First, we eschew surveys of existing renters in favor of using only newly-signed lease contracts. Second, we employ a weighted repeat rent estimator, that replicates for the rental market, as closely as possible, the weighted repeat sales estimator of Calhoun (1996), following Case and Shiller (1989) and Bailey et al. (1963). Third, we construct quarterly indexes for a larger number of cities than are available for the BLS, thus expanding the professions ability to make cross-sectional comparisons of housing markets, particularly in conjunction with FHA data. We provide explicit comparisons between our repeat rent index and the BLS index for 11 metropolitan areas.

Our general conclusions are that (a) there is considerable heterogeneity in the behavior of rents across cities over the 2000-2010 decade, but the number of cities and years for which nominal rents fell is substantial; (b) rents fell more, or rose more slowly, over the decade than would be inferred from the BLS data. In particular we find that rents fell in many cities following the onset of the housing crisis in 2007. This is not usually observed in the BLS data; (c) repeat rent indexes (RRI) are more volatile than the BLS indexes; (d) the BLS lags the repeat rent index by 2-4 quarters. The last two conclusions follow directly from the differences in sampling methods and index construction.

Key words: Residential Leases, Repeat Sales Index, Housing

JEL Classification: G2

1 Introduction

Studies of real estate markets have long been hamstrung by the lack of reliable information on the flow price of housing. In contrast to the voluminous information on constant-quality real estate sale prices (from e.g. the Federal Housing Finance Administration) comparable indexes for rents have not been available. The only widely-available data come from the Bureau of Labor Statistics (BLS), who compile survey data on rents. The BLS constructs rent indexes from these surveys for the nation, the census regions, but only for a limited number of metropolitan areas.

Additionally, the BLS series suffer from significant problems both in construction and coverage. For example, Crone et al. (2010) document the extensive revisions over the years in the BLS methodology that were designed to remove various biases and omissions in the consumer price index for tenant rents. Yet, even with these corrections, Crone et al. (2010) noted that the BLS tenant rent index remains biased due to missing rent increases that occurred when units experienced a change in tenant (what they characterize as a "nonresponse bias".) In addition, Gordon and vanGoethem (2007) note that the BLS rent series suffers biases from longitudinal changes in unit quality and they propose the use of hedonic models using Census of Housing data to construct a constant quality rent index.

Yet, for all its shortcomings, the BLS rent series has remained practically the only data source for information on the flow price of housing services. As a result, the BLS rent series often appears in various housing studies. For example, Dougherty and Order (1982) and Bajari et al. (2005) use rental indexes as proxies for price of housing service flows in common "user-costs" models. In another example, Sinai and Souleles (2005) rely on rent indexes to demonstrate that home ownership is an effective hedge for anticipated housing rent increases. Furthermore, rent indexes lie at the heart of the growing literature examining house price bubbles (e.g. McCarthy and Peach (2004), Himmelberg et al. (2005), Campbell and Shiller (1988, 2005), Brunnermeier and Julliard (2008), and Campbell et al. (2009), among many others.) Thus, to the extent that the BLS rent index fails to adequately capture changes in housing service price flows, new data tracking residential rents could have important impacts in a number of areas.

This research rectifies that data lacuna in three ways. First, we eschew surveys of existing renters in favor of using only newly-signed lease contracts. Such contracts are, by that fact, more reflective of current market conditions than are surveys of renters in the middle of leases; moreover, we use only leases that are signed by new tenants, in order to avoid possible tenure biases. Second, we employ a weighted repeat rent estimator, that replicates for the rental market, as closely as possible, the weighted repeat sales estimator of Calhoun (1996), following Case and Shiller (1989) and Bailey et al. (1963). This method of estimating house price indexes has become standard, primarily because of its use by the Federal Housing Finance Agency (FHFA) in constructing the widely-used repeat sales indexes for housing for every MSA in the US, and constructing a similar estimator for rents would seem a fitting addition. Third, we are able to construct quarterly indexes for a larger number of cities than are available for the BLS thus expanding the profession's ability to make cross-sectional comparisons of housing markets, particularly in conjunction with FHA data.

In this paper we present two sets of findings. First, we present the rent indexes for a large number of cities and describe in general terms their behavior over the past decade. Secondly, we provide explicit comparisons between our repeat rent index and the BLS index for 11 large metropolitan areas, and we compare the behavior of our repeat rent index to the BLS index. Our general conclusions are that (a) there is considerable heterogeneity in the behavior of rents across cities over the 2000-2010 decade, but the number of cities and years for which nominal rents fell is substantial; (b) rents fell more, or rose more slowly over the decade than would be inferred from the BLS data.

In particular we find that rents fell in many cities following the onset of the housing crisis in 2007. This is not usually observed in the BLS; (c) repeat rent indexes (RRI) are more volatile than the BLS indexes; (d) the BLS lags the repeat rent index by 2-4 quarters. Granger causality tests on that account indicate that RRI Granger-causes the BLS index. The last two conclusions follow directly from the differences in sampling methods and index construction.

The next section describes the data source and methods used to prepare for the repeat rent regressions. This regression is discussed in Section 3, and is compared to BLS index construction. Section 4 presents the results (summarized above), and Section 5 concludes.

2 Data

We utilize the residential rent transaction data compiled by Experian RentBureau for the period from January 1998 to December 2010. RentBureau maintains a national database on tenant rental payment performance collected from property management companies. The database contains lease characteristics (lease start date, lease termination date, renter move-in date, renter move-out date, last transaction date) and property location (city, state, and zip-code). To maintain privacy, limited information is disclosed on specific property locations and individual renters. The company updates lease records every month, noting whether rent was paid on time or not, the type of payment delinquency, and if applicable, the accrued number of late payments, along with any write-off on rental or non-rental payments due. Over time, RentBureau expanded its geographic coverage adding new properties and locations to the database.

Rent payments for each lease, whether active or closed, are recorded in a 24-digit vector representing the renter's payment performance over the previous 24 months from the month of reporting or the month the lease ended. Since RentBureau only maintains a 24-month payment record for each lease, lease payment records are therefore left censored. The rental data were last updated in November 2010, the last month of reporting. We restrict our analysis to lease observations with rent payments greater than $100 per month.

In addition to the 24-month vector of rental payment performance, each lease observation also reports the monthly rent for that lease. Since RentBureau maintains a unique identification number for each rental unit in each property, we are able to create a time-series of monthly rents on the same apartment units by linking observations by the unit identification number and utilizing information about the lease start and end dates. Since the majority of residential leases are 12-month contracts, the ability to link apartment units via the unique unit identification number allows us to create a rent series holding unit quality constant.

Our final data filter is to eliminate leases where the same tenant is renewing a lease. This is to avoid the sitting or tenure discount (or perhaps premium) that is available to renewing tenants (Guasch and Marshall, 1987; Goodman and Kawai, 1985; Kanemoto, 1990; Hubert, 1995; Raess and von Ungern-Sternberg, 2002). We do this by comparing the move-in date and the date of the first rental payment. For new tenants this first payment is normally at the time of the signing of the rental contract, and before the tenant physically takes possession of the unit, whereas for renewed leases the tenant has (except under quite unusual circumstances) already moved in, this comparison discriminates between new and renewed leases, whereupon we eliminate the latter from the sample.

After applying the above filters and removing observations with missing or incorrectly coded data (e.g. rents less than $100 per month, move-in dates after 2010, or incorrectly coded unit id numbers), the data set contains information on over 1.4 million individual lease contracts originated for 551,126 individual residential units in 2,934 multifamily properties (or complexes). On average, the database contains 2.7 lease contracts per individual apartment unit. Figure 1 shows the frequency distribution of the leases and rental properties per year. The yearly frequency count of leases in Panel (a) reveals how RentBureau significantly expanded its lease tracking activity during the previous decade. For example, RentBureau reported payment transaction data on 7,586 leases in 2000 and had expanded to 339,443 leases by 2009. Panel (b) reports the number of individual rental properties underlying the lease records. Again, we see a dramatic increase over time in the number of properties reporting to RentBureau. Table 1 reports the distribution of leases across states. The top five states represented in the data are Texas (17%), Georgia (15%), Florida (15%), California (12%) and Arizona (10%). Together, these five states account for approximately 69% of all the leases in the data set.

In the analysis below, we compare our repeat rent index to the BLS rental index for 11 large metropolitan areas. Thus, Table 2 reports the distribution of lease contracts across the MSAs that match with the markets covered by the BLS. We note that the RentBureau data contains information on 518,381 leases in the BLS markets, which represents approximately 35% of the national data set. Not surprising, since RentBureau began operations in the South, Atlanta has the largest representation in the database. Over the period from 2000 to 2010, RentBureau contains information on 170,046 lease contracts on 66,945 apartment units in Atlanta. This represents approximately 2.5 lease contracts per unit over the sample period. Furthermore, the 66,945 units are located in 326 different apartment complexes across the city. Other major markets with over 10,000 leases in the database include Dallas, Houston, Los Angeles, Miami, San Francisco, Seattle, and Washington, DC. Interestingly, San Francisco has the highest average number of leases per unit (3.1) with 18,225 leases on 5,803 units in 48 complexes.

Of the BLS markets, Detroit has the fewest number of leases (4,967) with an average of 2.2 leases per unit.

3 Methods

The repeat rent index is a quality-constant measure of rent changes in a particular market over time. In constructing this index we look to methods established for similar quality-constant indexes in the residential sales market. The obvious problem in both markets is that simple averages of transaction prices in each time period do not account for the changing (presumably rising) quality of the transacted units, and so will presumably overstate the rate of price increase. One could use these transacted units in the construction of an index if the quality of the units, as embodied by the characteristics of these units, was controlled for. The most common method of doing this is through hedonic regressions. With a database of transactions, their dates, and their characteristics, consider a regression of the form:

$$logP_{it} = X_i\beta + \gamma_1 T_1 + \cdots + \gamma_N T_N + \varepsilon_{it} \tag{1}$$

where P_{it} is the price of the ith housing unit at time t, X_i is a row vector of housing characteristics for the i^{th} unit and β is a column vector of regression coefficient/characteristic weights. $T_j, j = 1N$ are binary variables which equal one if the transaction took place during time period j, and γ_j are the associated coefficients. The error term ε_{it} is assumed to be a random walk plus noise (Case and Shiller, 1989). By virtue of including X in the regression, the γ_j terms represent the incremental value of transactions taking place in period j, holding quality constant. For convenience, we include an intercept term, and omit one of the T variables from the equation. Without loss of generality, we choose T_1 for that role. The equation becomes

$$logP_{it} = \beta_0 + X_i\beta + \gamma_2 T_2 + \cdots + \gamma_N T_N + \varepsilon_{it} \tag{2}$$

so that the intercept represents the price of housing in the first period and the sequence γ_2 through γ_N represents a constant-quality price index for housing for respective time periods.

A major difficulty is that not all of the quality measures may be recorded in the data; i.e. some of the Xs are unobserved. This can cause difficulties since the bias that arises in regression models when there are omitted variables can be severe. However, as pointed by Bailey et al. (1963), one can net out the effects of those omitted variables when differences (rates of change, to be precise) are considered. So consider two transactions at time periods s and t. According to the regression model, the predicted prices at those two time periods are:

$$logP_{it} = \beta_0 + X_i\beta + \gamma_t T_t + \varepsilon_{it} \tag{3}$$

and

$$logP_{is} = \beta_0 + X_i\beta + \gamma_s T_s + \varepsilon_{is} \tag{4}$$

Subtracting (4) from (3), we get:

$$logP_{it} - logP_{is} = \gamma_t T_t - \gamma_s T_s + \varepsilon_{it} - \varepsilon_{is}. \tag{5}$$

Thus the rate of change (over $t-s$ periods) is a function only of the time periods involved (and the change in the "noise") and not due to any quality variable (whether observed or unobserved). So, construct a data set comprised of sales of properties for which there are at least two sales observed. We reformulate the model into the *repeat sales regression* (Bailey et al., 1963):

$$logP_{it} - logP_{is} = \gamma_2 D_2 + \cdots + \gamma_N D_N + \varepsilon_{it} - \varepsilon_{is} \tag{6}$$

where

$$\begin{aligned} D_t &= 1 \text{ if the second sale in the pair took place at time period } t \\ D_t &= -1 \text{ if the first sale in the pair took place at time period } t \end{aligned}$$

By using log prices on the right hand side, the parameters represent percentage differences in prices from the base year. Note that the index for that base year is zero (since all $D_i = 0$).

Case and Shiller (1989) popularized this method of estimating house price indexes, however they noted that the error term in (6) is very likely heteroskedastic due to differences between transactions. Calhoun (1996) suggests the following three-stage procedure: first, estimate (6) using OLS; second, regress the squared residuals from that equation on $(t-s)$ and $(t-s)^2$ and collect the fitted values; and third, use the inverse of the square roots of those fitted values as weights in a weighted least squares regression of (6). The resulting γs form the *weighted repeat sales index*. We directly apply these methods. Instead of prices, we use *contract rents* evaluated at the time of lease signing to construct *(weighted) repeat rent indexes.*

Meese and Wallace (1997) and Gatzlaff and Haurin (1997) raised an objection to the use of repeat sales in that two sales are required for a houses inclusion in the sample, and such frequently traded units were perhaps not representative of the broader housing market. This is not so important in the case of rent indexes. The rental data are collected not whenever there is a sale, but whenever there is a lease, which occurs on a regular basis. Thus, no self-selection is involved.[1]

The Bureau of Labor Statistics publishes a rent index, constructed using surveys of renting households. Actually, BLS publishes both a rent index and an owners equivalent rent index. The latter measures rental rates of owner-occupied housing units, and so is not particularly germane to our inquiry here. Verbrugge and Poole (2010) discuss the differences between the two and their recent divergence. The BLS compiles six panels of households, each of whom is surveyed every six months on a rotating basis (e.g. panel 1 is surveyed in January and July, panel 2 in February and August, etc.) for 11 large metropolitan areas.[2] To simplify for the moment, assume just one such panel exists. An index for this panel is constructed from the percentage change of the aggregate rents of the panel. That is, the rent index at time t is

$$\delta_t = \delta_{t-6}\left[\frac{\Sigma_i \omega_i R_{it}}{\Sigma_i \omega_i R_{it-6}}\right]^{\frac{1}{6}} \tag{7}$$

where ω_i is the weight attached to the i^{th} unit to allow the sample of units to be representative of the population. Thus, given some base value for the first time period (just as in the repeat sales method) the rate of price increase is calculated as the rate of increase in the weighted sum of rents for the entire panel.

As noted above, there are six panels, so monthly data are available, and the updating each month is based on the previous months index (albeit from a different panel):

$$\delta_t = \delta_{t-1}\left[\frac{\Sigma_i \omega_i R_{it}}{\Sigma_i \omega_i R_{it-6}}\right]^{\frac{1}{6}} \tag{8}$$

[1]Another objection (Clapp and Giaccotto, 1998; Case and Quigley, 1991; McMillen and Thorsnes, 2006) is that the repeat sales model requires that the X vector (of observable and unobservable housing characteristics) and the β vector (of characteristic weights) to be constant. Since our primary goal in this paper is to mimic standard repeat sales indexes for the rental market we do not pursue models which allow such changes, but this remains a goal for future research.

[2]See Bureau of Labor Statistics (2007) for details on the BLS method.

Thus the BLS method also uses repeated observations on the same units to construct its index, and both BLS and our RRI take advantage of the fact that rents are more regularly observed than are prices.[3] However the sampling methods used by BLS indicate two differences between the two indexes. First and most importantly, our method reflects current market conditions. If all leases are annual, only $\frac{1}{12}^{\text{th}}$ of the BLS sample will reflect market conditions and some rents will reflect market conditions that are (nearly) a year old. Our use of rents from leases at the beginning of the rental contract insures that the data used in the construction of our indexes will reflect the contemporaneous market conditions, and suggests further that in times of market change, the RRI will lead the BLS index, since the BLS only reflects the conditions at time t at some future time period (depending on the distribution of renewal months.) This problem is of course exacerbated if sitting tenant discounts or premiums exist in the market. Second, the BLS method over smooths the rental index. Verbrugge (2008) discusses this in the context of the discrepancy in volatility between the estimated user cost of housing and the BLS index. As Verbrugge (2008) explains, there is implicit and explicit smoothing in the BLS index due to temporal aggregations. The implicit smoothing occurs because the index is an average of all extant leases, including newly renewed leases and previously renewed leases. The explicit smoothing occurs because the BLS index is constructed from overlapping semi-annual growth rates. As a result, the volatility of the BLS index underrepresents the actual volatility of rental prices.[4]

[3]Thus while the form of the BLS index is similar to that of a chain index, the regularity of its sampling makes it more immune to the criticism of chain indexes in Bailey et al. (1963).

[4]Another major difference between the two indexes is that the BLS constructs an arithmetic index, as opposed to a multiplicative one (Shiller, 1991). The distinction between the two is somewhat artificial, as one can be converted to the other with appropriate reweighting.

4 Results

Figure 2 shows the aggregate national repeat rent index and the national BLS rent index. In addition, for comparison, Figure 2 also shows the mean rent prevailing on leases in our data set. The BLS index indicates that national rents increased throughout the sample period with a brief pause in 2009. According to the BSL index, national housing costs (as approximated by aggregate rents) increased on average 3.1% per year between 1999 and 2010. In contrast, the national repeat rent index (RRI) indicates that rents were mostly constant during the first half of the sample period (1999 through 2004), increasing 2.8% in total or 0.48% per year, and actually ended 0.1% **lower** in 2010 than in 1999. A simple analysis of correlations between the respective indexes confirms that the RRI and BLS do not move together. For example, over the full sample period (1999 to 2010) the simple correlation coefficient between the quarterly change in the BLS and RRI indexes is 13%. However, over the period prior to the financial crisis, the simple correlation coefficient was -4%. We confirmed the simple correlations by estimating the following regression of the change in indexes:

$$\Delta RRI_t = \alpha + \beta_1 \Delta BLS_t + \beta_2 Crisis + \beta_3 Crisis * \Delta BLS_t + \varepsilon_t, \tag{9}$$

where ΔRRI_t and ΔBLS_t represent the quarterly change in the RRI and BLS indexes, respectively, and $Crisis$ is a dummy variable equal to one during the period of the financial crisis (2007 to 2010) and zero otherwise. The estimated coefficients are statistically insignificant, confirming our visual analysis that the RRI and BLS series do not track each other.

Figure 3 compares the repeat rent index (RRI) with the BLS rent index and the FHFA house price index (HPI), as well as the sample average rent for the eleven comparison MSAs. All of the indexes are normalized to 100 for the first time period for which the

RRI can be constructed for that MSA. Eleven panels are presented in ascending order of the average standard errors of the RRI index. A few observations are immediately evident. First, and somewhat surprisingly, the mean rent does not always rise at a higher rate than the indexes. The point of controlling for quality (through whatever method one chooses) is that unobserved quality improvements over time will cause mean rent index to rise even when there is no change in quality constant rental rates. In this sample, it is often the case that mean rent increases are less than index changes. While in some sense this blunts the need for quality-controlling indexes, it is also reassuring that our repeat sales indexes are perhaps not subject to the critique of Clapp and Giaccotto (1998), Case and Quigley (1991), and McMillen and Thorsnes (2006). Second, unlike the BLS, the RRI exhibits a sharp decline after 2007 in most cities. This is of interest because it contradicts, at least partially, one story about the financial crisis, which is that rent price ratios after 2007 were climbing, thus making owner-occupation a better financial decision (Yglesias, 2012). Third, and related to the previous two points, the average growth rate of the RRI is lower than that of the BLS index. Fourth, the RRI is, on visual inspection, more volatile than the BLS index. Fifth, the RRI tends to lead the BLS index roughly by one year. We provide statistical evidence on these last three points shortly.

As an example, in Dallas (Figure 3 (a)), the RRI shows small twin peaks in the first quarters of 2001 and 2002, and the BLS index exhibits twin peaks in the second quarters of 2002 and 2003. The RRI hit the bottom in the first quarter of 2004, and the BLS index hit the bottom in the second quarter of 2005. After four years of an upward trend, the RRI started to decrease in the second quarter of 2008, and the BLS index started to decrease in the fourth quarter of 2009. The recent decline is sharper in the RRI than in the BLS.

Similarly, in Seattle (Figure 3 (b)), the RRI marked a peak in the first quarter of 2008 after five years of steady appreciation, and the BLS index marked a peak in the first quarter of 2009. The RRI fell sharply until reaching a bottom in the third quarter of 2009. The BLS slightly decreased until reaching a bottom in the third quarter of 2010.

In Atlanta (Figure 3 (c)), the RRI reached a peak in the second quarter of 2001, and the BLS index reached a peak in the second quarter of 2002. The RRI bottomed in the fourth quarter of 2003, and the BLS index bottomed in the first quarter of 2005. The rent appreciation in the RRI continued for three years but ended around the second quarter of 2007. The appreciation in the BLS index ended around the first quarter of 2008. Both indexes exhibit declines thereafter.

Figure 4 depicts the lead-lag structure in the RRI and BLS indexes for Dallas, Seattle, and Atlanta, where the RRI is estimated with highest precision. The black solid line shows the RRI, while the red dotted line shows a one-year lead in the BLS index with an adjusted mean level. The two series clearly exhibit similarities in each MSA. However, after 2007, two series diverge from each other. Again, we observe a sharper decline in the RRI than in the BLS index and the growth rates of the RRI tend to exhibit a larger variance than those of the BLS index.

Table 3 presents the average quarterly growth rates of the RRI and the BLS index for the eleven MSAs and the results of the t-test for equal mean growth rates between the two indexes. The RRI exhibits a lower mean growth rate than the BLS index for all MSAs. The simple average of the mean growth rates for eleven MSAs is -0.04% (i.e. -0.17% per annum) for the RRI and 0.71% (i.e. 2.83% per annum) for the BLS index. This difference arises partly because the BLS index is inflated to adjust for the aging effect. However, this adjustment is approximately 0.2% per year and explains only a small fraction of the difference. The difference is largest for Washington, DC (-0.16%

for the RRI and 1.00% for the BLS index) and smallest for Seattle (0.43% for the RRI and 0.66% for the BLS index). The statistical significance for the difference between the indexes varies across MSAs. For example, in Dallas, Atlanta, Houston, Los Angeles, and Washington DC, we reject the null hypothesis of an equal mean at a 10% significance level against an alternative hypothesis that the mean growth rate of the RRI is lower than that of the BLS. For other MSAs, the difference is not statistically significant, possibly due to larger standard errors of estimates.

Table 4 presents the results for the F-test for equal variance of quarterly growth rates between the RRI and BLS. The RRI exhibits a larger variance in growth rates than the BLS index for all comparison MSAs. For example, in Dallas, the variance is 4.02×10^{-4} for the RRI and 6.50×10^{-5} for the BLS index. Because the F-statistics are large for all MSAs, we reject the null hypothesis of equal variance across all MSAs at a 5% or lower level of significance against an alternative that the variance of RRI is lower than that of the BLS, which is congruent with our expectations, given the discussion above of Verbrugge (2008).[5]

Figure 5 presents the RRI for fifteen other MSAs, for which the BLS index is not available. We present the RRI, the simple mean rent in our sample, and the FHFA HPI. Our index allows us to study rents in these relatively small MSAs. Despite their smaller sizes, they in general (but with some exceptions) all show slow rise in rents, and declines after 2007 in particular. Copies of the individual MSA level RRI indexes are available at the Penn State Institute for Real Estate Studies web-site (http://www.smeal.psu.edu/ires/repeat-rent-indexes.)

We now present evidence on the time series properties of these rental indexes. The first question is whether or not the new rental indexes share the same stationarity properties as the BLS. To that end, we present in Columns 1 and 2 of Table 5 the probability-

[5] We do not have BLS sample sizes for each city, so the difference in volatility may be due to the RRI using smaller samples than BLS. This is a topic of current research.

values of Phillips-Perron tests for unit roots for each of the individual RRIs and BLS indexes respectively. Recall that the null hypothesis in the Phillips-Perron test is that the series has a unit root. With only a few exceptions, the probability-values are greater (usually much greater) than 0.05, and so by usual criteria we do not reject the null hypothesis that these series are non-stationary. There are, to be sure, exceptions, so the conclusion is not universal. On that account we use a panel unit root test. The Im et al. (2003) test constructs critical values for the average Dickey-Fuller test statistic under the null that the panel data have a unit root. The probability-value for this test is 0.46 for the RRI, and 0.95 for the BLS panel, indicating a failure to reject that null.

Given these results it is natural to ask whether the two indexes, for any given MSA, are cointegrated. It is natural to suspect that they would be, since the unit root in both series could result from common stochastic trends resulting from permanent shocks to the MSAs housing market, and so would be reflected in both series, as opposed to permanent shocks in, say, the particular sampling pattern of either of the two indexes. We employ the standard Johansen-Juselius test, which uses the rank of the matrix of level coefficients in a vector error-correction model to assess cointegration. If the rank is one then cointegration exists, and if zero then the two series are not cointegrated. In Table 5 we present the trace test for the null that the rank is zero. Rejection (i.e. in favor of the alternative that the rank is 1) indicates cointegration exists. The results are split down the middle, with test statistics for five of the 11 MSAs indicating cointegration exists. Therefore we again have need for recourse to panel methods and here we employ the test of Westerlund (2007). There are multiple versions of the test, depending on the nature of the alternative hypothesis, but all of them have probability-values greater than 0.90 and so the conclusion we draw is that the series are (jointly considered) not pairwise cointegrated.[6]

[6] See Persyn and Westerlund (2008) for a description of the tests implementation in Stata.

We next turn to the Granger-causal relationship between the two series. Again, we consider this first on a city-by-city basis. Given the results on integration and cointegration it is appropriate to conduct the Granger tests using first differences. The tests thus regress the change in RRI on lagged changes on both RRI and BLS. A rejection of the F-test that the BLS coefficients are jointly zero indicates that BLS causes RRI. The roles of BLS and RRI are then reversed. We expect from these latter regressions a rejection, given our initial belief that RRI reflects current market conditions, but BLS does so only with a lag. We use four lags of each in these regressions and the probability-values are displayed in the last two columns of Table 5. In those columns, using a 5% critical value, we find that in five cases RRI causes BLS but not the reverse; in two cases BLS causes RRI but not the reverse, in three cases neither Granger-causes the other (although in two of those, the probability-value for RRI causing BLS is the lower of the two). Finally there is one case where there is mutual causality. In summary, the evidence is weighted, as we expected, toward the RRI index being causally prior to BLS. But once again we appeal to panel regressions to settle the question. We run a simple panel version of the above regressions, and include MSA dummies to capture fixed effects. The probability-value for the test of BLS causing RRI is 0.22, while for the reverse the probability-value is .03. Thus, jointly considered, our conclusion is unambiguously that RRI is causally prior to BLS and not the reverse.

5 Conclusions

We have constructed repeat rent indexes for a large number of cities, thus filling a hole in the current available data. We find that these series behave rather differently than BLS rent data. Our general conclusions are that the number of cities and years for which nominal rents fell is substantial, and by more than would be indicated by the BLS data particularly after the onset of the housing crisis in 2007. Repeat rent indexes (RRI) are more volatile than the BLS indexes, which is attributable to the smoothed nature of BLS sampling. Finally the BLS lags the repeat rent index, which is consistent with the idea that the BLS index is not indicative of current market conditions.

These differences in the path of rental is striking, and provide grounds for new research on rental markets and the relationship between rental and real estate markets. In particular, this will provide new perspectives on the contribution of rent to cost-of-living indexes, on the relative volatility of rent and housing prices, and the path of rent-price ratios and real estate capitalization, especially in the wake of the 2007 crash in prices. All of these are the object of current research.

References

Bailey, M. J., R. F. Muth, and H. O. Nourse. 1963. A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association* 58:pp. 933–942.

Bajari, P., C. L. Benkard, and J. Krainer. 2005. House prices and consumer welfare. *Journal of Urban Economics* 58:474–487.

Brunnermeier, M. K., and C. Julliard. 2008. Money Illusion and Housing Frenzies. *Review of Financial Studies* 21:135–180.

Bureau of Labor Statistics. 2007. The Consumer Price Index. In *BLS Handbook of Methods*, chap. 17. URL `http://www.bls.gov/opub/hom/`.

Calhoun, C. A. 1996. OFHEO House Price Indexes: HPI Technical Description. Working paper, Federal Housing Finance Agency. URL `http://www.fhfa.gov/webfiles/896/hpi_tech.pdf`.

Campbell, J. Y., and R. J. Shiller. 1988. Stock Prices, Earnings, and Expected Dividends. *Journal of Finance* 43:661–76.

Campbell, J. Y., and R. J. Shiller. 2005. Valuation Ratios and the Long-Run Stock Market Outlook: An Update. In R. Thaler (ed.), *Advances in Behavioral Finance, Volume II*, pp. 173–201. Princeton, NJ: Princeton University Press.

Campbell, S. D., M. A. Davis, J. Gallin, and R. F. Martin. 2009. What moves housing markets: A variance decomposition of the rent-price ratio. *Journal of Urban Economics* 66:90–102.

Case, B., and J. M. Quigley. 1991. The Dynamics of Real Estate Prices. *The Review of Economics and Statistics* 73:pp. 50–58.

Case, K. E., and R. J. Shiller. 1989. The Efficiency of the Market for Single-Family Homes. *The American Economic Review* 79:pp. 125–137.

Clapp, J. M., and C. Giaccotto. 1998. Price Indices Based on the Hedonic Repeat-Sales Method: Application to the Housing Market. *The Journal of Real Estate Finance and Economics* 16:5–26.

Crone, T. M., L. I. Nakamura, and R. Voith. 2010. Rents Have Been Rising, Not Falling, in the Postwar Period. *The Review of Economics and Statistics* 92:628–642.

Dougherty, A., and R. V. Order. 1982. Inflation, Housing Costs, and the Consumer Price Index. *The American Economic Review* 72:pp. 154–164.

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

Goodman, A. C., and M. Kawai. 1985. Length-of-Residence Discounts and Rental Housing Demand: Theory and Evidence. *Land Economics* 61:pp. 93–105.

Gordon, R. J., and T. vanGoethem. 2007. Downward Bias in the Most Important CPI Component: The Case of Rental Shelter, 1914-2003. In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, NBER Chapters, pp. 153–195. National Bureau of Economic Research, Inc.

Guasch, J. L., and R. C. Marshall. 1987. A theoretical and empirical analysis of the length of residency discount in the rental housing market. *Journal of Urban Economics* 22:291–311.

Himmelberg, C., C. Mayer, and T. Sinai. 2005. Assessing High House Prices: Bubbles, Fundamentals and Misperceptions. *Journal of Economic Perspectives* 19:67–92.

Hubert, F. 1995. Contracting with costly tenants. *Regional Science and Urban Economics* 25:631–654.

Im, K. S., M. H. Pesaran, and Y. Shin. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115:53–74.

Kanemoto, Y. 1990. Contract types in the property market. *Regional Science and Urban Economics* 20:5–22.

McCarthy, J., and R. W. Peach. 2004. Are home prices the next "bubble"? *Economic Policy Review* pp. 1–17.

McMillen, D. P., and P. Thorsnes. 2006. Housing Renovations and the Quantile Repeat-Sales Price Index. *Real Estate Economics* 34:567–584.

Meese, R. A., and N. E. Wallace. 1997. The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression and Hybrid Approaches. *The Journal of Real Estate Finance and Economics* 14:51–73.

Persyn, D., and J. Westerlund. 2008. Error-correctionbased cointegration tests for panel data. *Stata Journal* 8:232–241.

Raess, P., and T. von Ungern-Sternberg. 2002. A model of regulation in the rental housing market. *Regional Science and Urban Economics* 32:475–500.

Shiller, R. J. 1991. Arithmetic repeat sales price estimators. *Journal of Housing Economics* 1:110 – 126.

Sinai, T., and N. S. Souleles. 2005. Owner-Occupied Housing as a Hedge Against Rent Risk. *The Quarterly Journal of Economics* 120:763–789.

Verbrugge, R. 2008. The Puzzling Divergence Of Rents And User Costs, 1980-2004. *Review of Income and Wealth* 54:671–699.

Verbrugge, R., and R. Poole. 2010. Explaining the RentOER Inflation Divergence, 1999-2007. *Real Estate Economics* 38:633–657.

Westerlund, J. 2007. Testing for Error Correction in Panel Data. *Oxford Bulletin of Economics and Statistics* 69:709–748.

Yglesias, M. 2012. Trulia Says the Price:Rent Ratio is Below 15 in 98 out of 100 Metro Areas. *Slate Magazine,* Moneybox: A blog about business and economics, March 21. URL `http://www.slate.com`.

State	Lease Contracts	Individual Units	Apartment Complexes	Average Leases per Unit	Average Units per Complex
AL	8,639	4,191	22	2.1	191
AR	4,820	1,166	6	4.1	194
AZ	151,710	63,502	336	2.4	189
CA	173,135	61,886	454	2.8	136
CO	48,916	18,107	78	2.7	232
CT	1,706	681	8	2.5	85
DC	2,940	928	4	3.2	232
FL	215,510	72,694	338	3.0	215
GA	217,440	82,293	406	2.6	203
IA	4,885	1,749	14	2.8	125
ID	2,214	771	4	2.9	193
IL	11,240	4,669	25	2.4	187
IN	9,386	3,592	17	2.6	211
KS	5,210	1,266	6	4.1	211
KY	6,876	2,230	10	3.1	223
LA	4,437	1,824	9	2.4	203
MA	8,313	3,272	20	2.5	164
MD	7,655	3,032	24	2.5	126
ME	197	85	1	2.3	85
MI	21,958	6,980	47	3.1	149
MN	3,577	1,100	8	3.3	138
MO	2,005	550	7	3.6	79
MS	3,812	1,629	12	2.3	136
NC	58,614	22,372	108	2.6	207
NE	3,770	1,233	10	3.1	123
NH	1,478	495	2	3.0	248
NJ	85	62	1	1.4	62
NV	22,446	7,795	40	2.9	195
NY	6,466	2,754	20	2.3	138
OH	19,179	5,442	30	3.5	181
OK	17,166	5,237	27	3.3	194
OR	14,710	4,604	22	3.2	209
PA	76	66	5	1.2	13
SC	31,174	11,730	83	2.7	141
TN	38,325	14,072	66	2.7	213
TX	255,685	103,166	491	2.5	210
UT	6,019	2,852	13	2.1	219
VA	42,701	13,650	58	3.1	235
WA	47,714	17,175	100	2.8	172
WI	901	224	2	4.0	112
Total	1,483,090	551,126	2,934	2.7	169

Table 1: **Distribution of Leases Across States**

MSA	Lease Contracts	Individual Units	Apartment Complexes	Average Leases per Unit	Average Units per Complex
Atlanta	170,046	66,945	326	2.5	205
Boston	6,624	2,579	16	2.6	161
Dallas	94,400	39,786	190	2.4	209
Detroit	4,967	2,237	19	2.2	118
Houston	91,928	39,683	186	2.3	213
Los Angeles	33,296	13,055	145	2.6	90
Miami	20,361	7,096	27	2.9	263
New York	6,330	2,675	19	2.4	141
San Francisco	18,225	5,803	48	3.1	121
Seattle	43,488	15,634	89	2.8	176
Washington	28,716	10,561	42	2.7	251
Total	518,381	206,054	1,107	2.5	177

Table 2: **Distribution of Leases Across the Major MSAs**

City	Obs	Growth Rates of RRI(γ_{RRI}) Mean	Std. Err.	Growth Rates of BLS (γ_{BLS}) Mean	Std. Err.	t-test for an equal mean t-stat	$Pr.$
Dallas	42	-0.0004	0.0031	0.0047	0.0012	-1.52	0.067
Seattle	29	0.0043	0.0052	0.0066	0.0018	-0.42	0.338
Atlanta	43	-0.0035	0.0025	0.0037	0.0018	-2.33	0.011
Houston	40	-0.0009	0.0025	0.0063	0.0009	-2.68	0.005
San Francisco	20	-0.0014	0.0068	0.006	0.0013	-1.06	0.15
Los Angeles	25	-0.0017	0.004	0.0108	0.0015	-2.96	0.003
Boston	25	-0.0022	0.0106	0.0041	0.0012	-0.59	0.28
Washington DC	25	-0.0016	0.0084	0.01	0.0012	-1.36	0.092
Miami	26	0.0052	0.0129	0.0114	0.0019	-0.47	0.32
New York	22	0.0006	0.0089	0.0106	0.0009	-1.11	0.139
Detroit	32	-0.0032	0.0105	0.0035	0.0016	-0.63	0.267

Table 3: **Mean Growth Rates of the RRI and BLS indexes**

City	Obs	Growth Rates of RRI Variance	Growth Rates of BLS Variance	F-test for equal variance F-stat	Pr.
Dallas	42	4.02E-04	6.50E-05	6.18	0
Seattle	29	7.73E-04	9.47E-05	8.16	0
Atlanta	43	2.76E-04	1.37E-04	2.02	0.013
Houston	40	2.59E-04	3.14E-05	8.25	0
San Francisco	20	9.21E-04	3.59E-05	25.69	0
Los Angeles	25	3.91E-04	5.38E-05	7.27	0
Boston	25	2.82E-03	3.52E-05	79.92	0
Washington DC	25	1.78E-03	3.59E-05	49.41	0
Miami	26	4.33E-03	9.78E-05	44.31	0
New York	22	1.76E-03	1.81E-05	97.06	0
Detroit	32	3.53E-03	8.62E-05	41.01	0

Table 4: **Variance of Growth Rates of the RRI and BLS indexes**

	Phillips-Perron			Granger Causality	
	RRI	BLS	Johansen	RRI→BLS	BLS→RRI
Atlanta	0.83	0.39	17.7*	0.06	0.41
Boston	0.53	0.61	6.6	0.22	0.02
Dallas	0.31	0.85	9.9	0.05	0.45
Detroit	0.01	0.36	11.9	0.18	0.37
Houston	0.47	0.8	8.4	0	0.02
Los Angeles	0.81	0.04	32.0*	0	0.76
Miami	0.05	0.35	21.1*	0.86	0.04
New York	0.16	0.85	11.7	0.34	0.96
San Francisco	0.71	0.89	19.8*	0	0.09
Seattle	0.49	0.94	38.5*	0	0
Washington	0.52	0.96	8.4	0.66	0.1

Table 5: **Causality Tests**

Note: The first two columns entries are probability-values for the null hypothesis that the given index contains a unit root. The third column is the test statistic for the null hypothesis that the pair of indexes for a give MSA are cointegrated. Stars indicate rejection at the 5% level. The last two columns are probability-values for the null hypothesis that the first-named series does not Granger-cause the second-named series.

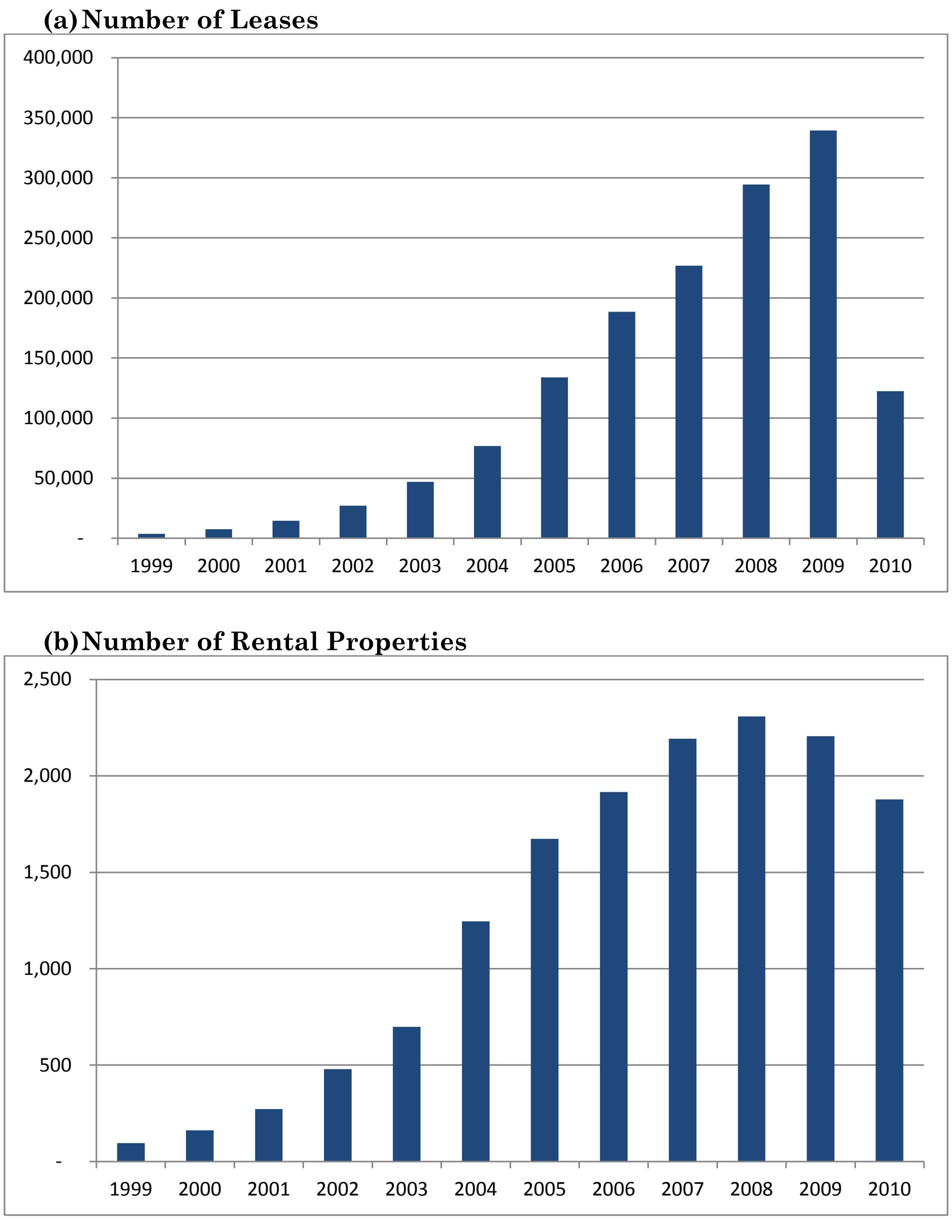


Figure 1: **Distribution of Leases and Rental Properties by Year**

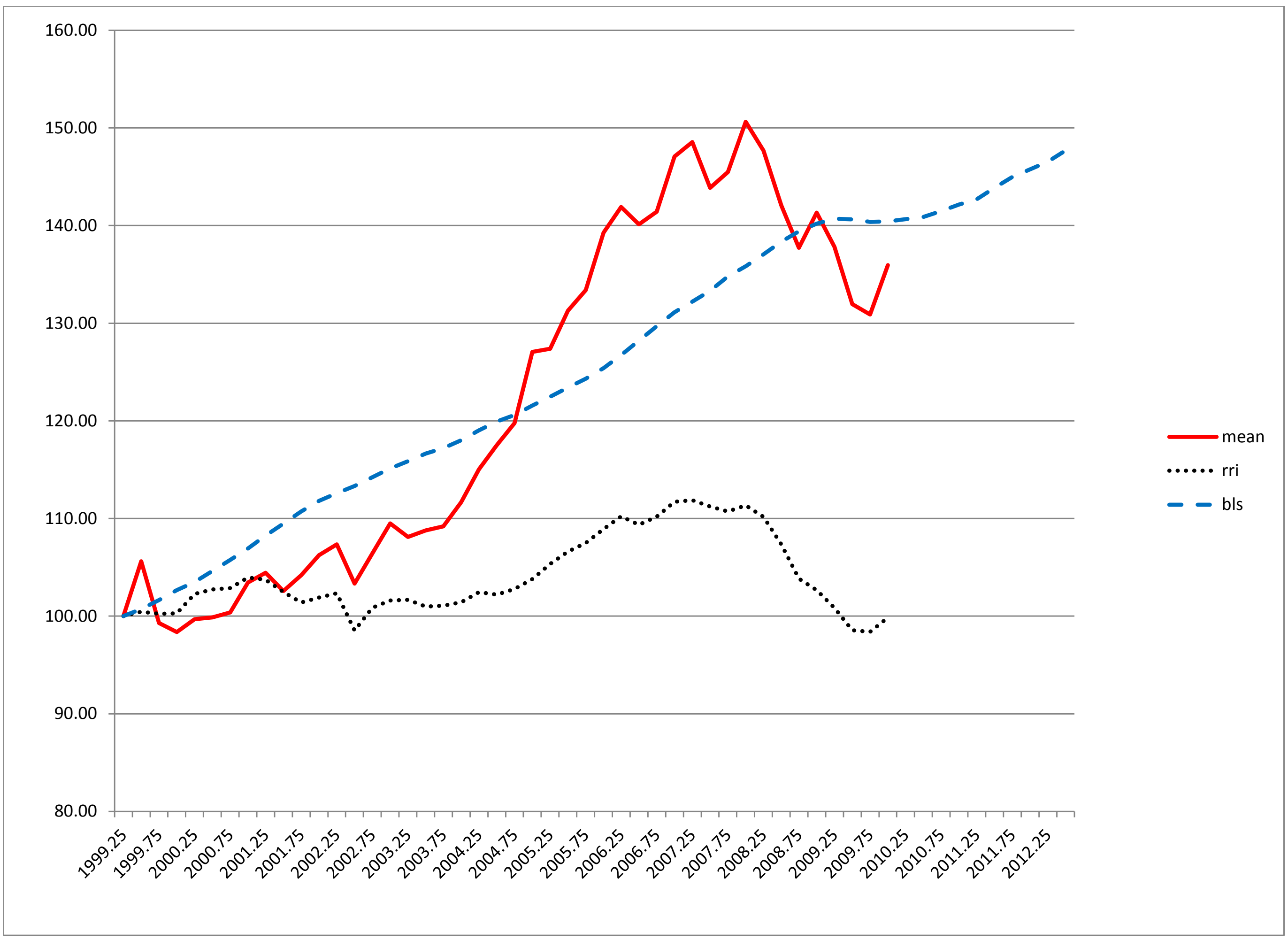
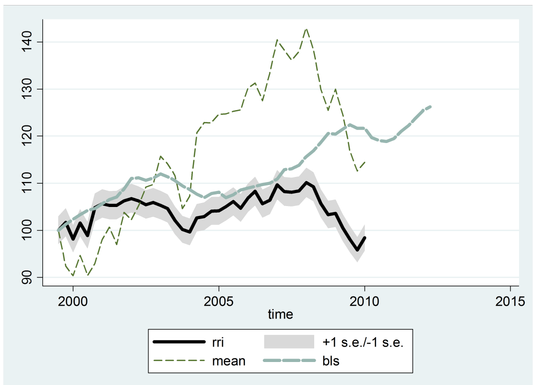
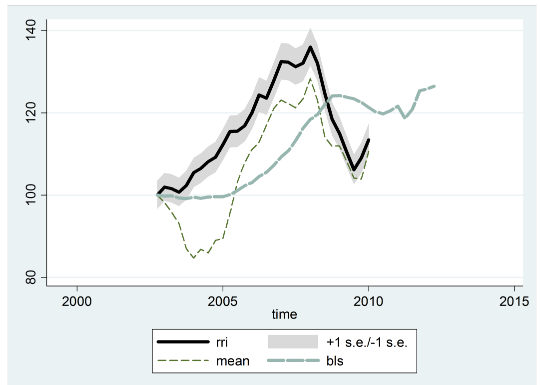


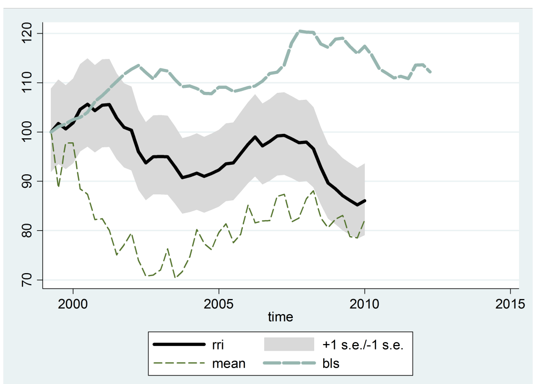
Figure 2: **National Repeat Rent and BLS Indexes**



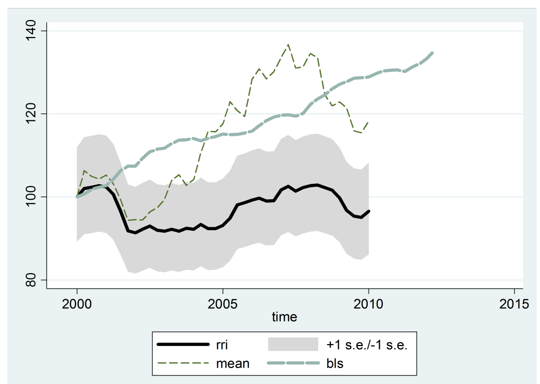
(a) Dallas



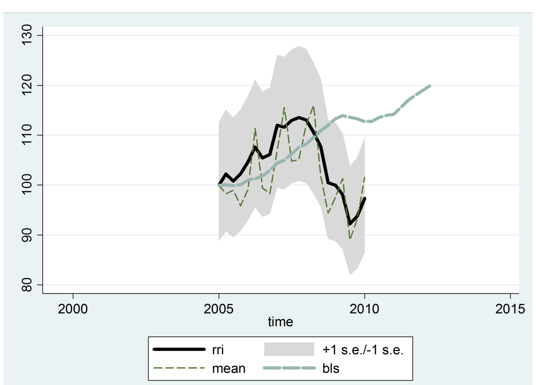
(b) Seattle



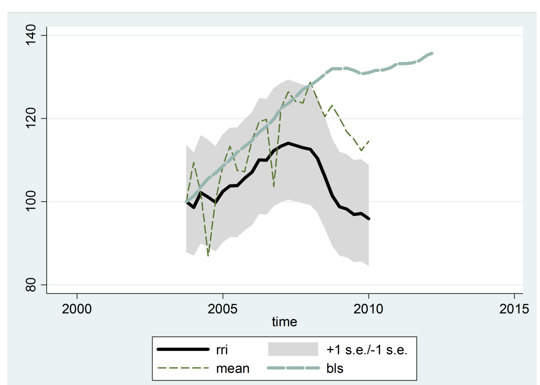
(c) Atlanta



(d) Houston

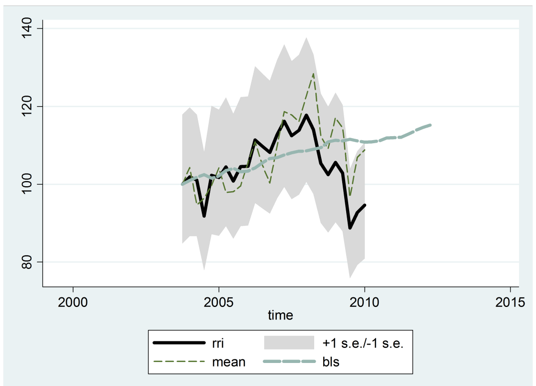


(e) San Francisco

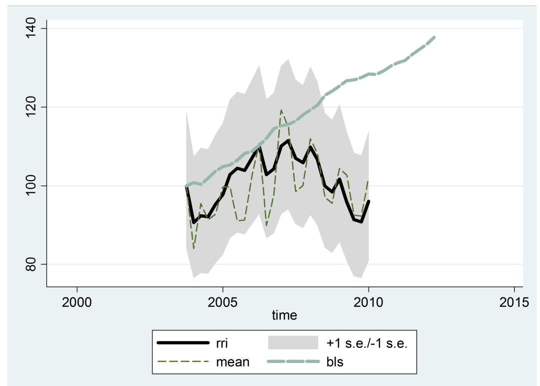


(f) Los Angeles

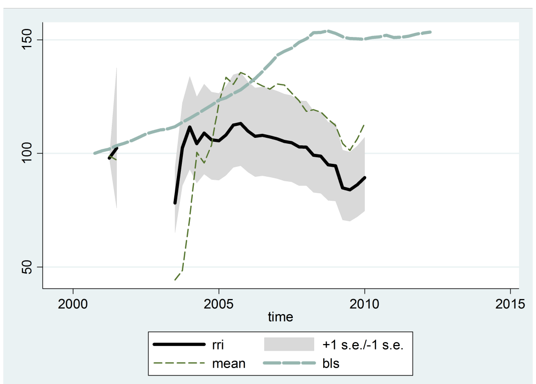
Figure 3: Comparison of Indexes for 11 MSAs



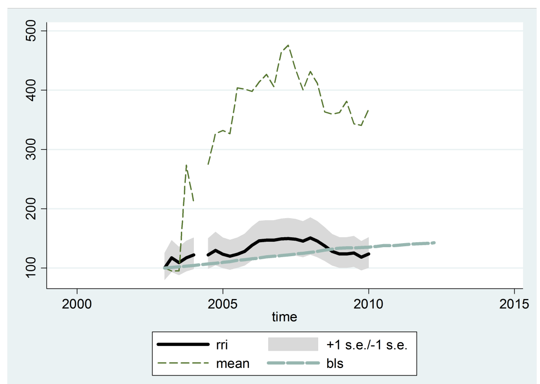
(g) Boston



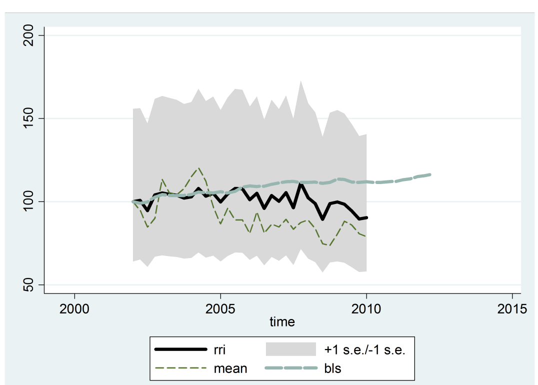
(h) Washington, DC



(i) Miami



(j) New York



(k) Detroit

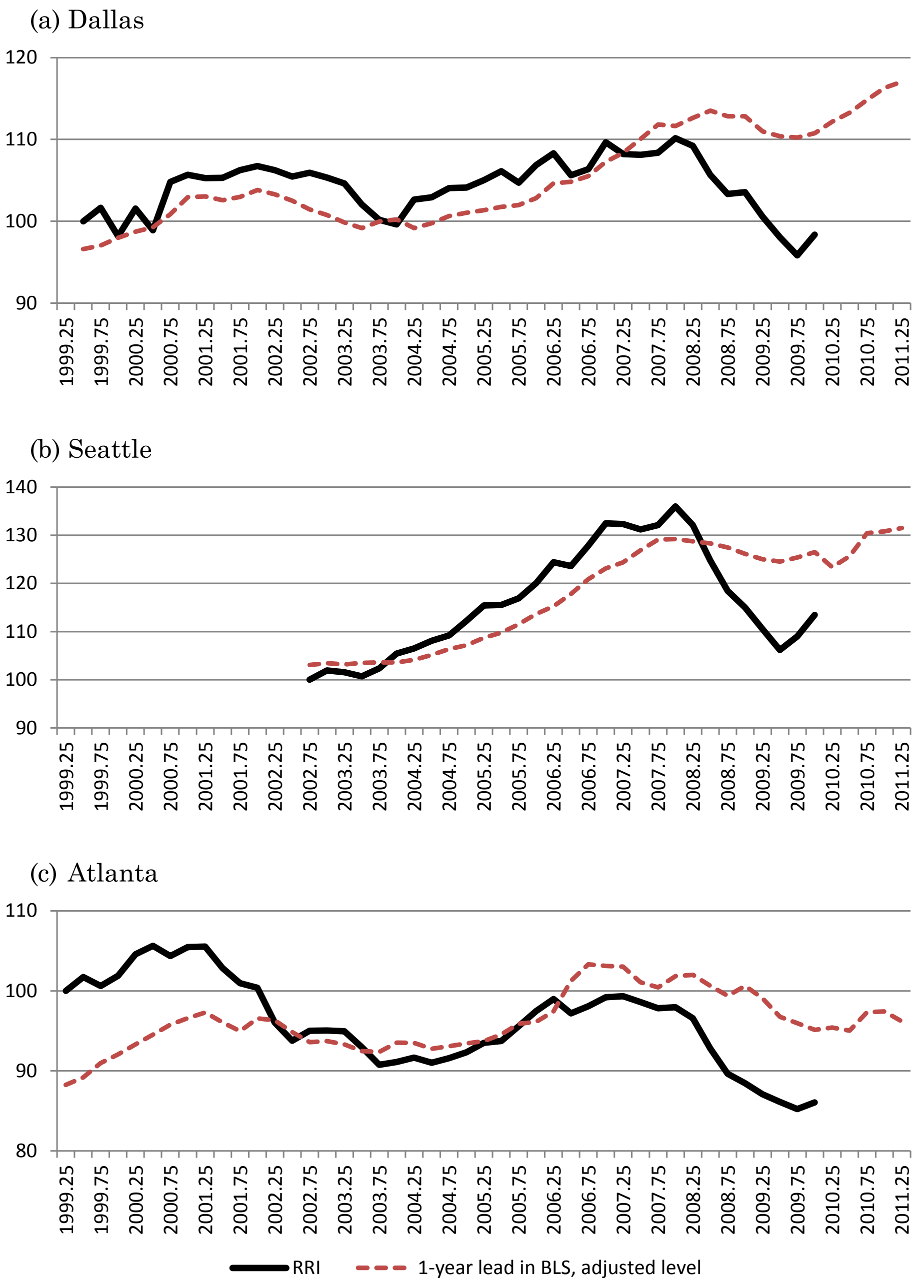


Figure 3: **The RRI and a one-year lead in the BLS index**

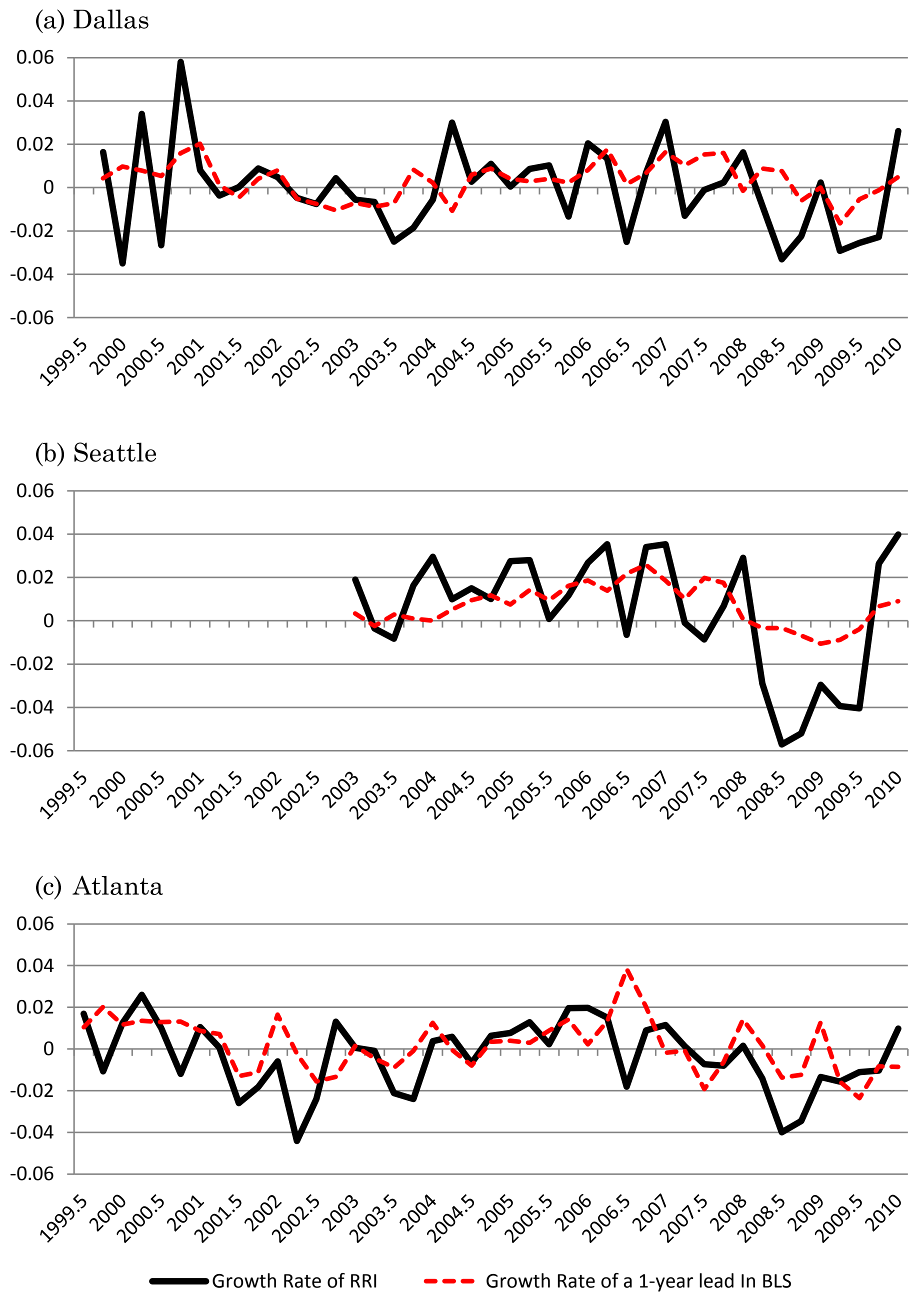
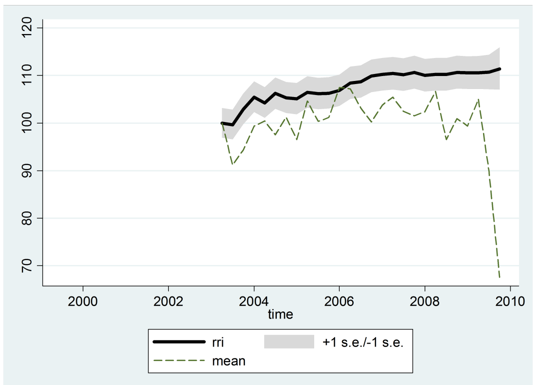
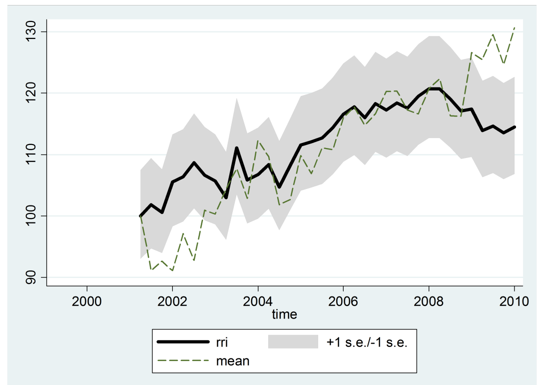


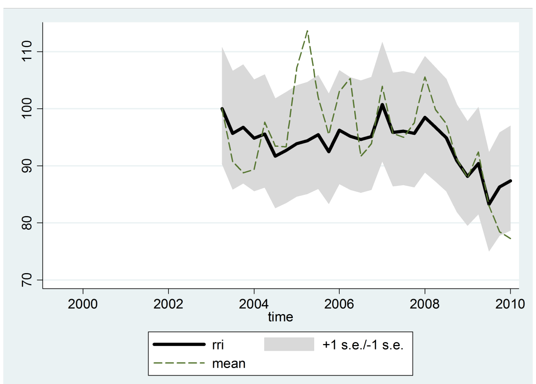
Figure 4: **Growth Rates of the RRI and a one-year lead in the BLS index**



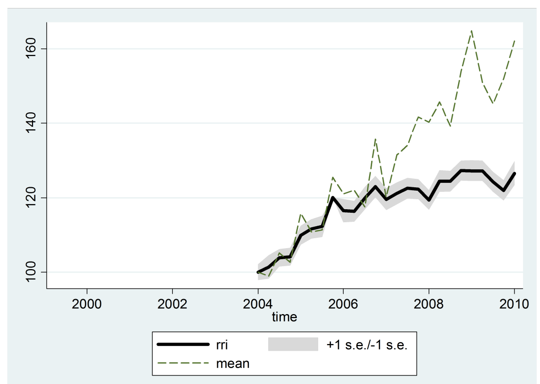
(a) Bloomington



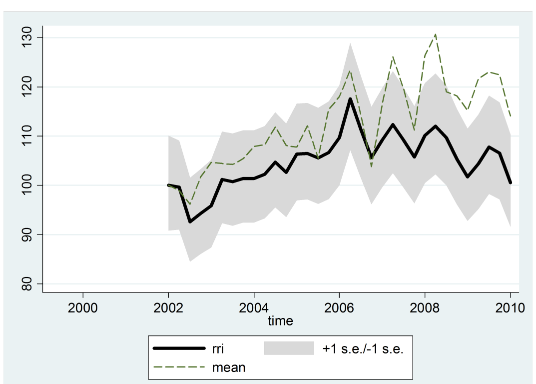
(b) Columbia



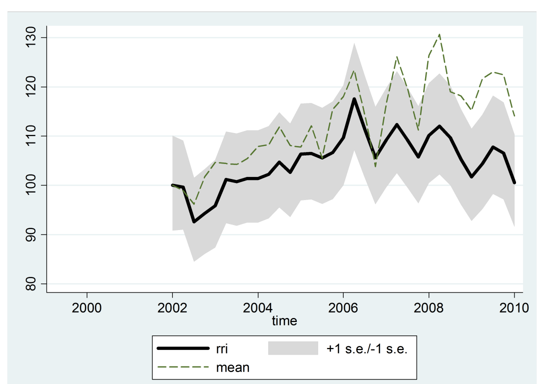
(c) Durham



(d) Crestview

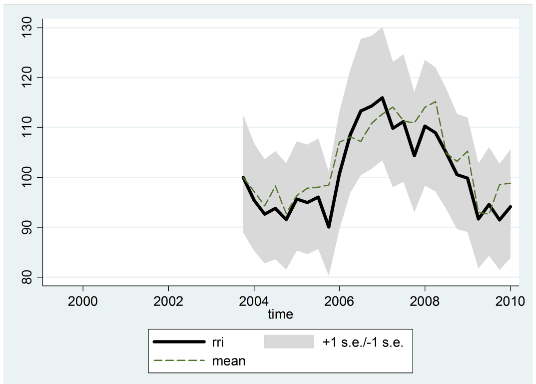


(e) Deltona

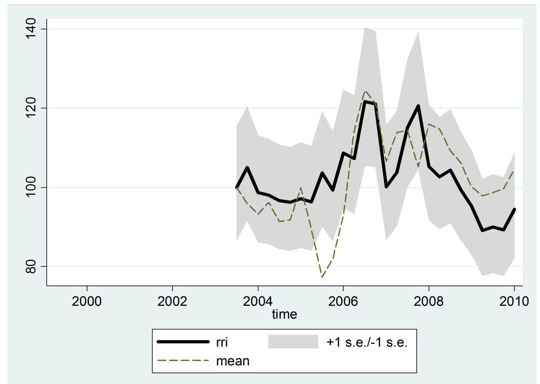


(f) Lansing

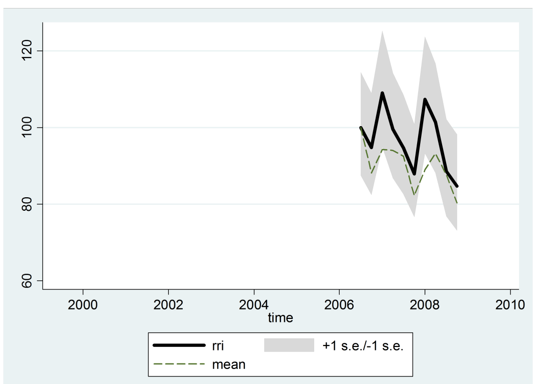
Figure 5: RRI for MSAs with no BLS index



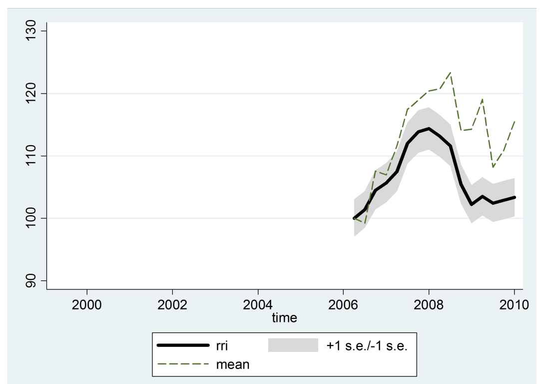
(g) Louisville



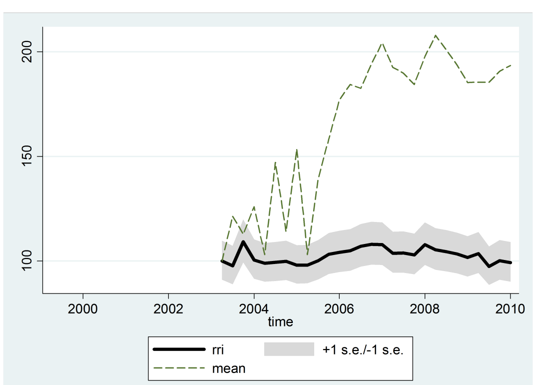
(h) Montgomery



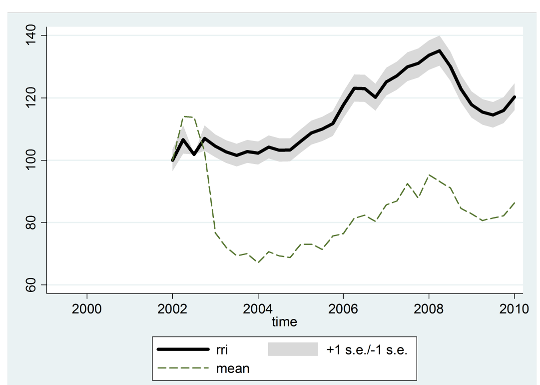
(i) Norwich



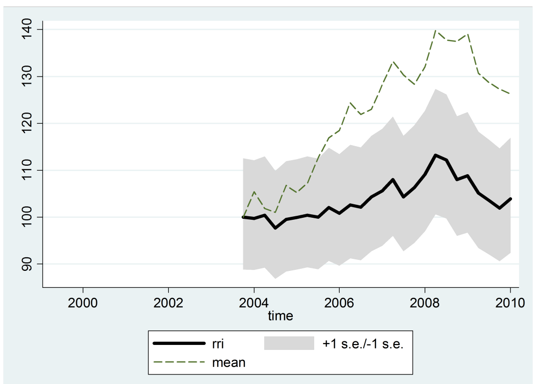
(j) Salt Lake City



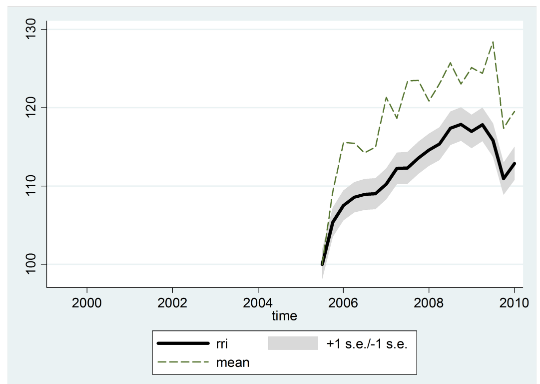
(k) Savannah



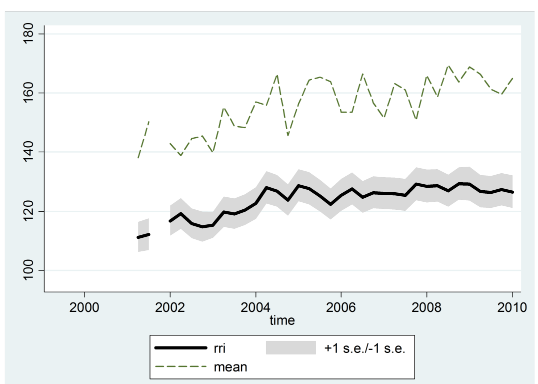
(l) Stockton



(m) Tulsa



(n) Wichita



(o) Worcester