Predicting returns with a co-fractional VAR model

Daniela Osterrieder

*CREATEs and Aarhus University*

and

Peter C. Schotman

*Maastricht University, CEPR and NETSPAR*

January 13, 2011

**Abstract:** We use co-fractional models to evaluate the predictive relations between returns and a valuation ratio. The co-fractional model can handle situations where financial returns are predicted using persistent valuation ratios, like dividend to price. For our application we consider very long time series, covering 355 years of real-estate returns and rent to price ratios. We find robust evidence of a fractional root of $d = 0.75$ in the rent to price ratio. The co-fractional model empirically outperforms the traditional triangular time-series model of return predictability. For annual data, the difference in predictive $R^2$ is about 8%. We conclude that the co-fractional VAR provides an alternative parsimonious model for the interaction between returns and valuation ratios. The long-memory properties of the fractional model have important implications for the fit of present-value models and the term structure of risk.

**JEL classifications:** C22, C32, C53, G12

**Keywords:** long-memory processes, fractional cointegration, return prediction, real-estate prices and rents

**Corresponding author:** D. Osterrieder, CREATEs, School of Economics and Management, Aarhus University, Bartholins Allé 10, DK-8000 Aarhus C, email: dosterrieder@creates.au.dk.
1 Introduction

The co-fractional vector autoregressive model of Johansen (2008) is very well suited for studying the relation between asset returns and a valuation ratio. Valuation ratios such as the dividend yield or, in a real estate context, the rent to price ratio, are known to be very persistent time series. Johansen’s model allows for the variables to have a fractionally integrated data generating process and hence accounts for a wide variety of persistence. The model differs from the Granger vector error correction model (Granger, 1986) in terms of the parameterization of the short-run dynamics. It is also much more flexible than the triangular specification of cointegrated systems that is typically used for predicting returns with valuation ratios. Despite its flexibility, the model is still parsimonious.

The theoretical properties of the co-fractional VAR process are well established; see Johansen (2008, 2009). There is also some research into the estimation of these processes and hypothesis testing. For instance, Johansen and Nielsen (2010) derive the asymptotic properties of the parameter’s maximum likelihood estimates and construct a likelihood ratio test for the null of a fractional unit root. Evidence on the empirical applicability of the model is still scarce, however.

Gil-Alana and Moreno (2007) suggest applying the co-fractional VAR to a multivariate extension of their yield curve model with fractional interest rates. Mills (2009) recommends the Johansen (2008) specification for modeling common trends and cointegration in economic time series, per se, and Rossi and Santucci de Magistris (2009) present it as an alternative to the Granger error correction model in their analysis of the range-based volatility.

We consider the co-fractional VAR model for asset return modeling and prediction. On the basis of a unique data set of 355 years of Dutch real-estate market observations, we model returns in the housing market. Our main findings are that the Johansen model fits the data very well, as opposed to the triangular model. Parameter estimates and goodness of fit are robust against the exact specification of the fractional integration parameter.

Why is the co-fractional VAR a good modeling choice for the prediction of asset returns? Firstly, we show in this paper that the predictor variable, in our case, the rent to price ratio, is a highly persistent process. Applying recent techniques due to Shimotsu and Phillips (2006), among others, to the long housing data series, we find that the logarithmic rent to price ratio is integrated of the order 0.75 and, hence, ex-
hibits long memory. The valuation ratio is thus a nonstationary, yet mean-reverting, series. This is in line with the stream of academic literature that observes the persistence of predictor variables. For instance, Stambaugh (1999) shows that in the stock market, the monthly dividend yield is highly persistent with an autoregressive parameter of 0.95 and higher. Campbell and Yogo (2006) report an 95%-confidence interval of \([0.96, 1.01]\) for the dividend-price ratio. Similarly, Ayuso and Restoy (2006) find an autoregressive parameter of around 0.9 and higher for rent to price ratios in the real-estate markets of Spain and the UK.

A second motivation for the Johansen model is the fractional cointegrating relation that we find. We provide evidence that the logarithmic rental, as well as the price series, follow unit-root processes. The difference between the two series is the logarithmic rent to price ratio, which is integrated of order 0.75. The dimensionality reduction resulting from the cointegrating relation is thus equal to 0.25. Real estate prices and rents are, hence, \textit{weakly fractionally cointegrated} in the terminology of Hualde and Robinson (2007). Weak cointegration in levels, in turn, implies a strong cointegrating relation between returns and the rent to price ratio.

Confirming cointegration between prices and cash flows in a financial market, we also contribute to the literature on empirical tests of present value models. Standard financial theory suggests that prices and cash flows follow a common stochastic trend. According to the present value model current prices are equal to discounted future cash flows and hence there must be an inherent long-run relation between both series. Yet, the empirical evidence in favor of a common trend is, at best, mixed. Campbell and Shiller (1987) find some support for the hypothesis in bond markets, but not in stock markets. More recently, Cochrane (2008) does not find evidence against cointegration in the stock market. In the real-estate literature, results for testing the cointegration hypothesis are more encouraging. Meese and Wallace (1994) find cointegration in the Californian housing market and Gallin (2008) generalizes this finding to the entire US housing market. We confirm these results for the Dutch housing market.

Finally, we add evidence to the discussion of predictability of asset returns. The strong support for the cointegration hypothesis and, as a result, the evidence in favor of the present value relation, implies that either asset returns or the growth in cash flows must be predictable to enforce the mean reversion of the valuation ratio after shocks to the system. Rejections of the forecastability of changes in cash flows are numerous. For instance, Cochrane (2008) firmly rejects the hypothesis
with respect to dividend growth and Plazzi et al. (2010) show that the rent to price ratio cannot forecast future rental growth in apartment, retail and industrial real-estate properties. Hence, returns must be predictable; and indeed, using the co-fractional VAR of Johansen to jointly model returns and the rent to price ratio, we capture a substantial fraction of the variation in annual returns. A standard predictive regression yields a much smaller predictable component.

2 Data properties

For our analysis, we use 355 years of annual housing market data from The Netherlands, covering the period 1650-2005. The data set has been constructed and analyzed by Ambrose et al. (2010). They construct a repeat sales index, reaching from 1650 until 1965, based on housing transactions on one of the canals in Amsterdam (Herengracht). The data for the remaining years, from 1966 to 2005, stem from the national median house price data from the national organization of Dutch realtors (NVM) and the Dutch Central Bureau of Statistics (CBS). The estimated rental index also has several sources. The Ambrose et al. (2010) index combines a repeated market rent index (1650-1850) of a broad spectrum of institutional housing property, a national house rent index (1851-1913), and various publications of the CBS for the remaining years. As we are ultimately interested in real data series, we further use the long-run consumer price index (CPI) constructed by Ambrose et al. (2010). The index is again a combination of sources and details can be found in Ambrose et al. (2010).

Ambrose et al. (2010) observe rents and prices for the same real estate object only in a very few years of the total time period. Hence, they scale the rent to price ratio to 4.5% in 2001, and adjust the price and rental series accordingly. The value 4.5% for 2001 is taken from the ROZ/IPD index, that is the Netherlands annual property index, for the corresponding year and resembles the annual rental yield direct return on Dutch residential real estate. Real prices and rents are obtained by dividing the respective rescaled index values at time \( t \) by the CPI at time \( t = 1, \ldots, T \).

Using the conventional notation, we define simple net returns \( R_{t+1} \) on an asset as

\[
R_{t+1} = \frac{P_{t+1} + F_{t+1}}{P_t} - 1,
\]

where \( F_{t+1} \) denotes the cash flow that the investor receives from his investment between \( t \) and \( t + 1 \) and \( P_t \) denotes the market price of the investment at time \( t \).
in Campbell et al. (2006), we proxy cash flows in the housing market by rents. We write the corresponding logarithmic measure as \( r_{t+1} \equiv \ln (1 + R_{t+1}) \), where lower case letters denote natural logarithms throughout the paper.

Table 1 shows summary statistics for the real housing series. The average rent to price ratio over the entire data set is 6%, with a standard deviation of approximately 2%. The average net real return on the real estate, as defined in (1), is 8.4% annually. The volatility of returns is around 22%. Returns on the housing market are thus substantially more volatile than the rent to price ratio. The first-order autocorrelation of returns is negative. This suggests a substantial measurement error in the return data. Motivated by the high volatility and possible measurement errors, we will not model the time-series of housing returns directly, but rather infer its dynamic properties from the rent to price ratio. The rental series, house prices, and the rent to price ratio, in levels as well as in natural logarithms, seem to exhibit a substantial degree of persistence. Their autocorrelation coefficients are all positive, large and decrease slowly.

| Table 1: Summary statistics for the real estate data: 1650-2005 |
|------------------------|--------|--------|--------|--------|--------|--------|--------|
|                        | Ave.   | StDev. | 1      | 2      | 3      | 12     | 24     |
| **Levels**             |        |        |        |        |        |        |        |
| Real prices \( P_t \)  | 64.528 | 26.919 | 0.836  | 0.755  | 0.710  | 0.383  | 0.229  |
| Real rents \( F_t \)   | 3.528  | 1.116  | 0.953  | 0.899  | 0.851  | 0.361  | 0.216  |
| Rent to price ratio \( F_t/P_t \) | 0.060 | 0.020  | 0.766  | 0.705  | 0.657  | 0.427  | 0.194  |
| Real returns \( R_t \) | 0.084  | 0.222  | -0.239 | 0.002  | -0.120 | 0.050  | 0.072  |
| **Natural logarithms** |        |        |        |        |        |        |        |
| Real prices \( p_t \)  | 4.080  | 0.429  | 0.877  | 0.821  | 0.768  | 0.489  | 0.245  |
| Real rents \( f_t \)   | 1.212  | 0.316  | 0.952  | 0.890  | 0.832  | 0.306  | 0.224  |
| Rent to price ratio \( (f – p)_t \) | -2.867| 0.317  | 0.801  | 0.746  | 0.706  | 0.440  | 0.208  |
| Real returns \( r_t \) | 0.060  | 0.202  | -0.276 | -0.009 | -0.101 | 0.041  | 0.078  |

The data set is especially attractive since such long historical data series are not readily available for other asset classes. For equity, for instance, there is the data set compiled by Shiller (2010) that reaches back until 1871 and Siegel (2010) covers bond and stock data for the 19th and 20th century. Dynamic properties of 355 years of real-estate returns and the rent to price ratio can thus provide a new perspective on the predictability, persistence, and mean reversion of asset returns.

The advantage of utilizing this very long real-estate data series is that it allows for a precise estimation of the dynamic properties of the real-estate market. Figure 1
shows a time series plot of the rent to price ratio over the entire data period. From a
graphical analysis of the figure, we conjecture that the ratio follows a very persistent
process. Especially in the years from 1770 to 1900, the series exhibits persistent
upward and downward trends. Reversions to the average of six per cent are only
intermittent over the complete period. In addition, the autocorrelation estimates from
Table 1 suggest that with a lag of 24 years, there is still approximately 20 percent of
a shock left in the process. Hence, autocorrelations of the rent to price ratio decrease
very slowly, motivating a long-memory time series model for the process.

Figure 1: Rent to price ratio (1650-2005) - The figure plots the rent to price
ratio in levels over 355 years. The black line corresponds to the annual percentage of
the ratio. The gray dotted line resembles the unconditional mean of the series.

3 Integration orders and fractional cointegration

We estimate a long-memory time-series model for the logarithmic rent to price ratio.
More precisely, we assume that the series \((f − p)_t\) is a fractionally integrated process,
defined as

\[(1 − L)^d(f − p)_t = u_t,\]  

where \(u_t\) is assumed to be stationary with zero mean and spectral density, \(f_u(\lambda)\),
satisfying \(f_u(\lambda) \sim \tau\) for \(\lambda \sim 0\) and \(\tau\) is a positive constant. \(L\) denotes the usual lag
operator and $d$ is the fractional integration parameter. The fractional filter $(1 - L)^d$ is defined as the infinite sum $(1 - L)^d = \Delta^d = \sum_{i=0}^{\infty} \theta_i L^i$, with $\theta_i = (-1)^i (\frac{d}{2})^i$.

If $d = 0$, the series follow a stationary $I(0)$ process; if $d = 1$, the series is nonstationary with a unit root. In general, a process with $d \in (-1/2, 1/2)$ is stationary and $d \geq 1/2$ defines a nonstationary series, yet mean reverting for $d \in (1/2, 1)$.

It is common to rely on semiparametric techniques for the estimation of the long-memory parameter, $d$, given that the short memory structure of the data is usually not known a priori. There are two commonly used classes of estimators to assess the long-memory behavior of the process close to frequency zero, while allowing for some unparameterized short-run dependence; the log-periodogram estimators (LP) introduced by Geweke and Porter-Hudak (1983) and the local Whittle estimators, originally developed by Künsch (1987). We utilize the latter class, since it is more robust and efficient, as pointed out by Henry and Zaffaroni (2002). Shimotsu and Phillips (2006) discuss three types of these estimators; the exact local Whittle (EW) due to Shimotsu and Phillips (2005), the local Whittle (LW), and the modified local Whittle (mLW) introduced by Phillips (1999) and Shimotsu and Phillips (2000). The EW, which is consistent and asymptotically normally distributed for any value of $d$, can be well approximated by the LW for $d \in (-1/2, 1/2)$ and by the mLW for $d \in (1/2, 1/2)$. The LW is consistent and asymptotically normal if $d \in (-1/2, 1/2)$. In addition, Robinson and Henry (1999) show that it is robust against conditional heteroskedasticity. The mLW is consistent and asymptotically normal for $d \in (1/2, 1/2)$ and is invariant to a linear trend. For robustness of the results, we consider each one of these estimators. The negative likelihood functions are defined as

$$Q_m^{(EW)} = \frac{1}{m} \sum_{j=1}^{m} \left( \ln \left( \tau \lambda_j^{-2d} \right) + \frac{1}{\tau} I_{(1-L)^d(f-p)}(\lambda_j) \right)$$

$$Q_m^{(LW)} = \frac{1}{m} \sum_{j=1}^{m} \left( \ln \left( \tau \lambda_j^{-2d} \right) + \frac{\lambda_j^{2d}}{\tau} I_{(f-p)}(\lambda_j) \right)$$

$$Q_m^{(mLW)} = \frac{1}{m} \sum_{j=1}^{m} \left( \ln \left( \tau \lambda_j^{-2d} \right) + \frac{\lambda_j^{2d}}{\tau} \left| w_{(f-p)}(\lambda_j) + \frac{e^{i\lambda_j} (f-p) \tau - (f-p) \tau}{\sqrt{2\pi T}} \right|^2 \right)$$

for $j = 1, \ldots, m$, which are minimized with respect to parameters $\tau$ and $d$. $I_{(1-L)^d(f-p)}(\lambda_j)$ and $I_{(f-p)}(\lambda_j)$ are the periodograms of the series $(1 - L)^d(f-p)_t$ and $(f-p)_t$, respectively; $\lambda_j$ are the harmonic Fourier frequencies; $m$ is a bandwidth parameter satisfying
\[
m = \frac{1}{T} + \frac{1}{m} \to 0 \text{ as } T \to \infty.
\]
\(w_{(f-p)}\) denotes the discrete Fourier transform of the series \((f-p)_t\).

The graphical analysis of the rent to price ratio and the inspection of the autocorrelations of the process hint at a persistent time series. Preliminary estimations of the fractional differencing parameter support the intuition that the ratio is nonstationary. Hence, we follow the suggestion of Shimotsu and Phillips (2006) and apply the LW estimation to the series in first differences, \(\Delta(f-p)_t\). The resulting estimator for the series in levels, \((f-p)_t\), is equal to \(\hat{d}_{LW} + 1\). The true mean of the data generating process is not known; this should be inconsequential for the consistency and asymptotic distribution of the LW, as we consider it in first differences. The same can be assumed for the mLW, since it is asymptotically equivalent to the LW in first differences. For the EW estimation, we consider the series detrended by its first observation, which ensures the consistency and asymptotic normality of the estimator for \(d \geq \frac{1}{2}\), as shown by Shimotsu (2008).

For the estimation, we have to select \(m\), the size of the spectral window. As all estimators are asymptotically distributed as \(\sqrt{m(\hat{d} - d)} \sim N(0, \frac{1}{4})\), the choice of \(m\) determines the speed of convergence. We select five different values for \(m \in [T^{0.6}, T^{0.75}]\), where the upper bound is the same as in Sun and Phillips (2004). It is common in this literature, to also consider \(m = T^{0.5}\) as a spectral window size (see, for instance, Schotman et al. (2008)). Yet, a visual inspection of the periodogram reveals that there are still substantial low-frequency movements at values \(m > T^{0.5}\).

Table 2 summarizes our results. The estimation is robust over the different estimators, but the results vary somewhat with the bandwidth. Overall, the rent to price ratio is integrated approximately within [0.65, 0.8], on average indicating a \(d\) equal to 0.75. The t-statistics suggest that we reject the null hypothesis of \(d = 0\) in all specifications. We also reject the hypothesis that the rent to price ratio is a unit root process with \(d = 1\), at a significance level of 5%.

To this point, we have assumed that the time-series properties of the rent to price ratio are stable over the entire data period of 355 years. That is, the low frequency behavior of the process is constant. However, there is a stream of the literature that claims that the detection of a long-memory structure in the data is often rather a spurious statistical artifact of a time series with structural breaks. For instance, Diebold and Inoue (2001) show that a regime-switching process can easily be perceived to be a fractionally integrated \(I(d)\) series. Given the ongoing debate, we consider the
Table 2: Estimates of $d$ for the rent to price ratio

The table reports the semiparametric estimators for the fractional differencing parameter, $d$, of the logarithmic rent to price ratio. $\hat{d}_{EW}$, $\hat{d}_{LW}$, and $\hat{d}_{mLW}$ denote the estimates resulting from the EW, the LW, and the mLW estimation, respectively. $m$ denotes the size of the spectral window. $s(d)$ is the asymptotic standard error, given by $1/\sqrt{4m}$.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\hat{d}_{EW}$</th>
<th>$\hat{d}_{LW}$</th>
<th>$\hat{d}_{mLW}$</th>
<th>$s(d)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>0.778</td>
<td>0.756</td>
<td>0.752</td>
<td>0.085</td>
</tr>
<tr>
<td>43</td>
<td>0.825</td>
<td>0.822</td>
<td>0.816</td>
<td>0.076</td>
</tr>
<tr>
<td>58</td>
<td>0.776</td>
<td>0.757</td>
<td>0.744</td>
<td>0.066</td>
</tr>
<tr>
<td>62</td>
<td>0.755</td>
<td>0.735</td>
<td>0.721</td>
<td>0.064</td>
</tr>
<tr>
<td>82</td>
<td>0.675</td>
<td>0.655</td>
<td>0.631</td>
<td>0.055</td>
</tr>
</tbody>
</table>

robustness of the long-memory properties of the rent to price ratio with respect to the sample period. Considering the time series in Figure 1 again, an obvious candidate for a structural break appears at the beginning of the 20th century. Whereas in the years before we observe very persistent and smooth upward and downward trends in the data, the fluctuation seems more volatile and less structured after the breakpoint. From Ambrose et al. (2010) we know that 1916 marks the start of a new regime in the rental market. The Dutch government exercised nominal rent control, which together with periods of high inflation and deflation caused extremely volatile real rents. Table 3 shows the results of estimating $d$ for both subsamples, i.e. 1650-1915 and 1916-2005, separately. Due to the smaller sample sizes, the estimates for $d$ vary somewhat more with the bandwidth length than in the entire sample. Yet, overall the estimated value for $d$ is high in both subsamples, with a minimum of 0.60 and a maximum of 0.87. Hence, the estimated non-stationarity of the entire data series does not seem to be caused by a structural change in the dynamics of the process.

Whereas the previous robustness check can be described as more intuitive, we continue with two statistical tests designed by Shimotsu (2006) to test for long memory versus structural breaks. The first test is based on the idea that if the estimates of the fractional differencing parameter are stable over $k$ subsamples, then the long memory cannot be spurious. Splitting our entire sample into $k = 4$ subsets of equal length, to ensure that there are still enough observations in each subsample and that the last subsample corresponds approximately to the period for which we assumed a structural break, we test the hypothesis that $H_0: d_{total} = d_1 = \ldots = d_k$. We calculate
The table reports the semiparametric estimators for the fractional differencing parameter, \( d \), of the logarithmic rent to price ratio in two different subsample periods. \( \hat{d}^{(1,2)}_{\text{EW}} \), \( \hat{d}^{(1,2)}_{\text{LW}} \), and \( \hat{d}^{(1,2)}_{\text{mLW}} \) denote the estimates for subsample 1 and 2 resulting from the EW, the LW, and the mLW estimation, respectively. \( m_1 \) denotes the size of the spectral window in the first subsample and \( m_2 \) in the second.

<table>
<thead>
<tr>
<th>subsample 1: 1650-1915</th>
<th>subsample 2: 1916-2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_1 )</td>
<td>( \hat{d}^{(1)}_{\text{EW}} )</td>
</tr>
<tr>
<td>29</td>
<td>0.870</td>
</tr>
<tr>
<td>36</td>
<td>0.808</td>
</tr>
<tr>
<td>47</td>
<td>0.784</td>
</tr>
<tr>
<td>51</td>
<td>0.741</td>
</tr>
<tr>
<td>66</td>
<td>0.710</td>
</tr>
</tbody>
</table>

the Wald statistic and the adjusted Wald statistic of Shimotsu (2006), both of which are distributed as \( \chi^2_{(k-1)} \). The second type of tests follows the intuition that if the original data series is differenced \( d \) times, where \( d \) has been previously estimated, then the resulting process should be \( I(0) \). Using the previously estimated values for \( d \), we prefilter the data series and conduct a Phillips-Perron (PP) unit-root test and a KPSS stationarity test on the residuals, for which Shimotsu (2006) provides critical values. If our results are robust, the former test should lead to a rejection and we should fail to reject the latter.

The results in Table 4 show that our long-memory structure of the rent to price ratio is very robust. Both, the Wald and the adjusted Wald test fail to reject the hypothesis that \( d \) is constant over all subsamples, at a 5% significance level\(^1\). The PP unit-root test on the residuals clearly rejects and the KPSS test for stationarity cannot reject. In sum, all the previous evidence taken together suggests that the (logarithmic) rent to price ratio is indeed a long-memory process with a fractional integration parameter of approximately 0.75.

The logarithmic rent to price ratio is by definition the sum of rents, \( f_t \), and (negative) prices, \( p_t \). Let rents be integrated of order \( I(d) \) and prices \( I(d) \). Then the sum of two integrated processes is a process that is integrated of the order \( \max(I(d), I(d)) \), unless the two time series cointegrate. To investigate the presence of a common

\(^1\)The only exceptions are the Wald tests using the LW and the mLW, with a bandwidth of 43 for the entire sample, and hence, a window of length 11 for each subsample. In these instances we fail to reject the null hypothesis only at a 1% significance level.
Table 4: Tests of long memory versus structural breaks

The table reports the values of the test statistics for the tests of long memory versus structural breaks following Shimotsu (2006). $k$ is the number of subsamples considered. $d_{EW}$, $d_{LW}$, and $d_{mLW}$ denote the semiparametric estimation technique, that is the EW, the LW, and the mLW estimation, respectively. $m_{total}$ denotes the size of the spectral window of the entire sample. Reported PP and KPSS critical values (CV) follow from Table 1 in Shimotsu (2006).

<table>
<thead>
<tr>
<th>$m_{total}$</th>
<th>Wald ($k = 4$)</th>
<th>adj. Wald ($k = 4$)</th>
<th>PP (w. intercept)</th>
<th>KPSS (w. intercept)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$W(\hat{d}_{EW})$</td>
<td>$W_c(\hat{d}_{EW})$</td>
<td>$PP(\hat{d}_{EW})$</td>
<td>$KPSS(\hat{d}_{EW})$</td>
</tr>
<tr>
<td>35</td>
<td>3.325</td>
<td>1.529</td>
<td>-23.731</td>
<td>0.085</td>
</tr>
<tr>
<td>43</td>
<td>5.789</td>
<td>2.922</td>
<td>-25.999</td>
<td>0.079</td>
</tr>
<tr>
<td>58</td>
<td>2.980</td>
<td>1.701</td>
<td>-23.703</td>
<td>0.085</td>
</tr>
<tr>
<td>62</td>
<td>3.834</td>
<td>1.976</td>
<td>-23.594</td>
<td>0.094</td>
</tr>
<tr>
<td>82</td>
<td>1.153</td>
<td>0.734</td>
<td>-21.604</td>
<td>0.125</td>
</tr>
<tr>
<td></td>
<td>$W(\hat{d}_{LW})$</td>
<td>$W_c(\hat{d}_{LW})$</td>
<td>$PP(\hat{d}_{LW})$</td>
<td>$KPSS(\hat{d}_{LW})$</td>
</tr>
<tr>
<td>35</td>
<td>6.546</td>
<td>3.010</td>
<td>-23.641</td>
<td>0.093</td>
</tr>
<tr>
<td>43</td>
<td>9.171</td>
<td>4.629</td>
<td>-25.247</td>
<td>0.081</td>
</tr>
<tr>
<td>58</td>
<td>5.209</td>
<td>2.973</td>
<td>-23.652</td>
<td>0.094</td>
</tr>
<tr>
<td>62</td>
<td>5.093</td>
<td>2.973</td>
<td>-23.138</td>
<td>0.092</td>
</tr>
<tr>
<td>82</td>
<td>2.518</td>
<td>1.603</td>
<td>-21.174</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>$W(\hat{d}_{mLW})$</td>
<td>$W_c(\hat{d}_{mLW})$</td>
<td>$PP(\hat{d}_{mLW})$</td>
<td>$KPSS(\hat{d}_{mLW})$</td>
</tr>
<tr>
<td>35</td>
<td>6.557</td>
<td>3.015</td>
<td>-23.528</td>
<td>0.095</td>
</tr>
<tr>
<td>43</td>
<td>9.157</td>
<td>4.622</td>
<td>-24.767</td>
<td>0.082</td>
</tr>
<tr>
<td>58</td>
<td>5.113</td>
<td>2.918</td>
<td>-23.352</td>
<td>0.096</td>
</tr>
<tr>
<td>62</td>
<td>5.045</td>
<td>2.945</td>
<td>-22.787</td>
<td>0.090</td>
</tr>
<tr>
<td>82</td>
<td>2.442</td>
<td>1.555</td>
<td>-20.687</td>
<td>0.134</td>
</tr>
<tr>
<td>sign. level</td>
<td>$\chi^2_{(3)}$ CV</td>
<td>$\chi^2_{(3)}$ CV</td>
<td>PP CV</td>
<td>KPSS CV</td>
</tr>
<tr>
<td>0.1</td>
<td>6.251</td>
<td>6.251</td>
<td>[-2.475,-2.568]</td>
<td>[0.321,0.348]</td>
</tr>
<tr>
<td>0.05</td>
<td>7.815</td>
<td>7.815</td>
<td>[-2.767,-2.855]</td>
<td>[0.419,0.463]</td>
</tr>
<tr>
<td>0.01</td>
<td>11.345</td>
<td>11.345</td>
<td>[-3.336,-3.430]</td>
<td>[0.661,0.743]</td>
</tr>
</tbody>
</table>
stochastic trend of real rents and prices, we continue to estimate the fractional integration parameter of both series in logarithms. Since both data series are denominated in real terms, we estimate $d$ by the univariate exact local Whittle estimator (UEW) of Sun and Phillips (2004). Sun and Phillips (2004) argue that by proxying unobservable expected inflation with ex post observed data, as we do in this study, one introduces weakly dependent dynamics (short memory) into the real rate. The EW estimator has unknown properties in the presence of additive perturbations, but the UEW explicitly accounts for the presence of additional short memory in the data series. In Appendix A, we reparametrize the negative Whittle likelihood of Sun and Phillips (2004). The concentrated negative log likelihood function for prices, $p_t$, is given by

$$
\hat{Q}^{(UEW)}_m = \ln \left( \frac{1}{m} \sum_{j=1}^m I_{(1-L)^d p} (\omega_j) \right) + \frac{1}{m} \sum_{j=1}^m \ln \left( \omega_j^{-2d} + \varphi \right) + 1,
$$

(6)

where $\varphi$ is a positive constant. We minimize (6) with respect to $d$ and $\varphi$. The likelihood for rents, $f_t$, follows analogously.

In the absence of a distribution theory for the UEW estimator, we consider an interval of asymptotic standard errors from the EW and the LP estimators, by which the standard deviation of the UEW is bounded, as Sun and Phillips (2004) show. Thus, standard errors of the UEW are bounded by $1/2\sqrt{m}$ from below, and by $\pi/2\sqrt{6m}$ from above (see, for instance, Robinson (1995)).

Table 5 summarizes the outcomes. The estimated fractional integration parameter for the real rents center around [0.89, 1.10] and for the house prices our estimates suggest $\hat{d} \approx 0.94$. We firmly reject the $I(0)$ hypothesis for both data series, yet more interestingly, we have substantial evidence that both series follow a unit-root process. For the rental index, the standard errors suggests a failure to reject the hypothesis, $d = 1$, at a 5% level. Even if t-statistics for the hypothesis are computed with the asymptotic variance of the EW estimator, we never reject the hypothesis. The same holds true for real house prices. To confirm this intuition, we conduct a standard unit root augmented Dickey-Fuller (ADF) test on both series. For both series we fail to reject the unit root hypothesis with a p-value of 0.11 for rents and a p-value of 0.16 for prices. If prices and rents are indeed $I(1)$ and the rent to price ratio is integrated with $d \approx 0.75$, rents and prices are weakly fractionally cointegrated. Hualde and Robinson (2007) define the term weak cointegration as the case when the difference between the
persistence of the respective individual series and their linear combination is less than \( \frac{1}{2} \). In our case this implies that there is a reduction in the dimensionality, but the common equilibrium-type explanation is not straightforward for such a cointegrating relation, as the linear combination is not covariance stationary.

Table 5: Estimates of \( d \) for the real rental and the price indices in logarithms

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \hat{d}_{UEW} )</th>
<th>( \hat{d}_{UEW} )</th>
<th>( s(d_{EW}), s(d_{LP}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>0.883</td>
<td>0.887</td>
<td>(0.085, 0.108)</td>
</tr>
<tr>
<td>43</td>
<td>0.987</td>
<td>0.896</td>
<td>(0.076, 0.098)</td>
</tr>
<tr>
<td>58</td>
<td>1.102</td>
<td>0.925</td>
<td>(0.066, 0.084)</td>
</tr>
<tr>
<td>62</td>
<td>1.024</td>
<td>0.953</td>
<td>(0.064, 0.081)</td>
</tr>
<tr>
<td>82</td>
<td>1.093</td>
<td>1.021</td>
<td>(0.055, 0.071)</td>
</tr>
</tbody>
</table>

4 Predicting and decomposing returns

If cash flows and prices follow a common stochastic trend on a financial market, valuation ratios, such as the rent to price ratio in the real estate market, should forecast future changes in prices or the cash-flow growth. This follows directly from the present value identity. Just like stock returns have a small but significant predictable component captured by dividend yields, real-estate returns may be predicted by the rent to price ratio. Whereas the former result is well established in the literature (see, for instance, Fama and French (1989), Lettau and Ludvigson (2004), and Cochrane (2008)), the predictability of real-estate returns has received less attention. Advocates of the latter are, for example, Case and Shiller (1990), who conclude that the rent to price ratio forecasts excess returns on the housing market, with a statistically significant coefficient, and Plazzi et al. (2010). To contribute to the discussion of predictability of returns on housing, we introduce a model that allows for fractional cointegration. The model representation is due to Johansen (2008, 2009). We apply this model to return modeling and discuss its relations to other well established models.
For logarithmic real returns, \( r_t \), we can write the well-known loglinear approximation for returns from Campbell and Shiller (1988a,b), given by

\[
r_t \approx \iota + (1 - \varsigma) (f_t - p_t) + \Delta p_t. \tag{7}
\]

From the previous section we know that the logarithmic rent to price ratio is an approximate \( I(0.75) \) process and we found evidence for a weak fractional cointegration between rents and prices. This in turn implies that returns and the rent to price ratio follow a common stochastic trend and that returns must be integrated of the same order as the ratio. The latter conclusion follows directly from Equation (7). Since returns are the sum of the rent to price ratio, \( I(0.75) \), and the change in prices, which is an \( I(0) \) process as established in the previous section, by definition the order of integration of returns is the higher of the two sum components. Hence, \( r_t \sim I(0.75) \).

Realized returns thus have a small persistent trend component, \( (f - p)_t \), that is masked by volatile price changes, \( \Delta p_t \). The \( \Delta p \)-part is mostly high frequency noise, while \( (f - p) \) determines the low-frequency properties.

If we solve Equation (7) for the growth in house prices, it follows that a linear combination of \( r_t \) and \( (f - p)_t \) is integrated of order \( I(0) \), because \( \Delta p_t \sim I(0) \). In sum, the weak cointegration between rents and prices implies a strong fractional cointegrating relation of returns and the rent to price ratio. The two series follow a common stochastic trend and a linear combination of the two is a stationary approximate \( I(0) \) process.

It is common in the financial asset pricing literature to predict asset returns by means of valuation ratios, like the dividend yield or price-earnings ratio (see, for instance, in Fama and French (1989). A natural candidate for conveying predictive information in the real estate market is the rent to price ratio, \( (f - p)_t \). We define the state vector \( z_t = [r_t, (f - p)_t]' \). Given the long memory in both elements of \( z_t \) and the cointegrating relation, we propose to represent the process by a co-fractional vector autoregressive model, \( \text{CFVAR}_d(p) \), where \( p \) denotes the number of lags included, due to Johansen (2008),

\[
\Delta^d z_t = \gamma \delta' (1 - \Delta^d) z_t + \sum_{i=1}^{p} \Gamma_i \Delta^d (1 - \Delta^d)^i z_t + \epsilon_t, \tag{8}
\]

where \( \epsilon_t \) are independently and identically distributed shocks with mean zero and covariance matrix \( \Sigma \). Equation (8) implies that \( z_t \) is fractional process of order \( d \),
whereas $\Delta^d z_t$ and $\delta' z_t$ are fractionally integrated of order zero. Hence, $\delta$ is the cointegrating vector and $\gamma$ contains the adjustment parameters of the error-correction term $(1 - \Delta^d)z_t$. If $\gamma\delta'$ was of full rank and $p = 0$, then model (8) would simply be a vector VAR(1) process in the lag operator $L_d = (1 - \Delta^d)$.

For our specification, the matrix $\gamma\delta'$ is of rank one, due to the cointegrating relation between $r_t$ and $(f - p)_t$. We can normalize the cointegration vector $\delta$ in such a way that the first element of the vector is equal to 1, i.e. $\delta = [1, -\tilde{\delta}]'$. As we impose cointegration on the model and we do not allow for complex coefficients, the conditions for inversion of the model derived in Theorem 8 of Johansen (2008) are satisfied and (8) has a solution for $z_t$. This solution is given by the sum of scaled past innovations of the process $\epsilon_t$, truncated at $t - 1$, plus a contribution of the initial condition, $\mu_t$.

If we assume further that $z_s = 0$ if $s \leq 0$, we can derive the infinite moving-average representation of the state vector $z_t$ as

$$z_t = \sum_{i=0}^{\infty} \Phi_i \epsilon_{t-i}. \quad (9)$$

The impulse responses $\Phi_i$ of the process can be derived recursively from

$$\begin{align*}
\Phi_0 &= I \\
\Phi_j &= \sum_{i=0}^{j-1} \Xi_{j-i} \Phi_i,
\end{align*} \quad (10)$$

where $\Xi_i = -I \theta_i - \gamma\delta' \theta_i - \Gamma_1 \sum_{l=0}^{i-1} \theta_{i-l} \theta_l$ for a CFVAR$_d(1)$ process. A full derivation of Equation (10) is given in Appendix B ((B1)-(B4)).

In order to establish what drives asset returns in the real estate market, we can decompose the variance of $r_t$ as in Campbell (1991). To do so, we first decompose the return shocks into a part resulting from news to the discount factor and the other part is due to innovations in the cash flow. As in Campbell and Shiller (1991), re-write Equation (7) as

$$r_{t+1} - E_t r_{t+1} = (E_{t+1} - E_t) \sum_{j=0}^{\infty} \varsigma^j \Delta f_{t+1+j} - (E_{t+1} - E_t) \sum_{j=1}^{\infty} \varsigma^j r_{t+1+j}. \quad (11)$$
and define shocks as

\[ \upsilon_{r,t+1} = \eta_{f,t+1} - \eta_{r,t+1}. \]  \hspace{1cm} (12)

\( \upsilon_{r,t+1} \) denotes the left-hand side of (11) and \( \eta_{f,t+1} \) and \( \eta_{r,t+1} \) are the expectation revisions about future cash-flow growth and returns, respectively. A textbook derivation of (11) and (12) from (7) can also be found in Campbell et al. (1997).

Now we separately solve for \( \upsilon_{r,t+1} \) and \( \eta_{r,t+1} \) using the process \( z_t \) following from the Johansen model and the corresponding impulse responses \( \Phi_i \). The solution for the unexpected shocks to returns is

\[ \upsilon_{r,t+1} = e^{1'} \epsilon_{t+1}, \]  \hspace{1cm} (13)

which is simply the innovation of the return process. For the revision in return expectations we find

\[ \eta_{r,t+1} = e^{1'} \sum_{j=1}^{\infty} \varsigma^2 \Phi_j \epsilon_{t+1}. \]  \hspace{1cm} (14)

Since \( |\varsigma| < 1 \) by construction, the infinite sum (14) converges as long as \( \Phi_i \) is bounded. The latter condition is satisfied, since the infinite moving-average representation of \( z_t \) has a solution, given by (8). Thus, (14) is well defined.

From the accounting identity \( \eta_{f,t+1} = \upsilon_{r,t+1} + \eta_{r,t+1} \), we further find that

\[ \eta_{f,t+1} = e^{1'} \sum_{j=0}^{\infty} \varsigma^2 \Phi_j \epsilon_{t+1} \]  \hspace{1cm} (15)

for the revision in cash-flow growth expectations. Appendix C contains a detailed derivation of the return decomposition (Equations (C1)-(C3)) and the corresponding variances and covariance, \( \text{Var}(\eta_r) \), \( \text{Var}(\eta_f) \), and \( \text{Cov}(\eta_r, \eta_f) \) (Equations (C4)-(C7)).

A further measure in this context that can easily be computed from the co-fractional VAR model is the term structure of risk, due to Campbell and Viceira (2005). We define the term structure, \( \sigma^{(k)} \), as the annualized standard deviation of future returns. More precisely, \( \sigma^{(k)} = \sqrt{\frac{1}{k} \text{Var}(\sum_{j=1}^{k} r_{t+j})} \). The conditional

\footnotetext{Note that the infinite moving-average representation of the process \( z_t \) from (9) is only needed for the derivation of future returns and their conditional expectations. It was derived under the assumption that \( z_s = 0 \) if \( s \leq 0 \). Yet, in this context the assumption is inconsequential and can be relaxed, because the contribution of the initial condition to the elements of the return decomposition is zero. This becomes obvious from the fact that \( \mu_t = E_0 z_t \), as outlined by Johansen (2008).}
variance is given by

\[ \text{Var}_t \left( \sum_{j=1}^{k} r_{t+j} \right) = \mathbb{E} \left( \sum_{j=1}^{k} e_1' z_{t+j} - \mathbb{E}_t e_1' z_{t+j} \right)^2 \]

\[ = \sum_{j=1}^{k} e_1' \left( \sum_{i=0}^{j-1} \Phi_i \right) \Sigma \left( \sum_{i=0}^{j-1} \Phi_i \right)' e_1. \]  

(16)

5 Relation to traditional models

The financial literature typically assumes that returns can be forecast by

\[ r_{t+1} = \alpha + \beta' x_t + u_t, \]  

(17)

where \( x_t \) is a vector of predictor variables. We call this the standard prediction model. Examples from the stock-market asset pricing literature include, among many others, Cochrane (1999, 2008).

If \( r_{t+1} \) and \( x_t \) are both nonstationary, Equation (17) is only a well-defined prediction model if the two cointegrate. In addition, it is customary to specify a process for the dynamics of the predictor variables. We can then rewrite the model in its triangular form (suppressing the constant), assuming that the prediction vector is univariate with \( x_t = (f - p)_t \), as

\[ r_t = \beta (f - p)_{t-1} + \nu_{t:t} \]  

(18)

\[ \Delta^d (f - p)_t = \nu_{2:t}, \]  

(19)

where \( \nu_t \) is assumed an \( I(0) \) process. \( \nu_{1:t} \) is integrated of order zero since the return and the rent to price ratio strongly cointegrate and their common trend is integrated of order zero. \( \nu_{2:t} \sim I(0) \), because the valuation ratio is integrated of order \( d \).

In (18) returns are assumed to be completely determined by the lagged predictor variable. To understand how specification (18) relates to the process for returns implied by the Johansen model, consider the CFVAR\(_d(p)\) in its simplest form without lagged dynamics, i.e. set \( p = 0 \). We can write the first element of \( z_t \) as

\[ r_t = (1 + \gamma_1) L_d r_t - \gamma_1 \delta L_d (f - p)_t + \epsilon_{1:t}. \]  

(20)

The triangular model, unlike the CFVAR, does not allow for the possibility that past
returns have predictive power for future returns. Imposing this restriction on (20) results in setting \( \gamma_1 = -1 \),

\[
    r_t = \tilde{\delta} L_d (f - p)_t + \epsilon_{1t}
    = -\tilde{\delta} \sum_{i=1}^{\infty} \theta_i (f - p)_{t-i} + \epsilon_{1t}.
\]  

(21)

Comparing (21) to the triangular model, we conclude that even under the restriction that past returns do not influence current and future returns, the co-fractional VAR is still a richer model in the sense that it assumes returns to be driven by the entire history of the predictor variable and not just the lagged level. The two models only coincide if \( d = 1 \). This unit root hypothesis was firmly rejected in Section 3.

We can also compare the two models in terms of their specification for the time series process of the predictor variable, \((f - p)_t\). For the rent to price ratio, the simplified CFVAR without lags implies that

\[
    \Delta^d (f - p)_t = \gamma_2 L_d r_t - \gamma_2 \tilde{\delta} L_d (f - p)_t + \epsilon_{2t}.
\]  

(22)

Again, the triangular model imposes an implicit restriction as compared to the Johansen model, as returns do not have a feedback effect on the predictor variable. We can restrict the co-fractional model similarly, by setting \( \gamma_2 = 0 \). Equation (22) becomes

\[
    \Delta^d (f - p)_t = \epsilon_{2t},
\]  

(23)

which is the same as in the triangular model (19). The model for the rent to price ratio is thus the same in both specifications, the Johansen approach and the triangular model, if the restriction \( \gamma_2 = 0 \) holds. Whereas the implied return process differs substantially between the two models, the model for the rent to price ratio is not very different.

Analogously to the CFVAR\(_d(p)\), we can decompose the unexpected return process that follows from the triangular model and compute its term structure of risk. Appendix D, Equations (D1)-(D11), provides the full decomposition when the rent to price ratio is assumed to follow an ARFIMA(1,d,0) process. The variance of unexpected returns, \( \nu_{r,t+1} \), resulting from the triangular model is \( \sigma^2_r = \text{Var}(\nu_{rt}) \), that is simply the variance of the error of the return prediction. If there is long-run risk in expected returns, this variance is expected to be larger than the corresponding
variance $\epsilon_1' \Sigma \epsilon_1$ from the co-fractional model, due to the dynamics of the predictors that are captured in the error term $\nu_t$. We further observe that since the triangular model does not allow for past returns to influence current and future returns, the innovations in the present value of future returns, $\eta_{r,t+1}$, are solely determined by the innovations to the rent to price ratio, $\nu_{2t}$. The same measure resulting from the Johansen model depends on both innovations, $\epsilon_{1t}$ and $\epsilon_{2t}$, by contrast.

We deduce from the preceding discussion that the triangular model is restricted in many ways. While the Johansen model imposes no restriction on the speed of adjustment vector $\gamma$, the triangular model implicitly assumes that it is known a priori, $\gamma = [-1 \ 0]$. In addition, the triangular model only considers the lagged level of the valuation ratio for return prediction, instead of the entire history of the process.

6 Estimation of the co-fractional VAR and the triangular model

Conditional on $d$, we estimate the $\text{CFVAR}_{0.75}(p)$ model, as specified in Equation (8), for the bivariate process of logarithmic real returns and the rent to price ratio, $\Delta^d z_t$.

The estimation proceeds in two stages. As a first step, we use the estimated fractional differencing parameter $\hat{d} = 0.75$ to construct the vector time series $\Delta^d z_t$, the filtered return and rent to price ratio. In the second stage, we obtain parameter estimates for the co-fractional model by iterated seemingly unrelated regression (SUR). As the CFVAR process is bivariate, there is only one cointegrating vector, $\delta$. Furthermore, we estimate the model with an intercept $\alpha$. For the selection of the lag length $p$, we employ the Bayesian-Schwarz information criterion (BIC).

We apply the fractional filters to the entire data series, starting in 1650, but the estimation of the coefficients only starts in 1660. Doing so, we correct for possibly inaccuracies at the initial observations that arise due to the application of the truncated fractional filter.

We obtain confidence intervals for the estimated coefficients by bootstrapping the residuals of the CFVAR. As we cannot exclude the presence of heteroskedasticity in the residuals, we apply the wild bootstrap. That is, we draw univariate series $\xi_t$ from a standard normal distribution and construct bootstrapped residuals by

$$
\epsilon_t^* = \xi_t \hat{\epsilon}_t, \quad (24)
$$
where $\hat{\epsilon}_t$ are the CFVAR residuals. Davidson and MacKinnon (2004) point out that the wild bootstrap is often the preferable method in the presence of a unknown non-constant error variance. We employ $\epsilon^*_t$ to re-build the process $z^*_t$ recursively, imposing fractional cointegration with $d = 0.75$.

We consider two different types of confidence intervals, an unconditional and a conditional on $d = 0.75$. To compute the first one, we estimate the fractional differencing parameter of the bootstrapped rent to price ratio, $d^*$, using the EW and setting $m = T^{0.6}$. From this, we can construct the new series $\Delta^d z^*_t$. For the conditional bootstrap confidence intervals we use the series $\Delta^d z^*_t$, that is we condition on $d = 0.75$.

Define $g$ as the vector of the model parameters of length $s$. The parameters $\hat{g}^*$ of the co-fractional VAR model for both $\Delta^d z^*_t$ and $\Delta^d z^*_t$ are estimated by SUR. We repeat this process $h = 10,000$ times, which results in a matrix $\hat{G}^*$ of bootstrapped coefficients of size $(s \times h)$. For all parameter estimates, we also compute the implied values for the elements of the return decomposition, $P_r$, $\text{Var}(\eta_r)$, $\text{Var}(\eta_{cf})$, and $\text{Cov}(\eta_r, \eta_{cf})$.

For 95%-confidence intervals, we first re-center the bootstrap parameters estimates at their sample estimates $\hat{g}$ and implied moments such that they have the same mean as the original estimates. For variance parameters the re-centering is done on their logarithms of the bootstrap estimates. We calculate the centered quantile intervals to avoid inaccurate confidence intervals due to asymmetric distributions of the parameter estimates, $\hat{g}$, following a suggestion of Hansen (2006).

For the triangular model, the estimation of the parameters of the return prediction, as well as the time-series process for the rent to price ratio, is conducted by ordinary least squares (OLS), including an intercept. Analogously to the co-fractional model, we start the estimation only in the year 1660. To include short-run dynamics of the rent to price ratio, we estimate an ARFIMA$(p, 0.75, q)$ on the series. We select $p$ and $q$ using the BIC.

As for the co-fractional VAR model, we compute bootstrapped confidence intervals for both equations of the triangular model. OLS standard errors, the corresponding confidence intervals, and t-statistics are known to be misleading due to the persistence of the predictor variable, as has been pointed out by Campbell and Yogo (2006). We use the wild bootstrap to obtain new residuals $\nu^*_1$ and $\nu^*_2$. We employ these residual series, the OLS estimated parameters, and the value $d = 0.75$ to construct the processes $r^*_t$ and $(f - p)^*_t$. Then we re-estimate and compute 95%-confidence
intervals in the same manner as for the CFVAR\(d(p)\) model.

The two models, the co-fractional VAR and the triangular model, are not nested. To test which one of the two models is preferable, we carry out a simulation-based version of the encompassing test due to Cox (1961, 1962), suggested by Coulibaly and Brorsen (1999). The test statistic that discriminates between two non-nested hypotheses \(H_0\) and \(H_1\) is defined as

\[
T = \frac{\sqrt{T}(\ell_{t_0} - E_0(\ell_{t_0}))}{\sqrt{\omega_0^2}},
\]

where \(\ell_{t_0} = \ell_0(\hat{g}_0) - \ell_1(\hat{g}_1)\) is the difference of the maximized log-likelihoods under \(H_0\) and \(H_1\). \(\hat{g}_0\) and \(\hat{g}_1\) are the estimated parameter values of the null and the alternative model. \(E_0(\ell_{t_0})\) is the expected value of \(\ell_{t_0}\) under the null hypothesis and \(\omega_0^2\) is the variance of \(\ell_{t_0} - E_0(\ell_{t_0})\) under \(H_0\).

To find the expected value of the log-likelihood, we use a bootstrap approach. We generate \(h = 200\) resampled residuals of the \(H_0\)-model using the wild bootstrap and build a new return process \(r^*_t\), using the estimated values \(\hat{g}_0\) and \(d = 0.75\). In each repetition of the bootstrap, we estimate the parameters of \(r^*_t\) under \(H_0\) and \(H_1\). The expected value of the log-likelihood follows from

\[
\hat{E}_0(\ell_{t_0}) = \frac{1}{h} \sum_{i=1}^{h} (\ell^*_{0,i}(\hat{g}_{0,i}) - \ell^*_{1,i}(\hat{g}_{1,i})).
\]

The computation of \(\omega_0^2\) follows Pesaran and Pesaran (1995), given by

\[
\hat{\omega}_0^2 = \frac{\sum_{t=1}^{T} (\ell^*_{01,t} - \tilde{\ell}^*_{01})^2}{T - 1},
\]

where \(\ell^*_{01,t} = \ell_{0,t}(\hat{g}_0) - \ell_{1,t}(\hat{g}_1)\), the log-likelihood ratio for observation \(t\), and \(\tilde{\ell}^*_{01}\) is its mean. The simulation estimator \(\hat{g}^*_1\) is defined as \(\frac{1}{h} \sum_{i=1}^{h} \hat{g}_{1,i}\). That is, as the average bootstrap parameter estimates of \(r^*_t\) under the model \(H_1\), where \(r^*_t\) has been generated under model \(H_0\). We obtain critical values for \(T\) using the simulations suggested by Coulibaly and Brorsen (1999).

Finally, to calculate the elements of the return decomposition, we need the coefficient of linearization, \(\varsigma\). It defined to be \(1/(1 + e^{(T-d)})\) by Campbell and Shiller (1988a,b). In our data set we find that \(\varsigma = 0.9462\). Yet, this specification is derived under the assumption that the rent to price ratio is a stationary process. Calculating
the linearization error \( r_t - \iota - (1 - \varsigma)(f - p)_t - \Delta p_t \) and the resulting \( R^2 \), we can show, however, that the approximation is extremely exact, capturing 99.95% of the variation in returns\(^3\).

7 Results

We start the discussion of our results by focusing on the question of whether there is predictability in returns. Estimating the co-fractional model for \( d = 0.75 \), we find substantial evidence that real-estate returns are predictable. Our model captures almost 60% of the variability of \( \Delta^d r_t \) and 20% of the logarithmic return series in levels. Case and Shiller (1990) find an \( R^2 \) of 0.109 in annual predictions of excess returns on housing with the rent to price ratio and Plazzi et al. (2010) explain respectively 10.3%, 11.6%, and 12.6% of the variation in one-year excess returns on apartment, industrial and retail properties in their fixed effects model.

Parameter estimates of the co-fractional VAR model are in Table 6. As the BIC is minimized at \( p = 1 \), we estimate a CFVAR\(_{0.75}(1)\) specification. The cointegrating parameter is equal to \( \hat{\delta} = -0.06 \). The value is statistically significant, no matter whether we consider the conditional or the more conservative unconditional confidence interval. Returns and the rent to price ratio thus have a small but significant long-run relation. The adjustment vector estimate is \( \hat{\gamma} = [-1.08, 0.39]' \). On the basis of both types of confidence intervals, we cannot reject the hypothesis that \( \gamma_1 = -1 \) and \( \gamma_2 = 0 \). This satisfies the restrictions that are imposed in the triangular return prediction model. Yet, the two models for return prediction still differ substantially, as long as \( d \neq 1 \).

Solving the co-fractional model in Equation (8) for \( z_t \) (see Appendix B) and plugging in the estimated parameter values, we find that returns are predicted by

\[
\begin{align*}
    r_t &= 0.27 + \sum_{i=1}^{\infty} \left( 0.09 \theta_i - 0.15 \sum_{l=0}^{i-1} \theta_{i-l} \theta_l \right) r_{t-i} \\
    &\quad + \sum_{i=1}^{\infty} \left( -0.07 \theta_i - 0.63 \sum_{l=0}^{i-1} \theta_{i-l} \theta_l \right) (f - p)_{t-i} + \epsilon_{1t},
\end{align*}
\]

where \( \theta_i = (-1)^i(0.75)^i \) are the coefficients of the fractional filter, as explained in Section 3. Equation (28) shows that returns are predicted by the entire history of the

\(^3\)We set \( \iota = -\ln(\varsigma) - (1 - \varsigma) \ln((1/\varsigma) - 1) = 0.2095 \), following Campbell and Shiller (1988a,b)
Table 6: SUR estimates for the co-fractional model, CFVAR_{0.75}(1)

The table reports estimates of CFVAR_{0.75}(1)

\[ \Delta^d z_t = \alpha + \gamma d' (1 - \Delta^d) z_t + \Gamma_1 \Delta^d (1 - \Delta^d) z_t + \epsilon_t. \]

Entries denote SUR parameter estimates and bootstrapped 95%-confidence intervals, where CI is unconditional and CI(d) is conditional on \(d = 0.75\). The values given for \(\text{Var}(\eta_r), \text{Var}(\eta_f), \text{and} -2\text{Cov}(\eta_r, \eta_f)\) are relative to the variance of the unexpected returns, \(\text{Var}(\nu_r)\).

<table>
<thead>
<tr>
<th>Est.</th>
<th>(\alpha)</th>
<th>(\gamma)</th>
<th>(\delta)</th>
<th>(\Gamma_1)</th>
<th>(\Sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.2647, -0.1138)</td>
<td>(-1.0847, 0.3902)</td>
<td>(1.0000, -0.0596)</td>
<td>(0.1449, 0.6291)</td>
<td>(0.0324, -0.0281)</td>
</tr>
<tr>
<td>CI</td>
<td>(0.1011, 0.3461)</td>
<td>(-1.5600, -0.3827)</td>
<td>(-1.4110, -0.0259)</td>
<td>(-0.4699, 0.7173)</td>
<td>(-0.9942, 0.2236)</td>
</tr>
<tr>
<td>CI(d)</td>
<td>(0.16215, 0.3414)</td>
<td>(-1.1540, 0.0259)</td>
<td>(0.0221, 0.7444)</td>
<td>(0.2974, 0.9670)</td>
<td>(0.3496, 0.3450)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Est.</th>
<th>(\text{Var}(\eta_r))</th>
<th>(\text{Var}(\eta_f))</th>
<th>(-2\text{Cov}(\eta_r, \eta_f))</th>
<th>(P_r)</th>
<th>(R^2_r)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.7896</td>
<td>0.3881</td>
<td>-0.1777</td>
<td>1.8736</td>
<td>0.1999</td>
</tr>
<tr>
<td>CI</td>
<td>[0.4802, 1.3736]</td>
<td>[0.2163, 0.7674]</td>
<td>[-0.5942, 0.5245]</td>
<td>[1.1600, 2.7081]</td>
<td></td>
</tr>
<tr>
<td>CI(d)</td>
<td>[0.5429, 1.2849]</td>
<td>[0.2206, 0.7091]</td>
<td>[-0.5429, 0.3766]</td>
<td>[1.2689, 2.6533]</td>
<td></td>
</tr>
</tbody>
</table>
processes $r_t$ and $(f-p)_t$. Plugging in the values for $\theta_i$, we find that all else equal the first few terms of the lagged rent to price ratio on returns are

$$(-0.07\theta_1 - 0.63\theta_0) (f - p)_{t-1} = 0.52(f - p)_{t-1}$$
$$(-0.07\theta_2 - 0.63 (\theta_2\theta_0 + \theta_1^2)) (f - p)_{t-2} = -0.29(f - p)_{t-2}$$
$$(-0.07\theta_3 - 0.63 (\theta_3\theta_0 + 2\theta_2\theta_1)) (f - p)_{t-3} = -0.06(f - p)_{t-3},$$

for the first, second, and third lag, respectively. The short-run effects of past returns on the current return are

$$(0.09\theta_1 - 0.15\theta_0) r_{t-1} = 0.05r_{t-1}$$
$$(0.09\theta_2 - 0.15 (\theta_2\theta_0 + \theta_1^2)) r_{t-2} = -0.08r_{t-2}$$
$$(0.09\theta_3 - 0.15 (\theta_3\theta_0 + 2\theta_2\theta_1)) r_{t-3} = -0.02r_{t-3},$$

showing that lagged returns hardly matter for return prediction.

Higher order effects have much smaller coefficients, but are collectively non-negligible. This long distributed lag is an important difference in comparison to the triangular model, which uses $(f-p)_{t-1}$ as the only predictor of returns. The complete picture of these dynamic effects is summarized by the impulse responses in Figure 2. The graph plots the first 50 impulse responses of returns with respect to shocks $\epsilon_t^{(r)}$ and $\epsilon_t^{(f-p)}$. The impulse responses are orthogonalized, with shocks to the rent to price ratio as the first element in the causal ordering. The size of shocks is scaled to one standard deviation. We can clearly see the effect of the long memory in the series. Whereas impulse responses of a stationary $I(0)$ process decay exponentially\(^4\), the fractional impulse responses decline slower. We observe that due to the long memory of both, the return and the rent to price ratio time series, impulse responses decay rather slowly and are still non-negligible after a lag of 50 years. Figure 2 also shows that as of the first lag, i.e. $i \geq 1$, shocks to the rent to price ratio have a larger (positive) effect than shocks to returns, on the return process.

Figure 3 shows the cumulative sum of the coefficients of the series $\eta_{r,t+1}$, which represent the shocks to the discount factor. More precisely, it plots the function $\sum_{j=1}^{i} \varsigma^j (e^1\Phi_j)$ at each $i = 1, \ldots, 199$, where $\Phi_j$ are the recursive impulse responses of $z_t$ defined in (10). Again, we observe that the convergence of both elements of the

\(^4\)if the process has a stationary and invertible ARMA representation

23
The figure plots the impulse responses of returns from the co-fractional model over 50 lags. The reported impulse responses are orthogonalized. A shock to the rent to price ratio occurs first. The size of shocks is scaled to one standard deviation. The sum is very slow. Between lag 50 and 100, both curves still slightly increase and only after 100 years they flatten out. The effect of the rent to price ratio on the revision of expectations about future discount rates is almost 4 times higher than the effect of returns. That is, $\eta_{r,t+1} \approx 0.31\epsilon_t^{(r)} + 1.16\epsilon_{t+1}^{(f-p)}$. Thus, a positive shock of 1% to the rent to price ratio in the remote past implies a positive shock of approximately 1.2% to the discount factor in the real-estate market. At the same time, it brings a 1.2% shock to expected rental growth with it. An increase in prices relative to rents in the past is associated with a capital loss today, due to a worse outlook on future cash-flow growth. This capital loss is, however, completely offset by the capital gain today that results from the increase in prices relative to rents. In contrast, a 1% shock to housing returns that happened long ago, induces a small shock of 0.3% on the discount factor, but a positive shock of 1.3% on the expected growth of rents. A rise in the market is warranted by improvements in expected future rents, but this capital gain is partly counterweighted by discounting capital streams at a higher rate. All else equal, an increase in the rent to price ratio implies an increase in rents and a small decrease in prices (see Equation (7)), whereas an increase in returns (with fixed $(f-p)_t$) involves prices and rents increasing by the same amount. It is therefore not
surprising that former effect entails substantially more discounting of future capital and slightly less expected future rental growth, than the latter shocks.

![Cumulative sum of coefficients](image)

**Figure 3:** Cumulative sum of coefficients $\sum_{j=1}^{i} \varsigma^j (e1'\Phi_j)$ from a CFVAR$_{0.75}(1)$.

- The figure plots the cumulative impulse responses of returns from the co-fractional model. More precisely, it plots the sum $\sum_{j=1}^{i} \varsigma^j (e1'\Phi_j)$ for each $i = 1, 2, \ldots, 199$.

For our further estimations, we truncate the sum $\sum_{j=1}^{\infty} \varsigma^j (e1'\Phi_j)$ at $j = 199$.\(^5\) Using the estimates of the co-fractional CFVAR$_{0.75}(1)$ model, we compute the variances and covariances of the elements of the return decomposition from Equation (13), (14), and (15). The fraction of the variance of unexpected return shocks that is due to future expected discount rates, $\text{Var}(\eta_{r,t+1})/\text{Var}(\upsilon_{r,t+1})$, is equal to 79%. Both 95% confidence intervals, the conditional and the unconditional one, support that this ratio is statistically different from zero, but not different from 100%. In contrast, the component of unexpected return variation that is captured by the variability of the innovations in the expected present value of future cash flows, $\text{Var}(\eta_{f,t+1})/\text{Var}(\upsilon_{r,t+1})$, is only 39% and statistically significant. Hence, the explanatory power of future discount rate news for the variation in unexpected returns, is twice as high as the one of

\[^5\]Since we know that the infinite sum converges, as shown in Section 6, and according to the observations from Figure 3, where the curves are flat at lag 199, this simplification is inconsequential for the precision of the results. In addition, we can compute the maximal truncation error that we can make, by observing that the following must hold: $\sum_{j=1}^{\infty} \varsigma^j e1'\Phi_j - \sum_{j=1}^{199} \varsigma^j e1'\Phi_j = \sum_{j=200}^{\infty} \varsigma^j e1'\Phi_j \leq \sum_{j=200}^{\infty} \varsigma^j e1'\Phi_{200} = \frac{\varsigma^{200}}{1-\varsigma} e1'\Phi_{200}$. This implies that the upper bound of the truncation error is given by $[1.65 \times 10^{-6} \quad 4.60 \times 10^{-6}]$, which is a negligibly small number.
future cash flows. These results are qualitatively in line with the findings of Campbell and Shiller (1991) and Campbell and Vuolteenaho (2004).

The housing market can typically be viewed as a capital market for conservative longer term investors. For such an investor, a negative shock to the discount factor implies a lower return now, but higher returns in the future. Campbell and Vuolteenaho (2004) therefore call these types of news transitory shocks to wealth. Negative news about future rental growth represent permanent shocks to wealth, however. A long-term investor will thus expect higher returns to compensate losses in the latter case; that is, he demands a higher risk premium. This in turn implies that the part of returns that is unexpected is less influenced by such permanent shocks, which is what we find.

Scaling the covariance between \( \eta_{r,t+1} \) and \( \eta_{f,t+1} \) by \( \text{Var}(\nu_{r,t+1}) \) as well, we find that \(-2\text{Cov}(\eta_{r,t+1}, \eta_{f,t+1}) = -0.18\). We cannot reject the hypothesis that the value is equal to zero, however. Hence, our methodology to identify discount-factor news and news about rental growth as two separate types of shocks is justified. The fact that the correlation between the two series is small but positive (\( \text{Corr}(\eta_r, \eta_f) = 0.16 \)) is a direct result of the finding that both time-series shocks in the system, \( \epsilon_t(r) \) and \( \epsilon_t(f_p) \), have a positive effect on discount-factor and cash-flow news.

To summarize, by means of the co-fractional VAR model of Johansen (2008), we can predict returns in the real estate market and decompose the innovations. Capturing 20% of the variation in real logarithmic returns, we find that the rent to price ratio has more predictive power than past returns. Furthermore, returns fluctuate mainly because of innovations to discount rates. Cash-flow news have a much smaller impact and the covariance effect is negligible.

### 7.1 Comparison with predictions from triangular model

To obtain a comparative measure for the performance of the CFVAR model in explaining and predicting returns on the real estate market, we continue to estimate returns by means of the triangular model.

For a given fractional integration parameter, \( d = 0.75 \), we estimate the return and the rent to price ratio equations by OLS. For the latter, the BIC is minimized at a ARFIMA(1,0.75,0) specification. The estimation result are in Table 7. The rent to price ratio predicts housing returns with a statistically significant coefficient of 0.22 and we find a positive significant intercept of 0.68. In total, the alternative model

26
captures 11.77% of the variance in the return series.

Table 7: OLS estimates for the triangular model for return predictions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>CI</th>
<th>CI(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.6801</td>
<td>[0.5855, 0.7031]</td>
<td>[0.5855, 0.7031]</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.2164</td>
<td>[0.1469, 0.2347]</td>
<td>[0.1469, 0.2347]</td>
</tr>
<tr>
<td>$c$</td>
<td>-0.0287</td>
<td>[-0.0535, 0.0692]</td>
<td>[-0.0480, -0.0096]</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.1893</td>
<td>[-0.5865, 0.0658]</td>
<td>[-0.3609, -0.0100]</td>
</tr>
<tr>
<td>$\sigma^2_{\nu_1}$</td>
<td>0.0357</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_{\nu_2}$</td>
<td>0.0321</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{\nu_1, \nu_2}$</td>
<td>-0.0274</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{Var}(\eta_r)$</td>
<td>2.1703</td>
<td>[0.4399, 7.5385]</td>
<td>[1.2085, 3.0231]</td>
</tr>
<tr>
<td>$\text{Var}(\eta_f)$</td>
<td>0.7760</td>
<td>[0.0911, 1.9174]</td>
<td>[0.3458, 1.2107]</td>
</tr>
<tr>
<td>$-2\text{Cov}(\eta_r, \eta_f)$</td>
<td>-1.9462</td>
<td>[-3.9009, 12.4515]</td>
<td>[-2.8096, 0.7416]</td>
</tr>
<tr>
<td>$R^2_r$</td>
<td>0.1177</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Analogously to the co-fractional VAR model, we can decompose unexpected returns into the expectation innovation in future discounted returns and changes in rents. The variance of unexpected returns $\text{Var}(\nu_{r,t+1})$ is approximately 0.04 and hence more than 10% larger than the corresponding figure for the co-fractional model. This result is a direct implication of the lower $R^2$. The triangular model explains less of the variation in returns than the co-fractional model and, hence, return innovations are more volatile.

The variance ratio of revised expectations about future discount rates scaled by unexpected returns is equal to 2.17. This value is much higher than what the co-fractional VAR model implied and it is statistically significant. Yet, the confidence interval of this estimate, especially the unconditional one, is extremely wide, indicating substantial uncertainty. The variance of innovations is future cash flow changes
divided by the variance of return shocks is 0.78. These results point into the same
direction as the analogous findings from the co-fractional model, yet values are of sub-
stantially larger magnitude with more statistical uncertainty. The reason for these
high resulting ratio values lies in the high correlation between both innovations, $\eta_{r,t+1}$
and $\eta_{f,t+1}$; the triangular model suggests that the negative double covariance scaled
by the variance of unexpected returns is equal to -1.95. We would have to multiply
the analogous covariance from the co-fractional VAR model by approximately 10 to
obtain the same result. However, the strong positive correlation questions the valid-
ity of the decomposition of unexpected returns into discount-factor and rental-growth
news in this case.

At this stage, we have compared the co-fractional VAR to the standard prediction
model and pointed out the differences. So far we have not been able to judge which
model performance is better in the statistical sense; that is which model attains a
higher goodness of fit. The reported $R^2$ values have to be interpreted with care.
Since returns in our models are not covariance stationary, it holds that asymptoti-
cally $R^2 \rightarrow 1$. To have a more robust instrument for statistical inference, we compare
the likelihood values of both models and, to correct for the different number of pa-
rameters, information criteria. The log likelihood value of the triangular prediction
system is 378.49, whereas the corresponding value from the Johansen model is 468.27.
Furthermore, the former model estimation results in a AIC of -2.16 and BIC of -2.12.
The respective values for the co-fractional model are -2.65 and -2.55. Hence, the
likelihood is maximized for the co-fractional model and the information criteria are
minimized. This gives us a good indication for the superior fit of the latter model.

What drives the outperformance of the Johansen model over the triangular model?
Can we predict returns better, or do we have a better model to explain the time-
series process of the rent to price ratio? As we explain in Section 5, we expect
a more accurate return prediction, since the co-fractional VAR allows for returns
to be determined by the entire history of the predictor process and not simply by
its lag. The estimation results confirm this intuition. The log likelihood value of
the return equation implicit in the CFVAR$_{0.75}(1)$ framework equals 103.19 (AIC=-
0.57 and BIC=-0.51). In contrast, the corresponding log likelihood value from the
triangular model is 86.22 (AIC=-0.49 and BIC=-0.46), and thus much lower. Hence,
the fractional filters seems to play an important role in return predictions.

On the other hand, the movements of the rent to price ratio are captured equally
well by both models. The log likelihood values are 107.78 and 107.00 for the co-
fractional and the triangular model, respectively. Since we estimate substantially more parameters in the second equation of the Johansen model than in the predictor specification in the triangular model, AIC and BIC are even minimized for the latter. More precisely, the AIC is equal to -0.61 (BIC=-0.58) in the rent to price ratio equation of the triangular model, whereas the corresponding values for the co-fractional model are AIC=-0.59 and BIC=-0.54. The superior fit of the Johansen model is, hence, entirely due to the improved return prediction.

An additional interesting aspect for comparison arises, when contrasting the implications of both models with respect to the log-linearization of returns. If the triangular model as well as the co-fractional VAR model were perfect predictors of asset returns, this would imply that the second element of the cointegrating vector, $\tilde{\delta}$, is equal to $-(1 - \varsigma)$, since for both models $\delta'z_t \sim I(0)$. For the triangular model we further know that $\tilde{\delta} = -\beta$. If we conduct simple t-tests on this hypothesis, that is $\hat{\beta} = (1 - \varsigma) = 0.05$ for the triangular model, we firmly reject. Yet, for the Johansen model, we fail to reject the hypothesis. Again this indicates that the Johansen framework provides a better means to model housing returns, as compared to the traditional approach.

7.2 Encompassing test of return predictions

The previous section suggested that the co-fractional model predicts returns more accurately than the triangular model. Can we, thus, discard the triangular model in favor of Johansen’s model? A formal test of this hypothesis is not straightforward, as the two models are not nested. Yet, we can test whether the CFVAR model encompasses the triangular one, as outline in Section 6. If we can account for all estimation outcomes of the latter by assuming that the former is the true model, the CFVAR model encompasses the triangular.

Our results suggest that the co-fractional VAR model, indeed, encompasses the triangular model for return prediction. If we test the hypothesis that

$$H_0 : \Delta^d r_t = e 1' \left( \alpha + \gamma \delta' (1 - \Delta^d) z_t + \sum_{i=1}^{p} \Gamma_i \Delta^d (1 - \Delta^d)^i z_t + \epsilon_t \right)$$

$$H_1 : r_t = \alpha + \beta (f - p)_t + \nu_{1t},$$

we find that the p-value is equal to 0.124. Hence, we cannot reject the $H_0$-hypothesis.
Alternatively, if we reverse the order of the hypotheses, such that the co-fractional model is the alternative, the resulting p-value is equal to 0.001. This implies that the standard return prediction model does not encompass the co-fractional VAR model. The latter model is clearly preferred.

### 7.3 Robustness of the co-fractional VAR

All previously reported estimates are based on the assumption that the true long memory parameter of the return and the rent to price series is given by \( d = 0.75 \). In Section 2 we estimate the fractional integration parameter and conclude that an exact determination of its value is difficult, since the estimates depend on the selection of the bandwidth of the spectral window under consideration. In this section, we therefore conduct a sensitivity analysis of the outcomes of the return prediction on the estimate of \( d \).

From the results in Section 2 it follows that the rent to price ratio and the return series are fractionally integrated, with \( \hat{d} \in [0.65, 0.8] \). In the robustness analysis, we therefore consider values \( d = \{0.65, 0.7, 0.75, 0.8\} \). For a general comparison, we further allow for \( d = 0.5 \) and \( d = 1 \). Note that the latter specification is just the typical vector error correction model (VECM) for \( z_t \). The BIC criterion is again utilized for selection of the lag length \( p \). If \( d = 0.5 \), we specify a CFVAR\(_{0.5}\)(2) model and for all remaining values of \( d \), we select a specification with only one lag.

The results are summarized in Table 8. Overall, we find that the results of the estimation are very robust to the fractional differencing parameter, \( d \). Within the estimated range of values for \( \hat{d} \in [0.65, 0.8] \), the fit of the model for predicting returns, measured in terms of BIC\(_r\), only varies marginally between -0.48 and -0.51\%. The value declines in \( d \) until \( d = 0.75 \) and increases beyond that point. The goodness of fit of the VECM is substantially lower than all other models, BIC\(_r\) = -0.48\%. This is not surprising, as the CFVAR\(_1\)(1) implicitly assumes that there is no cointegration between rents and house prices and, as a consequence, violates the present value relation. Hence, this model should be less accurate according to theory, which is supported by our statistical evidence.

---

\(^6\)Note that this is not the case for the unconditional version of the bootstrapped confidence intervals. Since we re-estimate \( d \) in every bootstrap repetition, the intervals are not conditional on the fractional differencing parameter.
The table reports estimates of different model specifications of the co-fractional VAR model for returns and the rent to price ratio, given by

$$\Delta^d z_t = \alpha + \gamma \delta^i \left(1 - \Delta^d\right) z_t + \sum_{i=1}^{p} \Gamma_i \Delta^d \left(1 - \Delta^d\right)^i z_t + \epsilon_t,$$

where $\gamma = [\gamma_1 \gamma_2]'$ and $\delta = [1 - \tilde{\delta}]'$. Entries denote SURE parameter estimates and and bootstrapped 95% confidence intervals, where the first row contains the unconditional and the second row the conditional on $d = 0.75$ interval. On the basis of the BIC criterion, we select $p = 2$ for the model specification if $d = 0.5$. If $d \geq 0.65$, the BIC supports a CFVAR$_d(1)$. The estimates $\text{Var}(\eta_r)$, $\text{Var}(\eta_f)$, and $-2\text{Cov}(\eta_r, \eta_f)$ are scaled by $\text{Var}(\epsilon_{r,t+1})$.

<table>
<thead>
<tr>
<th>Specification</th>
<th>$\alpha_1$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$-\tilde{\delta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR$_{0.5,0.5}(2)$</td>
<td>0.3617</td>
<td>-1.7695</td>
<td>0.9588</td>
<td>-0.0491</td>
</tr>
<tr>
<td></td>
<td>[0.1983, 0.5598]</td>
<td>[-2.7995, -0.4872]</td>
<td>[0.2940, 2.0158]</td>
<td>[-0.0709, -0.0142]</td>
</tr>
<tr>
<td></td>
<td>[0.1546, 0.4864]</td>
<td>[-2.5669, -0.6932]</td>
<td>[0.0187, 1.9350]</td>
<td>[-0.0709, -0.0177]</td>
</tr>
<tr>
<td>VAR$_{0.65,0.65}(1)$</td>
<td>0.3075</td>
<td>-1.2780</td>
<td>0.5997</td>
<td>-0.0561</td>
</tr>
<tr>
<td></td>
<td>[0.1239, 0.4077]</td>
<td>[-1.9226, -0.4029]</td>
<td>[-0.0235, 1.3997]</td>
<td>[-0.0809, -0.0183]</td>
</tr>
<tr>
<td></td>
<td>[0.1861, 0.3957]</td>
<td>[-1.6843, -0.7708]</td>
<td>[0.1515, 1.6557]</td>
<td>[-0.0753, -0.0327]</td>
</tr>
<tr>
<td>VAR$_{0.7,0.7}(1)$</td>
<td>0.2824</td>
<td>-1.1660</td>
<td>0.4869</td>
<td>-0.0576</td>
</tr>
<tr>
<td></td>
<td>[0.1065, 0.3719]</td>
<td>[-1.7258, -0.3773]</td>
<td>[-0.0933, 1.1891]</td>
<td>[-0.0818, -0.0167]</td>
</tr>
<tr>
<td></td>
<td>[0.1709, 0.3641]</td>
<td>[-1.5252, -0.7129]</td>
<td>[0.0832, 0.8856]</td>
<td>[-0.0784, -0.0333]</td>
</tr>
<tr>
<td>VAR$_{0.75,0.75}(1)$</td>
<td>0.2647</td>
<td>-1.0847</td>
<td>0.3902</td>
<td>-0.0596</td>
</tr>
<tr>
<td></td>
<td>[0.1011, 0.3461]</td>
<td>[-1.5600, -0.3827]</td>
<td>[-0.1434, 1.0040]</td>
<td>[-0.0860, -0.0177]</td>
</tr>
<tr>
<td></td>
<td>[0.1621, 0.3414]</td>
<td>[-1.4110, -0.6820]</td>
<td>[0.0221, 0.7444]</td>
<td>[-0.0826, -0.0333]</td>
</tr>
<tr>
<td>VAR$_{0.8,0.8}(1)$</td>
<td>0.2518</td>
<td>-1.0351</td>
<td>0.3207</td>
<td>-0.0607</td>
</tr>
<tr>
<td></td>
<td>[0.1029, 0.3298]</td>
<td>[-1.4515, -0.4535]</td>
<td>[-0.1559, 0.8415]</td>
<td>[-0.0918, -0.0193]</td>
</tr>
<tr>
<td></td>
<td>[0.1568, 0.3249]</td>
<td>[-1.3312, -0.6743]</td>
<td>[-0.0148, 0.6444]</td>
<td>[-0.0865, -0.0325]</td>
</tr>
<tr>
<td>VAR$_{1,1}(1)$</td>
<td>0.2165</td>
<td>-0.1219</td>
<td>0.2728</td>
<td>-0.0529</td>
</tr>
<tr>
<td></td>
<td>[0.1377, 0.2801]</td>
<td>[-1.2853, -0.7109]</td>
<td>[-0.0464, 0.5394]</td>
<td>[-0.0730, -0.0266]</td>
</tr>
<tr>
<td></td>
<td>[0.1524, 0.2694]</td>
<td>[-1.2413, -0.7650]</td>
<td>[0.0281, 0.5112]</td>
<td>[-0.0721, -0.0320]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Specification</th>
<th>$\text{Var}(\eta_r)$</th>
<th>$\text{Var}(\eta_f)$</th>
<th>$-2\text{Cov}(\eta_r, \eta_f)$</th>
<th>BIC$_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR$_{0.5,0.5}(2)$</td>
<td>0.8468</td>
<td>0.3260</td>
<td>-0.1728</td>
<td>-0.4900</td>
</tr>
<tr>
<td></td>
<td>[0.4897, 1.6931]</td>
<td>[0.1619, 0.6903]</td>
<td>[-0.6272, 0.6724]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.5888, 1.4460]</td>
<td>[0.1709, 0.7636]</td>
<td>[-0.6018, 0.3616]</td>
<td></td>
</tr>
<tr>
<td>VAR$_{0.65,0.65}(1)$</td>
<td>0.8233</td>
<td>0.3777</td>
<td>-0.2009</td>
<td>-0.5023</td>
</tr>
<tr>
<td></td>
<td>[0.5425, 1.4698]</td>
<td>[0.2181, 0.8244]</td>
<td>[-0.6514, 0.3796]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.5693, 1.3023]</td>
<td>[0.2036, 0.7332]</td>
<td>[-0.5937, 0.3901]</td>
<td></td>
</tr>
</tbody>
</table>
The cointegrating variable, $-\hat{\delta}$, increases only marginally in absolute value as $d$ increases from 0.5 to 0.8, from -0.05 to -0.06. Only for the unit root case, it is again smaller, equal to -0.05. In all instance, we find that the coefficient is statistically different from zero, given both, the unconditional and the conditional versions of the 95%-confidence intervals. The magnitude of the effect of returns on the rent to price ratio, $\gamma_2$, changes a little bit more with $d$, yet with the exception of $d = 0.5$, we can never reject that the parameter is equal to zero, on the basis of the unconditional confidence intervals.

The measures that result from the decomposition of unexpected asset returns also show little sensitivity to the model specification. The ratio of the variances of innovations in future discount rates over unexpected returns marginally fluctuates within the interval $[0.79, 0.86]$. The fraction decreases as $d$ increases up to 0.8 and it increases beyond that value; it is minimized at $d = 0.8$ and maximized at $d = 1$. Furthermore, the ratio is statistically different from zero for all $d$. In addition, the variation with $d$ of the measures $\text{Var}(\eta_f)/\text{Var}(\nu_r)$ and $-2\text{Cov}(\eta_f, \eta_r)/\text{Var}(\nu_r)$ is small, lying within $[0.33, 0.39]$ and $[-0.22, -0.17]$, respectively. The former is statistically significant in all instances and the latter is never distinguishable from zero.

The speed of adjust parameter of the return-prediction equation, $\gamma_1$, of the return prediction changes slightly. Its estimate decreases in absolute value as $d$ becomes larger, from -1.77 to -1.02. It is always different from zero and never different from

<table>
<thead>
<tr>
<th>$\text{VAR}_{0.7,0.7}(1)$</th>
<th>0.8018</th>
<th>0.3886</th>
<th>-0.1904</th>
<th>-0.5075</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[0.5117,1.4113]</td>
<td>[0.2209,0.8076]</td>
<td>[-0.6260,0.4458]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.5550,1.2769]</td>
<td>[0.2154,0.7300]</td>
<td>[-0.5669,0.3775]</td>
<td></td>
</tr>
<tr>
<td>$\text{VAR}_{0.75,0.75}(1)$</td>
<td>0.7896</td>
<td>0.3881</td>
<td>-0.1777</td>
<td>-0.5105</td>
</tr>
<tr>
<td></td>
<td>[0.4802,1.3736]</td>
<td>[0.2163,0.7674]</td>
<td>[-0.5942,0.5245]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.5429,1.2849]</td>
<td>[0.2206,0.7091]</td>
<td>[-0.5429,0.3766]</td>
<td></td>
</tr>
<tr>
<td>$\text{VAR}_{0.8,0.8}(1)$</td>
<td>0.7877</td>
<td>0.3803</td>
<td>-0.1679</td>
<td>-0.5101</td>
</tr>
<tr>
<td></td>
<td>[0.4540,1.3941]</td>
<td>[0.2050,0.7170]</td>
<td>[-0.5710,0.6048]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.5258,1.3497]</td>
<td>[0.2191,0.6796]</td>
<td>[-0.5327,0.3944]</td>
<td></td>
</tr>
<tr>
<td>$\text{VAR}_{1.1}(1)$</td>
<td>0.8641</td>
<td>0.3569</td>
<td>-0.2210</td>
<td>-0.4809</td>
</tr>
<tr>
<td></td>
<td>[0.3512,1.5259]</td>
<td>[0.1447,0.5907]</td>
<td>[-0.6145,1.5271]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.5464,1.5748]</td>
<td>[0.2024,0.6048]</td>
<td>[-0.6224,0.4567]</td>
<td></td>
</tr>
</tbody>
</table>
minus one, in a statistical sense. The same is true for $-\gamma \tilde{\delta}$; the estimates depreciates from 0.09 to 0.05. As the derivation of the representation of the return process and its impulse responses in the Appendix shows, this reduction of the parameter estimates should, however, be offset by the increase in the coefficients of the fractional filter, $\theta_i$. Figure 4, which plots the impulse responses of returns with respect to shocks in the return (4(a)) and the rent to price (4(b)) processes, confirm this intuition. Whereas there are some discrepancies in the initial impulse responses, the differences disappear at higher lags. The only exception is the VECM, where per definition shocks to the system do not die out, which Figure 4(b) clearly shows.

![Figure 4](image.png)

(a) w.r.t. shocks $\epsilon_t^{(r)}$

(b) w.r.t. shocks $\epsilon_t^{(f-p)}$

Figure 4: **Impulse responses of real log returns from a CFVAR$_d$(p)** - The figure plots the impulse responses of returns from the co-fractional model for different $d$ over 20 lags. The reported impulse responses are orthogonalized. A shock to the rent to price ratio occurs first. The size of shocks is scaled to one standard deviation.

While the co-fractional model is hence very robust to the fractional differencing parameter, $d$, this can not be generalized to the triangular model. Within the estimated range of $d = \{0.65, 0.7, 0.75, 0.8\}$, the portion of the variance of unexpected returns that is due to the variation in the revision in expectations about the present
value of future returns is equal to \{1.53, 1.79, 2.17, 2.70\}. In comparison, the analogue measure resulting from the co-fractional VAR decreased slightly within the same range for \(d\). Similarly, the ratio of cash-flow news variance relative to unexpected return variation varies within \{0.49, 0.59, 0.75, 1.01\} with \(d\). We conclude that the sensitivity of the co-fractional VAR to the fractional differencing parameter is very limited, but the robustness of the triangular return prediction model is questionable.

8 The term structure of risk

The term structure of risk measures the volatility of expected returns over different investment horizons. We find that returns on the housing market have a predictable component. This implies that the perceived risk for an investor varies over time. If this risk, measured as the annualized standard deviation, decreases as the investment horizon increases, expected returns exhibit mean reversion. In turn, if the risk of a real-estate investment increases with the horizon, returns are mean averting.

Figure 5 plots the term structure of risk for the housing market, resulting from different model specifications. Similar to the findings of Campbell and Viceira (2005) for the stock market, we find that risk at the one-year horizon is approximately 18%. In line with these authors, we also confirm mean-reversion in the housing market, but only for short-horizon returns. The CFVAR\(_{0.75}(1)\) suggests mean reversion for a horizon up to 20 years and mean aversion thereafter. The same is true for the CFVAR\(_{1}(1) = \text{VECM}\). The triangular model implies an even shorter mean-reversion horizon of approximately 12 years. Only if returns are modeled by a stationary VAR(2) we do not find mean aversion after the initial mean-reverting period, since the term structure of risk in this case is almost flat after 20 years.

We first discuss the difference between the stationary VAR(2) model and the other three models\(^7\). In a VAR model, the effect of a shock decays exponentially, which explains why there is no long-run risk associated with it. Schotman et al. (2008) show that if \(d = 0\) the term structure of risk, \(\sigma^{(k)}\), is bounded. In contrast, if \(d > 0\) the variance of expected returns and the covariance between expected and unexpected returns diverge in \(k\) and the former will dominate asymptotically, which explains the observed mean aversion in our fractional models.

Schotman et al. (2008) also show that the variance of expected returns that results

\(^7\)Note that the BIC of the VAR(2) is -0.5063. Hence, the model has a better fit than the triangular model and the VECM, but fits the data worse than the CFVAR\(_{0.7-0.8}(1)\) specifications.
Figure 5: **The term structure of risk for the housing market** - The figure plots the term structure of risk, $\sigma^{(k)}$, for real log returns. Parameter estimates are based on the prediction estimates from the co-fractional CFVAR$_{0.75}(1)$, the triangular model (where the rent to price ratio is modeled as an ARFIMA(1,0.75,0)), a stationary unrestricted VAR(2), and a VECM for $[r_t, (f-p)_t]'$, respectively. The y-axis measure is the annualized standard deviation. The x-axis contains $k$ in years.

from a system specifications diverges faster with increasing $d$. As this variance term dominates $\sigma^{(k)}$ as $k \to \infty$, it follows that the degree of mean aversion is more pronounced for the VECM than for the CFVAR$_{0.75}(1)$ at long horizons. This is confirmed by our empirical results. If the rent to price ratio and returns are modeled as unit-root nonstationary processes as opposed to mean-reverting nonstationary processes, real-estate investments are riskier in the long run.

In the triangular model, we predict returns with only one lag of the rent to price ratio, whereas in the CFVAR$_{0.75}(1)$, we consider the entire history of the predictor variable. As a result, there is less predictability in housing returns in the triangular model. This implies that there is more uncertainty associated with expected returns in this case. As this uncertainty is the driving force of the term structure of risk at long horizons, it is logical consequence that the mean aversion is more pronounced for the triangular model than for the co-fractional model.

In the previous sections we found evidence that the CFVAR$_{0.75}(1)$ outperforms all other specifications considered here in terms of modeling real-estate returns. We thus conclude this section with the observation that the risk associated with housing-
market returns decreases up to an investment horizon of 20 years, from an initial level of 18% to approximately 11%. For longer investment horizon, risk increases again. This increase is rather moderate, however. The risk level equals the initial risk level \((k = 1)\) only after 100 years. With an investment horizon of 200 years, the annualized standard deviation is 31%.

9 Conclusion

We have analyzed the implications of the co-fractional vector autoregressive model of Johansen (2008) as a model for returns in financial markets and considered its empirical performance for long time series of real-estate returns and rents. We find that we can predict a substantial portion of future movements of 355 years of annual housing returns, using only past returns and the rent to price ratio. In contrast, the standard univariate return prediction model captures much less of the variance of returns from the same data set. In addition, we find that the estimates resulting from the co-fractional model are highly robust to misspecifications in the low frequency behavior of the respective series. Standard models are, however, much more sensitive to changes in the order of fractional integration.

We attribute the improved return prediction of the co-fractional model over the triangular model to the fact that the former predicts returns with the entire history of the rent to price ratio, whereas the latter considers only the first lag. The idea to smooth a valuation ratio in a predictive equation is not new. For instance, Campbell and Shiller (1988b) already consider a 30-year moving average of earnings relative to prices as a state variable. In contrast to most studies, however, we do not use predetermined weights in the long distributed lag of our predictor. Rather, the weights in our fractional prediction framework result from a fully parameterized model. In our co-fractional model, housing returns are not only predicted by a valuation ratio, but also a weighted sum of past returns. This specification thus has a second advantage, since smoothing is also known as a means to cope with measurement errors (see, for instance, Graham and Dodd (1934)), which is most likely present in our return data.

As a motivation for the application of Johansen’s model, we name the persistence of the rent to price ratio and the cointegration between prices and rents. With semi-parametric Whittle estimation, we show that the rent to price ratio is a nonstationary long memory process, while rents and prices have a unit root. From the resulting weak fractional cointegration, we conclude that returns and the rent to price ratio strongly
cointegrate and that returns exhibit long memory as well.

As a byproduct of the estimations of the co-fractional VAR, we confirm several data properties that are interesting from an asset pricing perspective. Both, past realizations of the rent to price ratio as well as the innovations to the process, have a larger influence on present returns than the counterpart of past returns. Innovations to the valuation ratio also dominate the part of unexpected returns that is due to discount factor news. In line with Campbell and Shiller (1991) and Campbell and Vuolteenaho (2004), we find that the variance of the latter makes up for the largest portion of the variability of unexpected returns. Similar to Campbell and Vuolteenaho (2004) the correlation between cash flow news and discount rate news in our model is small an positive, as opposed to Campbell and Shiller (1991).

Returns on the housing market exhibit mean reversion at low investment horizons up to 20 years. At longer horizons, the term structure of risk is increasing, implying that long-run investment have mean-averting returns. The degree of long-horizon mean-aversion is relatively smooth in the co-fractional model. In comparison, other prediction model understate or overstate the increase in risk.

The results in this study are based on historic real estate data from the Dutch market. This market is relatively small in terms of size as well as significance, as compared to the US housing market, for instance, and it is certainly less well-researched than stock and bond markets. Nevertheless, we assume that our conclusions that the co-fractional model opens a very promising path for future return predictions in financial markets, are generalizable. One the one hand, generalizability to the US real estate market can be expected, since the properties of our price and rent series are very similar to the CMPHI index and the tenants’ rent component of the CPI, respectively. We observe the same almost monotonic increase of both series in real term after 1970, where the latter movements are smoother, as in Gallin (2008). The properties of the rent to price ratio are also comparable, as both series exhibit a trough in the late 1970s, for example. On the other hand, we are confident that our outcomes also hold in markets with other asset classes. According to the study of Leamer (2002) there is a clear analogy between the rent to price ratio in the housing market and the price-earning, or similarly the dividend yield, in the stock market. In addition, Caporale and Gil-Alana (2004) provide empirical evidence of weak fractional cointegration between prices and dividends, corresponding to the common stochastic trend of housing prices and rents in our work. The hypothesized generalizability provides an interesting avenue for future research.
Appendix

A Derivation of the concentrated likelihood of the UEW in (6)

We reparametrize the negative likelihood function of the UEW estimator of Sun and Phillips (2004), given by

\[ Q_m^{(UEW)} = \frac{1}{m} \sum_{j=1}^{m} \left( \ln \left( \tau \lambda_{j}^{-2d} + \kappa \right) + \frac{1}{\tau + \kappa \lambda_{j}^{2d}} I_{(1-L)^{d}y} (\lambda_{j}) \right). \]  

(A1)

for some process \( y_t \). We let \( \varphi = \frac{\pi}{\tau} \) and re-write Equation (A1) as

\[ Q_m^{(UEW)} = \frac{1}{m} \sum_{j=1}^{m} \left( \ln (\tau) + \ln (\lambda_{j}^{-2d} + \varphi) + \frac{1}{\tau} I_{(1-L)^{d}y} (\lambda_{j}) \right). \]  

(A2)

In Section 3 we define \( \tau \) as a positive constant to which the spectral density of the innovations of a long-memory series \( y_t \) converges as approaching frequency zero. \( \kappa \) multiplied by \( 2\pi \) denotes the variance of the high-frequency component of the process \( y_t \). As a consequence, \( \varphi \geq 0 \).

The partial derivative with respect to \( \tau \) of (A2) is computed as

\[ \frac{\partial Q_m^{(UEW)}}{\partial \tau} = \frac{1}{m} \sum_{j=1}^{m} \left( \frac{1}{\tau} - \frac{1}{\tau^2} I_{(1-L)^{d}y} (\lambda_{j}) \right). \]  

(A3)

We can obtain the solution of the first order condition of the optimization with respect to \( \tau \) by setting \( (\partial Q_m^{(UEW)}/\partial \tau) = 0 \). As \( \tau > 0 \), there is one feasible solution, given by the following equation

\[ \tau = \frac{1}{m} \sum_{j=1}^{m} \frac{I_{(1-L)^{d}y} (\lambda_{j})}{1 + \varphi \lambda_{j}^{2d}}. \]  

(A4)

Plugging the optimal value for \( \tau \) from Equation (A4) into the likelihood function (A2) gives the concentrated negative likelihood in Equation (6).

B Derivation of Impulse Responses of CFVAR\(_d\)(1) in (10)

We exemplify the transformation process and the derivation of the recursive impulse responses for the CFVAR\(_d\)(1). The solution for any other lag length selection \( p \) follows
analogously.

We re-write the process for $z_t$, specified in (8) in the following way. Recall that $\Delta^d z_t = \sum_{i=0}^{\infty} \theta_i(d) = L^i z_t$, with $\theta_i = \theta_i(d) = (-1)^i \theta_i$. Then,

$$
\Delta^d z_t = \gamma \delta'(1 - \Delta^d) z_t + \Gamma_1 (1 - \Delta^d) \Delta^d z_t + \epsilon_t.
$$

(B1)

Expanding the $\Delta^d$ operators, we have

$$
\sum_{i=0}^{\infty} \theta_i L^i z_t = \gamma \delta' \left(1 - \sum_{i=0}^{\infty} \theta_i L^i\right) z_t + \Gamma_1 \left(1 - \sum_{i=0}^{\infty} \theta_i L^i\right) \sum_{i=0}^{\infty} \theta_i L^i z_t + \epsilon_t. 
$$

(B2)

Moving all lags to the right-hand side gives

$$
z_t = \sum_{i=1}^{\infty} \left(- I - \gamma \delta'\right) \theta_i L^i z_t - \Gamma_1 \sum_{i=1}^{\infty} \theta_i L^i \sum_{j=0}^{\infty} \theta_j L^j z_t + \epsilon_t
$$

$$
= \sum_{i=1}^{\infty} \left(- I - \gamma \delta'\right) \theta_i L^i z_t - \Gamma_1 \sum_{i=1}^{\infty} \left(\sum_{l=0}^{i-1} \theta_i \theta_l\right) L^i z_t + \epsilon_t
$$

$$
= \sum_{i=1}^{\infty} \left(- I \theta_i - \gamma \delta' \theta_i - \Gamma_1 \sum_{l=0}^{i-1} \theta_i \theta_l\right) L^i z_t + \epsilon_t
$$

$$
= \sum_{i=1}^{\infty} \Xi_i L^i z_t + \epsilon_t. 
$$

(B3)

The coefficient $\Xi_i$ can be calculated recursively and form the starting point of the computations of impulse responses, $\Phi_i$. We find that

$$
\Phi_0 = I
$$

$$
\Phi_1 = \Xi_1
$$

$$
\Phi_2 = \left(\Xi_2 + \Xi_1^2\right)
$$

$$
: :
$$

$$
\Phi_j = \sum_{i=0}^{j-1} \Xi_{j-i} \Phi_i. 
$$

(B4)

We further need an causal ordering of the shocks. We assume that shocks to the rent to price ratio are first, i.e. we start with fundamentals. Such an ordering is also consistent with the triangular model. The orthogonalized impulse responses are
\[ \Psi_i = \Phi_i P, \text{ where } P \text{ is the upper diagonal matrix} \]

\[
P = \begin{pmatrix}
\sqrt{\Sigma_{11} - \Sigma_{12}} \\
\Sigma_{12} \\
\sqrt{\Sigma_{22}} \\
\Sigma_{11} \\
\Sigma_{12} \\
\sqrt{\Sigma_{22}}
\end{pmatrix}
\]  \quad (B5)

C Derivation of return decomposition for the CFVAR\(_d(p)\)

The return process resulting from the co-fractional model is given by

\[
r_t = e1' \sum_{i=0}^{\infty} \Phi_i \epsilon_{t-i},
\]  \quad (C1)

as outlined in Section 4.

From this we can derive the return decomposition \(v_{r,t+1} = \eta_{f,t+1} - \eta_{r,t+1}\).

\[
v_{r,t+1} = r_{t+1} - E_t r_{t+1}
\]

\[
= e1' \sum_{i=0}^{\infty} \Phi_i \epsilon_{t+1-i} - e1' \sum_{i=1}^{\infty} \Phi_i \epsilon_{t+1-i}
\]

\[
= e1' \epsilon_{t+1}.
\]  \quad (C2)

For the innovations in the present value of future returns, we find that the model implies

\[
\eta_{r,t+1} = \sum_{j=1}^{\infty} \varsigma^j (E_{t+1} - E_t) r_{t+1+j}
\]

\[
= \sum_{j=1}^{\infty} \varsigma^j \left( e1' \sum_{i=j}^{\infty} \Phi_i \epsilon_{t+1+j-i} - e1' \sum_{i=j+1}^{\infty} \Phi_i \epsilon_{t+1+j-i} \right)
\]

\[
= e1' \sum_{j=1}^{\infty} \varsigma^j \Phi_j \epsilon_{t+1}.
\]  \quad (C3)

Hence, the variances of the elements of the unexpected return process are

\[
\text{Var}(v_{r,t+1}) = e1' \Sigma e1
\]  \quad (C4)

and

\[
\text{Var}(\eta_{r,t+1}) = e1' \left( \sum_{j=1}^{\infty} \varsigma^j \Phi_j \right) \sum \left( \sum_{j=1}^{\infty} \varsigma^j \Phi_j \right)' e1.
\]  \quad (C5)
Using the equality \( \nu_{r,t+1} + \eta_{r,t+1} = \eta_{f,t+1} \) to back out the innovations to expected future discounted changes in cash flows, we find that

\[
Var(\eta_{f,t+1}) = e_1' \left( \sum_{j=0}^{\infty} \zeta^j \Phi_j \right) \Sigma \left( \sum_{j=0}^{\infty} \zeta^j \Phi_j \right)' e_1. \quad \text{(C6)}
\]

The covariance is given by

\[
\text{Cov}(\eta_{r,t+1}, \eta_{f,t+1}) = e_1' \left( \sum_{j=1}^{\infty} \zeta^j \Phi_j \right) \Sigma \left( \sum_{j=0}^{\infty} \zeta^j \Phi_j \right)' e_1. \quad \text{(C7)}
\]

D Derivation of the return decomposition of the triangular model

The triangular model for our the return process, defined in Section 4, is given by

\[
\begin{align*}
r_t &= \beta (f - p)_{t-1} + \nu_{1t} \quad \text{(D1)} \\
\Delta^d (f - p)_t &= \rho \Delta^d (f - p)_{t-1} + \nu_{2t}, \quad \text{(D2)}
\end{align*}
\]

suppressing the constants for convenience. Let the operator \( \Delta_+^d \) denote \( \sum_{i=0}^{t-1} (-1)^i (\frac{-d}{i}) L^i \).

The ARFIMA process for the rent to price ratio can be transformed into

\[
\begin{align*}
(f - p)_t &= \rho (f - p)_{t-1} + \Delta_+^d \nu_{2t} \\
(1 - \rho L) (f - p)_t &= \Delta_+^d \nu_{2t} \\
(1 - \rho L) (f - p)_t &= \sum_{j=0}^{t-1} \theta_j (-d) L^j \nu_{2t} \\
(f - p)_t &= \sum_{i=0}^{\infty} \rho^i L^i \sum_{j=0}^{t-1} \theta_j (-d) L^j \nu_{2t} \\
(f - p)_t &= \sum_{i=0}^{\infty} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l (-d) \right) L^i \nu_{2t}. \quad \text{(D3)}
\end{align*}
\]

It follows that returns can be written as the following process

\[
r_t = \beta \sum_{i=0}^{\infty} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l (-d) \right) L^i \nu_{2t-1} + \nu_{1t}. \quad \text{(D4)}
\]
From this we can derive the return decomposition $\nu_{t+1} = \eta_{f,t+1} - \eta_{r,t+1}$.

$$
\nu_{t+1} = r_{t+1} - Er_{t+1} \\
= \beta \sum_{i=0}^{\infty} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l(-d) \right) L^i \nu_{2t} + \nu_{1t+1} - \beta \sum_{i=0}^{\infty} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l(-d) \right) L^i \nu_{2t} \\
= \nu_{1t+1}.
$$

(D5)

For the innovations in the present value of future returns, we find that the triangular model implies

$$
\eta_{t+1} = \sum_{j=1}^{\infty} \varsigma^j \left( E_{t+1} - E_t \right) r_{t+1+j} \\
= \sum_{j=1}^{\infty} \varsigma^j \left( \beta \sum_{i=j-1}^{\infty} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l(-d) \right) \nu_{2t+j-i} - \beta \sum_{i=j}^{\infty} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l(-d) \right) \nu_{2t+j-i} \right) \\
= \sum_{j=1}^{\infty} \varsigma^j \beta \left( \sum_{l=0}^{j-1} \rho^{j-1-l} \theta_l(-d) \right) \nu_{2t+1}.
$$

(D6)

Hence, the variances of the elements of the unexpected return process are

$$
\text{Var}(\nu_{r,t+1}) = \sigma_1^2
$$

(D7)

and

$$
\text{Var}(\eta_{r,t+1}) = \beta^2 \left( \sum_{j=1}^{\infty} \varsigma^j \left( \sum_{l=0}^{j-1} \rho^{j-1-l} \theta_l(-d) \right) \right)^2 \sigma_2^2,
$$

(D8)

where $\sigma_1^2$ and $\sigma_2^2$ are the variance of $\nu_{1t}$ and $\nu_{2t}$, respectively.

Using the equality $\nu_{r,t+1} + \eta_{r,t+1} = \eta_{f,t+1}$ to back out the innovations to expected future discounted changes in cash flows, we find that

$$
\text{Var}(\eta_{f,t+1}) = \sigma_1^2 + \beta^2 \left( \sum_{j=1}^{\infty} \varsigma^j \left( \sum_{l=0}^{j-1} \rho^{j-1-l} \theta_l(-d) \right) \right)^2 \sigma_2^2 \\
+ 2\beta \sum_{j=1}^{\infty} \varsigma^j \left( \sum_{l=0}^{j-1} \rho^{j-1-l} \theta_l(-d) \right) \sigma_{1,2},
$$

(D9)

where $\sigma_{1,2}$ is the covariance between $\nu_{1t}$ and $\nu_{2t}$. The covariance between the elements
of the return decompositions is given by

\[
\text{Cov}(\eta_{r,t+1}, \eta_{f,t+1}) = \beta^2 \left( \sum_{j=1}^{\infty} \varsigma^j \left( \sum_{l=0}^{j-1} \rho^{j-1-l} \theta_l (-d) \right) \right)^2 \sigma_2^2 + \beta \sum_{j=1}^{\infty} \varsigma^j \left( \sum_{l=0}^{j-1} \rho^{j-1-l} \theta_l (-d) \right) \sigma_{1,2}. \tag{D10}
\]

For the term structure of risk, we compute the conditional variance of future returns as

\[
\text{Var}_t \left( \sum_{j=1}^{k} r_{t+j} \right) = E \left( \sum_{j=1}^{k} r_{t+j} - E_t(r_{t+j}) \right)^2 = E \left( \sum_{j=1}^{k} \nu_{t+j} + \sum_{j=2}^{k} \left[ \beta \sum_{i=0}^{j-2} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l (-d) \right) \right] \nu_{2t+j-1-i} \right)^2 = k \sigma_1^2 + \sum_{j=2}^{k} \left[ \beta \sum_{i=0}^{j-2} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l (-d) \right) \right]^2 \sigma_2^2 + 2 \sum_{j=2}^{k} \beta \sum_{i=0}^{j-2} \left( \sum_{l=0}^{i} \rho^{i-l} \theta_l (-d) \right) \sigma_{1,2}. \tag{D11}
\]

The term structure is computed as \( \sqrt{\frac{1}{k} \text{Var}_t \left( \sum_{j=1}^{k} r_{t+j} \right)} \).
References


