Liquidity and Return Reversals

Preliminary: Please do not quote

Pierre Collin-Dufresne and Kent Daniel[†]

- Abstract -

We estimate a short term reversal process for daily US equity returns. Over our primary sample period of 1972-2014, and for our sample of the 100 largest traded firms, on average approximately 90% of idiosyncratic price shocks are permanent. The remaining 10% is temporary, and decays exponentially toward zero, with a half life of about 2.5 days. While the rate of decay (the half life) is relatively constant over time, the magnitude decay varies considerably over the sample. Our findings are consistent with the slow movement of capital (Duffie 2010). Also, in contrast with previous literature, we find no evidence that this rate of mean reversion is related to market-wide measures of illiquidity, such as the VIX. Our results are thus also consistent with a lack of integration across capital markets.

Introduction

The cross section of individual stock returns over the coming week or month is strongly negatively related to the past returns of the same firms over the past week or month (Jegadeesh 1990, Lehmann 1990). This negative serial correlation is generally interpreted as evidence consistent with incomplete liquidity provision, and much empirical evidence is consistent with this: Chan (2003) and Tetlock (2011) show that large stock price moves exhibit more reversal if they are not associated with news and Da, Liu, and Schaumburg (2013) show that within industry/residual based reversals have a far higher Sharpe ratio.[1]

In addition, Avramov, Chordia, and Goyal (2006) argue that the reversal effect is present only in small, illiquid stocks with high turnover, and Khandani and Lo (2007) document a dramatic decline over time in the efficacy of the strategy. Finally, the strength of the reversal strategy appears to depend on arbitrageurs ability to access capital: Nagel (2012) documents a strong positive correlation between the return of a \$1-long/\$1-short short-term-reversal and the level of the VIX. He attributes this co-variation to being consistent with the "... withdrawal of liquidity supply and an associated increase in the expected returns from liquidity provision ... during times of financial market turmoil." Overall, he interprets his results as being consistent with limited investor capital, but implicitly also with strong integration of capital markets.

We make several contributions to this literature. First, show that in fact the reversal effect is remarkably strong in even the largest and most liquid stocks. Second, we more accurately estimate the nature of the reversal effect: we show that the effect is well captured by a model in which a individual stock's expected return is an exponentially weighted function of it's past residual returns, consistent with the slow-moving capital hypothesis of Duffie (2010). Finally we show that, in contrast with evidence presented elsewhere, while there is considerable variation over time in the both the rate mean reversion, and the Sharpe-ratio of the short-term-reversal strategy, this time variation is unrelated to standard measures of liquidity such as the VIX. It is, however, related to the volatility of the short-term reversal strategy itself. Thus, this evidence is suggestive of somewhat limited capital, but a lack of integration of financial markets.

---

[1]Da, Liu, and Schaumburg (2013) form portfolios based on past one-month returns. They show that the alpha of their residual-reversal strategy has a large and statistically significant alpha *w.r.t* a "five-factor" version of the Fama-French model.These results are summarized in their Table 2. The model has a factors the standard Fama-French four factors plus the short-term reversal factor on Ken French's website.

We proceed as follows: In Section 1 we describe our data sources and data selection procedure, and in Section 2 we do basic analysis of the data that motivates our later tests. In Section 3 we discuss our basic model, the estimation of this model and some issues related to this estimation. and Section 4 concentrates specifically on the estimation of the str process using the UC-ARIMA process implemented via a Kalman filter, and the Hodrick and Prescott (1980) filter. Section 5 examines the covariation of the returns to short-term liquidity provision with the VIX, and Section 6 concludes.

1   Data Description

Our primary sample is the 100 largest firms by market capitalization in the universe of US common stocks, over the period from January 1927 through March 2013.[2] Specifically, at close on the last trading day of each calendar year we select the 100 firms with the largest market capitalization. We trade only those firms over the next calendar year. What this means is that our cross-section is the very largest firms in the CRSP universe. We do this to ensure that each and every one of the stocks in our sample can be traded each trading day.

Our data primarily comes from CRSP, and includes price, trading volume, and shares outstanding. In addition, we get earnings announcement dates from COMPUSTAT.

1.1   Non-Trading:

Even though we restrict our sample to the 100 largest firms by market capitalization at the beginning of each year, there is still a substantial amount of non-trading in the pre-1950 period. Specifically, before 1950 about half of the firms in our sample have at least one day where they do not trade, and about a quarter of the sample has more than 10 non-trading days. One reason for this is that, pre-1950, the number of traded firms is considerably smaller. A second more important reason is that share turnover is considerably smaller pre-1960.

However, post-1960, the number of non-trading days falls dramatically. As is seen in Figure 1, the median number of firms which have at least one non-trading day in the post-1960s sample is zero. In fact, in the last 54 years of our sample (1960-2013) there

---

[2]However, in our robustness checks, we also examine the largest 500 stocks, and the 100 and 500 largest firms where the sample is restricted to NYSE firms and to NASDAQ firms.
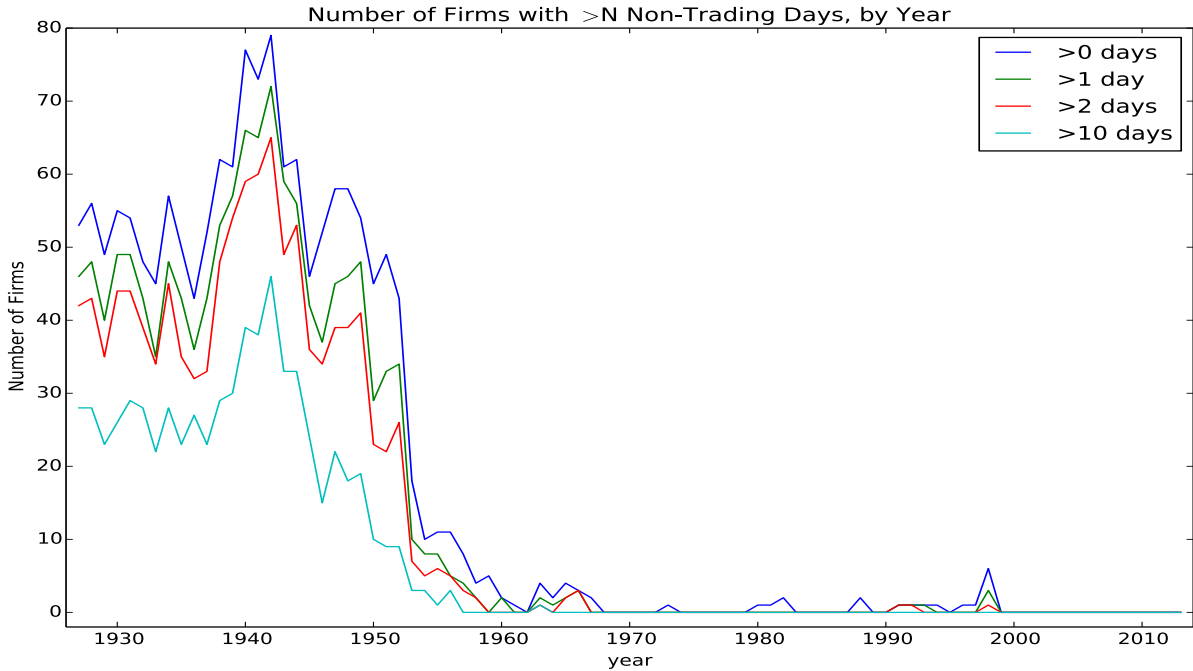
This plot shows the number of firms in our sample of 100 firms which experienced at least N non-trading days, by year. The first four lines in the plot give the number of firms that experienced more than 0, 1, 2, and 10 non trading days, by calendar year.

Figure 1: **Non-Trading Analysis**

are only 6 firm-years (out of 5400) where there are more than 2 non-trading days, and only one firm-year in which there are more than 10 non-trading days.

1.2   Share Turnover:

We define share turnover as trading volume in shares dividend by the total number of shares outstanding. Both quantities are reported by CRSP on a daily basis.

Figure 2 plots the 42-day (2-month) rolling average of the annualized, value-weighted turnover for our sample of the 100 largest firms. It is clear from this plot that share turnover varies dramatically over our sample period, something we will exploit later in our analysis. Turnover is over 100%/year in the late 1920's, and then falls over the 1930 by about a factor of 20. Annualized turnover averages only 6.72% from 1940 through 1965, but begins a fairly steady rise in the late 1960s. It peaks in late 2008, during the financial crisis, and reaches a 42-day average of 367% at the peak of the financial crisis.[3]

---

[3]The peak is the 42-day period ending on November 4, over which the average turnover is 367.3%. The average value-weighted annualized turnover over the one-year period from 2009:06-2010:05 is 291%. The highest one-day turnover in our sample is 2.59% on October 10, 2008, and the lowest is 0.003% on August 24, 1940.
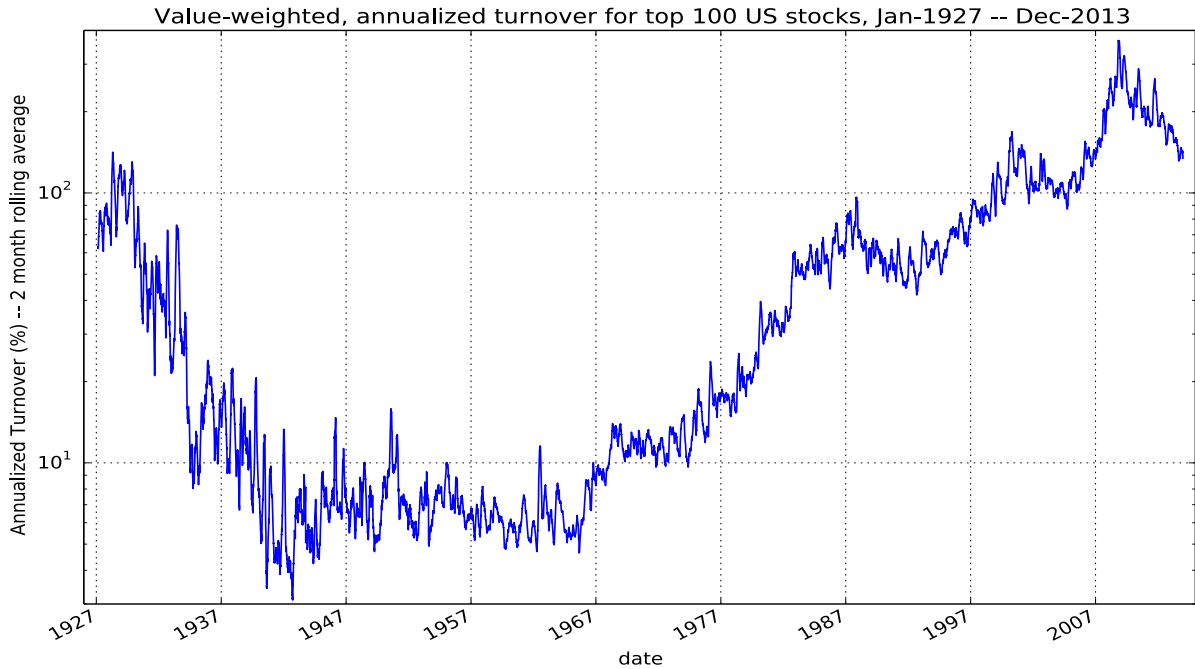
This figure plots the 42-day (2-month) moving average of the annualized value-weighted turnover of the 100 stocks in our sample, in percent/year. Turnover is defined as trading volume dividend by the total number of shares outstanding.

Figure 2: **Average Annualized Turnover**

1.3   Delistings:

The median number of delistings per year in our sample is zero. However, some delistings do occur. The maximum is in 1998, when there are six. Over half are in the two decades between 1980 and 2000, and resulted from mergers. Hence, the delisting return supplied by CRSP is probably reliable. If a firm is delisted, up until the delisting takes place, we trade as if we are unaware that the delisting will take place. On the delist date, we assume that any holdings, long or short, of that firm's shares earn the return on that date plus the CRSP delisting return.

Note the the universe of firms in our sample often changes on the last trading day of each calendar year. If a firm leaves our sample (because it is no longer one of the 100 largest firms) we close out our position at the closing price on that day. Similarly, if a firm enters the sample of the 100 largest firms, we start trading into that firm starting at close on the last trading day of the year.

Estimated short-term reversal effect (1972-2013)
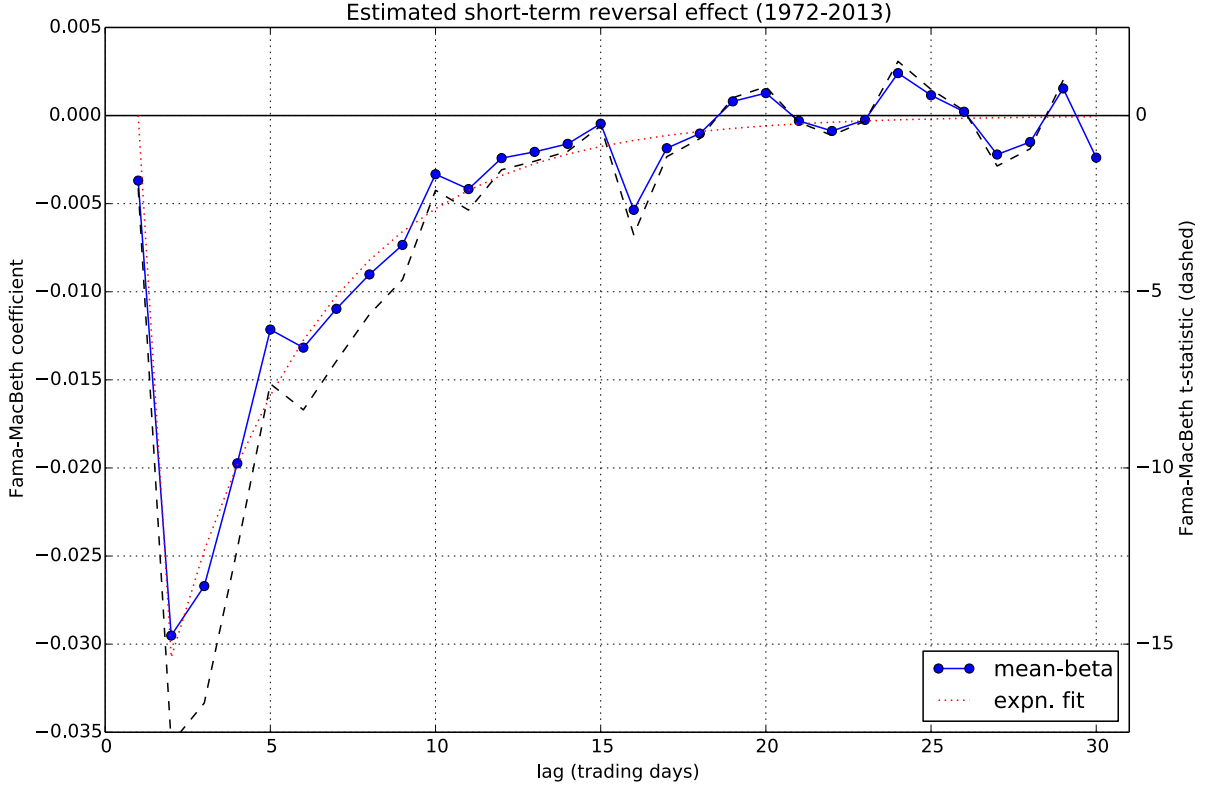
This figure plots the average Fama-MacBeth coefficients, the Fama-MacBeth t-statistics for a Fama-MacBeth regression of the residual returns (calculated as described in the text) on the lagged residual returns for the 100 largest common stocks traded on US exchanges. Lagged dependent returns which occur within one day of the an earnings announcement date (per COMPUSTAT) are set to zero.

Figure 3: **Fama-MacBeth Regression Results**

1.4   Earnings Announcement Dates:

Our earnings announcement data is from COMPUSTAT. These data begin with earnings announcements after September 1971 (*i.e.,* third quarter earnings announcements, for firms with a December fiscal year end), and continue through the end of our sample in March 2014, giving us 170 quarters ($\sim$42.5 years) of earnings announcement dates.

2   Data Characterization

2.1   Return Decay Following Price Shocks

We begin with some simple empirical tests that help to characterize the patterns in our data.

First, we run a daily Fama and MacBeth (1973) regression for the largest 100 firms,

measured at the beginning of each year. We perform our analysis over the period from January 1972 through the end of March 2014. Our choice of starting date is primarily motivated by the fact that there are almost no non-trading days post-1972, and secondarily by the fact that our COMPUSTAT data contains earnings announcement dates for almost all firms (among the top 100) after 1972.

We first calculate residual returns for each of our 100 firms, for each day. We do this by first estimating the market beta for each firm using daily data over the two years leading up to the start of each calendar year for each of the 100 firms in our sample. We shrink these estimated betas using the equation:

$$\beta_i^s = 0.23 + 0.77 \cdot \hat{\beta}_i \tag{1}$$

The coefficients in this shrinkage regression are the average coefficients (intercept and slope) from a pooled regression of the *ex-post* daily betas on the on the $\hat{\beta}_i$s (in equation (1)).

We then calculate the residual return for each firm as:

$$\epsilon_{i,t} \equiv (r_{i,t} - r_{f,t}) - \beta_i^s(r_{m,t} - r_{f,t}).$$

For our baseline Fama-MacBeth regression, our dependent variable is the cross-section of residual returns for our 100 largest firms, and the vector of independent variables is the lagged cross-section of residual returns, for lags between 1 and 30 days. That is, we run daily cross-sectional regressions with the specification:

$$\epsilon_{i,t} = \gamma_{0,t} + \sum_{\tau=1}^{30} \gamma_{\tau,t} \cdot \epsilon_{i,t-\tau}^* \tag{2}$$

where the $*$ superscript in $\epsilon_{i,t-\tau}^*$ denotes that the ϵ is set to zero if it is within one day of a COMPUSTAT earnings announcement date.

To be able to assess the efficacy of zeroing out earnings-announcement date (EAD) returns, we also run the regression:

$$\epsilon_{i,t} = \gamma_{0,t}^a + \sum_{\tau=1}^{30} \gamma_{\tau,t}^a \cdot \epsilon_{i,t-\tau} \tag{3}$$

where, here, the residuals are not zeroed out on Earnings Announcement Dates.

Figure 4: **Fama-MacBeth Results - Earnings Announcement Dates Only**
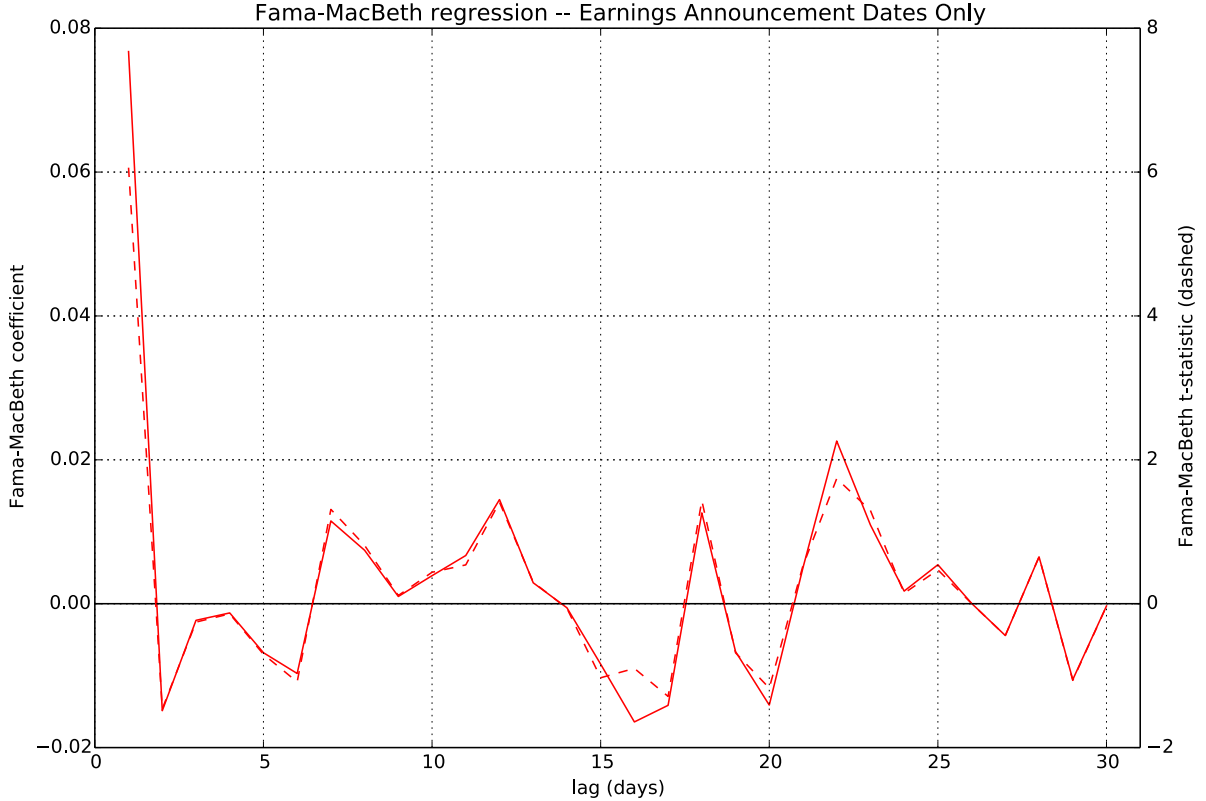
This figure plots the average Fama-MacBeth coefficients and FM t-statistics for a FM regression of the residual returns (calculated as described in the text) on the lagged residual returns for the 100 largest common stocks traded on US exchanges. Here, the coefficients are each for univariate regressions which are run only on dates, for the dependent variable, there are four or more earnings announcements.

We discuss the motivation for zeroing out the earnings announcement date returns in Section 2.2. Figure 3 presents the mean coefficient as a function of τ, and the coefficient t-statistic ($i.e.$, $\frac{1}{\sqrt{T}}\hat{\gamma}_\tau/std(\hat{\gamma}_\tau)$). In addition we plot an exponential fit to the Fama-MacBeth coefficients.

Several things are of note here. First, even for these large market capitalization firms, there is substantial evidence of reversal. The t-statistic on the two-day lagged return is above 17, and the t-statistic on each of the returns from a two-day lag up to 12-days lag is greater than 2.0.

It is notable that the magnitude of the Fama-MacBeth coefficients as a function of lag are well captured by a simple exponential function, except at the first lag,

2.2 Earnings Announcement Dates

As noted above, we zero out the lagged returns in the forecasting regression when those returns occur on earnings announcement dates. We also do this later when we construct our trading strategy. We do this because we find that there is no evidence of reversal on earnings announcement dates.

We show this in three ways. First, Figure 4 presents the results of regression returns on lagged returns – similar to the regression of Figure 3 – except here the regression is run only for dates on which the lagged returns are on earnings announcement dates.[4] Note that, with the exception of the one-day lag return, all of the coefficients except one are statistically indistinguishable from zero at the 5% level.

Second, we run the Fama-MacBeth regression in equation (2) in two ways, with the dependent lagged residual returns set to zero on earnings-dates, and without. We can then examine the efficiency of each of the coefficient portfolios in the baseline regression (i.e., with the EAD returns zeroed out) and without. We do this by regressing the time series of coefficients, for a given lag, on the corresponding coefficients with the alternative regression. Recall that, if a portfolio is efficient with respect to another portfolio, the intercept of such a time series projection will be zero. The intercepts from these regressions, and for the reverse regressions, for lags 1-10, are presented in Table 1. Note that most of the coefficients in the reverse regression are significantly different from zero, but none of the coefficient for the corresponding reverse regression are.

Our third test involves the construction of an alternative trading strategy. As we discuss below, our baseline trading strategy constructs a portfolio where the weight on each asset in the portfolio is an exponentially weighted sum of past residual returns. Note that the weights on the past residual returns are *negative* – so the strategy is designed to have a positive Sharpe ratio. In constructing those weights, we zero out any lagged residual returns within one day of an earnings announcement. Over the 1972-2013 period, our baseline strategy has an unconditional annualized Sharpe ratio of 3.08.

To test this approach, we construct a alternative trading strategy where we ignore the earnings-announcement date. Over the same period, this strategy has an annualized unconditional Sharpe ratio of 2.91, and is 98.9% correlated with the baseline strategy.

To further then compare the baseline strategy returns to the alternative strategy

---

[4]Or, rather, when the returns occur within one day of the earnings-announcement date in COMPUS-TAT

Table 1: **Fama-MacBeth (1973) Earnings Announcement Date Analysis**

This table gives the intercept (α_τ) and the intercept t-statistics for time-series regressions of the baseline FM-coefficient portfolio returns ($\hat{\gamma}_{t,\tau}$ from equation (2)) on the corresponding FM-coefficient portfolio returns for the cross-sectional regression in equation (3) where the lagged residuals returns $\epsilon_{i,t-\tau}$ are not set to zero on EADS. We also present the intercept and the t-statistic for the reverse regression – which we denote

τ	α_τ	α_τ^a	$t(\alpha_\tau)$	$t(\alpha_\tau^a)$
1	0.00449	-0.004794	6.695	-6.974
2	-0.001005	-0.00313	-1.673	-4.999
3	-0.001048	-0.002741	-1.799	-4.508
4	-0.0002399	-0.002517	-0.4115	-4.175
5	-0.000506	-0.001316	-0.8575	-2.135
6	0.0006172	-0.00256	1.054	-4.19
7	0.0002584	-0.001872	0.4434	-3.085
8	-0.0001569	-0.001059	-0.279	-1.805
9	0.0002197	-0.001261	0.3853	-2.114
10	0.0002006	-0.000662	0.3512	-1.121

returns. To test the efficiency of the two strategies we run the two regressions:

$$r_t^{str} = \alpha + \beta r_t^{str-a} + u_t$$
$$r_t^{str-a} = \alpha^a + \beta^a r_t^{str} + u_t$$

We find that $\alpha = 93$ bps/day, with a t-statistic of 7.0. In contrast, $\alpha^a = -36$ bps/day, with a t-statistic of -2.8.

Note, however, that this simple test suggests that a more efficient portfolio actually *shorts* a small amount of the alternative trading strategy. Over this sample, the *ex-post* efficient combination puts *positive* weight on firms which experience positive returns on earnings announcement dates – consistent with the presence of post-earnings announcements drift. However, the Sharpe ratio of this MVE combination is only 3.11. In the rest of our analysis here, we simply zero out the lagged residual returns of firms. This also means that our strategy is really just a short-horizon reversal strategy, and doesn't profit from post-earnings announcement drift.

3 A Model of Return Reversals

To motivate the construction of our test portfolio, we here develop a model for individual firm excess returns ($\tilde{r}_{i,t+1}$). Our baseline model is:

$$\tilde{r}_{i,t+1} = \beta_{i,m}\tilde{r}_{m,t+1} + \underbrace{B_{i,t}\lambda_t + \sigma_i h_{\epsilon,t}\tilde{\epsilon}_{i,t+1}}_{\equiv \tilde{u}_{i,t+1}} \qquad\qquad \tilde{\epsilon}_{i,t+1} \sim \mathcal{N}(0,1) \qquad (4)$$

where $\mathbb{E}_t[\tilde{\epsilon}_{i,t+1}\tilde{\epsilon}_{j,t+1}] = 0 \ \forall \ i \neq j$, and where $\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 = 1$.

Firm i's excess return $\tilde{r}_{i,t+1}$ has a loading of $\beta_{i,m}$ a single common factor (the market). $u_{i,t+1}$ denotes firm i's residual return at time $t+1$.

3.0.1 The market return

While it is no particularly important for our estimation here, we nonetheless model that process for the excess market return ($\tilde{r}_{m,t+1}$). Our model is:

$$\tilde{r}_{m,t+1} = \mu + \underbrace{h_{m,t}\tilde{v}_{t+1}}_{\equiv \tilde{u}_{m,t+1}}, \qquad\qquad \tilde{v}_{t+1} \sim \mathcal{N}(0,1) \qquad (5)$$

where $\mathbb{E}_t[\tilde{v}_{t+1}\tilde{\epsilon}_{i,t+1}] = 0 \ \forall \ i$. The market return $\tilde{r}_{m,t+1}$ has a constant mean return, and variance that follows a Glosten, Jagannathan, and Runkle (1993) GARCH process:

$$h_{m,t}^2 = \omega_m + \beta_m h_{m,t-1}^2 + (\alpha_m + \gamma_m \mathrm{I}[(\tilde{r}_{m,t} - \mu) < 0])\, u_{m,t}^2. \qquad (6)$$

Maximum likelihood estimates of the parameters in equations (5) and (6) are given in Table 2.

3.0.2 Residual return specification

Expected return: As noted earlier, we specify that a firm's residual return $u_{i,t+1}$ is not mean zero, but rather has a conditional expected return that is negatively correlated with it's lagged residual returns for lags of up several weeks. Specifically, equation (4) specifies that firm i's conditional expected residual return $\mathbb{E}_t[\tilde{u}_{i,t+1}] = B_{i,t}\lambda_t$ is the product of a firm specific exposure $B_{i,t}$ and a common premium λ_t. The exposure ($B_{i,t}$) is governed

11

by an autoregressive process:

$$
\begin{aligned}
B_{i,t} &= \beta_r B_{i,t-1} + (1 - \beta_r)\tilde{u}_{i,t-1} \\
&= (1 - \beta_r)\sum_{l=1}^{\infty} \beta_r^l u_{i,t-l} \tag{7}
\end{aligned}
$$

Firm i's time t exposure is thus an exponentially weighted sum of its lagged daily residuals, starting at time $t-1$. Note that our specification does include the day t return in the expected return for day $t+1$ – we skip a day to avoid various econometric problems. [5] For our sample of the largest 100 firms, over the 1972-2014 time period, the average λ_t is about -0.12. This means that, following a residual shock of 1%, the price of a firm i will fall (starting in one day) by about 12 basis points over the next few weeks. Our estimation of equation (7) for daily returns gives a $\hat{\beta}_r = 0.720$, corresponding to a half-life of this price decay of 2.4 days.

Variance Process Specification: Equation (4) specifies that the volatility of firm i's residual return is $\sigma_i h_{\epsilon,t}$ – the product of a time-invariant but firm-specific term σ_i, and the (common) level of cross-sectional volatility, $h_{\epsilon,t}$. Our specification is consistent with Kelly, Lustig, and Van Nieuwerburgh (2012), who argue that time variation in individual firm idiosyncratic volatilities are largely captured by a single factor structure. Here, all changes in idiosyncratic volatility a result of changes in the common level $h_{\epsilon,t}$.

Equation (4) specifies that the shock to firm i's residual returns is

$$
(\tilde{u}_{i,t+1} - B_{i,t}\lambda_t) = \sigma_i h_{\epsilon,t}\tilde{\epsilon}_{i,t+1}
$$

---

[5]With this specification, the expected residual return at time $t+1$ is a function of residuals at times $t-1$ and earlier, skipping day t. This makes the strategy somewhat more implementable, as there is a one-trading day gap between the point where the residual returns are observed and the portfolio weights determined, and when the trades must be executed. This also avoid any potential bid-asked bounce effects. Note that (almost) every firm in our sample trades (almost) every day.

This means that:

$$\mathbb{E}_t\left[\frac{1}{n}\sum_{i=1}^{n}(\tilde{u}_{i,t+1}-B_{i,t}\lambda_t)^2\right] = h_{\epsilon,t}^2\mathbb{E}_t\left[\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\tilde{\epsilon}_{i,t+1}^2\right]$$

$$= h_{\epsilon,t}^2\left(\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\right)$$

$$= h_{\epsilon,t}^2 \tag{8}$$

That is, given our restrictions that $\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2=1$ and that the ϵ's are $i.i.d.$, unit-variance, $h_{\epsilon,t}^2$ is the conditional cross-sectional variance (as defined by the LHS of equation (8).)

The dynamics of $h_{\epsilon,t}^2$ are captured by an GARCH(1,1) process (Bollerslev 1986):

$$h_{\epsilon,t}^2 = \kappa_\epsilon + \alpha_\epsilon h_{\epsilon,t-1}^2 + \mu_\epsilon \sigma_{xs,t}^2, \tag{9}$$

where $\sigma_{xs,t}^2$ denotes the realized cross-sectional variance in period t:

$$\sigma_{xs,t}^2 \equiv \frac{1}{n}\sum_{i=1}^{n}(\tilde{u}_{i,t}-B_{i,t-1}\lambda_{t-1})^2. \tag{10}$$

For an individual firm, equation (4) specifies that $\mathbb{E}_t\left[(\tilde{u}_{i,t+1}-B_{i,t}\lambda_t)^2\right]=\sigma_i^2 h_{\epsilon,t}^2$. That is, σ_i^2 is the ratio of an individual firm's residual variance to the average residual variance of the n firms in our sample. As noted earlier, in our empirical implementation, we estimate σ_i at the beginning of each year, and then hold it fixed from the first trading day of the year through the last, consistent with the specification in equation (4). Empirically, we find considerable cross-sectional variation in σ_i.

Premium Specification: From equation (4), λ_t is the time-t expectation of the premium per unit of exposure that will be earned over period $t+1$ (*i.e.*, between t and $t+1$). The updating rule for λ_t is:

$$\lambda_{t+1} = \alpha_\lambda \lambda_t + (1-\alpha_\lambda)\tilde{u}_{B_1,t+1} \tag{11}$$

13

where $\tilde{u}_{B_1,t+1}$ denotes the period $t+1$ residual return of a portfolio with unit exposure to the str factor (that is, $B_{p,t} = 1$), and specifically the portfolio with time-t weights:[6]

$$w_{i,t}^{B_1} = \frac{B_{i,t}}{\sum_i B_{i,t}^2} \tag{12}$$

Here, λ_t should be interpreted as the econometrician's estimate of the latent period-$t+1$ premium λ_t^* – i.e., $\lambda_t = E_t[\lambda_t^*]$. Similarly, equation (11) should be interpreted as describing the evolution of this expectation. It is a reduced form for the Kalman filter solution to the system of equations:

$$
\begin{aligned}
\lambda_{t+1}^* &= \lambda_t^* + \tilde{v}_{t+1} \\
\tilde{u}_{B_1,t+1} &= \lambda_t^* + \tilde{\epsilon}_{B_1,t+1}
\end{aligned}
$$

where λ_{t+1}^* is the latent/unobserved premium. In a later Section, we will model the underlying premium λ_t^* as following a random walk.

In summary, according to this specification, when a firm experiences a large positive residual-return, its expected residual return falls in response. Absent future shocks, the expected return converges back towards zero with a rate governed by the parameter β_r in equation (7). As we will see in a moment, our estimated $\beta_r = 0.75$, corresponding to a half-life for this decay of 2.4 days. In contrast, the *magnitude* of the price decline varies slowly.

3.1 Model Estimation

In this section we present the results of the estimation of the reversal model described in Section 3 for our sample of the 100 largest CRSP firms, and over our sample period of 1974:01-2013:03.

We begin with the estimation of the market process (equations (5) and (6)), which we estimate jointly using a numerical maximum likelihood procedure. The results of this estimation are reported in Table 2.

We also estimate the parameters of the individual firm process (equations (7), (11) and (9)).

---

[6]Note that $\sum_i w_{i,t}^{B_1} B_{i,t} = 1$. If the residuals were uncorrelated and with uniform variance, this would be the portfolio with minimum variance portfolio subject to the constraint that $B_{p,t} = 1$, and consequently the minimum variance estimator of λ_t^*, conditional on only time $t+1$ returns.

Table 2: **Maximum Likelihood Estimates of Market Process Parameters**
This table presents the estimates, standard errors and t-statistics from the joint estimation GJR-GARCH process (equations (5) and (6)) over the period from 1972:01-2014:03. All parameter estimates are obtained from an iterative ML procedure run on daily market returns, where the market is defined as the equal-weighted average of the returns of the 100 largest firms, measured at the beginning of each year.

Joint Estimation of Equations (5) and (6)			
Variable	ML. Est.	std. err.	t-stat
$\mu^{\dagger}$	2.6119	1.0658	2.4505
$\omega_m^{\dagger}$	1.5646	0.3555	4.4008
α_m	0.0152	0.0039	3.9286
γ_m	0.1006	0.0171	5.8844
β_m	0.9207	0.0102	90.3474

†The coefficients and std. errors for μ are $\times 10^4$, and for ω are $\times 10^6$.

We estimate this in the following way. First, using the set of Fama-MacBeth regression results plotted in Figure 3, we estimate β_r. Specifically, fitting an an exponential decay specification over lags between $k = 2$ days and 15 days (3 weeks) yields:

$$\hat{\boldsymbol{\beta}}^{(k)} = -\gamma e^{-\lambda \cdot k} \tag{13}$$

where $\gamma = 0.061$ and $\lambda = 0.292$, for $k \geq 2$, This gives a $\beta_r = e^{-\lambda} = 0.715479$, and a half-life of 2.4 days. We use this specification in our subsequent tests.

Note that given this value of β_r, and with a history of residual returns ($u_{i,t}$ in equation (4)) for each firm, we can calculate using equation (7) the loading $B_{i,t}$ for each firm. With these loadings, we can then calculate the weights in equation (12) for a portfolio with unit loading on the factor.[7]

Next, we simultaneously estimate equations (9) and (11) using maximum likelihood, using the calculated values of $\tilde{u}_{B1,t+1}$, and the calculated cross-sectional variance (in equation (10). The results of this estimation are presented in Table 3.

While the parameters here mostly appear reasonable, and the *ex-ante* ML estimates appear to capture the time-series and cross-sectional variation pretty well.

For the cross-sectional variability, Figure 5 shows the daily realized cross-sectional volatility and the estimated *ex-ante* $h_{\epsilon,t}$. Rather than just presenting this over the period for which we estimate the model, this shows the forecast and realized cross-sectional

---

[7]Also, see footnote 6.

Table 3: **Maximum Likelihood Estimates of Individual Firm Model Parameters**

This table presents the estimates, standard errors and t-statistics for the model parameters in equations (7), (9) and (11). All estimates are obtained from an iterative ML procedure run on daily returns from the 100 largest market capitalization at the start of each year. The sample period is January 2, 1974 through March 28, 2013. Note that $\hat{\lambda}_0$ is the starting value of λ_t that maximizes the the likelihood function.

param.	ML-est.	std. err	t-stat
$\hat{\beta}_r$	0.751470		
$\hat{\lambda}_0$	0.084259	0.046678	1.805094
α_λ	0.997765	0.001298	768.672674
$\kappa_\epsilon^\dagger$	4.741034	1.109324	4.273805
α_ϵ	0.723503	0.021953	32.956604
μ_ϵ	0.260749	0.019914	13.093699

$\dagger$The coefficient and std. error for κ_ϵ are $\times 10^6$.

volatility over the full period for which we have data. Figures 6 and 7 zoom in on realized cross sectional volatility and $h_{\epsilon,t}$ in two extreme periods, the financial crisis and around the market crash of 1987.

The result of the joint estimation of λ_t, and of the volatility of λ_t are presented in Figure 8.

Figure 9 presents – again over the full sample for which we have data – the rolling 252-day mean return, along with the model-forecast mean. Figure 10 presents the 252-day rolling Sharpe-ratio (i.e., the annualized mean return scaled by the volatility. Figure 11 presents the rolling Sharpe-ratios for the top-100 NYSE firms by market capitalization, and for the top-100 NASDAQ firms by market cap.

One concern is the value of α_λ, which is close to 1, suggesting the presence of a unit root. For this reason, in the next section, we explore two alternative ways of estimating λ_t: first, modeling it with a UC-ARIMA (Watson 1986) process and estimating this with a Kalman filter, and second estimating this with a Hodrick and Prescott (1980) filter.

4   Estimating Reversal Magnitude Time Variation

Fitting a stationary process to the expected return of the str strategy suggests the presence of a unit root. We stochastically detrend the returns of the str strategy using two
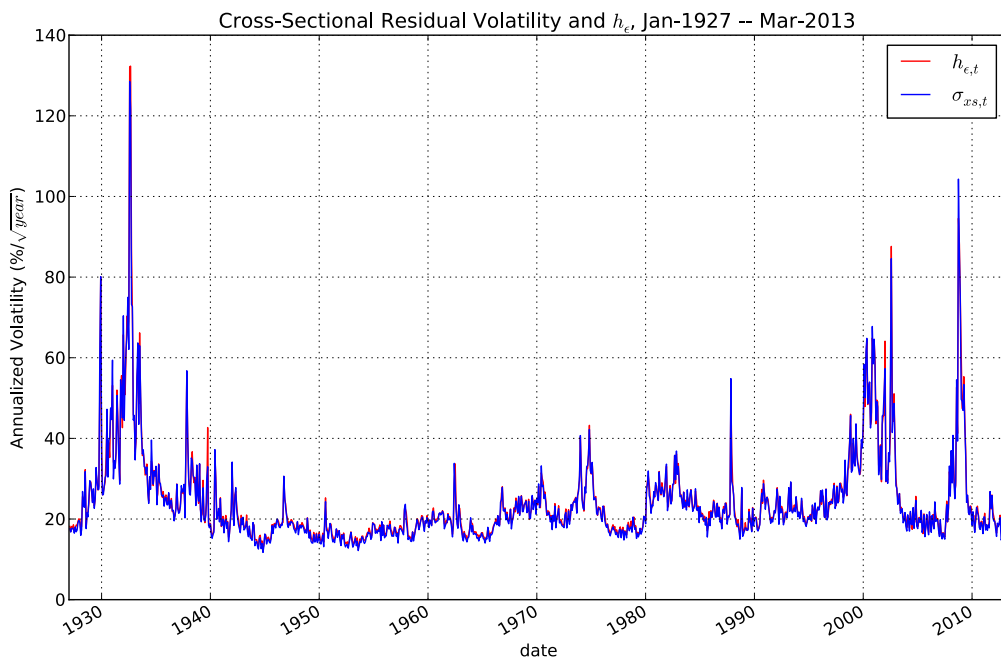
Figure 5: **Realized Cross-Sectional Vol and $h_{\epsilon,t}$**

approaches: the UC-ARIMA approach advocated by Watson (1986) and Stock and Watson (1988), and a one sided-Hodrick and Prescott (1980) filter. What we find is that both capture the slow variation in expected returns over both the post-1974 sample and the full 1929-2014 sample.

4.1 UC-ARIMA

4.1.1 UC-ARIMA Specification

Based on this analysis, our premise is that the expected return of the str portfolio obeys the following stochastic process:

$$\lambda_t = \delta + \lambda_t^{\tau} + \lambda_t^{c} \tag{14}$$

$$\lambda_t^{\tau} = \lambda_{t-1}^{\tau} + u_t^{\tau} \tag{15}$$

$$\lambda_t^{c} = \sum_{k=1}^{p} \phi_k \lambda_{t-k}^{c} + u_t^{c} \tag{16}$$

That is, we assume the time-t premium λ_t is integrated, and can be decomposed into two components: an integrated "trend" and a stationary "cycle." (our terminology here
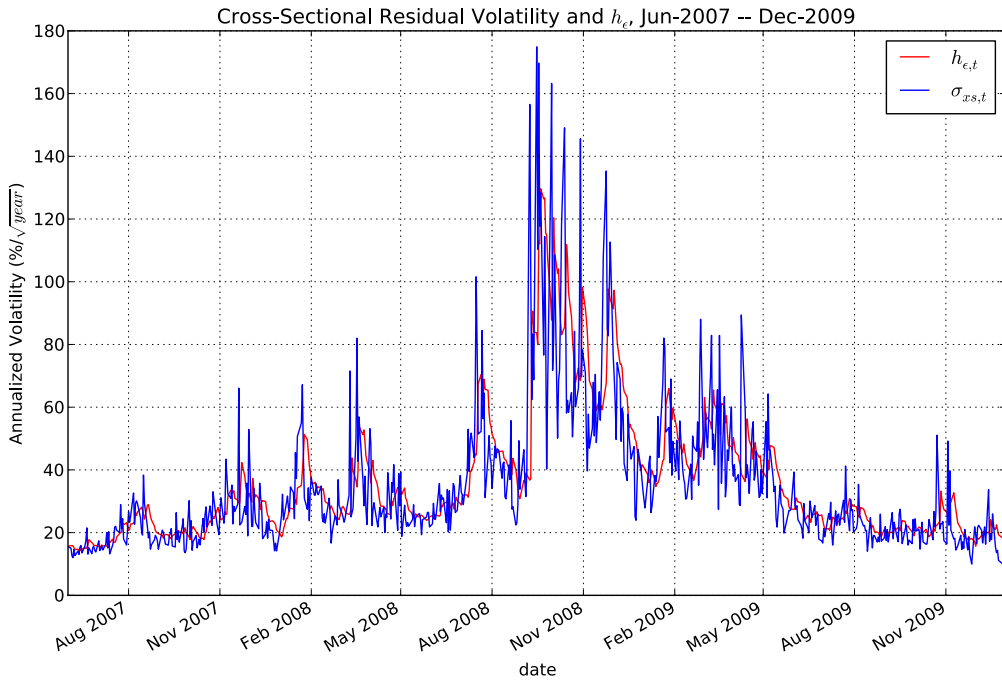
17

Figure 6: **Realized Cross-Sectional Vol and $h_{\epsilon,t}$ in the Financial Crisis**
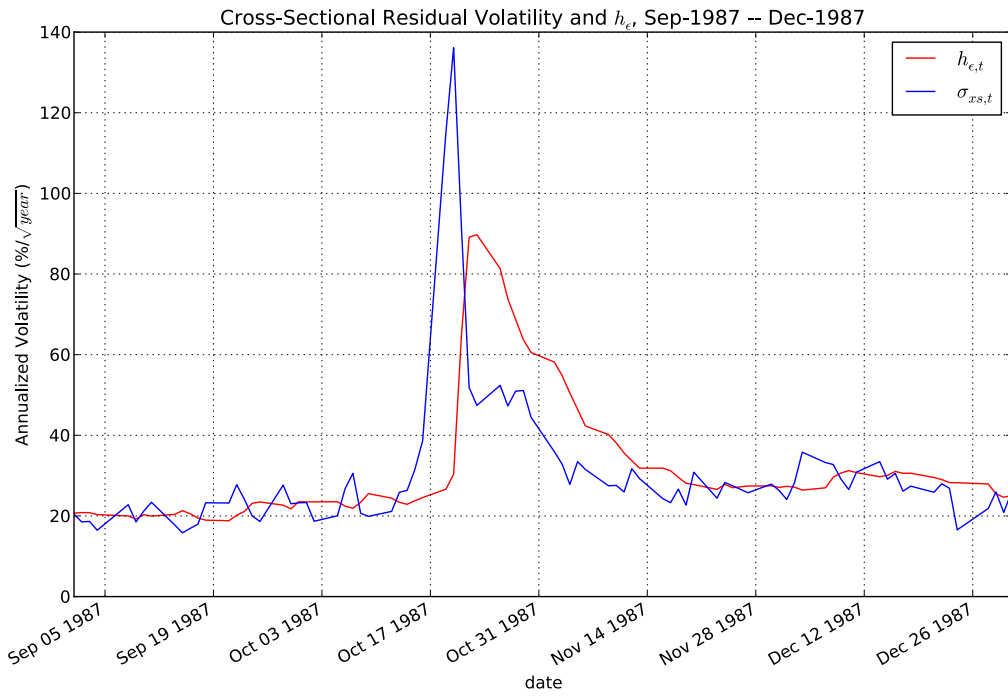

Figure 7: **Realized Cross-Sectional Vol and $h_{\epsilon,t}$ – September-December 1987**
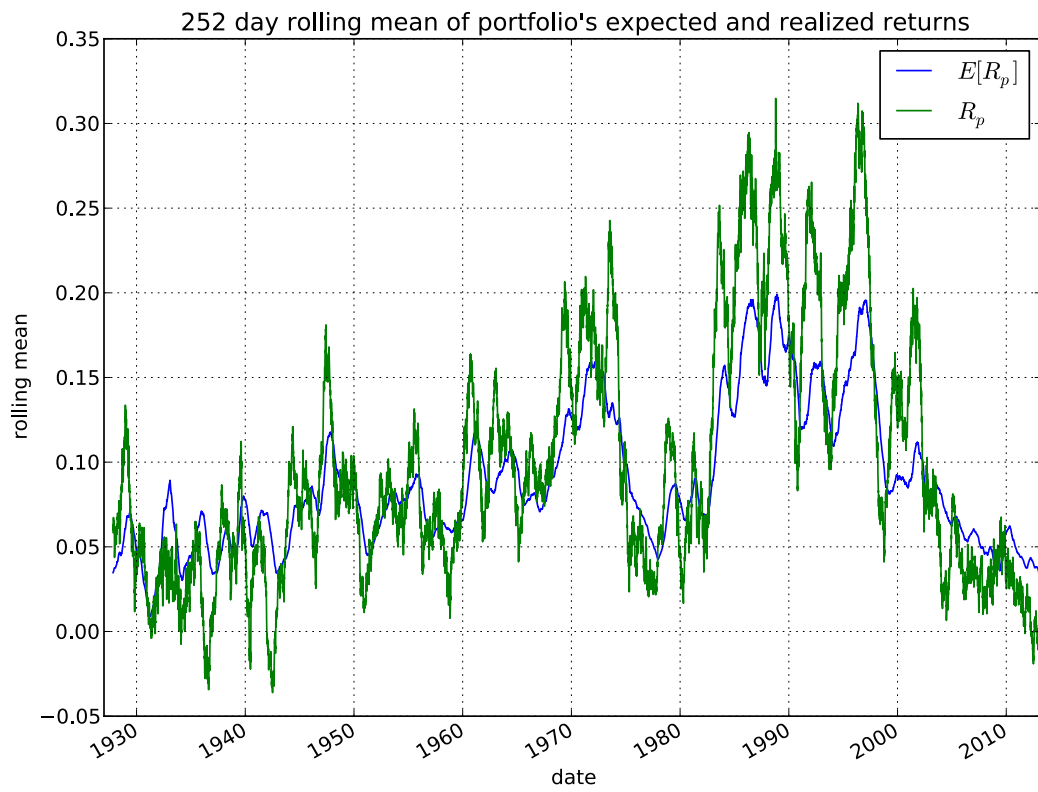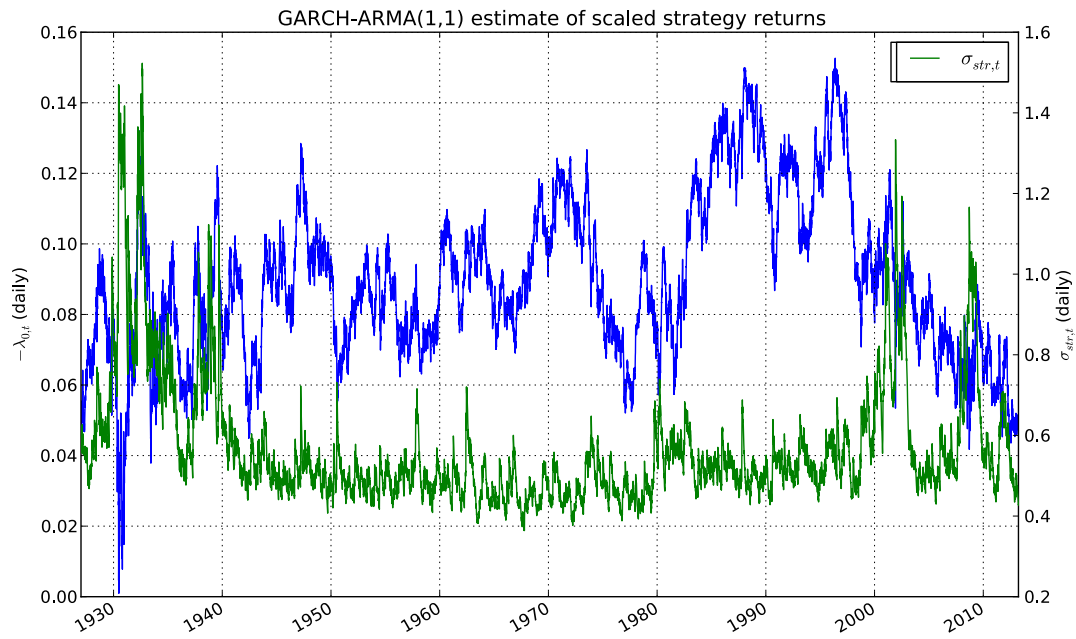
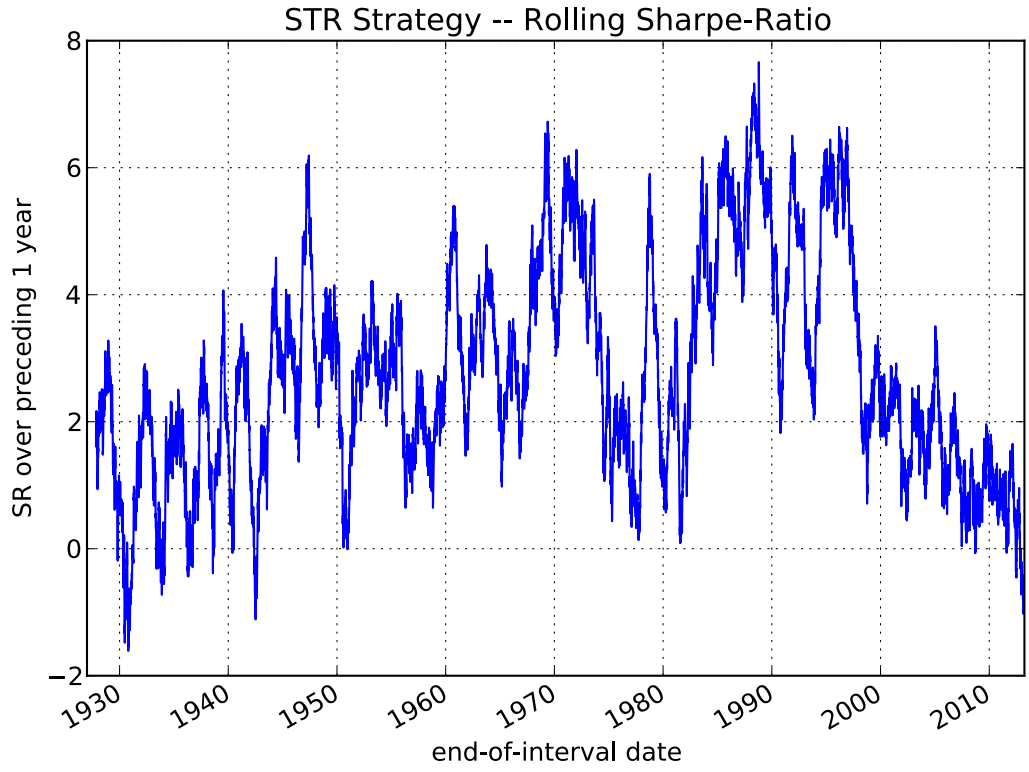Figure 9: **Time variation in the forecast and realized strategy returns**

Figure 10: **STR Strategy − Rolling Sharpe-Ratio**
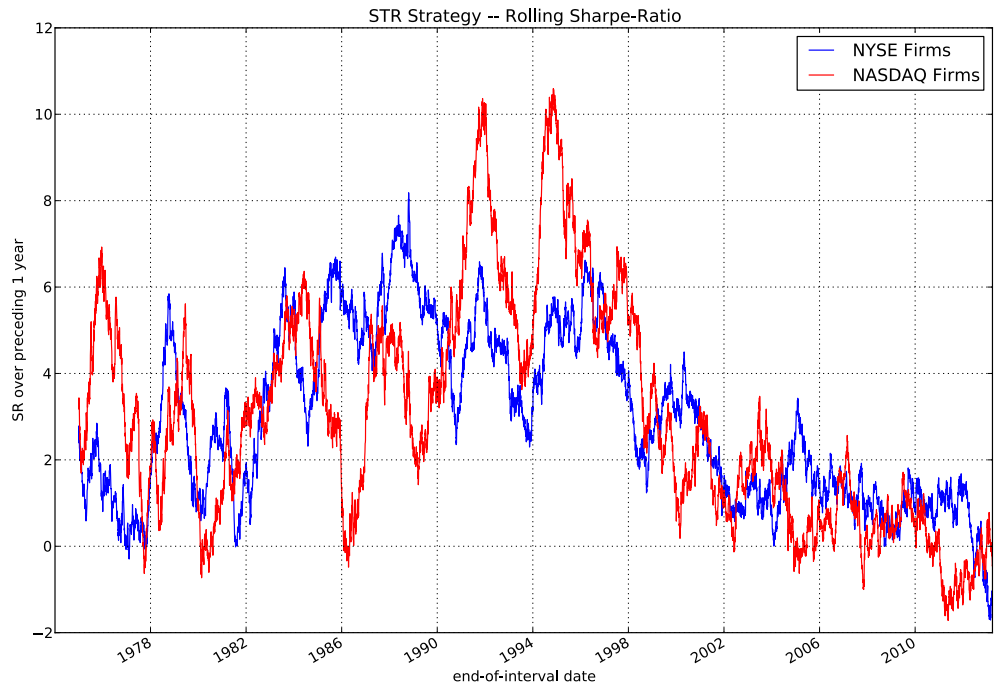

Figure 11: **STR Strategy − Rolling Sharpe-Ratio: NYSE and NASDAQ firms**

comes from the business cycle literature). The trend component λ^τ follows a random walk with drift. The stationary component is here specified to follow a stationary $\mathrm{AR}(p)$ process.

We don't observe either the two components of the expected return – λ_t^τ and λ_t^c – or their sum λ_t. Instead we observe the return to a portfolio for which the expected return is λ_t:

$$r_t^{str} = \lambda_t + \epsilon_t \tag{17}$$

4.2 Kalman Filter Estimation

We use a Kalman filter to estimate the system of equations (14)-(17), and to provide *ex-ante* estimates of λ_t^τ and λ_t^c.

To estimate this we first write equations (14)-(17) in state-space form. Here, we lay this out for the case where λ_t^c follows an $AR(2)$ process.

The state-space formulation has two governing equations. The *state-evolution* equation describes the evolution of the state variables:

$$\mathbf{x}_t = \mathbf{A} x_{t-1} + \mathbf{w}_t \tag{18}$$

where, for our $\mathrm{AR}(2)$ setting:

$$\mathbf{x}_t = \begin{bmatrix} \lambda_t^\tau \\ \lambda_t^c \\ \phi_2 \lambda_{t-1}^c \end{bmatrix}, \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \phi_1 & 1 \\ 0 & \phi_2 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_t = \begin{bmatrix} u_t^\tau \\ u_t^c \\ 0 \end{bmatrix}$$

The first two state variables are the trend and cycle component of the premium. The final state variable, $\phi_2 \lambda_{t-1}^c$, is essentially a "place-keeper" for the lagged component of the premium; note that the final equation is of the system is an identity.

The shocks are necessarily mean zero and conditionally normal:

$$E_{t-1}[\mathbf{w}_t] \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}\right).$$

21

We assume the trend and cycle shocks are uncorrelated:

$$\mathbf{Q} = \begin{bmatrix} \sigma_\tau^2 & 0 & 0 \\ 0 & \sigma_c^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We specify the relation between the state variables and the (observable) return to the str portfolio r_t^{str} with the *measurement equation*:

$$r_t^{str} = \mathbf{G}\mathbf{x}_t + \epsilon_t. \tag{19}$$

where $\mathbf{G}$ is specified to make the conditional expected return the sum of the two components. The unexpected return ϵ_t is uncorrelated with any of the elements of $\mathbf{w}_t$ and is here specified to have constant variance R, though it is straightforward to incorporate a GARCH process for the variance.

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \epsilon_t \sim \mathcal{N}(0, R)$$

4.2.1 Kalman Filter Updates

The filtering step: Suppose that the distribution of the state variable $\mathbf{x}_t$, conditional on having observed all returns up through time $t-1$ (*i.e.*, $r_1^{str}, r_2^{str}, \ldots, r_{t-1}^{str}$) but *not* r_t^{str}, is given by:

$$\mathbf{x}_{t|t-1} \sim \mathcal{N}\left(\hat{\mathbf{x}}_{t|t-1}, \boldsymbol{\Sigma}_{t|t-1}\right)$$

The posterior distribution, conditional of r_t^{str}, is given by:

$$\mathbf{x}_{t|t} \sim \mathcal{N}\left(\hat{\mathbf{x}}_{t|t}, \boldsymbol{\Sigma}_{t|t}\right) \tag{20}$$

where:

$$\begin{aligned} \hat{\mathbf{x}}_{t|t} &= \hat{\mathbf{x}}_{t|t-1} + \boldsymbol{\Sigma}_{t|t-1}\mathbf{G}'\left(\mathbf{G}\boldsymbol{\Sigma}_{t|t-1}\mathbf{G}' + R\right)^{-1}\left(r_t^{str} - \mathbf{G}\hat{\mathbf{x}}_{t|t-1}\right) \\ \boldsymbol{\Sigma}_{t|t} &= \boldsymbol{\Sigma}_{t|t-1} - \boldsymbol{\Sigma}_{t|t-1}\mathbf{G}'\left(\mathbf{G}\boldsymbol{\Sigma}_{t|t-1}\mathbf{G}' + R\right)^{-1}\mathbf{G}\boldsymbol{\Sigma}_{t|t-1}. \end{aligned}$$

The forecast step: With the distribution of $\mathbf{x}_{t|t}$ from equation (20) and the state evolution equation (18) we can construct the distribution for the str portfolio return for

22

period $t+1$ by first "updating" the forecast of the state variable x_t to create a distribution for x_{t+1}, and then based on this distribution calculating the distribution for r_{t+1}^{str}.

First, the distribution of $\mathbf{x}_{t+1|t}$ is:

$$\mathbf{x}_{t+1|t} \sim \mathcal{N}\left(\hat{\mathbf{x}}_{t+1|t}, \mathbf{\Sigma}_{t+1|t}\right) \tag{21}$$

where the mean and the covariance matrix come directly from equation (18):

$$\hat{\mathbf{x}}_{t+1|t} = \mathbf{A}\hat{\mathbf{x}}_{t|t} \quad \text{and} \quad \mathbf{\Sigma}_{t+1|t} = \mathbf{A}\mathbf{\Sigma}_{t|t}\mathbf{A}' + \mathbf{Q}$$

From equations (19) and (21), the distribution of the time $t+1$ str return is:

$$r_{t+1}^{str} \sim \mathcal{N}\left(\mathbf{G}\hat{\mathbf{x}}_{t+1|t}, \mathbf{G}\mathbf{\Sigma}_{t+1|t}\mathbf{G}' + R\right) \tag{22}$$

4.2.2   Estimating the Kalman Filter Parameters

We begin with a prior distribution for the state variables. Using equation (20) applied to the parameters of the prior and the first return, we update the state variable distribution parameters. We then apply equations (21) and (22) to forecast the next return. We apply this method for each subsequent return.

Note that sequentially applying the Kalman filter gives us a time series of *ex-ante* distributions (mean and variance) for r_t^{str}. To estimate the parameters of the Kalman filter, we simply apply maximum likelihood to the series of forecasts and realizations, selecting the the set of process parameters that maximize the log-likelihood function. For the AR(2) example laid out above, the parameter-vector has elements: $\{\phi_1, \phi_2, \sigma_\tau^2, \sigma_c^2, R\}$.

We apply the Kalman filter to returns from 1970:01-2014:03. However, the log-likelihood calculation use only returns from 1974:01-2014:03. That is, we discard first 4 years of forecasts and returns. The logic for this is that we don't want our prior to have a strong effect on the estimated parameters. Via trial and error, we found that after 1000 observations the effect of the prior was minimal.

The MLE parameters, along with standard errors and t-statistics are presented in Table 4, and a plot of the *ex-ante* expected return and its components, along with the 252-day rolling mean of the r_t^{str} series, are plotted in Figure 12. 12.

Several things are of interest. First, while the parameters of the estimated cyclical component appear reasonable, the magnitude is small. Consequently, it doesn't provide
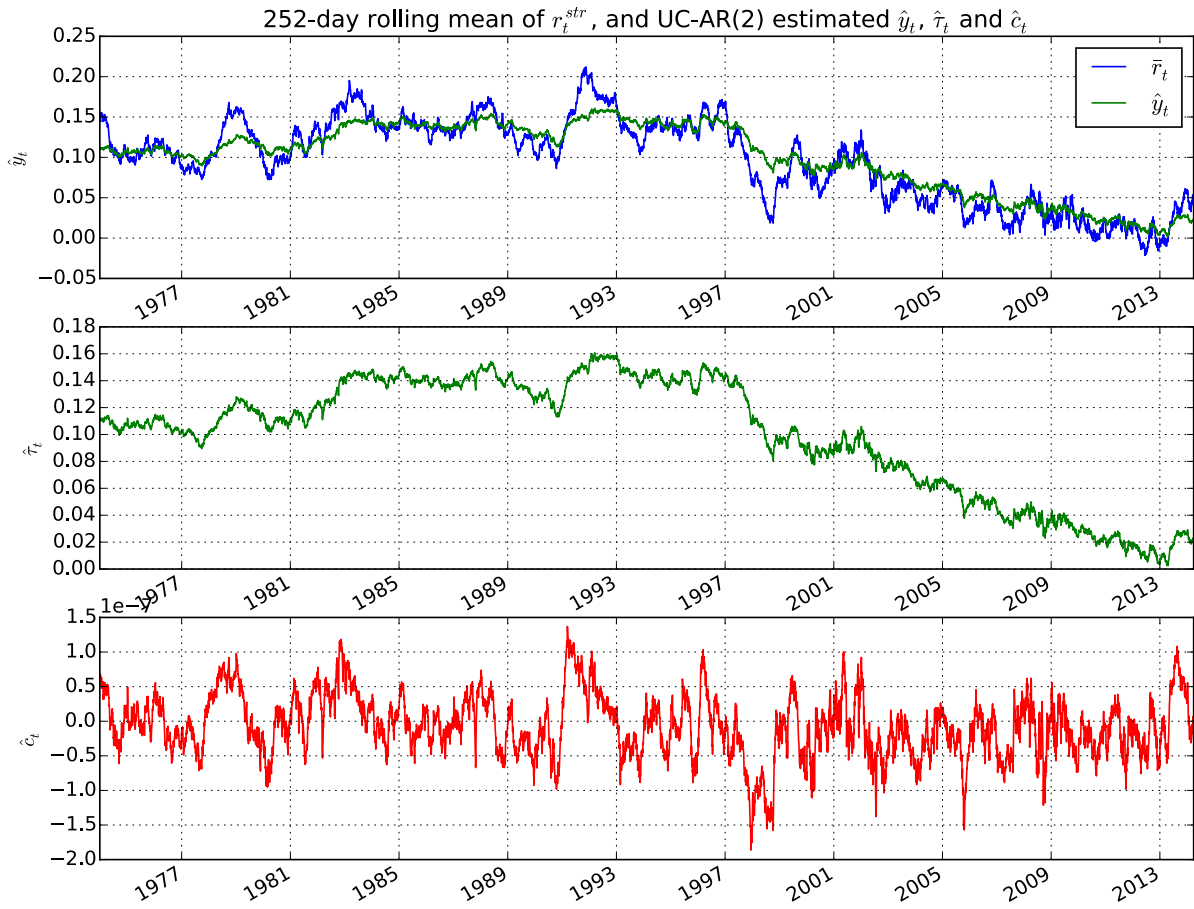
23

Figure 12: Kalman-Filter Estimated Components of the Expected Return of the r_t^{str} Portfolio, for a UC-AR(2) Model

24

Table 4: **Results of Estimation of UC-AR(2) Process for r_t^{str}**

param.	ML-est.	std. err	t-stat
ϕ_1	0.9862	5.10e-04	1930.7
ϕ_2	0.0031	4.58e-06	675.2
$\sigma_\tau^\dagger$	0.0803	5.50e-03	14.6
σ_c	0.0008	1.24e-05	61.9
σ_ϵ	38.33	4.71e-01	81.4

† ML estimates and standard errors for σ_τ, σ_c and σ_ϵ are $\times 100$.

much help in forecasting str returns: in either univariate regressions or multivariate (with τ_{t-1} as the other RHS variable), it is statistically insignificant.

Secondly, based on Figure 12, it looks like the estimated stochastic trend captures variation in the average returns of the str portfolio well. However, since what is plotted in Figure 12 is the rolling mean return of the str portfolio, this could be because of the averaging. To verify this, if we run a regression of r_t^{str} on τ_{t-5}, gives a t-statistic of 11.28.[8]

4.3 Hodrick-Prescott Filter

As a robustness check, we also use a one-sided version of the Hodrick and Prescott (1980, 1997) filter to estimate the expected return of the str portfolio. The HP filter has frequently been used in the macro-economics literature as a technique for extracting the stochastic-trend of integrated business-cycle variables. The HP-trend component of a time-series (τ_t) is the solution to:

$$\min_{\tau_t} \left(\sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} \left[(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}) \right]^2 \right) \tag{23}$$

That is, the HP-filter selects the trend τ_t so as to find the best fit to the series y_t, subject to a penalty series of second derivatives of the trend. The penalty coefficient λ is chosen so as to achieve a sufficiently smooth. (King and Rebelo 1993) note that the HP filter is a low-pass filter, and HP-trend-component contains the low-frequency components of the original series y_t. Consistent with this Ravn and Uhlig (2002) show the value of λ will be dependent on the sampling frequent for y_t, and should be proportional to the fourth

---

[8]Note, if we use the constant volatility strategy, the t-statistics rises to 15.78.

power of the observational frequency. We select λ based on spectral analysis, detailed below.

4.3.1 Calculating the one-sided HP Filter Weights

A matrix representation of equation (23) is:

$$\min_{\tau'}(\mathbf{y} - \boldsymbol{\tau})'\mathbf{I}(\mathbf{y} - \boldsymbol{\tau}) + \lambda\boldsymbol{\tau}'\mathbf{A}'\mathbf{A}\boldsymbol{\tau} \tag{24}$$

where $\mathbf{y}$ and $\boldsymbol{\tau}$ are the $T \times 1$ vectors of time-series observations and the extracted trend. $\mathbf{I}$ is a $T \times T$ identity matrix, and $\mathbf{A}$ is a block diagonal $(T-2) \times T$ matrix with the form:

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & & \ldots & 1 & -2 & 1 \end{bmatrix}$$

Differentiating equation (24) with respect to $\boldsymbol{\tau}'$ and rearranging gives the first order condition for the HP-trend $\boldsymbol{\tau}$ that satisfies (24):

$$\boldsymbol{\tau} = (\mathbf{I} + \lambda\mathbf{A}'\mathbf{A})^{-1}\mathbf{y}$$

A straightforward way to solve this problem is to invert the block diagonal matrix $(\mathbf{I} + \lambda\mathbf{A}'\mathbf{A})$. Note that each row of the inverted matrix represents the coefficients of a moving-average filter that can be applied to the original series y_t to obtain the HP-trend τ_t.

Two rows of the matrix $(\mathbf{I} + \lambda\mathbf{A}'\mathbf{A})^{-1}$ are plotted in the upper panels of Figures 13 and 14: the "middle" row (here $N/2$ for $N = 10000$) and the final row. In addition, in the lower panels of the two Figures, we plot the spectrum of the two filters. Recall that, since a convolution in the time domain is a multiplication in the frequency domain. To HP-filter our str portfolio return data, we effectively convolve the original return series with the filter weights. An equivalent way to do this in the frequency domain is to the the Fourier transform of the return series by the Fourier transform of the filter, and then inverse Fourier transform. This will give exactly the HP filtered series. Therefore, to see the effect of the filtering, it can be useful to see the filter spectrum.

For both the two sided filter and the one sided filter, the smoothing parameter λ was
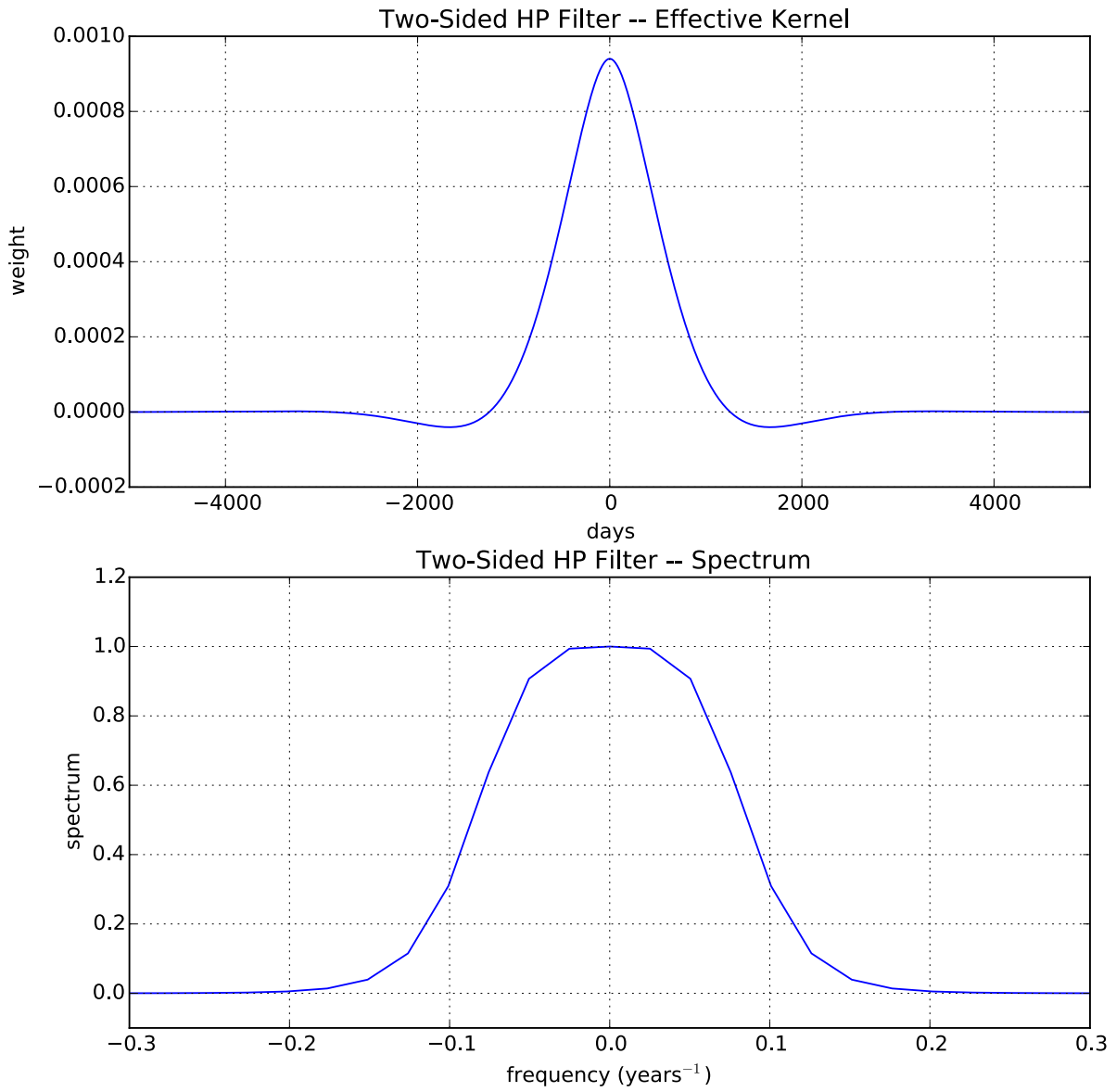
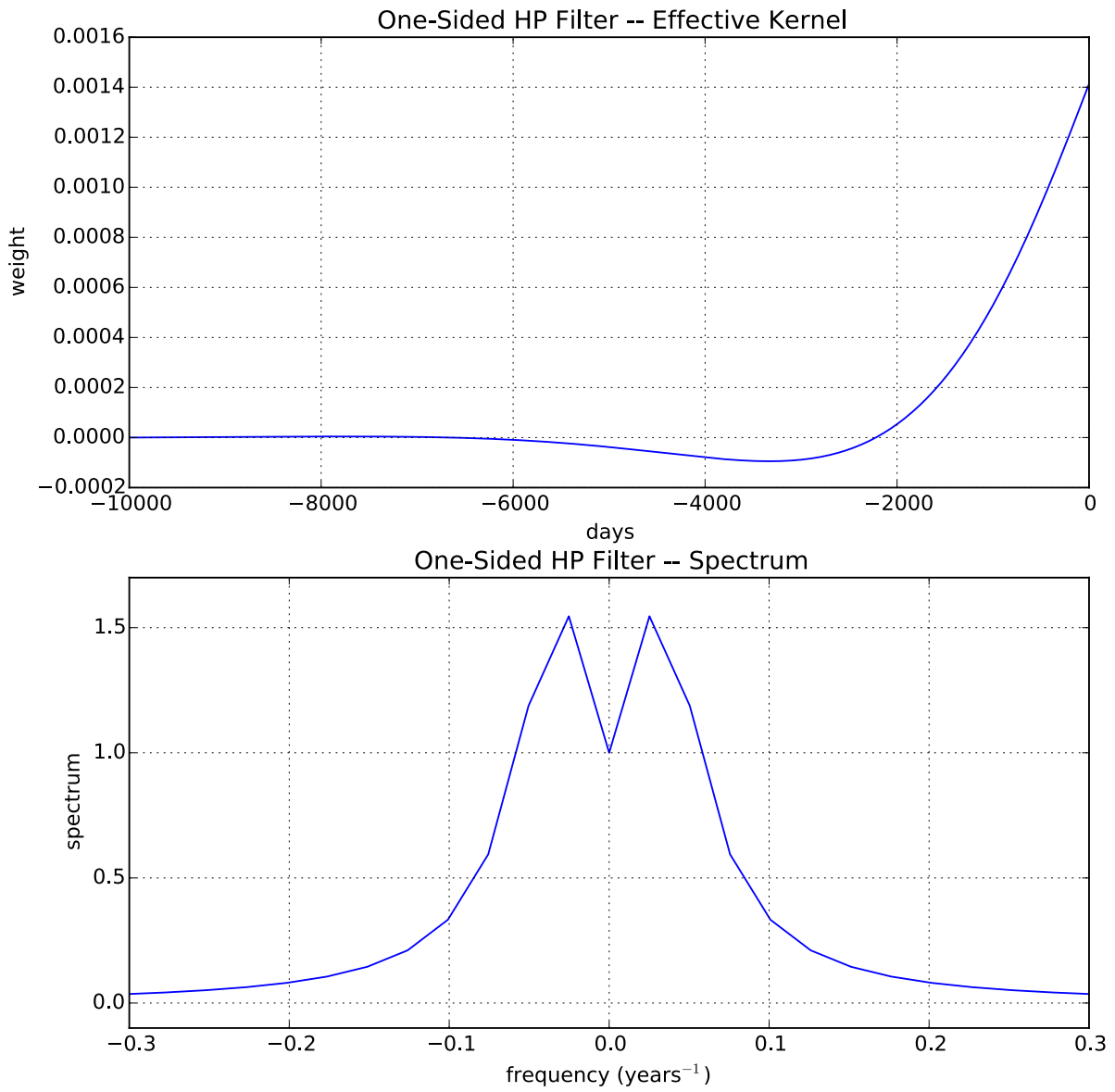Figure 13: **HP filter – Effective Kernel and Spectrum**

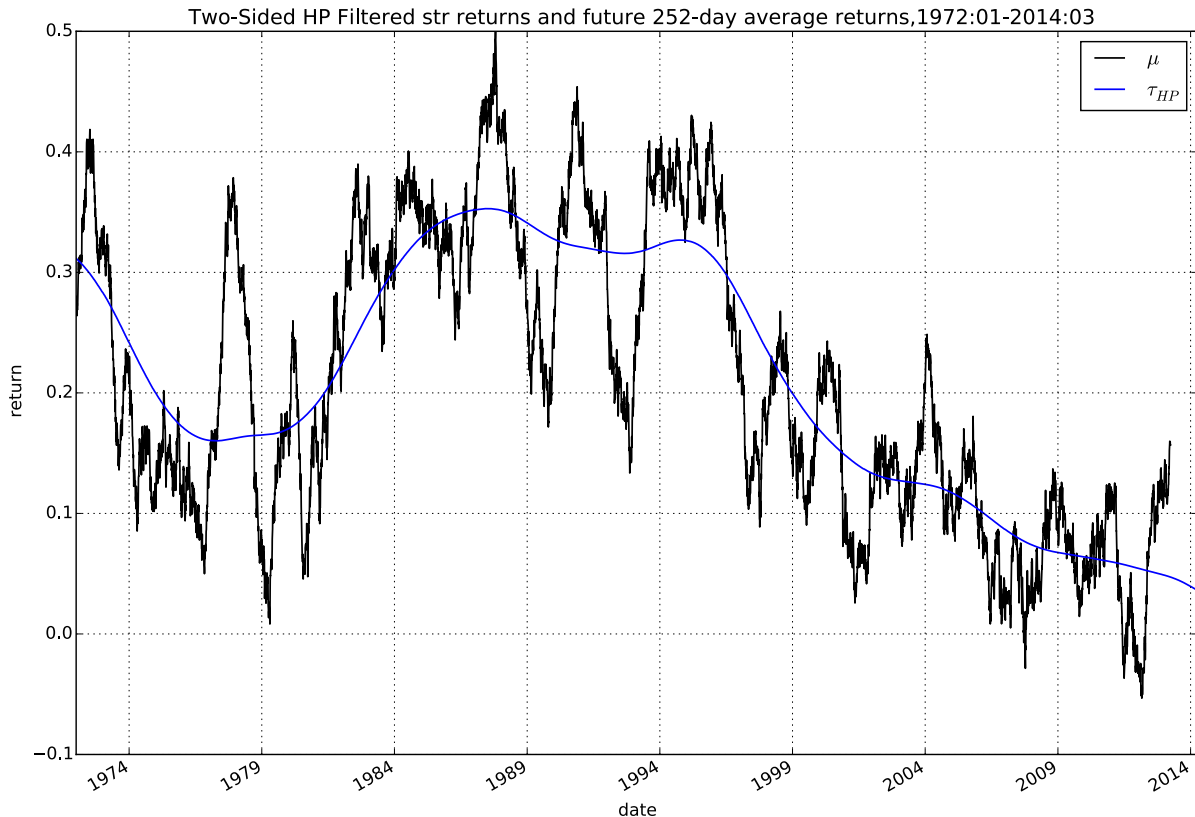Figure 14: **One-Sided HP filter – Effective Kernel and Spectrum**

Figure 15: **HP filtered str Returns**

chosen so the the returns components below a frequency of $(1/5)$ years^{-1} are filtered out. Note that the one-sided filter (the last row of the matrix) has a less appealing spectrum (*i.e.*, it has less of the box shape of a band-pass filter), and also suffers from phase distortions (not plotted). However, because it is one-sided, it generates estimates of the underlying stochastic trend based only on lagged returns, while the two-sided filter uses both past and future return. For this reason, for our application we will use the one-sided filter.

4.4   Results

5   Strategy Time Variation

While most of the empirical work in Nagel (2012) concentrates on time variation of the average *return* of the short-horizon reversal strategy with predictive variables, Nagel notes that his model predicts that the conditional Sharpe ratio associated with liquidity provi-
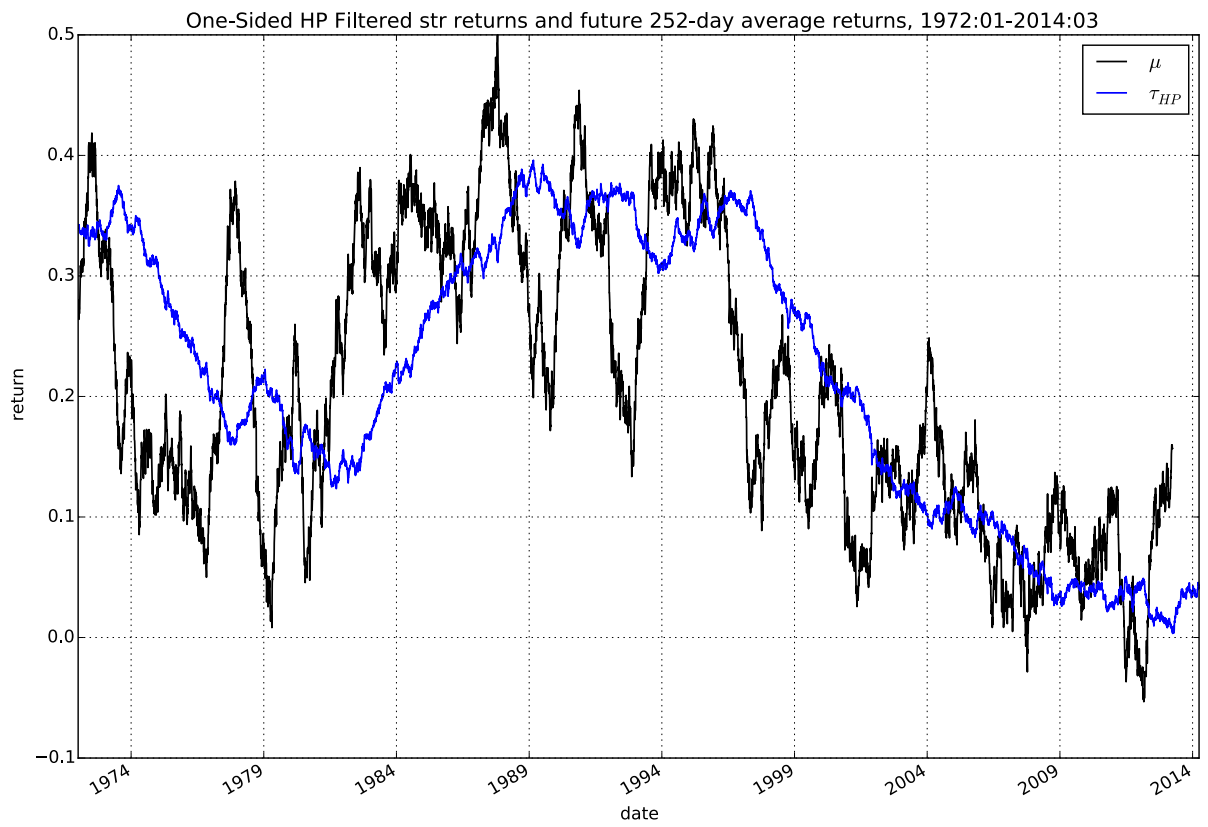
Figure 16: **One-Sided HP filtered str Returns**

sion strategies – such as the short-term reversal strategy that he explores – is likely to be affected by shocks which affect financially constrained intermediaries.

He therefore also examines time variation in the *compensation for risk* (*i.e.,* the Sharpe Ratio) of his 5-day STR strategy. He proceeds in the following way:[9]

First, he estimates time variation in conditional volatility associated with the VIX by running the regression specified in his equation (22):

$$|\tilde{L}_t^R|\kappa = a_0 + a_1 \cdot \text{VIX}_{t-5} + a_2 \cdot d_{t-5} + u_t \tag{25}$$

where $\tilde{L}_t^R$ is the residual from the regression of daily reversal strategy returns on VIX_{t-5} and d_{t-5} (his decimalization dummy).[10] The fitted values from this regression at each time are used as estimator of σ_t, which is used in the regression

$$\frac{L_t^R}{\sigma_t} = b_0 + b_1 \cdot \text{VIX}_{t-5} + b_2 \cdot d_{t-5} + e_t, \tag{26}$$

and the time series of fitted values from this regression are interpreted as the conditional Sharpe ratio of the reversal strategy return.

This is clearly a much different approach than the one that we take here. First, our baseline str portfolio is much different than the one Nagel employs: . Second, to estimate the volatility of our str process, we fit a GARCH process to the residuals of the

This is an important distinction. Recently, Kelly, Lustig, and Van Nieuwerburgh (2012) document that cross-sectional volatility and market volatility are not particularly highly correlated, suggesting that the VIX is probably not a particularly good forecast of the volatility of the short-term reversal strategy – something we will show directly below.

Second, if the VIX is only poor proxy for the true volatility, this will lead to an upward biased estimate of the correlation between the Sharpe ratio of a return process and the VIX. Intuitively while the fitted regression in equation (25) may be unbiased, when we scale the strategy return by the fitted σ_t (in equation (26) we are scaling by a convex function of the estimated volatility. Thus errors in estimated volatility result in a upward biased estimate of the return of the scaled process (that is, $E\left[\frac{1}{\sigma+\tilde{\epsilon}}\right] > \frac{1}{\sigma}$). When the error variance is higher for larger levels of the VIX, this will result in a positive estimated relation between the level of the VIX and the return Sharpe ratio, even the true Sharpe

---

[9]See Nagel (2012), p. 2025-26.

[10]d_t takes a value of one prior to decimalization (April 9, 2001) and a value of zero thereafter
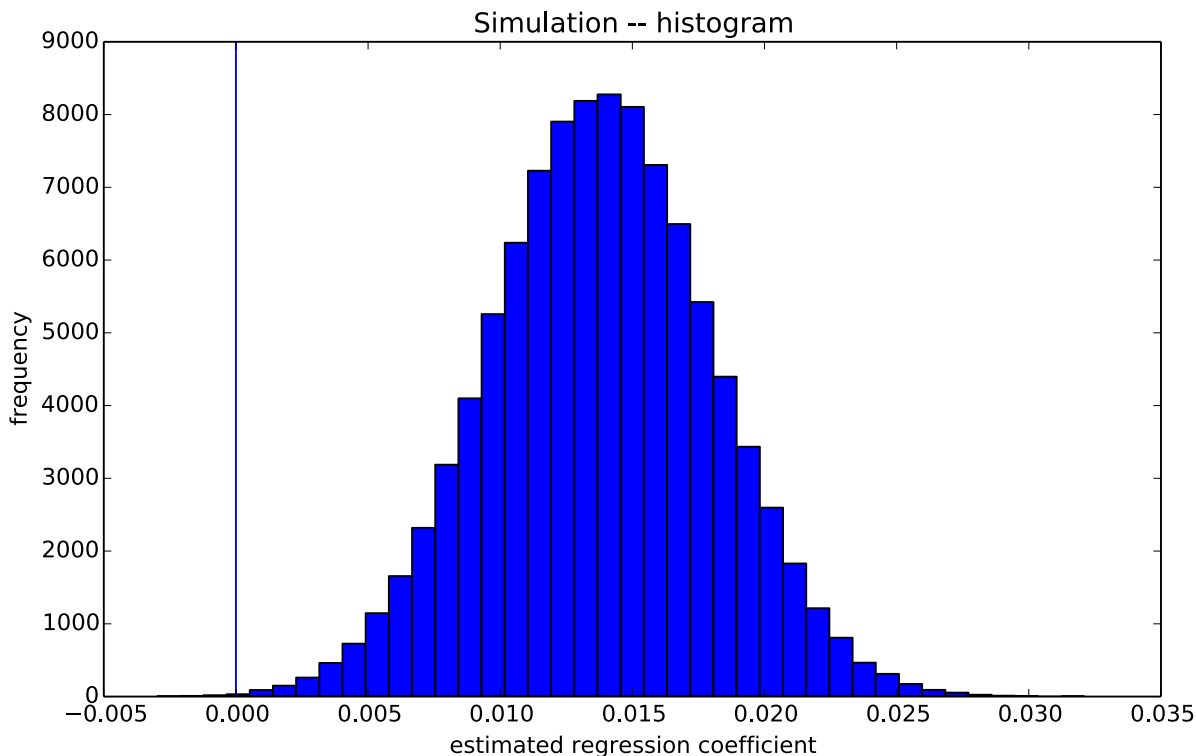
This histogram is for the coefficients obtained in regressing the estimated Sharpe-ratio, using the technique described above, on the simulated level of the VIX, which in the simulation is a noisy proxy for the true return volatility. There were 100,000 runs in the simulation, and 3276 (= 13×252) time periods. The correlation between the simulated VIX and the true volatility was 61%.

Figure 17: **Simulation Histogram**

ratio is actually constant.

In the Appendix, we demonstrate the existence of this bias via simulation. Figure 17 plots the histogram of the estimated relationship between the VIX and the strategy Sharpe-ratio (using the simulation procedure outlined above), when the true Sharpe ratio is constant.

Next, here are several regressions that further illustrate our points. Here, more consistent with Nagel, we utilize a strategy constructed from the 500 largest capitalization firms each year. The strategy we examine is an equal-weighted strategy which goes long the $\sim$250 firms with negative weighted residuals, and short the $\sim$250 firms with positive weighted residuals. Also, we match Nagel's sample period..

First we regress the returns of the strategy on the dummy variable and on the VIX.

$$r_t^{str-ew} = a + b \cdot \text{VIX}_{t-5} + c \cdot d_{t-5} + e_t$$

Variable	Coef	Std Err	t-stat
d	0.0010	0.0003	3.19
VIX	0.0001	0.0000	4.11
intercept	-0.0008	0.0004	-2.11

Consistent with Nagel, we see a the strongly significant relation between the strategy returns and the VIX. Next, we regress the absolute values of the residuals from this regression on the same two RHS variables:

Variable	Coef	Std Err	t-stat
d	0.0025	0.0002	11.21
VIX	0.0002	0.0000	18.51
intercept	-0.0002	0.0003	-0.96

Note another strong relation between the absolute residuals and the VIX.

Next, we regress the scaled strategy return on the two RHS variables:

Variable	Coef	Std Err	t-stat
d	0.0999	0.0579	1.73
VIX	0.0060	0.0028	2.16
intercept	-0.0150	0.0672	-0.22

Note that the scaled portfolio returns continue to be positively related to the VIX.

Next we examine the ability of the VIX to forecast the volatility of the str series. We begin by regressing the absolute value of the residual returns of our EW strategy on its own lagged one-week, one-month, and 6-month historical volatilities, each lagged 5 days. The regression R^2 is 7.8%, and the coefficients on each of the three variables are significant. Henceforth we use the square-root of the estimated variance, and label it as $\hat{\sigma}_{t-5}$

Next, to illustrate our point we regress first the absolute value of the residual returns on the VIX, and on the level of the VIX and on $\hat{\sigma}$. The regression of the $|r_t^{str-ew}|$ on the VIX yields:

```
         Variable      Coef     Std Err      t-stat
        ------------------------------------------------
              VIX     0.0002      0.0000       19.75
        intercept    -0.0000      0.0003       -0.07
```

with an R^2 of 10.6%. However, when we include $\hat{\sigma}_{t-5}$ as a RHS variable, we obtain:

```
         Variable      Coef     Std Err      t-stat
        ------------------------------------------------
            sigma     0.0005      0.0000       20.91
              VIX     0.0000      0.0000        1.17
        intercept    -0.0012      0.0003       -4.58
```

Consistent with Kelly, Lustig, and Van Nieuwerburgh (2012), the $\hat{\sigma}_{t-5}$ is strongly significant, and the VIX becomes insignificant. The VIX is not a good proxy for cross-sectional volatility, or for the

We now repeat the set of regressions above, but including $\hat{\sigma}_{t-5}$ as a RHS variable in each step. We present only the final regression in which the scaled strategy return is the dependent variable:

```
         Variable      Coef     Std Err      t-stat
        ------------------------------------------------
                d     0.0103      0.0646        0.16
            sigma     0.0177      0.0067        2.63
              VIX    -0.0008      0.0036       -0.23
        intercept    -0.0547      0.0651       -0.84
```

What this regression shows is that, after controlling for the volatility of the str strategy, the VIX no longer has any power to forecast Sharpe-ratio. Interestingly, for this EW strategy, we see that the historical volatility-based estimator $\hat{\sigma}_{t-5}$ does have power to forecast the mean of the scaled strategy.

Finally, Figure 18 plots the VIX and our x-sectional volatility estimator. While the high correlation ($\approx 60\%$) between the two are evident, there are clearly distinct differences in the behavior of the two series. For example, the cross-sectional volatility declines far more quickly after the market decline in 2002-3, and following the financial crisis.

6   Conclusions

We estimate a short term reversal process for daily US equity returns. Over our primary sample period of 1972-2014, and for our sample of the 100 largest traded firms, on average
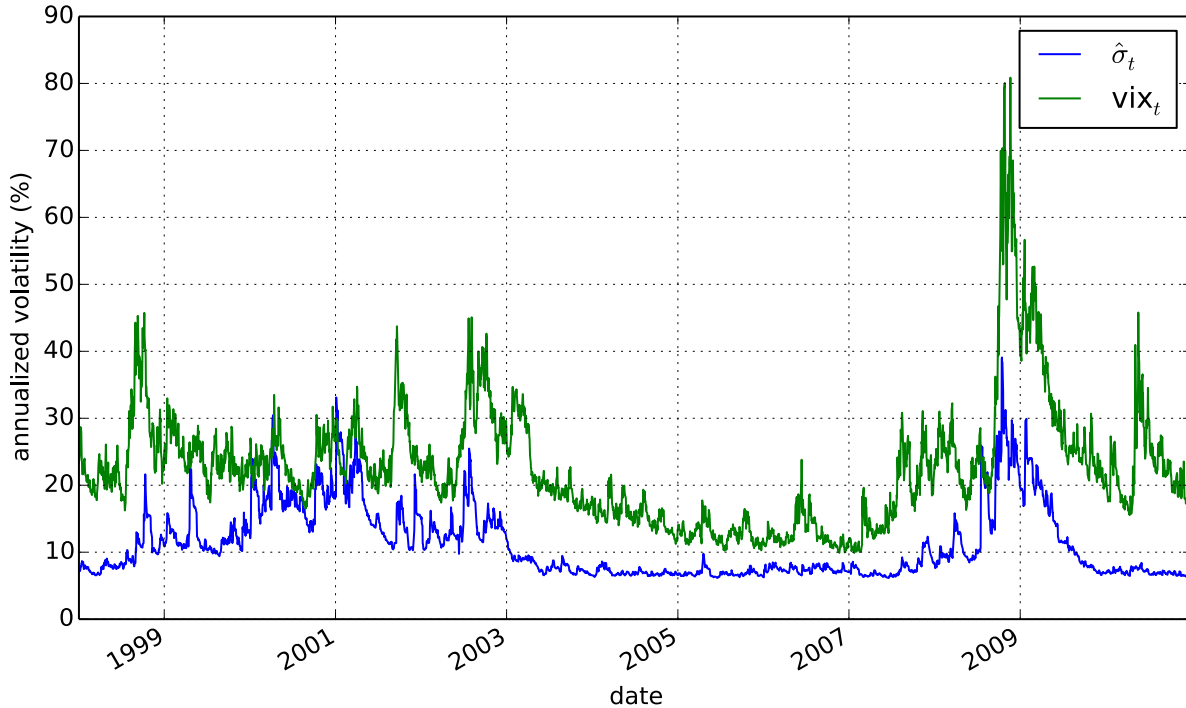
Figure 18: **VIX and Historical Estimates of cross-sectional volatility**

approximately 90% of idiosyncratic price shocks are permanent. The remaining 10% is temporary, and decays exponentially toward zero, with a half life of about 2.5 days. While the rate of decay (the half life) is relatively constant over time, the magnitude decay varies considerably over the sample. Our findings are consistent with the slow movement of capital (Duffie 2010). Also, in contrast with previous literature, we find no evidence that this rate of mean reversion is related to market-wide measures of illiquidity, such as the VIX. Our results are thus also consistent with a lack of integration across capital markets.

The evidence presented here suggests several promising areas for future research. First, it would be interesting to better understand the frictions behind the slow moving capital that appear to result in the slow movement of prices back towards equilibrium levels following an idiosyncratic price shock. Second, we should aim to better understand the nature of the barriers that prevent a flow of capital into, and out of the strategies that attempt to arbitrage the short-term reversal patterns we document here. What is the nature of the barriers that prevent this flow of capital? Also, the very different patterns in cross-sectional and time-series volatility – which has been noted elsewhere in the literature

35

– suggest that a successfully modeling capital flows between sectors/markets may also require better modeling of the volatility transmission mechanisms.

References

Avramov, Doron, Tarun Chordia, and Amit Goyal, 2006, Liquidity and autocorrelations in individual stock returns, *The Journal of Finance* 61, 2365–2394.

Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307–327.

Chan, Wesley S., 2003, Stock price reactions to news and no-news: Drift and reversal after headlines, *Journal of Financial Economics* 70, 223–260.

Da, Zhi, Qianqiu Liu, and Ernst Schaumburg, 2013, A closer look at short-term return reversal, *Management Science,* forthcoming.

Duffie, Darrell, 2010, Presidential Address: Asset price dynamics with slow-moving capital, *Journal of Finance* 65, 1237–1267.

Fama, Eugene F., and James MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Frazzini, Andrea, and Lasse H. Pedersen, 2013, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Glosten, Lawrence R., Ravi Jagannathan, and David Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

Grossman, Sanford J, and Merton H Miller, 1988, Liquidity and market structure, *the Journal of Finance* 43, 617–633.

Hodrick, Robert, and Edward C. Prescott, 1980, Postwar U.S. business cycles: An empirical investigation, Discussion Paper 451, Carnegie-Mellon University.

Hodrick, Robert J., and Edward C. Prescott, 1997, Postwar us business cycles: an empirical investigation, *Journal of Money, Credit, and Banking* pp. 1–16.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Kelly, Bryan, Hanno N. Lustig, and Stijn G. Van Nieuwerburgh, 2012, The volatility factor structure, Chicago Booth Research paper 12-54.

——— , 2013, Firm volatility in granual networks, Chicago Booth Research paper.

Khandani, Amir E., and Andrew W. Lo, 2007, What Happened To The Quants in August 2007?, *Journal of Investment Management* 5, 5–54.

King, Robert G., and Sergio T. Rebelo, 1993, Low frequency filtering and business cycles, *Journal of Economic Dynamics and Control* 17, 207–231.

Lehmann, Bruce N., 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1–28.

Nagel, Stefan, 2012, Evaporating liquidity, *Review of Financial Studies* 25, 2005–2039.

Ravn, Morten O, and Harald Uhlig, 2002, On adjusting the Hodrick-Prescott filter for the frequency of observations, *Review of Economics and Statistics* 84, 371–376.

Richardson, Scott, Pedro Saffi, and Kari Sigurdsson, 2014, Deleveraging risk, London Business School.

Stock, James H., and Mark Watson, 1988, Variable trends in economic time series, *Journal of Economic Perspectives* 3, 147–174.

Tetlock, Paul C., 2011, All the news that's fit to reprint: Do investors react to stale information?, *Review of Financial Studies* 24, 1481–1512.

Watson, Mark W., 1986, Univariate detrending methods with stochastic trends, *Journal of Monetary Economics* 18, 9–75.