Liquidity and Return Reversals

Pierre Collin-Dufresne[†] and Kent Daniel[‡]

Preliminary Draft: October 20, 2013
This Draft: May 28, 2015

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 reversal magnitude 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 the magnitude of the temporary component (and the profitability of the reversal strategy) is related to market-wide measures of illiquidity, such as the VIX. Thus, our results are 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 (Lehmann 1990, Jegadeesh 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. The evidence of Da, Liu, and Schaumburg (2013) suggests that industry returns exhibit weaker reversals than firm-specific reversals.[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 argues that this covariation is 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, and implicitly with strong integration of capital markets.

We make several contributions to this literature. First, we show that the reversal effect is remarkably strong in even the largest and most liquid stocks. Second, we estimate an impulse response function for price shocks, and show that this function is well captured by a model in which, in response to a shock to demand, a individual stock's price overshoots,

---

[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.

and then exponentially decays to the new equilibrium price with a half-life of about 2.5 days. This slow reversion of the price is consistent with slow-moving capital (Duffie 2010). Finally we show that while there is considerable variation over time in the both the magnitude of the overshooting, and in the magnitude of the Sharpe-ratio of the associated short-term-reversal strategy. This time variation is unrelated to standard measures of market turmoil such as the VIX, but are related to market specific measures such as the volatility of the short-term reversal strategy itself. Thus, this evidence is suggestive of a model in which arbitrageurs have limited capital, but where arbitrage capital does not freely flow across markets in response to changing investment opportunities.

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 of the short-term-reversal process, the estimation of this model, and some econometric issues related to this estimation. Section 4 concentrates specifically on the estimation of this 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 measures of market turmoil. Section 6 concludes.

1   Data Description

Our sample is the 100 largest firms by market capitalization in the universe of US common stocks, over the period from January 1972 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. Both the choice

---

[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.
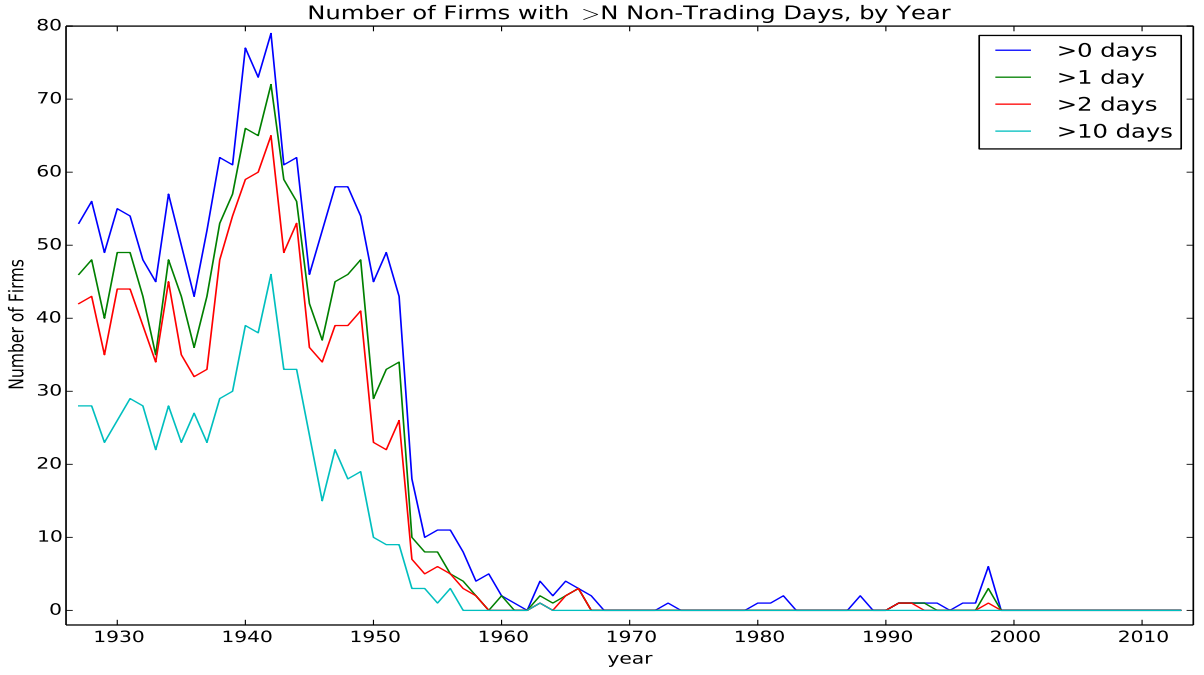
This plot shows the number of firms in our sample of 100 firms which experienced at least N non-trading days, by year. The first four lines in the plot give the number of firms that experienced more than 0, 1, 2, and 10 non trading days, by calendar year.

Figure 1: **Non-Trading Analysis**

of the starting date for sample, and the restriction of our universe to the 100 largest firm are chosen to ensure that each 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 per year where they do not trade, and about a quarter of the sample has more than 10 non-trading days per year. One reason for this is that, pre-1950, the number of traded firms is considerably smaller. A second and more important reason is that share turnover is considerably smaller pre-1960.

4

Value-weighted, annualized turnover for top 100 US stocks, Jan-1927 -- Dec-2013
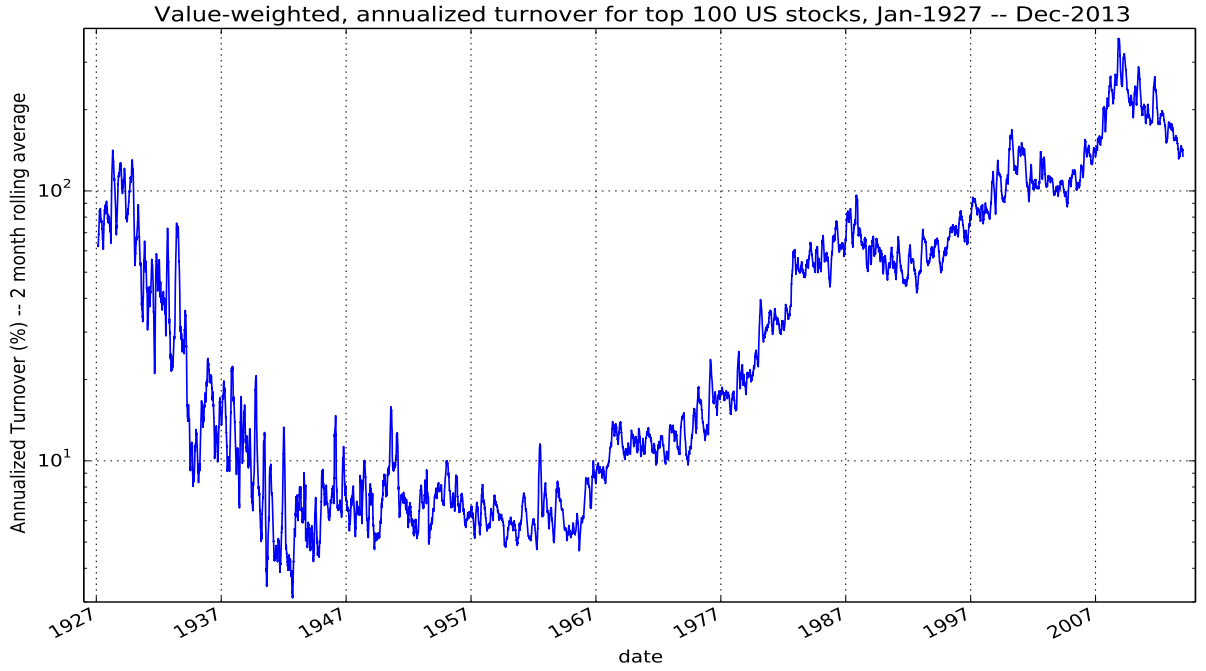
This figure plots the 42-day (2-month) moving average of the annualized value-weighted turnover of the 100 stocks in our sample, in percent/year. Turnover is defined as trading volume dividend by the total number of shares outstanding.

Figure 2: **Average Annualized Turnover**

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 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]

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 of the total number of delistings from 1927-2013 are in the two decades between 1980 and 2000, and most of these are a result of a merger. 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.

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 (∼42.5 years) of earnings announcement dates.

---

[3]The peak is the 42-day period ending on November 4, 2008, 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.

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 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 define the residual return for firm i on trading day t as:

$$\tilde{u}_{i,t} \equiv \tilde{r}_{i,t} - \beta_i^s \, \tilde{r}_{m,t},$$

where $\tilde{r}_{i,t}$ and $\tilde{r}_{m,t}$ denote the excess returns of firm i and the market, respectively, on trading day t, and β_i^s is the shrunk beta for that year, calculated according to equation

(1).

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:

$$u_{i,t} = \gamma_{0,t} + \sum_{\tau=1}^{30} \gamma_{\tau,t} \cdot u^*_{i,t-\tau} + \tilde{\nu}_{i,t} \tag{2}$$

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

$$u^*_{i,t-\tau} = \begin{cases} 0 & \text{if} \quad t - \tau \in [t_i^{EAD}-1, t_i^{EAD}+1] \\ u_{i,t-\tau} & \text{otherwise} \end{cases} \tag{3}$$

where t_i^{EAD} is any COMPUSTAT earnings announcement date for firm i.

We zero out the firm residual returns on earnings announcement dates for both theoretical and empirical reasons. Our hypothesis is that short-term reversal results from reversal of price movements resulting from liquidity-based trades. From a theoretical perspective, price movements on earnings-announcement-dates are largely information driven, so we would expect weaker reversals, or even continuation. Also, this is consistent with the results of the empirical analysis as documented in Section 2.2 below.

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

$$u_{i,t} = \gamma_{0,t}^{\text{NZ}} + \sum_{\tau=1}^{30} \gamma_{\tau,t}^{\text{NZ}} \cdot u_{i,t-\tau} + \tilde{\nu}_{i,t}^{NZ} \tag{4}$$

where here, the NZ superscript denotes that the residual returns used as independent variables are not set to zero in the 3-day window around earnings announcement dates.
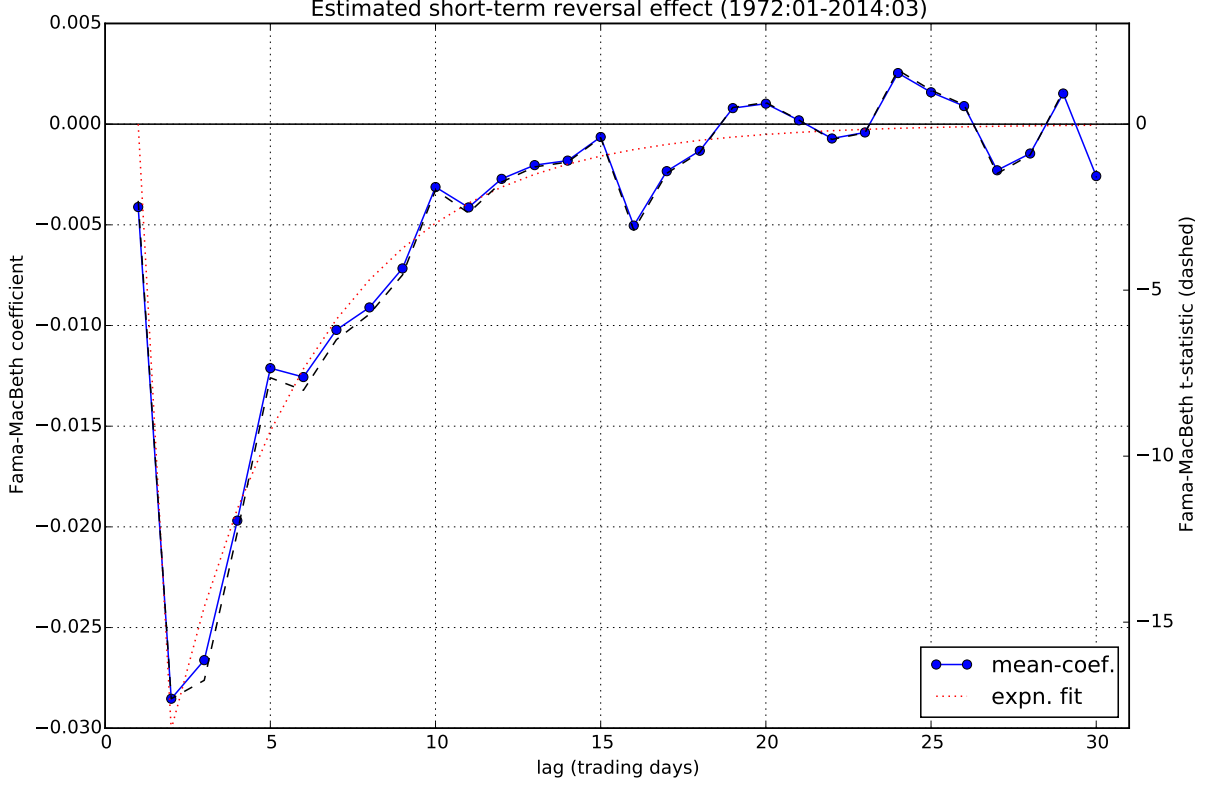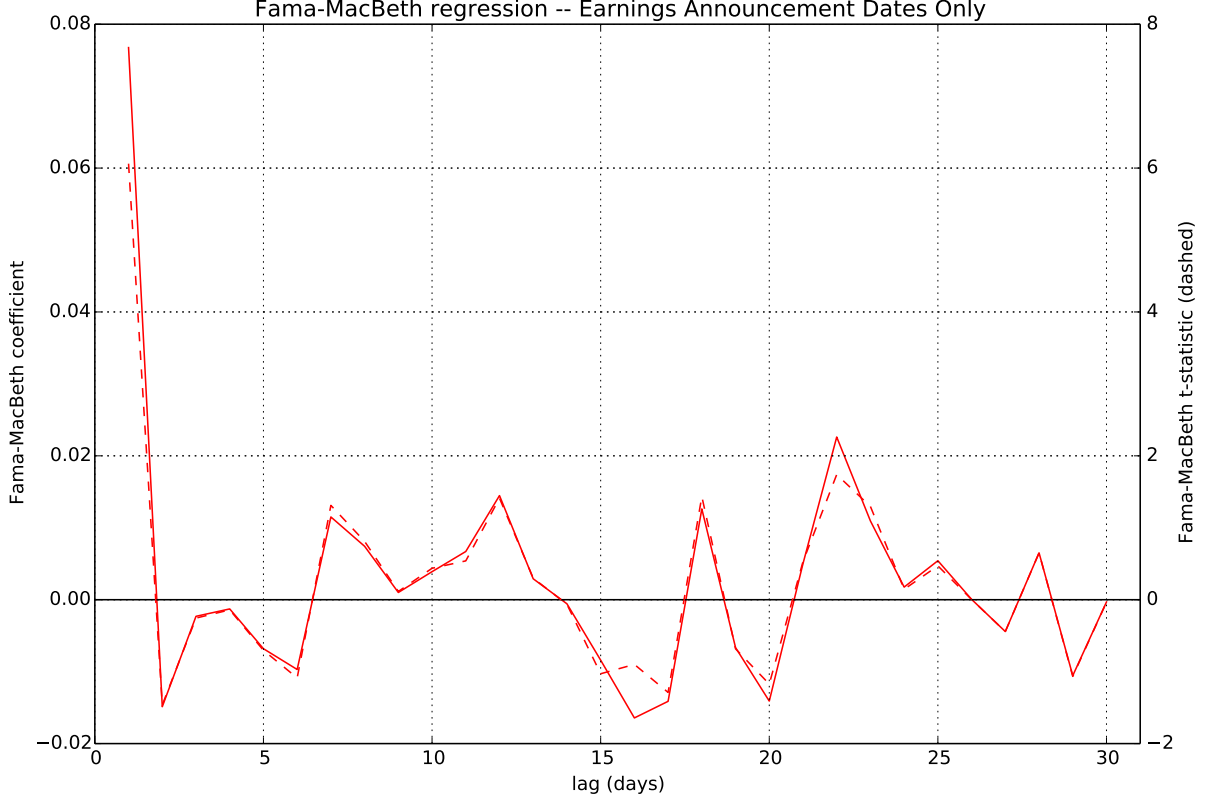
This figure plots the average Fama-MacBeth coefficients, the Fama-MacBeth t-statistics for a Fama-MacBeth regression of the residual returns (calculated as described in the text) on the lagged residual returns for the 100 largest common stocks traded on US exchanges. Lagged dependent returns which occur within one day of the an earnings announcement date (per COMPUSTAT) are set to zero.

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

Figure 3 plots the time-series mean of the cross-sectional regression coefficients $\hat{\gamma}_{\tau,t}$ in equation (2) as a function of τ, and also plots 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. Also, 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.

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.

Figure 4: **Fama-MacBeth Results - Earnings Announcement Dates Only**

2.2 Earnings Announcement Dates

As noted above, we zero out the lagged returns in the forecasting regression when the lagged return occurs on within 1 trading day of an earnings announcement date. There is both a theoretical and a empirical motivation for this. First, from a theoretical perspective, there should be no reversal for purely information based price moves. Since on earnings announcement dates price moves are more directly linked to the new information released with the earnings announcement, and therefore proportionately less associated with liquidity shocks, there should be proportionately less reversal. Second, we show via several empirical analyses that, consistent with this hypothesis, return reversal is indeed smaller 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 within 1 trading day of a COMPUSTAT earnings announcement date, as specified above. Figure 4 shows that, with the exception of the one-day lag return, each of the coefficients one are statistically indistinguishable from zero at the 5% level. Interestingly, the one-day lagged return is *positive* and highly statistically significant.

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 from the equation (2) cross-sectional regressions that, for a given lag, on the corresponding coefficients with the alternative equation (4) cross-sectional regressions. 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 in more detail below, our baseline short-term-reversal portfolio – the returns of which we label r_t^{str} – weights each asset in proportion to an exponentially weighted sum of that asset's lagged residual returns. However, if any of those lagged returns are within one trading day of an earnings announcement, the weight on the lagged return is set to zero. As will be discussed below, the weights are selected with the objective of constructing a mean-variance efficient portfolio, and zeroing out the EAD-returns is consistent with this objective if EAD returns have no power to forecast future returns. Over the 1972-2013 period, our baseline strategy has an unconditional annualized Sharpe ratio of 3.08.

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 (4) where the lagged residuals returns $u_{i,t-\tau}$ are not set to zero on EADS. We also present the intercept and the t-statistic for the reverse regression – which we denote α_τ^{NZ}.

τ	α_τ	$t(\alpha_\tau)$	α_τ^{NZ}	$t(\alpha_\tau^{\text{NZ}})$
1	0.449	6.70	-0.479	-6.97
2	-0.101	-1.67	-0.313	-5.00
3	-0.105	-1.80	-0.274	-4.51
4	-0.024	-0.41	-0.252	-4.18
5	-0.052	-0.86	-0.132	-2.14
6	0.062	1.05	-0.256	-4.19
7	0.026	0.44	-0.187	-3.09
8	-0.016	-0.28	-0.106	-1.81
9	0.022	0.39	-0.126	-2.11
10	0.020	0.35	-0.066	-1.12

To test this hypothesis that EAD-returns are don't forecast future returns, we construct a alternative trading strategy with returns $r_t^{str-\text{NZ}}$ where we do not zero out EAD-returns. 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 directly test the mean-variance efficiency of the two strategies, we run the regressions:

$$r_t^{str} = \alpha + \beta \cdot r_t^{str-\text{NZ}} + u_t$$

$$r_t^{str-\text{NZ}} = \alpha^{\text{NZ}} + \beta^{\text{NZ}} \cdot r_t^{str} + u_t$$

We find that $\alpha = 93$ bps/day ($t = 7.0$). and $\alpha^{\text{NZ}} = -36$ bps/day ($t = -2.8$). This simple test confirms that zeroing out past-residual returns around EADs leads to a more efficient portfolio. However, the fact that α^{NZ} is negative suggests that a still more efficient portfolio can be obtained by actually *shorting* a small amount of the alternative NZ trading strategy. Over this sample, the *ex-post* efficient combination would put a small *positive*

12

weight on firms which experience positive returns on earnings announcement dates –
consistent with the presence of a small post-earnings announcements drift. 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.[4]

3   A Model of Return Reversals

In this section, we construct an econometric model of short-term reversal, where the
model specification is consistent with the empirical findings of our data characterization
in Section 2.

3.1   Model Specification

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

$$\tilde{r}_{i,t+1} = \beta_{i,m}\tilde{r}_{m,t+1} + \underbrace{B_{i,t}\lambda_t + \sigma_i h_{u,t}\tilde{\epsilon}_{i,t+1}}_{\equiv \tilde{u}_{i,t+1}} \qquad\qquad \tilde{u}_{i,t+1} \sim \mathcal{N}(0,1) \qquad (5)$$

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 $\beta_{i,m}$ on a single common factor (the
market). $u_{i,t+1}$ denotes firm i's residual return at time $t+1$.

3.1.1   Expected return specification:

As noted earlier, we specify that a firm's residual return $u_{i,t+1}$ is not conditionally mean
zero. Rather, the has a conditional expected residual return is negatively correlated with

---

[4]The strategy which puts the ex-post optimal (positive) weight on returns around EADs has an
annualized Sharpe-ratio of 3.11.

its lagged residual returns for lags of up several weeks. Equation (5) 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. Consistent with our analysis in Section 2, we specify that the exposure $(B_{i,t})$ is governed by an autoregressive process:

$$
\begin{aligned}
B_{i,t} &= \beta_r B_{i,t-1} + (1 - \beta_r)\tilde{u}^*_{i,t-1} \\
&= (1 - \beta_r)\sum_{l=1}^{\infty} \beta_r^l u^*_{i,t-l}
\end{aligned}
\tag{6}
$$

Firm i's time t exposure is thus an exponentially weighted sum of its lagged daily residuals, starting at time $t-1$.[5] 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.[6] For our sample of the largest 100 firms, over our 1972-2014 time period, our estimation in Section 3.2 yields an average λ_t of -0.12. However, as we will see shortly, there is considerably time variation in λ_t This means that, following a residual shock of +1%, the price of a firm will fall (starting in one day) by about 12 basis points over the next few weeks. Our estimation of equation (6) for daily returns gives a $\hat{\beta}_r = 0.720$, corresponding to a half-life of this price decay of 2.4 days.

3.1.2 Variance Process Specification:

Equation (5) specifies that the volatility of firm i's residual return is $\sigma_i h_{\epsilon,t}$ – the product of a time-invariant 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

---

[5]Consistent with equation (3), $u^*_{i,t-l}$ is set to zero whenever $t-l$ is within a three day window around an COMPUSTAT earnings announcement date for firm i, consistent with our analysis in Section 2.2.

[6]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 (See the analysis in Section 1).

argue that time variation in individual firm idiosyncratic volatilities are largely captured by a single factor structure. Here, each firm's idiosyncratic volatility is a product of a firm-specific time-invariant term σ_i and a common level term $h_{\epsilon,t}$.

Equation (5) specifies that the shock to firm i's residual returns is

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

This means that:

$$
\begin{aligned}
\mathbb{E}_t \left[\frac{1}{n} \sum_{i=1}^{n} (\tilde{u}_{i,t+1} - B_{i,t}\lambda_t)^2 \right] &= h_{\epsilon,t}^2 \mathbb{E}_t \left[\frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 \tilde{\epsilon}_{i,t+1}^2 \right] \\
&= h_{\epsilon,t}^2 \left(\frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 \right) \\
&= h_{\epsilon,t}^2
\end{aligned}
\tag{7}
$$

That is, given our restriction that $\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 = 1$, and the restriction 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 (7).)

We specify that $h_{\epsilon,t}^2$ follows a GARCH(1,1) process (Bollerslev 1986):

$$h_{\epsilon,t}^2 = \kappa_\epsilon + \alpha_\epsilon h_{\epsilon,t-1}^2 + \mu_\epsilon \sigma_{xs,t}^2, \tag{8}$$

where $\sigma_{xs,t}^2$ denotes the realized cross-sectional variance in period t:[7]

$$\sigma_{xs,t}^2 \equiv \frac{1}{n} \sum_{i=1}^{n} (\tilde{u}_{i,t} - B_{i,t-1}\lambda_{t-1})^2, \tag{9}$$

---

[7]Note that this specification is equivalent to:

$$h_{\epsilon,t}^2 = \kappa_\epsilon + (\alpha_\epsilon + \mu_\epsilon)h_{\epsilon,t-1}^2 + \mu_\epsilon(\sigma_{xs,t}^2 - h_{\epsilon,t-1}^2).$$

where now the last term is the time-t cross-section variance minus the time-$t-1$ expectation of the cross-sectional variance. From Table 2, the point estimate for $\alpha_\epsilon + \mu_\epsilon$ is 0.98425, implying a half-life for cross-sectional variance shocks of 44 trading days.

and, from equation (7), $h_{\epsilon,t}^2 = \mathbb{E}_t\left[\sigma_{xs,t+1}^2\right]$. For an individual firm, equation (5) 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 (5). Empirically, we find considerable cross-sectional variation in σ_i.

3.1.3 Premium Specification:

From equation (5), 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:

$$\lambda_{t+1} = \alpha_\lambda \lambda_t + (1-\alpha_\lambda)\tilde{u}_{B_1,t+1} \tag{10}$$

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, with $B_{p,t} = 1$), and specifically the portfolio with time-t weights:[8]

$$w_{i,t}^{B_1} = \frac{B_{i,t}}{\sum_i B_{i,t}^2} \tag{11}$$

λ_t can 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 (10) can be interpreted as describing the evolution of this expectation. It is a reduced form for the Kalman filter solution to

---

[8]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.

the system of equations:

$$\lambda^*_{t+1} = \lambda^*_t + \tilde{v}_{t+1}$$

$$\tilde{u}_{B_1,t+1} = \lambda^*_t + \tilde{\epsilon}_{B_1,t+1}$$

where λ^*_{t+1} is the latent/unobserved premium.

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 (6). 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, λ_t is the magnitude of the reversal that is expected to follow a unit residual return. λ_t varies over time, and the AR(1) parameter α_λ governs how quickly shocks to λ_t mean revert. As we discuss below, our MLE estimate of α_λ (in equation (10)) is 0.9978, corresponding to a half-life of approximately 1 year. So the magnitude of the firm-specific residual effect moves very slowly in comparison to the price reversal itself.

3.2 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 estimate of the parameters of the individual firm process (equations (6), (10) and (8)) in the following way. First, using the set of mean Fama-MacBeth coefficient and the associated standard errors (as plotted in Figure 3) we estimate β_r. We do this by fitting an an exponential decay specification over lags between $k = 2$ days and 15 days (3 weeks) yields:

$$\hat{\boldsymbol{\beta}}^{(k)} = -ae^{-b \cdot k} \tag{12}$$

17

Table 2: **Maximum Likelihood Estimates of Individual Firm Model Parameters**

This table presents the estimates, standard errors and t-statistics for the model parameters in equations (6), (8) and (10). 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 $\times 10^6$.

where $a = 0.061$ and $b = 0.332$, for $k \geq 2$, This gives a $\beta_r = e^{-b} = 0.715470$, 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 (5)) for each firm, we can calculate using equation (6) the loading $B_{i,t}$ for each firm. With these loadings, we can then calculate the weights in equation (11) for a portfolio with unit loading on the factor.[9] Next, we simultaneously estimate equations (8) and (10) using maximum likelihood, using the calculated values of $\tilde{u}_{B1,t+1}$, and the calculated cross-sectional variance (in equation (9). The results of this estimation are presented in Table 2. The parameters here appear reasonable, and the *ex-ante* ML estimates appear to capture the time-series and cross-sectional variation pretty well.

Figure 5 plots the daily realized cross-sectional volatility and the ML-estimated $h_{\epsilon,t}$. Rather than just presenting this over the period for which we estimate the model, this shows the forecast and realized cross-sectional volatility over the 1927-2013 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. $h_{\epsilon,t}$
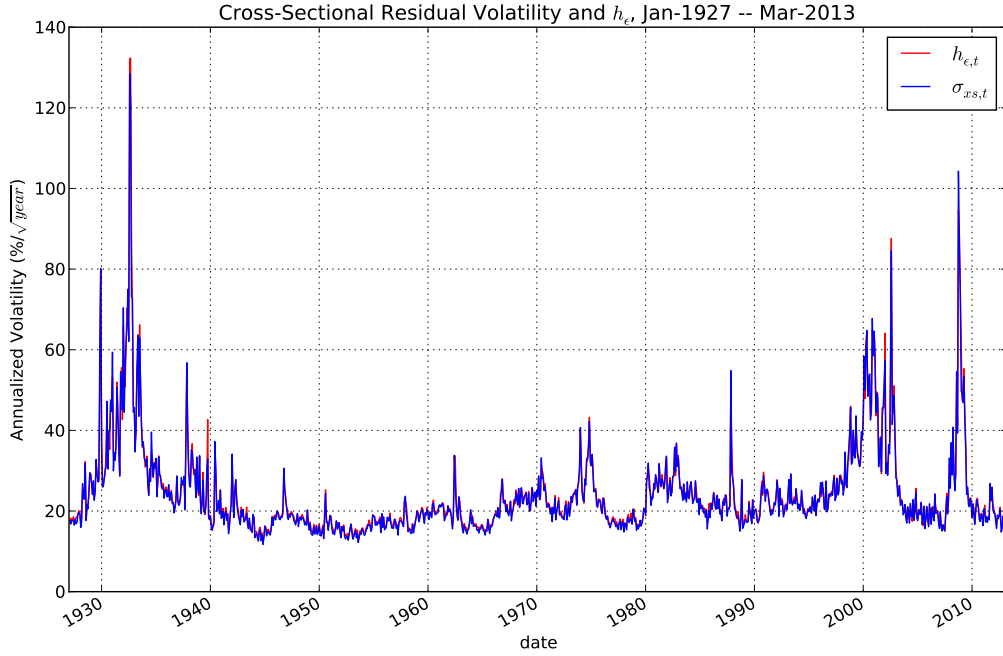
---

[9]Also, see footnote 8.

Figure 5: **Realized Cross-Sectional Vol and** $h_{\epsilon,t}$

seems to capture movements in cross-sectional volatility fairly well, albeit with a lag of a few days.

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 λ_t with a Hodrick and Prescott (1980) filter.
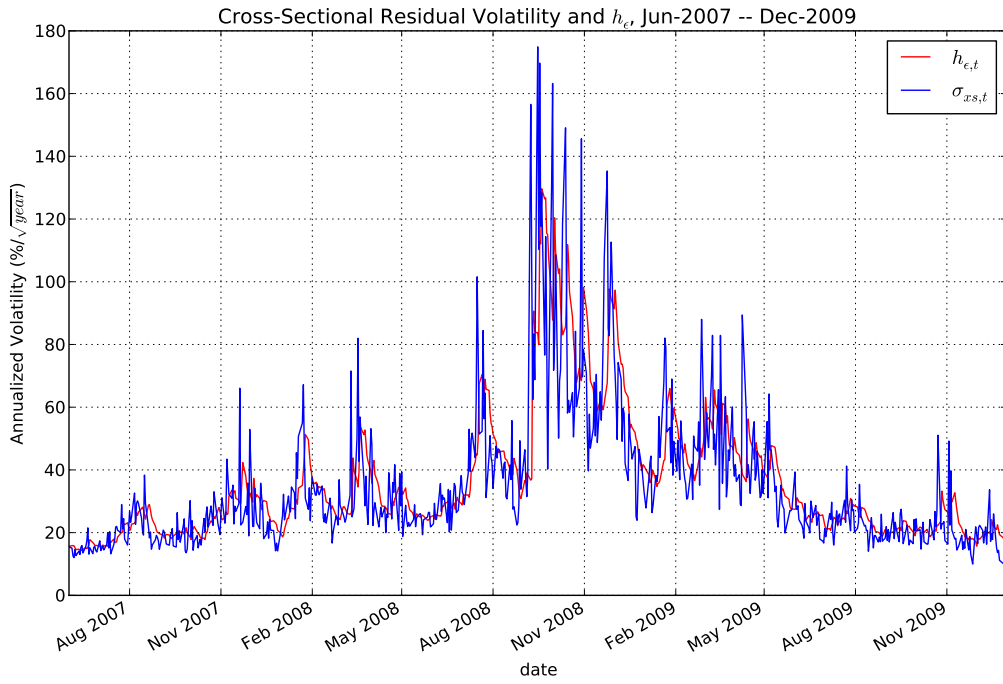
Figure 6: **Realized Cross-Sectional Vol and $h_{\epsilon,t}$ in the Financial Crisis**
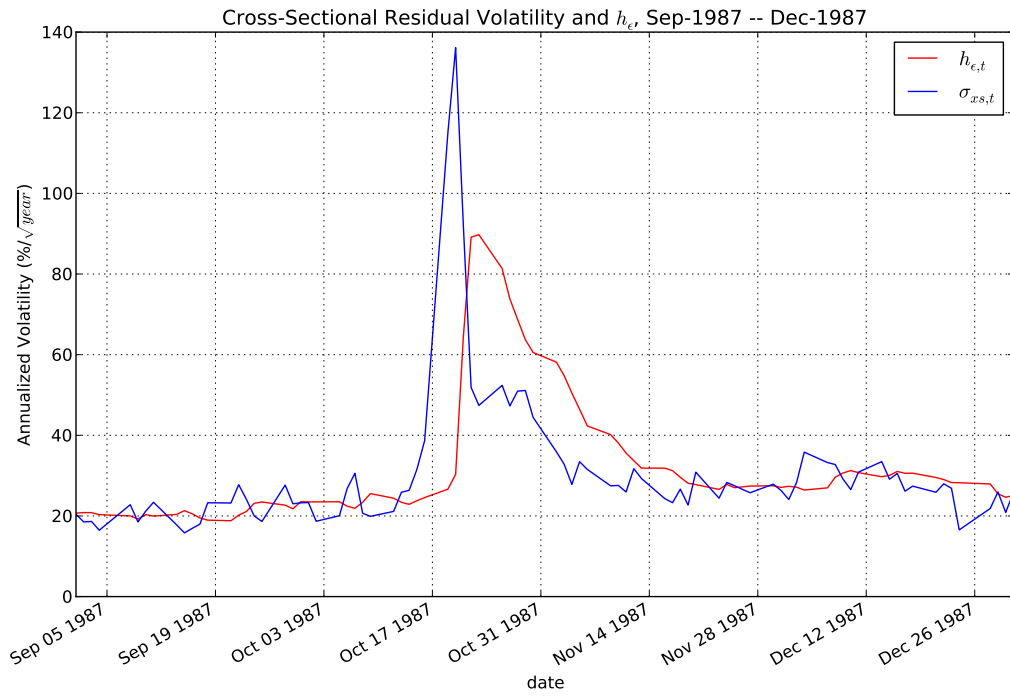

Figure 7: **Realized Cross-Sectional Vol and $h_{\epsilon,t}$ – September-December 1987**
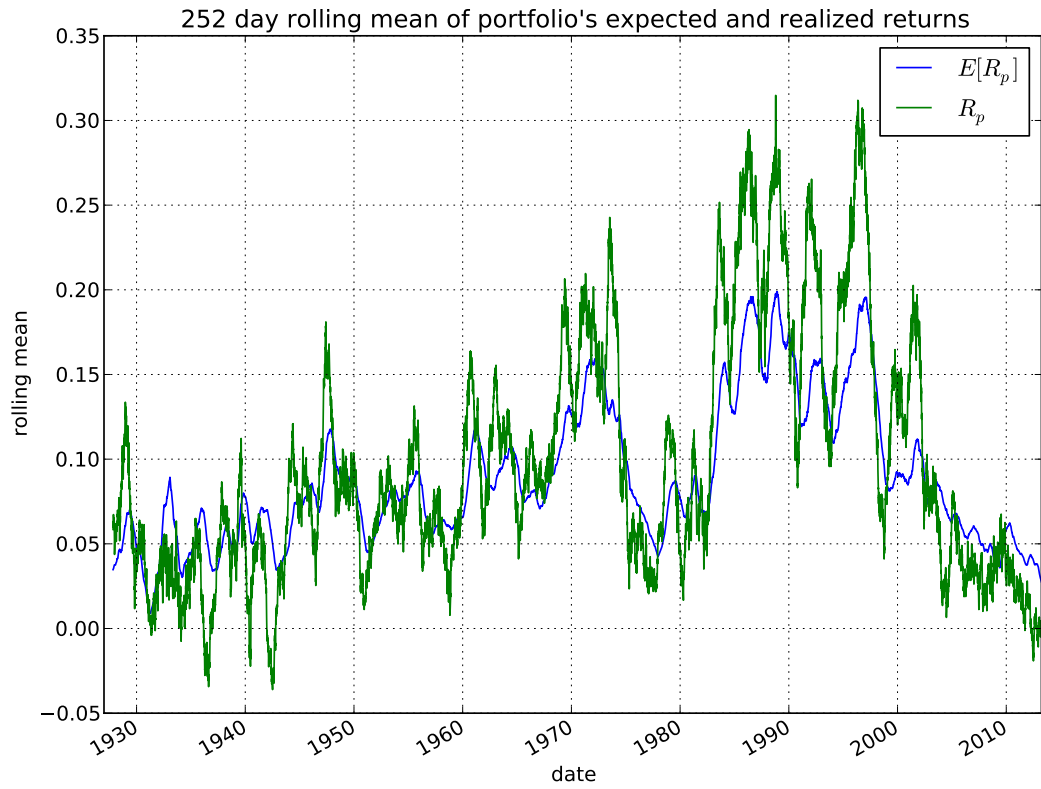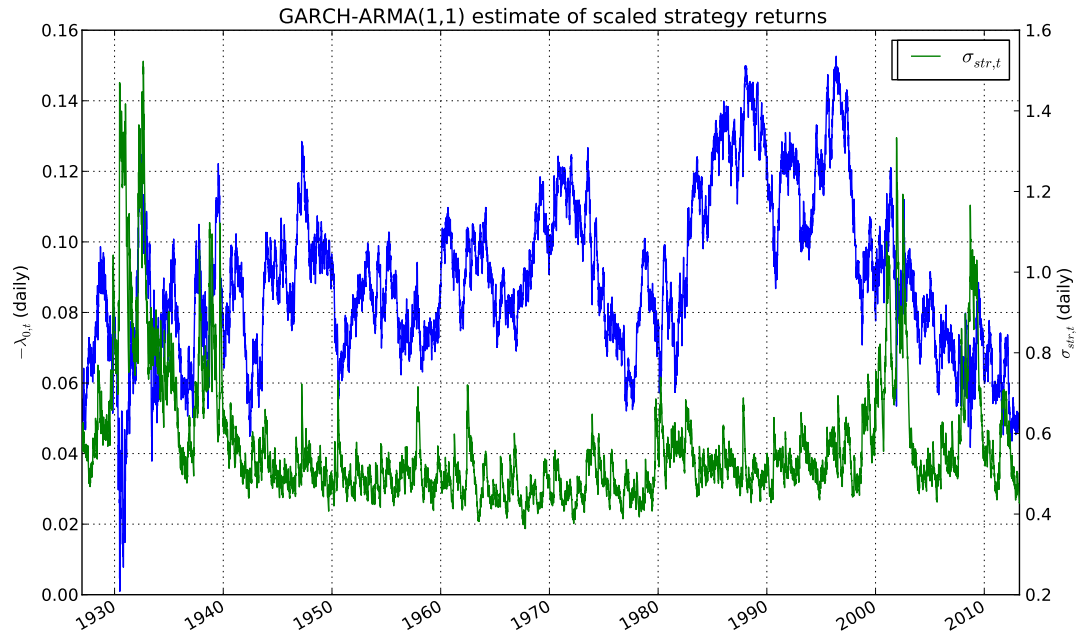
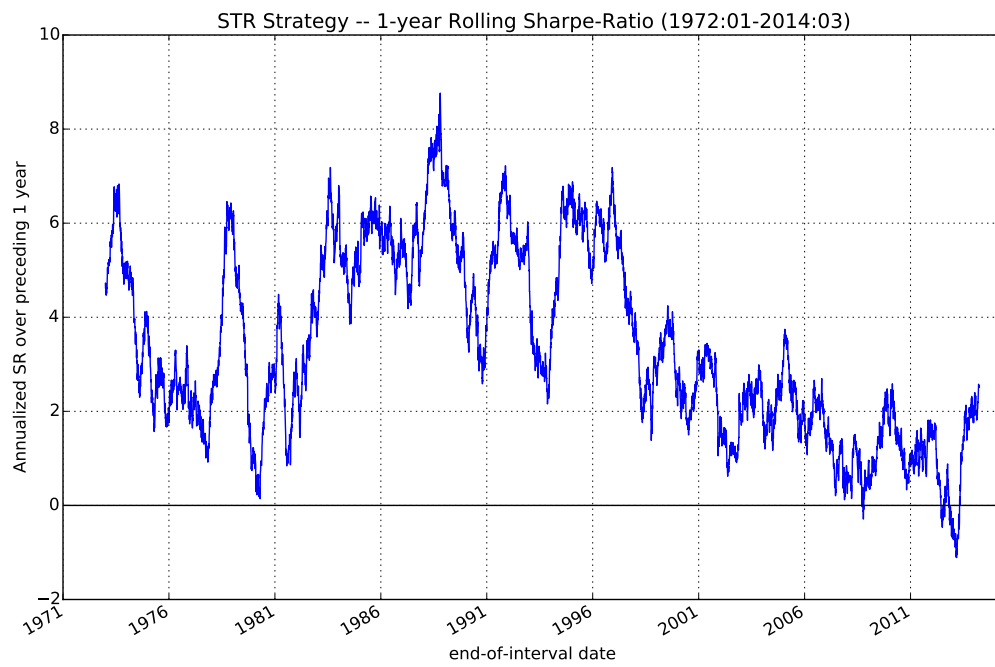Figure 9: **Time variation in the forecast and realized strategy returns**

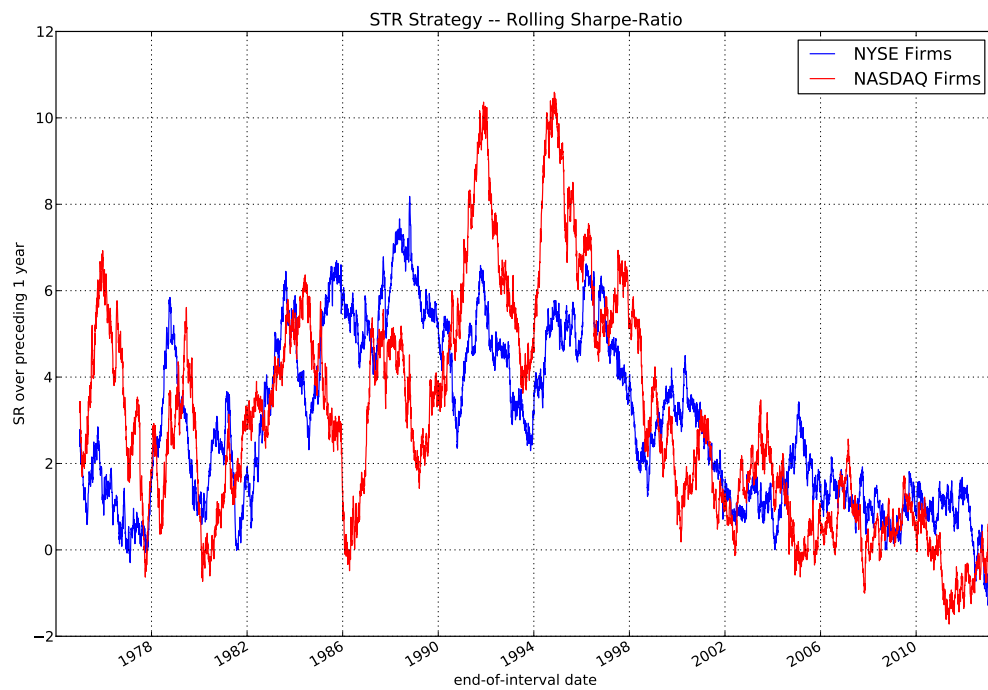Figure 10: **STR Strategy – Rolling Sharpe-Ratio**

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

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 approaches: the UC-ARIMA approach (Watson 1986, Stock and Watson 1988), and a one sided-Hodrick and Prescott (1980) filter.

4.1 UC-ARIMA

4.1.1 UC-ARIMA Specification

The premise of the UC-ARIMA model we employ in this section is that the expected return of the str portfolio – λ_t – is governed by the following stochastic process:

$$\lambda_t = \lambda_t^\tau + \lambda_t^c \tag{13}$$

$$\lambda_t^\tau = \lambda_{t-1}^\tau + u_t^\tau \tag{14}$$

$$\lambda_t^c = \sum_{k=1}^{p} \phi_k \lambda_{t-k}^c + u_t^c \tag{15}$$

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 comes from the business cycle literature). The trend component λ^τ follows a random walk with drift. The stationary component follows a stationary AR(p) process.

The two components of the expected return – λ_t^τ and λ_t^c – and their sum λ_t are all latent variables unobservable to the econometrician. Instead we observe the return to a portfolio for which the expected return is λ_t:

$$r_t^{str} = \lambda_t + \epsilon_t \tag{16}$$

4.2   Kalman Filter Estimation

We use a Kalman filter to estimate the system of equations (13)-(16), and to provide *ex-ante* estimates of λ_t^τ and λ_t^c.

To estimate this we first write equations (13)-(16) 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:

$$\mathbf{x}_t = \mathbf{A} x_{t-1} + \mathbf{w}_t \tag{17}$$

where, for our $AR(2)$ specification:

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

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:

$$\mathbf{w}_t \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}\right).$$

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*:

$$r_{t+1}^{str} = \mathbf{G}\mathbf{x}_t + \epsilon_{t+1}. \tag{18}$$

where $\mathbf{G}$ is a row-vector:

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

and where R is a time-invariant scalar. Note that specification for $\mathbf{G}$ ensures that the expected return of the str portfolio (which, as discussed above, has a unit factor loading), is equal to the sum of the trend and cycle components: $\lambda_t^\tau + \lambda_t^c$. The unexpected return ϵ_t is assumed to be uncorrelated with each of the shocks in $\mathbf{w}_t$.[10]

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:

$$\mathbf{x}_{t|t} \sim \mathcal{N}\left(\hat{\mathbf{x}}_{t|t}, \boldsymbol{\Sigma}_{t|t}\right) \tag{19}$$

---

[10]ϵ_t is here specified to have constant variance R, though it is straightforward to incorporate a GARCH process for the variance.

where:

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

$$\mathbf{\Sigma}_{t|t} = \mathbf{\Sigma}_{t|t-1} - \mathbf{\Sigma}_{t|t-1}\mathbf{G}'\left(\mathbf{G}\mathbf{\Sigma}_{t|t-1}\mathbf{G}' + R\right)^{-1}\mathbf{G}\mathbf{\Sigma}_{t|t-1}.$$

The forecast step: With the distribution of $\mathbf{x}_{t|t}$ from equation (19) and the state evolution equation (17) we can construct the distribution for the str portfolio return for 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:

$$\mathbf{x}_{t+1|t} \sim \mathcal{N}\left(\hat{\mathbf{x}}_{t+1|t}, \mathbf{\Sigma}_{t+1|t}\right) \tag{20}$$

where the mean and the covariance matrix come directly from equation (17):

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

From equations (18) and (20), the distribution of the time $t+1$ str return is:

$$r_{t+1}^{str} \sim \mathcal{N}\left(\mathbf{G}\hat{\mathbf{x}}_{t+1|t}, \mathbf{G}\mathbf{\Sigma}_{t+1|t}\mathbf{G}' + R\right) \tag{21}$$

4.2.2 Estimating the Kalman Filter Parameters

We begin with a prior distribution for the state variables. Using equation (19) applied to the parameters of the prior and the first return, we update the state variable distribution parameters. We then sequentially apply equations (20) and (21) 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*
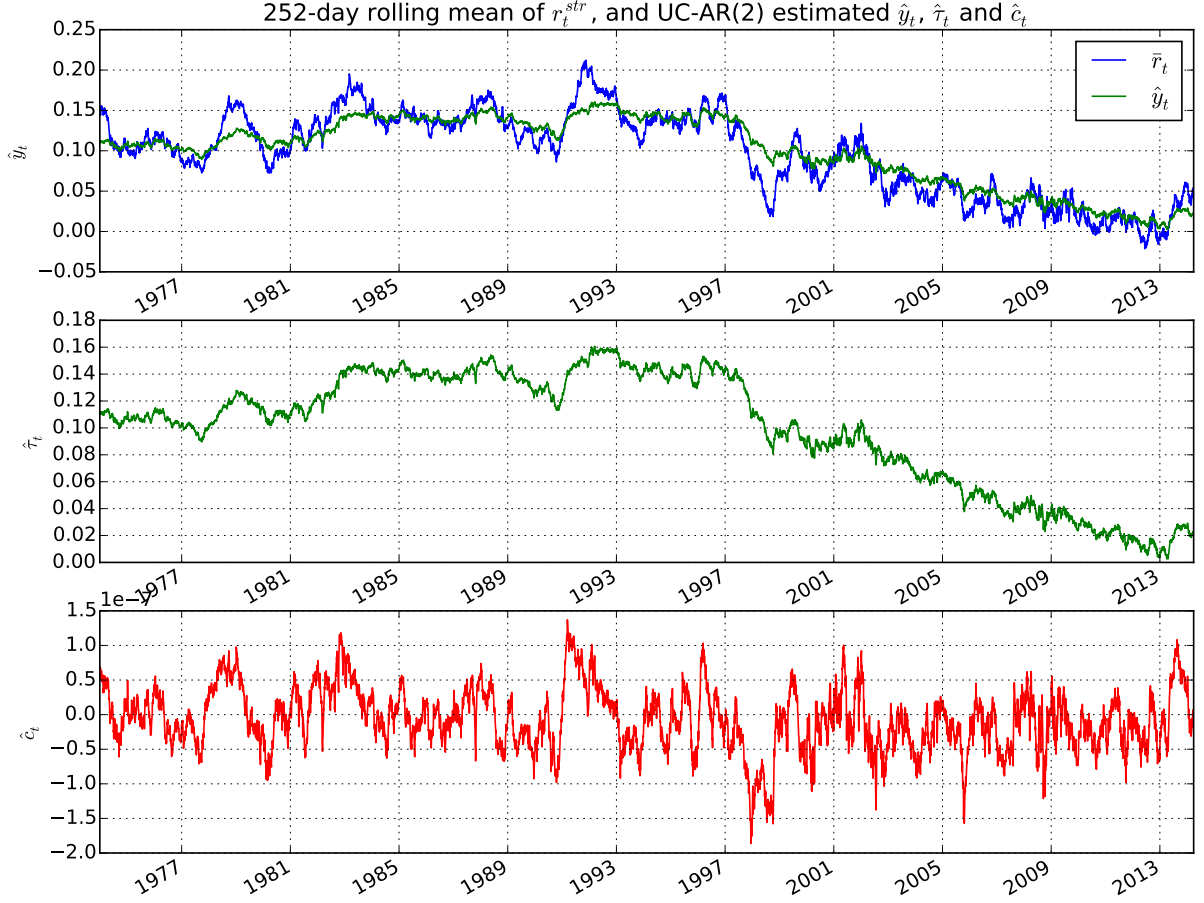
Figure 12: Kalman-Filter Estimated Components of the Expected Return of the r_t^{str} Portfolio, for a UC-AR(2) Model

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 3, 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.

Table 3: **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 $\times 100$.

Several things are of interest. First, while the parameters ϕ_1 and ϕ_2 of the estimated cyclical component appear reasonable, the magnitude of the estimated temporary component is small relative to the magnitude of the trend component (note that $\sigma_\epsilon \gg \sigma_c$). Consistent with this, λ_t^c doesn't provide 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 λ_{t-5}^τ, gives a t-statistic of 11.28.

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:

$$\min_{\tau_t} \left(\sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right) \tag{22}$$

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 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 (22) is:

$$\min_{\tau'} (\mathbf{y} - \boldsymbol{\tau})' \mathbf{I} (\mathbf{y} - \boldsymbol{\tau}) + \lambda \boldsymbol{\tau}' \mathbf{A}' \mathbf{A} \boldsymbol{\tau} \tag{23}$$

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:

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

Differentiating equation (23) with respect to $\boldsymbol{\tau}'$ and rearranging gives the first order condition for the HP-trend $\boldsymbol{\tau}$ that satisfies (23):

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

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 a convolution in the time domain is a multiplication in the frequency domain. Thus, to HP-filter the time series of str portfolio returns r_t^{str}, we convolve the original return series with the filter weights. This is the time domain analog of the frequency domain operation of multiplying the Fourier transform of the return series by the Fourier transform of the filter, and then inverse Fourier transforming. 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 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 returns. For this finance application we want an *ex-ante* estimate of λ_t, so we use the one-sided filter. The one-sided HP filtered str return series is plotted in Figure 15
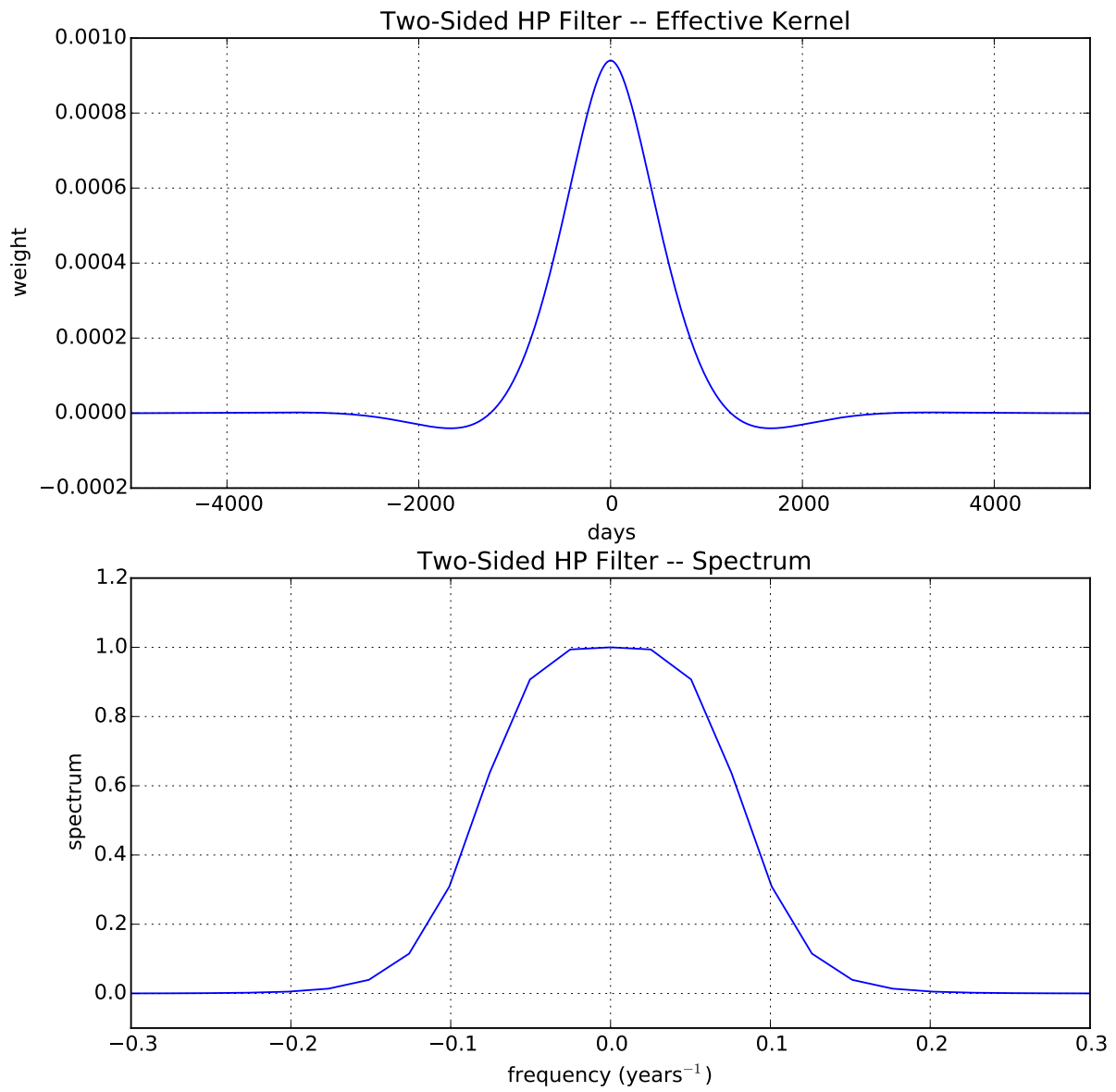
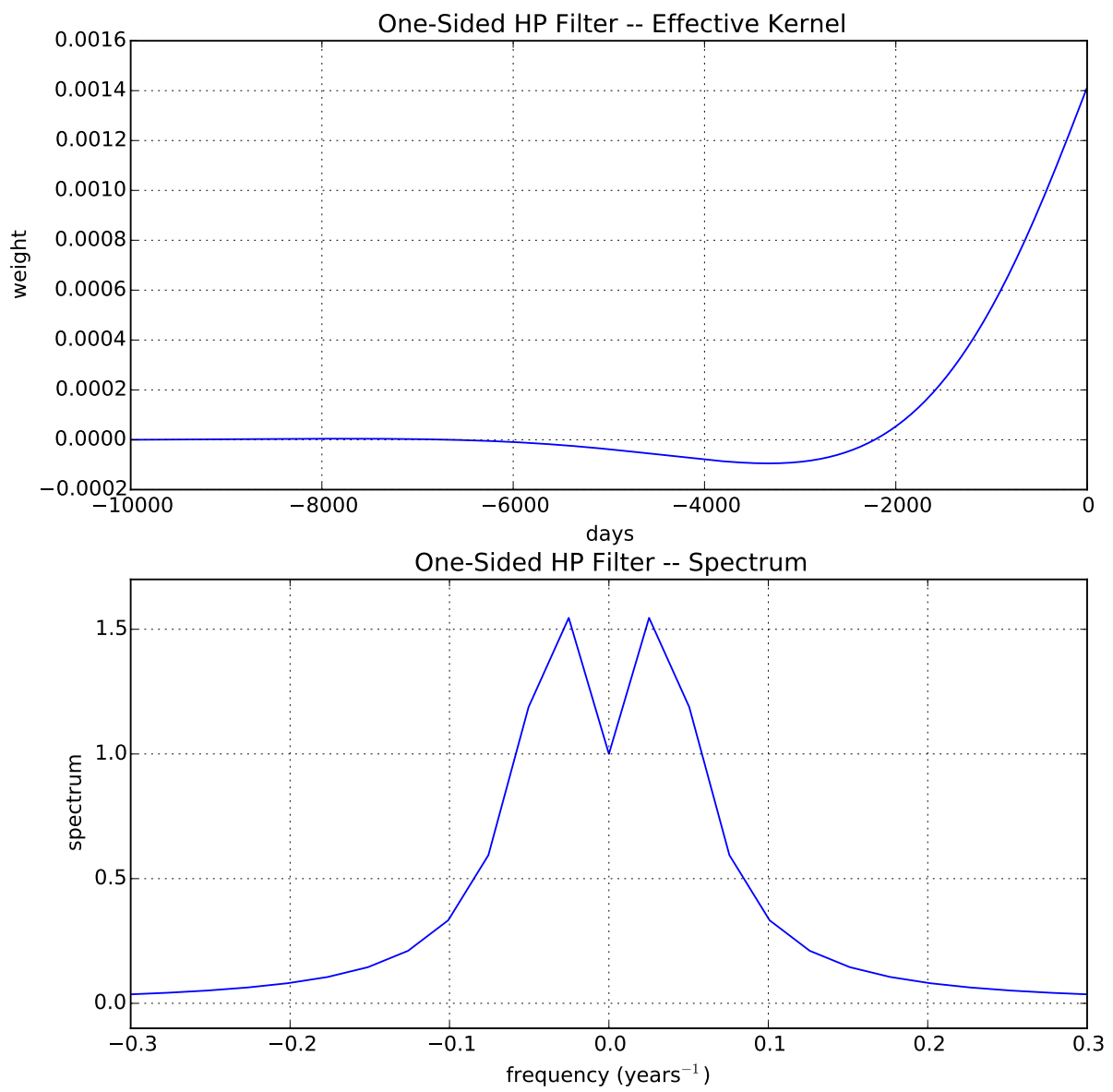Figure 13: **HP filter – Effective Kernel and Spectrum**

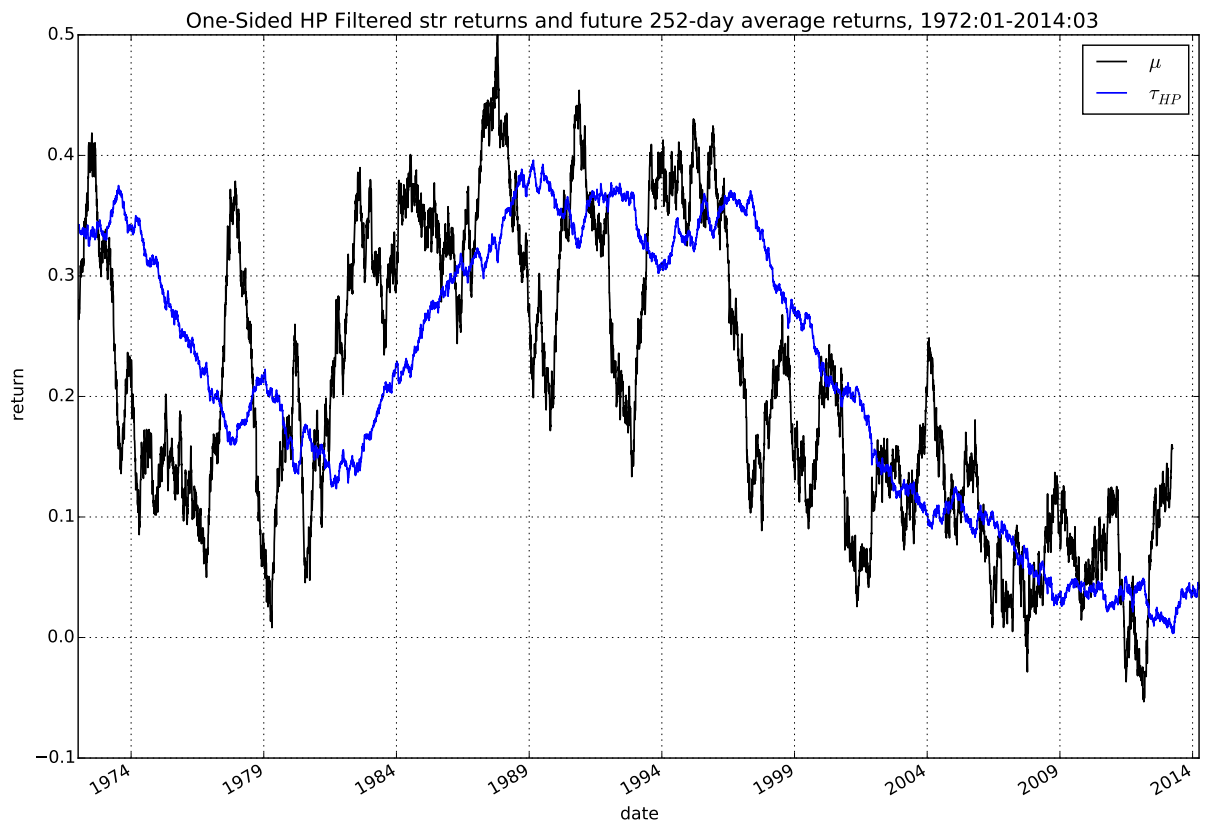Figure 14: **One-Sided HP filter – Effective Kernel and Spectrum**

Figure 15: **One-Sided HP filtered str Returns**

5 Strategy Time Variation

We now analyze the time variation in the returns of r_t^{str} portfolio with the goal of determining the sources of variation in λ_t. We begin by contrasting our results with those of Nagel (2012).

Most of the empirical work in Nagel (2012) concentrates on forecasting the *return* of the short-horizon reversal strategy with predictive variables, including the level of the VIX. In particular Nagel establishes a strong relationship between the level of the VIX and the return of his short-term reversal strategy. This evidence suggests that when capital is scarce – for example during times of financial turmoil when the VIX is high – strategies such as this str strategy which require capital will necessarily earn a higher premium.

However, one problem with the use VIX as a proxy as a proxy for financial turmoil and capital scarcity is that the VIX is potentially correlated with a number of other factors which are unrelated to capital scarcity, but which may still affect the returns on a short-term-reversal strategy. For example, when the VIX is high, cross-sectional volatility is also likely to be high, and this will lead to a higher risk for a \$1-long/\$1-short reversal strategy. Thus, the higher str strategy return when the VIX is higher could be simply reflect that the average returns of the str strategy rises when the risk rises.

To address this question, Nagel also examines time variation in the *compensation for risk (i.e.,* the Sharpe Ratio) of his 5-day STR strategy.[11] He proceeds in the following way:[12] First, he estimates time variation in conditional volatility associated with the VIX by running the regression specified in his equation (22):

$$|\tilde{L}_t^R|\kappa = a_0 + a_1 \cdot \text{VIX}_{t-5} + a_2 \cdot d_{t-5} + u_t \tag{24}$$

---

[11]Note that the theoretical model used in Nagel (2012) predicts that the conditional Sharpe ratio associated with liquidity provision strategies – such as the short-term reversal strategy that he explores – is likely to be affected by shocks which affect financially constrained intermediaries.

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

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).[13] The fitted values from this regression at each time are used as estimator of σ_t, which is used in the regression

$$\frac{L_t^R}{\sigma_t} = b_0 + b_1 \cdot \text{VIX}_{t-5} + b_2 \cdot d_{t-5} + e_t, \tag{25}$$

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: we use only firms which are trad each day, and skip one day between the return measurement and portfolio formation. This ensure that our results are not contaminated by non-trading effects. Second, we exponentially weight lagged returns, consistent with our model in Section 3 and our empirical findings in Section 2. Also, our portfolio construction incorporates the GARCH modeling of the residual variation modeled in Section 3.

This last distinction is important. 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 a 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 (24) may be unbiased, when we scale the strategy return by the fitted σ_t (in equation (25) 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

---

[13]d_t takes a value of one prior to decimalization (April 9, 2001) and a value of zero thereafter

Simulation -- histogram
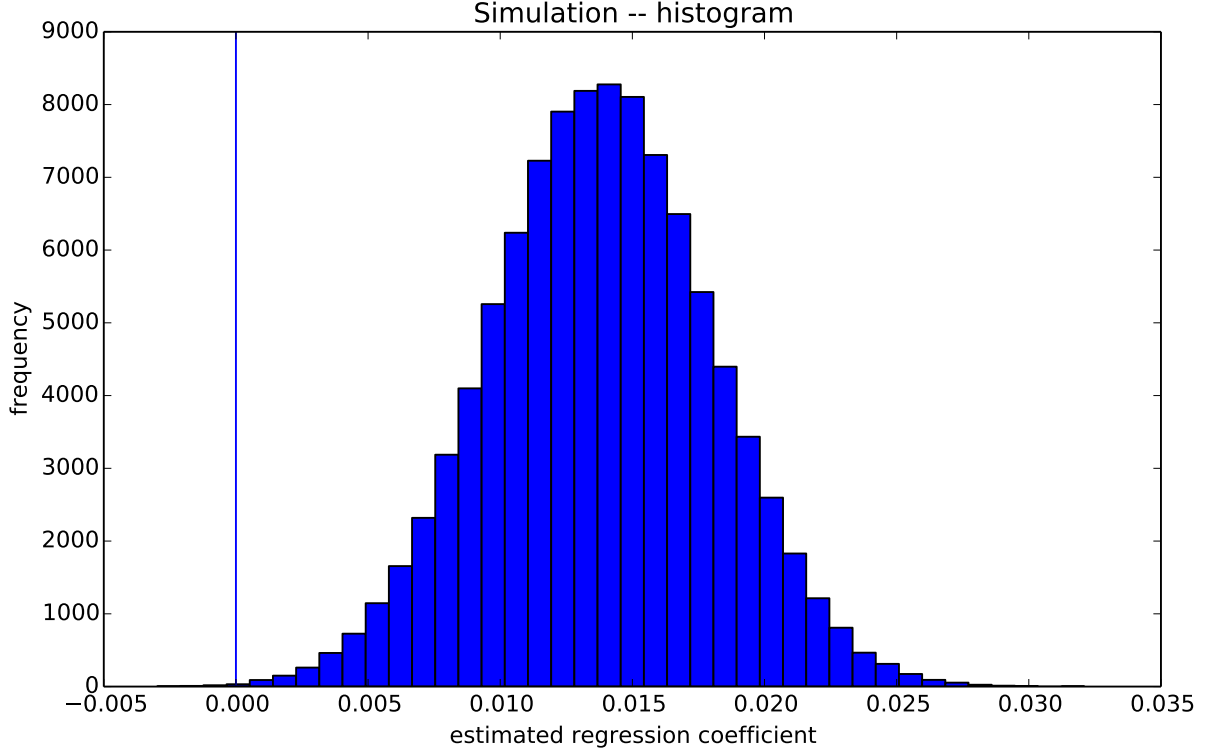
This histogram is for the coefficients obtained in regressing the estimated Sharpe-ratio, using the technique described above, on the simulated level of the VIX, which in the simulation is a noisy proxy for the true return volatility. There were 100,000 runs in the simulation, and 3276 ($= 13 \times 252$) time periods. The correlation between the simulated VIX and the true volatility was 61%.

Figure 16: **Simulation Histogram**

estimated relation between the level of the VIX and the return Sharpe ratio, even the true Sharpe ratio is actually constant.

In the Appendix, we demonstrate the existence of this bias via simulation. Figure 16 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 ~250 firms with negative weighted residuals, and short the ~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 observe a the strong and 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.

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.
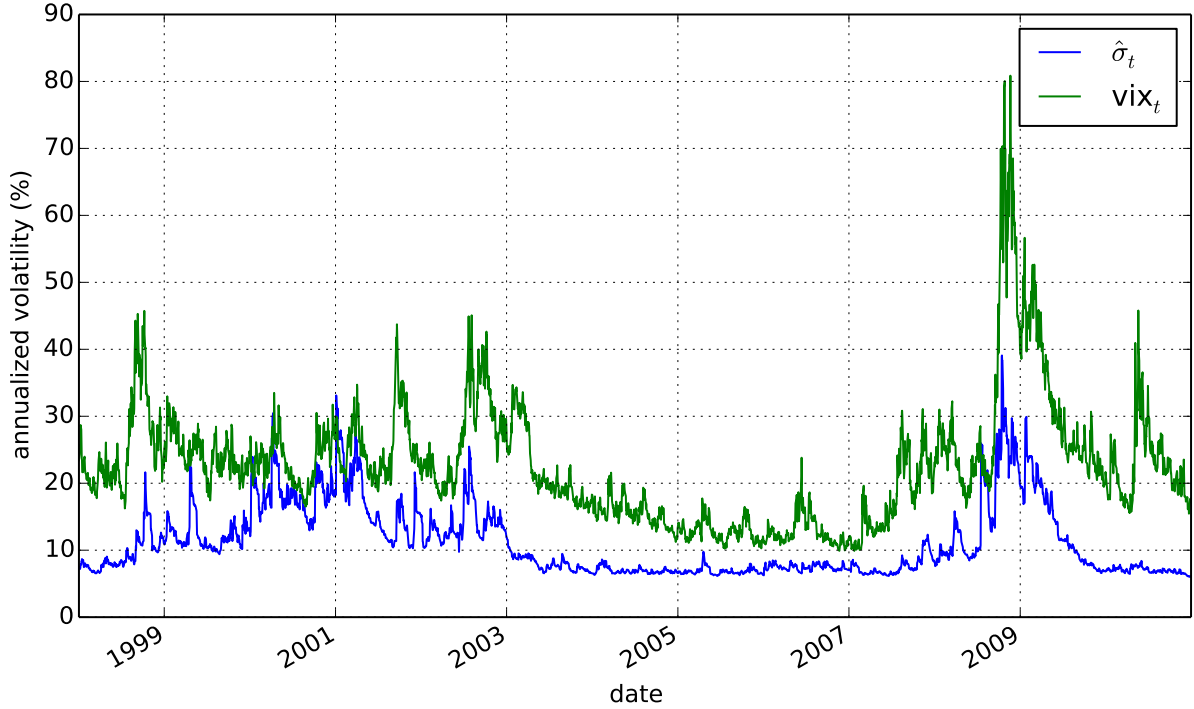
Figure 17: **VIX and Historical Estimates of cross-sectional volatility**

Finally, Figure 17 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 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 – 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.

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.

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.

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.