Common Factors in Stock Market Seasonalities*

Matti Keloharju

Aalto University School of Business and CEPR

Juhani T. Linnainmaa
University of Chicago Booth School of Business

Peter Nyberg
Aalto University School of Business

April 18, 2013

Abstract

Well-diversified portfolios of stocks formed by either characteristics or factor loadings have relatively high or low returns every year in the same calendar month. These common seasonalities account for at least 80% of the seasonalities in individual stock returns. The source of seasonalities matters: a trading strategy that tries to profit from seasonalities in individual stock returns is necessarily exposed to common return factors. We develop a model in which seemingly firm-specific seasonalities emerge from stocks amplifying and aggregating small seasonalities in common factors. We demonstrate that this mechanism is a pervasive feature of asset returns at monthly and daily frequencies.

*Corresponding author: Juhani Linnainmaa, juhani.linnainmaa@chicagobooth.edu. We thank John Cochrane, Mark Grinblatt, Chris Hansen, Steven Heston (discussant), Toby Moskowitz, Stefan Nagel, Ľuboš Pástor, Tapio Pekkala, and Ruy Ribeiro for insights that benefited this paper, and seminar participants at Maastricht University, Nanyang University of Technology, National University of Singapore, Singapore Management University, University of Arizona, University of Chicago Booth School of Business, University of New South Wales, University of Sydney, University of Technology in Sydney, Aalto University, Luxemburg School of Finance, INSEAD, and 2013 FSU SunTrust Beach Conference for valuable comments, and Yongning Wang for invaluable research assistance. An earlier version of this paper was circulated under the title "The Sum of All Seasonalities."

1 Introduction

One of the most intriguing stock market anomalies is the seasonality in individual stock returns. Heston and Sadka (2008) show that a trading strategy that chooses stocks based on their historical same-month returns earns excess returns as high as 12% per year. This pattern is puzzling. It survives tests that control, one at the time, for earnings announcements, dividends, firm size, industry, exposures to common risk factors, and calendar month, and it has persisted for decades. Historical same-month returns up to 20 years old are informative about the cross-section of returns today, and the same pattern also exists in international markets.[1] This paper explains the source of this pattern. We show that the seasonalities in individual stock returns are a necessary consequence of seasonalities in *common* return factors. Common factors explain at least 80% of the seasonalities in individual stock returns.

We show that portfolios of stocks formed by sorts on size, book-to-market ratios, estimated factor loadings, and other attributes display return seasonalities similar to those displayed by individual stocks, both in January and in other months. Given that these portfolios are well diversified, the seasonalities must be related to the characteristics by which the portfolios are formed. We also show that the seasonalities share commonalities: the seasonalities of book-to-market- and size-sorted portfolios, for example, span the seasonalities of dividend-to-price- and earnings-to-price-sorted portfolios. This spanning behavior is similar to that found in unconditional stock returns; Fama and French (1992), for example, show that the size and value effects subsume the dividend-to-price and earnings-to-price anomalies.

The commonalities in seasonalities are important. A strategy that tries to take advantage of the seasonalities in stock returns *has to* remain exposed to the common return factors—an attempt to

[1] See Heston and Sadka (2010).

hedge these exposures would eliminate the seasonalities as well. An investor who tries to profit from January seasonalities, for example, must remain exposed to shocks to small- and large-firm returns. This paper does not take a stance on whether these common factors represent risk or mispricing. What we show is that the set of factors that explains differences in average stock returns is, by and large, the same as the one that explains seasonalities in stock returns.

Return seasonalities and the differences in average returns are intimately linked to one another. If the returns on factors (or characteristics) accrue unevenly over the calendar year, stocks *must* display seasonalities in the cross-section by the virtue of loading differently on these factors. If the average market return, for example, is significantly higher during winter months than summer months[2], high-beta stocks have relatively higher returns during winter months. The seasonalities in individual stock returns may thus be an inevitable consequence of there being multiple common return factors in the data, each of which exhibits some variation in return premia over the calendar year.

The seasonalities in individual stock returns are prevalent because historical same-month returns *amplify* and *aggregate* common seasonalities. To illustrate the aggregation mechanism, consider the seasonality in stock returns as a function of firm size. Small stocks tend to outperform large stocks in January[3], so firms' historical January returns are noisy signals of their sizes. A sort of stocks into portfolios by their past January returns thus predicts variation in future January returns because it correlates with firm size. The same intuition applies when many stock attributes are associated with different return seasonalities. For example, industries related to consumer consumption have relatively low winter returns[4], so a sort on past returns picks up industry seasonalities as well. A cross-sectional regression of returns on past same-month returns is equivalent to a regression of returns on a noisy

[2]See Kamstra, Kramer, and Levi (2003).
[3]See, for example, Keim (1983) and Reinganum (1983).
[4]See Jacobsen and Visaltanachoti (2009) and Choi (2008, chapter 1, pp. 62–63).

combination of attributes associated with return seasonalities. The existence of multiple common seasonalities also explains why the seasonal patterns in individual stock returns survive a battery of controls: none of these factors alone is responsible for the pattern. An amplification mechanism is also important. For example, even if the size-related seasonality is small for most stocks, the gap in seasonalities between the firms at the extreme ends of the size distribution is large. In the terminology of factor models, heterogeneity in factor loadings amplifies small seasonalities in risk premia.

If the factors generating unconditional differences in average returns and the seasonalities in returns are the same, why cannot standard asset pricing models explain any of the returns on the strategy that trades on seasonalities in individual stock returns? The reason is that the seasonality strategy dynamically shifts factor exposures from month to month. This strategy, for example, may be long small value stocks with high market betas in a particular industry in one month, but the next month it is long big growth stocks with low market betas in a different industry. As a result of these shifts, the strategy's *unconditional* covariances with the usual asset pricing factors are close to zero.

The importance of common factors as an explanation to the seasonalities can be inferred in at least two ways. First, the returns on a strategy that trades on the seasonalities in individual stock returns are far too volatile to be explained by idiosyncratic seasonalities alone. We estimate that the volatility of the strategy would be less than half of what it is, were it to pick up just idiosyncratic seasonalities in expected returns. Second, we can measure the influence of common factors by studying comovement within portfolios formed by sorting on stocks' average same-month returns. We find that portfolios formed based on same-month returns display about as much comovement as portfolios formed by industry. This excessive comovement suggests that the seasonality strategy groups together stocks with similar characteristics or factor loadings.

We formalize the transmission of common seasonalities into individual stock returns by developing a factor model in which factor risk premia exhibit seasonal variation. Although we use the factor terminology, the model can also be interpreted in terms of characteristics and return premia associated with these characteristics. The model shows that even if the seasonalities in the risk premia are small, the seasonalities in individual stock returns can be large for two reasons. First, individual stocks aggregate seasonalities across the factors. Second, because stocks can load very differently on the factors, individual stock returns amplify seasonalities in risk premia. In the model the seasonality strategy can thus have a high mean because it sums up amplified seasonal differences in factor risk premia. The model fits the data. Our estimates suggest that small variation in monthly risk premia is enough to generate the seasonalities in individual stock returns.

In our model lagged same-month returns include incremental information about the returns today even in multivariate regressions. Their significance does not imply that stocks "repeat" specific events that took place exactly 12, 24,... months ago; they are incrementally important only because every additional data point gives an independent signal about the product of a stock's factor loadings and its same-month factor risk premia. Our model thus predicts the optimal estimate of this product to subsume individual past returns. Our empirical results are consistent with this prediction. We show that when we control for stocks' long-run same-month returns—which are the optimal estimates of the product of factor loadings and risk premia when stocks' factor loadings are relatively stable—individual past same-month returns no longer explain variation in this month's returns.

Although the seasonalities in individual stock returns remain largely intact when controlling for different stock attributes one by one, they fare less well when confronted with multiple characteristics at the same time. Characteristics such as industry, size, and book-to-market ratio *jointly* explain four-fifths of the cyclical coefficient pattern. The remaining seasonalities in stock returns are concentrated at shorter lags. We arrive at the same conclusion not only by controlling for stock characteristics but also by showing that simple seasonality-mimicking factors constructed from portfolio returns explain more than half of the seasonalities in individual stock returns.

Amplification and aggregation of seasonalities should not be confined to the monthly frequency. Motivated by research finding common seasonalities at daily frequency[5], we illustrate the pervasiveness of the aggregation mechanism using daily returns. We find a strong seasonal pattern in the data. In regressions of today's stock returns against lagged same-weekday returns, that control for one-year momentum, almost every weekly lag up to the one-year mark is positive, and over half of them are statistically significant. Past same-weekday returns serve as signals of firm characteristics such as size and institutional ownership, so a regression of returns on past returns is analogous to a regression of returns on a combination of characteristics. These results indicate that seasonality amplification and aggregation figure importantly in asset returns and that they generalize to other frequencies.

Past reseach has extensively studied the origins of seasonalities in factor premia. Kamstra, Kramer, and Levi (2003, 2012) and Garrett, Kamstra, and Kramer (2005), for example, attribute the seasonalities in equity risk premium and Treasury returns to seasonal variation in the price of risk: investors are more risk averse during the winter. Seasonalities could also arise from seasonal variation in the quantities of risk or mispricing, including that induced by tax-loss selling.[6] Our paper does not take a stand on the origins of the seasonalities. However, it shows that *if* seasonalities exist, they can aggregate into large and seemingly stock-specific seasonalities in individual stock returns.

[5]See, for example, Keim and Stambaugh (1983) and Chan, Leung, and Wang (2004).

[6]Kamstra, Kramer, and Levi (2012) test whether seasonalities in mispricing or quantity of risk could explain the seasonal patterns in Treasury returns. They find seasonal variation in many macroeconomic variables such as the GDP and unemployment growth rates. While this variation could generate a seasonality in the quantity of risk, it does not explain the seasonality in Treasury returns. Kamstra, Kramer, Levi, and Wang (2012) calibrate an asset pricing model to match the seasonalities in equity risk premium and Treasury returns, and find that variation in investors' willingness to substitute consumption across periods (the elasticity of intertemporal substitution) is also important in explaining the changes in risk premia. The turn-of-the-year seasonalities in the U.S. are often attributed to December tax-loss selling and the rebound that follows—see, for example, Wachtel (1942), Rozeff and Kinney (1976), Dyl (1977), Reinganum (1983), and, more recently, Grinblatt and Moskowitz (2004) and Kang, Pekkala, Polk, and Ribeiro (2011). Heston and Sadka (2008) note that although the turn-of-the-year effect is important, economically significant seasonalities abound even when controlling for it.

The rest of the paper is organized as follows. Section 2 measures common seasonalities in stock returns. Section 3 develops a model of stock market seasonalities, and illustrates how common seasonalities aggregate into large seasonalities in stock returns. Section 4 shows that the data are consistent with the model's predictions and documents large seasonalities also in daily returns. Section 5 concludes.

2 Return Seasonalities in Well-diversified Portfolios

2.1 Data

Our tests use monthly and daily return data on NYSE-, AMEX-, and Nasdaq-listed stocks from the Center for Research in Securities Prices (CRSP). We exclude securities other than ordinary common shares. We use CRSP delisting returns; if a delisting return is missing, and the delisting is performance-related, we impute a return of -30%.[7] We use returns from January 1963 through December 2011 to compute portfolio returns and as dependent variables in cross-sectional regressions. However, for right-hand side returns, used in regressions of today's returns against historical returns, we use monthly returns going back to January 1943. We also take industry classifications and market values of equity from the CRSP tapes.

Accounting data and earnings announcement dates are from Compustat. We define book value of equity as stockholders' equity plus balance sheet deferred taxes and investment tax credit (if available) minus the book value of preferred stock. We use redemption, liquidation, or par value, in that order, depending on availability, for the book value of preferred stock. Stockholders' equity is primarily that reported in Compustat. If the Compustat item is missing, we measure stockholders' equity as the

[7] See Shumway (1997) and Shumway and Warther (1999).

sum of book value of common equity and the book value of preferred stock. If common equity is not available, we define stockholders' equity as the book value of assets minus total liabilities.[8]

2.2 Measuring return seasonalities

We measure seasonal differences in returns on well-diversified portfolios using information in past *same-month* returns. The simplified return process we consider is

$$r_{p,t} - \bar{r}_t = \mu_{p,m(t)} + e_{p,t}, \tag{1}$$

where $r_{p,t}$ is the return on portfolio p in month t, $\bar{r}_t$ is the equal-weighted average return across the portfolios, $\mu_{p,m(t)}$ is the expected return of portfolio p in calendar month $m(t)$ (1 = January, ..., 12 = December), and the return innovation $e_{p,t}$ is mean zero. The model in equation (1) is the same as that considered in Heston and Sadka (2008) except that we apply it here to portfolios. We model individual stocks in Section 3.

If equation (1) describes portfolio returns, then past same-month returns ($t - 12$, $t - 24$, and so forth) are informative about differences in expected returns across assets in month t. Any market-adjusted same-month return, or the average of same-month returns, is an estimate of the $\mu_{p,m(t)}$ term. Similar to Heston and Sadka (2008), we use cross-sectional regressions and portfolio sorts to assess the magnitude of return seasonalities. In the cross-sectional regressions we regress month-t returns against various past same-month returns. In the portfolio sorts, in which we sort *portfolios* into portfolios, we form the sorts based on average market-adjusted same-month returns.

[8] These tiered definitions for the book value of equity are the same as those used by Ken French at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`. We use the Davis, Fama, and French (2000) data to fill in the gaps in the book values of equity in the pre-1963 Compustat data.

2.3 Portfolios

We form the test portfolios by sorting stocks by either firm characteristics or estimated factor loadings. The firm characteristics are firm size, book-to-market ratio, dividend-to-price ratio, earnings-to-price ratio, and industry membership. The estimated factor loadings are the market, SMB, and HML betas from the Fama and French (1993) three-factor model.

We construct the portfolios in June of year t and then compute equal-weighted returns on these portfolios from year-t July to year-$t+1$ June.[9] Except for industry, for which we use the Fama-French 48 industry classification, we form 25 portfolios based on each variable using NYSE breakpoints. We follow the usual conventions to time the variables that use accounting information.[10] The book value of equity, for example, is from the fiscal year ending in calendar year $t-1$ and, in computing the book-to-market ratio, this book value is divided by the market value of equity at the end of December of year $t-1$. We use up to five years of monthly data to estimate the three-factor model loadings, and we require stocks to have at least two years of data to estimate loadings. We use return data to the end of June of year t in these computations.

Stocks move across portfolios from year to year because we use returns on annually rebalanced portfolios in our tests. We discuss the difference between rebalanced and fixed-weight portfolios at the end of the next section. The portfolio return data run from January 1943 through December 2011 for all portfolios except for those based on the earnings-to-price ratio. Because Compustat has comprehensive earnings data only starting in 1951, we start computing E/P portfolio returns in 1953, and use them as the dependent variable after January 1973.

[9] We use equal-weighted portfolios throughout this paper. The Internet appendix, available on the corresponding author's website, provides value-weighted versions of all tables that use portfolios, and shows that the results are not sensitive to the choice between equal- and value-weighting.

[10] See Fama and French (1992).

2.4 Estimates of return seasonalities

Figure 1 sets the stage by plotting the average coefficients $\hat{b}$ from cross-sectional regressions of month-t returns against month-$t-k$ returns:

$$r_{i,t} = a_t + b_t \, r_{i,t-k} + e_{i,t}, \tag{2}$$

in which the lag k ranges from one to 120 months. The data in Panel A are the monthly returns on individual stocks from 1963 through 2011. This figure is the same as that in Heston and Sadka (2008) except that we include Nasdaq stocks in addition to AMEX and NYSE stocks, and our sample period is a few years longer. The data in Panel B are the monthly returns on 124 portfolios: 25 portfolios sorted by market capitalization (ME), 25 book-to-market ratios (BE/ME), 25+1 dividend-to-price portfolios (D/P) with $D = 0$ stocks in their own portfolio, and the 48 Fama-French industry portfolios. We do not include the E/P portfolios in this test only because that return series starts 10 years after the other return series.[11]

The patterns in average coefficients are strikingly similar in the two panels. The peaks in the regression coefficients every 12 months indicate that both individual stocks and well-diversified portfolios of stocks have relatively high or low returns every year in the same calendar month. The cyclical coefficient pattern in portfolio returns is as impressive as that in individual stock returns. The average coefficient is positive in the portfolio regressions at all annual lags up to 20 years, and 17 of these averages are associated with t-values greater than 2. In the individual-stock regressions, 19 of 20 coefficients are positive and 18 of them are statistically significantly greater than zero at the 5% level.

[11] The results are almost unchanged if the E/P portfolios are included and the sample period is shortened by 10 years. We show below that the seasonalities in E/P portfolios, even though they exist, are subsumed by the seasonalities displayed by the size- and BE/ME-sorted portfolios, so it does not make a difference whether they are included in the set of portfolios in Figure 1.

Table 1 measures seasonalities in individual stock returns and in returns on portfolios sorted by various stock attributes. The first column in Panel A, for example, summarizes information on the individual stock coefficients plotted in Figure 1 Panel A. Row "Average annual" reports that the average annual peak in the figure, if we were to extend its x-axis up to 20 years, is 0.97; the average *non-annual* coefficient is -0.26, and the difference between these two averages is significant with a t-value of 9.48. The other columns in Panel A estimate the regressions using portfolio returns. Column "All" uses the 124 ME, BE/ME, D/P, and industry portfolios from Figure 1. The estimates show that *all* well-diversified portfolios, including those constructed by sorting on estimated factor loadings, have relatively high or low returns every year in the same calendar month.

Panel B sorts stocks or portfolios of stocks into portfolios by the 20-year average same-month or other-month return. In January 1963, for example, we sort on either the average January return in 1943–1962 or the average non-January return in 1942–1961. We skip months 1–11 in the non-annual portfolio sorts because of the one-year momentum in stock returns. In the first column we use the equal-weighted CRSP index to demean the returns before taking the average because stocks differ in the availability of historical return data. When we sort $N = 25$ or $N = 124$ portfolios into portfolios, the first decile contains $\lfloor \frac{N}{10} \rfloor$ portfolios. The seasonalities in individual stock returns are economically significant: the equal-weighted high-minus-low portfolio based on the past same-month return earns a monthly excess return of 1.11% ($t = 8.99$). The portfolio results are similar. A strategy that constructs the same long-short portfolio from the 124 ME, BE/ME, D/P, and industry portfolios, for example, earns a monthly excess return of 1.02% ($t = 7.87$). The last row shows that the three-factor model cannot account for the seasonalities in stock or portfolio returns. This result extends Heston and Sadka's (2008) findings on the effect of risk adjustment from individual stocks to well-diversified portfolios.

The results in Table 1 speak to the pervasiveness of seasonalities in stock market returns. Not only do individual stocks earn relatively high or low returns every year in the same calendar month, but so do well-diversified portfolios of stocks. The effect exists not only for characteristics-sorted portfolios but also for portfolios sorted by estimated factor loadings. Panel B shows that a strategy that trades on the seasonality in portfolios constructed from market betas earns an excess return of 0.46% per month ($t = 5.42$). This result is consistent with the presence of seasonal variation in market returns (Kamstra, Kramer, and Levi 2003). If the average market return is higher during winter months, high-beta stocks must have relatively higher returns in these months. The seasonalities are stronger for the HML- and SMB-sorted portfolios and, as we show below, these seasonalities are intertwined with those displayed by the size- and BE/ME-sorted portfolios.

Table 1 uses rebalanced portfolios, i.e. stocks are reassigned into portfolios every June. Thus, even though the Fama-MacBeth regressions explain returns on various value/growth portfolios with past returns on these portfolios, the portfolios on the left- and right-hand side of the equation can consist of very different stocks. This rebalancing is important: the results are very different when we repeat the analyses using fixed portfolio weights. We re-estimate the Fama-MacBeth regressions by regressing equal-weighted portfolio returns in month t on month- $t-k$ average returns on the *same* stocks. We discuss the results on the BE/ME-sorted portfolios to illustrate the difference. Table 1 shows that the average annual regression coefficient peak is 6.08 ($t = 7.47$) when computed from the rebalanced BE/ME portfolios. This average decreases to just 1.36 ($t = 1.91$) when using fixed portfolio weights. The dampening effect occurs for all but industry portfolios, which already are fixed-weight portfolios.

These computations confirm that the seasonalities in portfolio returns emanate from common attributes. Given that the portfolios are well-diversified, the portfolio-level seasonalities cannot be driven just by individual stock seasonalities.[12] Because we rebalance the portfolios annually, our tests measure seasonalities in, for example, the value, size, and industry effects. They do not measure seasonalities in returns of stocks that *today* have a particular attribute.

The results in Table 1 are stronger in the month of January. Consider, for example, the average annual slope coefficients in the "stocks" and "all" columns in this Table. These slope estimates decrease from the full-sample estimates of 0.97 and 3.94 to 0.60 and 2.03 when we exclude the January estimates.[13] The seasonalities, however, remain significant. The t-values for the annual-minus-non-annual estimates are 9.15 and 6.06 in the non-January sample. Panel B's high-minus-low portfolio returns are affected as well. The average monthly premia in the "stocks" and "all" columns decrease from 1.11% and 1.02% per month to 0.67% and 0.62% per month, but the t-values associated with these non-January averages are still 6.34 and 5.4. Our results thus agree in full with those in Heston and Sadka (2008): the return seasonalities are stronger in the months of January—both for individual stocks and portfolios of stocks—but are by no means limited to them.

2.5 Commonalities in return seasonalities

Table 2 evaluates the extent to which the seasonal patterns in well-diversified portfolios are independent of each other. We estimate regressions of high-minus-low seasonality portfolio returns (from Table 1 Panel B) on the returns on all other seasonal strategies. These spanning tests measure whether the seasonalities in, for example, BE/ME-sorted portfolios are subsumed by the seasonalities in the other portfolios.

The results reveal significant commonalities in seasonalities. The seasonalities in the D/P- and

[12] The average number of stocks in each of the 25 BE/ME-sorted portfolios is 25.9 in January 1963. This average rises above 50 in 1967, and is consistently above 100 after 1973.

[13] These non-January results are reported in the Internet appendix.

E/P-sorted portfolios, for example, are entirely spanned by those found in the other portfolios. The seasonal D/P strategy's return drops from 0.42% ($t = 3.9$) to -0.01% ($t = -0.12$) when this strategy is regressed against the other seasonal strategies. The return on the E/P strategy decreases from 0.78% ($t = 5.28$) to 0.06% ($t = 0.75$). This spanning behavior is very similar to that found in unconditional risk premia: the size and value effects, for example, subsume the returns on the E/P and D/P strategies (Fama and French 1992, 1996). Given that the commonalities in seasonalities have the same appearance as the commonalities in average returns, the set of factors that explain differences in average returns may be the same as the one that explains return seasonalities. If there are multiple return factors, and the premia on these factors accrue unevenly over the calendar year, stocks *must* display seasonalities in the cross-section by the virtue of loading differently on these factors.

The seasonalities in the portfolios formed from estimated factor loadings are also subsumed by other seasonalities, but not vice versa. The seasonalities in BE/ME portfolios, for example, remain significant whereas the seasonalities in portfolios sorted by HML loadings disappear after controlling for the other seasonalities. The finding that characteristics drive out estimated factor loadings may be due to the noisiness of the factor-loading estimates.[14] Not surprisingly, the returns on the HML- and SMB-based seasonality strategies have the highest correlations with the returns on the BE/ME- and ME-based seasonalities.

Even though the seasonalities in portfolio returns are correlated, they are not captured by any one attribute alone. Table 2 shows that the returns on the size, book-to-market, and industry seasonality strategies remain economically and statistically significant after controlling for all other seasonality strategies. The intercept in the BE/ME regression, for example, is 0.59% per month ($t = 5.51$), which is close to the excess return on this strategy in Table 1, 0.91% per month ($t = 7.06$).

[14]See Daniel and Titman (1997) and Davis, Fama, and French (2000) for discussions on the problem of disentangling factor loadings from characteristics.

2.6 Comovement within average-return-sorted portfolios

The seasonalities displayed by well-diversified portfolios suggest that common seasonalities may be responsible for a significant fraction of individual stock seasonalities. Before formalizing the mechanism through which these seasonalities are transmitted into individual stock returns, we first look for direct evidence of such commonalities. We do so by measuring the extent to which a sort of stocks into portfolios by their past same-month returns groups stocks with similar attributes together. If stocks ending up in a given portfolio are similar, they should display excessive comovement with other stocks in this portfolio. Our null hypothesis is that the seasonalities in individual stock returns are *stock specific*. Under the null hypothesis, stocks should comove as much with all other stocks, regardless of whether we measure covariation within or across average-return-sorted portfolios.

We test this hypothesis as follows. Every month t we sort stocks into ten portfolios based on their average historical same-month return over the prior 20-year period. In January 1999, for example, the average historical return is computed using every January from 1980 through 1999. Next, we extract for every stock their monthly returns over the following three-year period, excluding the calendar month on which the sort is based. In the ongoing example, the post-sort monthly returns would run from February 1999 through December 2001, excluding January 2000 and January 2001. A stock is included if it has at least five years of historical data and 12 months of post-sort data at the time of the sort.

We measure comovement between stocks by estimating ten regressions for every stock j using monthly post-sort returns. The first regression is against the average return on decile-1 stocks, the second regression is against the average return on decile-2 stocks, and so forth. In every regression we also control for the average return on all other stocks. We exclude stock j from the universe of stocks before computing the average returns on the portfolios that appear on the regression's right-hand side. We compute the average regression slopes for every decile-decile combination after estimating these regressions for all stocks. The first row in Table 3, for example, reports the average betas against deciles 1, ..., 10 for stocks that are in decile 1 at time t.

The diagonal elements in Table 3, which are extracted from the regressions against the other stocks in the stock's own decile, are larger than the off-diagonal elements. On the first row, for example, the own-portfolio beta in column 1 is 0.72 while the average of the nine other betas (columns 2–10) is -0.29. The last two columns report the difference between the own beta and the average and the t-value associated with this difference.[15] The large positive differences for the extreme portfolios show that stocks that have historically earned high or low returns in month m comove significantly in the *future*, even when month m is excluded from the analysis.

The own-minus-other beta differences are positive and large in all deciles in the data. The amount of comovement within industry-sorted portfolios is comparable to that within industry-sorted portfolios. When we apply the same methodology to measure comovement inside the 12 Fama-French industry portfolios, the own-minus-other differences range from 0.31 to 1.16 with an average of 0.73. These comovement results are consistent with the view that a sort on historical same-month returns groups together stocks with similar characteristics. These characteristics may, for example, be related to firm size, BE/ME ratios, and industry memberships. Whatever these characteristics are, they collectively generate a remarkable amount of comovement between stocks. These results reject the alternative hypothesis that past same-month returns measure either stock-specific seasonalities, or that stocks repeat high or low idiosyncratic returns from the past.

[15] We sort stocks into portfolios and run regressions in months $t = 1, \ldots, 12, 37, \ldots, 48, 73, \ldots$ in Table 3 to estimate the betas and t-values. That is, after running twelve regressions for one calendar year, we skip over the next two years before running any new regressions. This procedure avoids the use of overlapping post-sort return data for the regressions.

This extraordinary amount of comovement in the average-return-sorted portfolios is also reflected in the volatility of the strategy that trades on the seasonalities in individual stock returns. Table 1 shows that a strategy that is long the top decile of stocks of average same-month returns, and short the bottom decile, earns an average return of 1.11% per month. The annualized standard deviation of returns on this long-short strategy is 11.13%. If this strategy indeed picks up firm-specific seasonalities in expected returns, the long and short legs of this strategy should have very similar characteristics and factor loadings; any differences in them should wash out in the long-short strategy. We test whether the volatility of 11.13% is consistent with this conjecture. Every month when stocks are assigned to portfolios by their average same-month returns, we also assign them separately into ten *random* portfolios, and use these portfolios to generate one random long-short strategy. We then repeat this process 100 times. The average annualized volatility on this random strategy—which uses the same universe of stocks and the same time period as the true seasonality strategy—is 4.1%. The excessive volatility of the true seasonality strategy suggests that when we sort stocks by their average same-month returns, different portfolios are populated with stocks with similar characteristics or factor loadings.

3 Model of Stock Market Seasonalities

3.1 Model

We use a factor model to illustrate how small common seasonalities aggregate into larger seasonalities in stock returns. In the model N stocks follow a J-factor model:

$$r_{i,t} = \beta_{i,t}^1 F_t^1 + \beta_{i,t}^2 F_t^2 + \cdots + \beta_{i,t}^J F_t^J + \varepsilon_{i,t}, \tag{3}$$

where $r_{i,t}$ is stock i's excess return in month t, $\beta^j_{i,t}$ is stock i's sensitivity to the jth factor in month t, F^j_t is the realized factor return on factor j in month t, and $\varepsilon_{i,t}$ is the firm-specific shock with mean zero and variance $\sigma^2_\varepsilon < \infty$. We use the standard arbitrage pricing theory[16] terminology throughout our discussion of this model. The model can, however, also be interpreted in terms of characteristics and the return premia associated with these characteristics.

Month-t return on factor j consists of its risk premia, which varies by calendar month, and a factor shock:

$$F^j_t = \lambda^j_{m(t)} + \xi^j_t, \tag{4}$$

where $m(t)$ is the calendar month corresponding to date index t. The risk premia $\lambda^j_{m(t)} \sim N(0, \sigma^2_\lambda)$ are drawn once in the beginning by nature. We assume that $\sigma^2_\lambda > 0$, which means that the factor risk premia vary over the calendar year. The risk premia draws are independent across calendar months and factors: $\mathrm{E}(\lambda^j_m \lambda^j_{m'}) = 0$ for $m \neq m'$ and $\mathrm{E}(\lambda^j_m \lambda^{j'}_{m'}) = 0$ for $j \neq j'$. Factor shocks $\xi^j_t \sim N(0, \sigma^2_\xi)$ are similarly independent with $\mathrm{E}(\xi^j_t \xi^{j'}_{t'}) = 0$ if $t \neq t'$ or $j \neq j'$.

Firms' betas can change over time. We assume that betas follow mean zero AR(1) processes that are independent across stocks and factors,

$$\beta^j_{i,t} = (1 - \delta)\beta^j_{i,t-1} + \eta^j_{i,t}, \tag{5}$$

where $0 < \delta \leq 1$, $\mathrm{E}(\eta^j_{i,t}) = 0$, $\mathrm{E}((\eta^j_{i,t})^2) < \infty$, $\mathrm{E}(\eta^j_{i,t}\eta^{j'}_{i',t'}) = 0$ for $i \neq i'$, $j \neq j'$ or $t \neq t'$. We denote the cross-sectional variance of betas by σ^2_β. The case of constant betas is a special case with $\sigma^2_\beta > 0$, $\delta = 0$, and $\mathrm{E}((\eta^j_{i,t})^2) = 0$. We assume that the shocks to betas are independent of both the idiosyncratic shocks and factor returns. We set the unconditional means of the betas and risk premia to zero to

[16] See Ross (1976).

focus on the cross-sectional predictability in returns. We consider the case in which the parameters characterizing the factor and beta processes are the same for all J factors and N firms.

We compute the theoretical slope coefficient from a cross-sectional regression of month-t returns on month-$t-k$ returns to examine how the parameters of the economy influence the seasonality of stock returns. This slope coefficient is

$$\begin{aligned}
\hat{b}_k &= \frac{\sum_{i=1}^N r_{i,t} r_{i,t-k}}{\sum_{i=1}^N r_{i,t-k}^2} = \frac{\frac{1}{N}\sum_{i=1}^N \left[\left(\sum_{j=1}^J \beta_{i,t}^j F_t^j + \varepsilon_{i,t}\right)\left(\sum_{j=1}^J \beta_{i,t-k}^j F_{t-k}^j + \varepsilon_{i,t-k}\right)\right]}{\frac{1}{N}\sum_{i=1}^N \left[\left(\sum_{j=1}^J \beta_{i,t-k}^j F_{t-k}^j + \varepsilon_{i,t-k}\right)^2\right]} \\
&= \frac{\left\{\begin{array}{c} \frac{1}{N}\sum_{i=1}^N \left[\left(\sum_{j=1}^J \left((1-\delta)^k \beta_{i,t-k}^j + \sum_{s=1}^k (1-\delta)^{k-s} \eta_{i,t-k+s}^j\right)(\lambda_{m(t)}^j + \xi_t^j) + \varepsilon_{i,t}\right)\right. \\ \left. * \left(\sum_{j=1}^J \beta_{i,t-k}(\lambda_{m(t-k)}^j + \xi_{t-k}^j) + \varepsilon_{i,t-k}\right)\right] \end{array}\right\}}{\frac{1}{N}\sum_{i=1}^N \left[\left(\sum_{j=1}^J \beta_{i,t-k}(\lambda_{m(t-k)}^j + \xi_{t-k}^j) + \varepsilon_{i,t-k}\right)^2\right]}, \quad (6)
\end{aligned}$$

where the last line uses equation (5) to rewrite a stock's beta at time t as its beta k periods ago plus all intervening shocks. Using the independence of the firm-level shocks (the variables subscripted by i) from all other shocks, this regression coefficient converges in probability to

$$b_k \equiv \operatorname*{plim}_{N\to\infty} \hat{b}_k = (1-\delta)^k \frac{\sigma_\beta^2 \sum_{j=1}^J (\lambda_{m(t)}^j + \xi_t^j)(\lambda_{m(t-k)}^j + \xi_{t-k}^j)}{\sigma_\beta^2 \sum_{j=1}^J (\lambda_{m(t)}^j + \xi_{t-k}^j)^2 + \sigma_\varepsilon^2}. \quad (7)$$

Equation (7) is the value of the regression coefficient conditional on the values of $\lambda_{m(t)}^j$, $\lambda_{m(t-k)}^j$, ξ_t^j, and ξ_{t-k}^j. These values are random variables, so the expected regression coefficient is the expected value of b_k taken over the distributions of these variables. We show in the appendix that this expectation equals

$$\mathrm{E}(b_k) = \begin{cases} (1-\delta)^k \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\xi^2}\left(1 - \frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}\mathrm{E}\left[\frac{1}{Q_J+\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}}\right]\right) & \text{if } m(t) = m(t-k), \\ 0 & \text{if } m(t) \neq m(t-k), \end{cases} \tag{8}$$

where $Q_J \sim \chi^2(J)$. The expected slope coefficient is zero at non-annual lags because the seasonal differences in risk premia are uncorrelated; it is strictly positive at annual lags if $\sigma_\lambda^2 > 0$ and $\delta < 1$. By Jensen's inequality, the lower bound on the expected annual slope coefficient is[17]

$$\mathrm{E}(b_k) > (1-\delta)^k \frac{J\sigma_\beta^2\sigma_\lambda^2}{J\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)+\sigma_\varepsilon^2}. \tag{9}$$

3.2 Comparative statics

The regression coefficient in equation (8) illustrates how the parameters of the economy affect the seasonality in individual stock returns as measured by the cross-sectional regressions of month-t returns on month-$t-k$ returns. First, the speed at which the annual slope coefficients decrease as lag k increases depends on how persistent the betas are. The slower the mean reversion in betas, the more persistent the coefficient pattern. Second, the seasonalities in common factors *aggregate* into larger seasonalities in individual stock returns. The amount of seasonalities apparent in stock returns is strictly increasing in the number of common factors J. Increasing the number of factors from J to $J+1$ in equation (8) is equivalent to adding the square of a standard normal variable into the denominator inside the expectation, strictly increasing $\mathrm{E}(b_k)$.

Third, the dispersion in betas *amplifies* common seasonalities. The term $\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}\mathrm{E}\left[\frac{1}{Q_J+\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}}\right]$

[17]Because $\frac{1}{\tilde{x}}$ is convex in $\tilde{x} > 0$, $\mathrm{E}\left(\frac{1}{\tilde{x}}\right) > \frac{1}{\mathrm{E}(\tilde{x})}$. Jensen's inequality also implies that as J increases, the gap between the expected slope coefficient and the lower bound decreases.

can be rewritten as $\frac{\sigma_\varepsilon^2}{\sigma_\lambda^2+\sigma_\xi^2}\mathrm{E}\left[\frac{1}{\sigma_\beta^2 Q_J+\frac{\sigma_\varepsilon^2}{\left(\sigma_\lambda^2+\sigma_\xi^2\right)}}\right]$, which is strictly decreasing in σ_β^2. We note that as either the number of factors or the dispersion in betas increases, $\mathrm{E}(b_k)$ converges towards $(1-\delta)^k\frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\xi^2}$. This upper bound on the seasonality in individual stock returns depends on how large the seasonal differences in factor risk premia are relative to the factor shocks. The seasonality in individual stock returns is strictly decreasing in σ_ε^2, so an increase in idiosyncratic volatility decreases the seasonalities in stock returns.

The aggregation and amplification mechanisms in this model are of particular interest. Even if every common factor exhibits only modest seasonal variation in its risk premium, the seasonalities in individual stock returns can be prevalent because individual stocks aggregate these common seasonalities. Individual stock returns also amplify differences in risk premia through their factor loadings, that is, the dispersion in betas can blow up even small differences in risk premia. These results are important. Seasonalities in stock returns would emerge even if the factors were not directly observable, and even if these factors generated so small differences in calendar-month risk premia that we lacked the statistical power to estimate them from the data. In spite of our inability to directly observe such seasonalities, the seasonal pattern would emerge in cross-sectional regressions.

3.3 Average return as a sufficient statistic

A cross-sectional regression of today's stock returns against historical same-month stock returns results in a positive slope coefficient because historical returns serve as noisy signals of the products of stocks' betas and factor risk premia. Equation (8) shows that if there is any persistence in betas, all historical same-month returns are at least to some extent informative about today's returns. In the model the weighted-average same-month return subsumes the cyclical coefficient pattern in the data. The predictable component in month-t returns is the sum $\sum_{j=1}^{J} \beta_{i,t}^{j} \lambda_{m(t)}^{j}$. If betas are constant, the average historical same-month return converges to this same quantity,

$$\operatorname*{plim}_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} r_{i,t-12k} = \sum_{j=1}^{J} \beta_{i}^{j} \lambda_{m(t)}^{j}, \tag{10}$$

because both the firm-specific shocks and factor shocks are independent in the time series. The average same-month return is thus a sufficient statistic for predicting variation in expected returns under the constant-beta case. This implies that individual same-month returns should not explain variation in today's returns when the regression controls for the average same-month returns.

If betas are time-varying, the optimal predictor of $\sum_{j=1}^{J} \beta_{i,t}^{j} \lambda_{m(t)}^{j}$ is the weighted average of historical same-month returns, with greater weight placed on more recent observations. The stability of the betas determines how quickly the optimal weights decrease in k. The weights decrease in k because, the greater the k, the more likely a stocks' loadings have changed between $t-k$ and t. Older same-month returns thus have a lower signal-to-noise ratio than more recent returns. In particular, if there is any time variation in betas (that is, $\delta > 0$), historical same-month returns eventually become uninformative about returns today because all betas mean revert in expectation to the common mean of zero, $\lim_{k \to \infty} \mathrm{E}(\beta_{i,t-k} \mid \beta_{i,t}) = 0$.

4 Empirical Tests

4.1 Model estimation

We estimate the model using the Generalized Method of Moments. We set the number of factors and dispersion in betas to values based on literature before estimating the remaining four model parameters from the data. We set the number of factors to $J = 5$. Estimates between two and seven are typical in the literature on the number of common factors in returns.[18] Based on the estimates in Ang, Liu, and Schwarz (2010), we consider two different values for σ_β: 0.5 or 1. Ang et al. provide estimates of the cross-sectional dispersion in the Fama and French (1993) three-factor model loadings. In individual stock returns the standard deviations of the beta estimates are 0.68 ($\hat{\beta}_{\text{mkt}}$), 1.04 ($\hat{\beta}_{\text{smb}}$), and 1.04 ($\hat{\beta}_{\text{hml}}$). These estimates decrease to 0.61, 0.95, and 0.94 when stocks are first grouped into portfolios to decrease estimation noise. We choose the values of $\sigma_\beta = 0.5$ and 1 as approximate lower and upper bounds for reasonable variation in factor betas.

We use 13 moment conditions to estimate the model. The first ten conditions are based on the average cross-sectional regression coefficients b_k for lags $k = 12, 24, \ldots, 240$. We estimate these coefficients from the data and compare these values to their theoretical counterparts in equation (8). To reduce estimation uncertainty, we group these coefficients into ten two-year groups ($k = 12$ and 24, $k = 36$ and 48, and so forth).

The remaining moment conditions inform GMM about the magnitudes of the σ_λ^2, σ_ξ^2, and σ_ϵ^2 shocks in the model. The first of these additional conditions matches the cross-sectional return variance between the data and model. As the number of stocks increases, the expected variance of returns is

$$\operatorname*{plim}_{N\to\infty} \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N} r_{i,t}^2\right) = J\sigma_\beta^2\left(\sigma_\lambda^2 + \sigma_\xi^2\right) + \sigma_\varepsilon^2. \tag{11}$$

The second additional moment condition matches the covariance between month-t and month-$t-k$ returns. We take averages of these covariances over the first five annual lags ($k = 12, 24, \ldots, 60$). The

[18] A number of studies, such as those by Roll and Ross (1980), Brown (1989), and Connor and Korajczyk (1993), examine the covariance structure of returns to infer the number of common factors, and discuss the problems inherent in this task.

expectation of this covariance in the model at an annual lag is

$$\operatorname*{plim}_{N\to\infty} \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N} r_{i,t}r_{i,t-k}\right) = (1-\delta)^k J\sigma_\beta^2\sigma_\lambda^2. \tag{12}$$

The expected covariance is zero at non-annual lags. The third and last additional moment condition relates to the average variance of residuals from the cross-sectional regressions of month-t returns on month-$t-k$ returns. The expected variance of the residuals in the model is[19]

$$\begin{aligned}\operatorname*{plim}_{N\to\infty} \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}(r_{i,t}-\hat{b}_k r_{i,t-k})^2\right) &= \\ & J\sigma_\beta^2\left(\sigma_\lambda^2+\sigma_\xi^2\right)+\sigma_\varepsilon^2-(1-\delta)^{2k}\Bigg(\sigma_\beta^2\left[\Lambda^2(J-1)+1\right](\sigma_\lambda^2+\sigma_\xi^2) \\ & -\Lambda^2\sigma_\varepsilon^2+\sigma_\varepsilon^2\left[\Lambda^2\left(1+\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}\right)-1\right]\mathrm{E}\left(\frac{1}{Q_J+\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}}\right)\Bigg),\end{aligned} \tag{13}$$

where $\Lambda \equiv \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\xi^2}$. We average variances of the residuals over the first five annual regressions to reduce estimation uncertainty. In computing the data moments, we exclude stocks that have fewer than five years of historical data available at the time of each cross-sectional regression. We also scale the slope coefficient estimates by $\frac{\mathrm{var}(r_{i,t-k})}{\mathrm{var}(r_{i,t})}$ to adjust for changes in the cross-sectional variance of stock returns.[20]

The estimated model is characterized by four parameters: δ is the persistence in factor loadings; σ_λ^2 is the variance of the monthly differences in risk premia; σ_ξ^2 is the variance of the monthly factor return shocks; and σ_ε^2 is the firm-specific variance. Our 13 moment conditions identify the model parameters. The regression slope coefficients are informative about the speed of mean reversion in

[19] See the Internet appendix for details.

[20] See, for example, Campbell, Lettau, Malkiel, and Xu (2001). The slope coefficient moment conditions in Table 4 are thus not comparable with those in Table 1.

betas, and about the relative magnitudes of σ_λ^2- and σ_ξ^2-shocks; the cross-sectional variance of returns is informative about the sum of all shocks in the model; the covariance moment is informative about the magnitude of σ_λ^2; and the variance of the residuals is informative about the relative magnitudes of the σ_ξ^2-shocks relative to σ_ε^2-shocks. If σ_ξ^2 is low, most of the unpredictable variation in returns is due to the firm-specific shocks. Under this scenario, the variation of residuals is low because firm-specific shocks diversify more effectively in the cross-section than factor shocks.[21]

Table 4 Panel A shows the parameter estimates and Panel B shows the moments in the data and model. The assumed dispersion of betas does not influence the fit of the model. In the model the variance of betas always multiplies either σ_λ^2 or the sum $\sigma_\lambda^2 + \sigma_\xi^2$, so proportional changes in σ_λ^2 and σ_ξ^2 perfectly subsume a change in σ_β^2. Panel B shows that the model describes the data well. The two marginally significant differences between the moments in the data and model are the first and last regression coefficients. The first annual regression coefficient may be high in the data because, at this lag, returns may reflect not only the seasonality mechanism but also momentum. The model also matches the levels of return variances, covariances and variances of residuals.

The parameter estimates in Panel A indicate that the factor loadings are quite persistent. The estimated rate of mean reversion is $\delta = 0.0053$. This estimate implies that if a stock's beta today is 1, its expected value in one year is 0.938, and it is 0.725 in five years. This conclusion about the persistence in factor loadings is largely model-free. The fact that the seasonalities in stock return persist up to 20 years imposes a bound on how quickly the betas can change. The assumed dispersion of betas affects the estimates of the risk factor parameters. If the dispersion of betas is low, a standard deviation of 1.4% in the monthly risk premia matches the data; if dispersion is high, a standard

[21]The expected regression coefficient in the model in equation (8) is not equal to $\frac{\mathrm{E}(\mathrm{cov}(r_{i,t}, r_{i,t-k}))}{\mathrm{E}(\mathrm{var}(r_{i,t-k}))}$—which equals the lower bound in equation (9)—because factor returns enter both variances and covariances, and they do not disappear in the probability limit. Thus, even with expected regression coefficients, variances, and covariances as moment conditions, none of these moment conditions are redundant.

deviation of half this amount, 0.7%, suffices. Hence, the data are consistent with a model in which a few common factors exhibit small seasonal differences in risk premia. The empirical counterpart of σ_λ is the standard deviation of a factor's average January, February, ..., December return premia. The standard deviation for the average return difference between the biggest- and smallest-stock deciles, computed across months, is 2.16%. This standard deviation is 1.45% for portfolios formed by book-to-market ratios.[22]

The solid line in Figure 2 plots the theoretical slope coefficient at the estimated parameter values. The cyclical coefficient pattern matches that observed in the data (see Figure 1) both in magnitude and persistence. The heights of the peaks decrease slowly (from 1.09 at $k = 12$ to 0.61 at $k = 120$) because of the estimated slow rate of mean reversion in factor loadings. Because we do not build momentum or long-term reversals into the model, it does not match these features of the data. The dashed line in Figure 2 simulates return data from the model. We run this simulation side by side with the actual regressions to match the number of stocks in each cross-section to the number in the CRSP data set. We also equate the survival of stocks between the simulations and real data. The simulation in Figure 2 indicates that the firm-specific and factor shocks generate estimation errors that are comparable to those in the data.

4.2 Average same-month return and the seasonality in returns

In the model the average same-month return subsumes individual monthly returns in explaining variation in today's returns. Table 5 Panel A simulates data from the model and estimates the cross-sectional regressions with and without a control for the average same-month return. The first two

[22] These computations use the value-weighted returns on the size- and BE/ME-sorted portfolios supplied by Ken French from 1926 through 2011. The standard deviations of the average SMB and HML returns, computed by calendar month, are 0.75% and 0.74% over the same period.

columns report average Fama-MacBeth regression coefficients and t-values from baseline regressions of month-t returns against month-$t-k$ returns. These averages are taken from simulations similar to those plotted in Figure 2. The average coefficient is 1.16 at lag $k = 12$ (t-value $= 3.58$), decreasing to 0.32 (t-value $= 1.88$) when moving to lag $k = 240$.

The other columns of Table 5 Panel A show estimates from regressions that include the average historical same-month return as an additional regressor. The individual same-month historical returns are no longer significant in these regressions except at the longest lags with the opposite sign. The coefficient on month-$t-12$ return, for example, decreases from 1.16 to 0.39 (t-value $= 1.24$) when controlling for the average same-month return. The coefficient on the average historical same-month return ranges from 7.7 to 11.8, depending on lag k, and this variable's t-value is around 9. The equal-weighted average return subsumes almost all individual same-month returns because the factor loadings in the model are quite stable.

Table 5 Panel B estimate the same regressions using monthly returns on NYSE, AMEX, and Nasdaq stocks from January 1963 through December 2011. Inclusion of the average return control changes the coefficient pattern at annual lags significantly. While 18 of the 20 coefficients for lags $k = 12, 24, \ldots, 240$ in the first column (without the control) are statistically significant, only lags $k = 12$ and $k = 24$ are significant in the controlled regressions. The fact that $k = 12$ remains very significant—its coefficient and t-value decrease from 2.05 ($t = 6.90$) to 1.35 ($t = 4.67$)—suggests that returns are informative at the one-year mark, not only because of characteristics-induced seasonalities, but also because momentum still exists at this horizon.

The insignificance of individual same-month returns is important. The seasonality in stock returns does not arise because stocks today repeat an event that occurred precisely k years ago. Heston and Sadka (2008) show that when month-t returns are regressed against all annual-lag returns in a multivariate regression, many of the coefficients remain significant. The same is true in our model. The reason the returns, say, from 15 years ago, are *incrementally* informative about the cross-section of returns today, is that they are independent signals of cross-sectional differences in expected returns. The (weighted) average same-month return combines these independent signals into one signal.

The coefficients on the average same-month return behave similarly between the simulations and data as a function of lag k. The estimated coefficients stay flat until $k = 60$ after which they trace out an increasing and concave pattern. The reason for this pattern is that the sample becomes more selective for lags $k > 60$. Although every firm must have survived for at least five years to be included in the analysis—which is why the coefficient pattern remains flat up to the five-year mark—at lags $k > 60$ the sample firms have survived longer, and the average is computed using more data. The average becomes more precise as the amount of data increases. The precision increases proportional to the square root of k, so the coefficients increase at a decreasing rate.

4.3 Additional tests of the model

We use two additional tests to compare data simulated from the model to actual data. These results are reported in full in the Internet appendix. In the first test we show that portfolio sorts support the conclusions of Table 5's Fama-MacBeth regressions. Both in the data and model, a strategy that sorts stocks into portfolios based on the 20-year average same-month return dominates strategies that use shorter sub-sample strategies (such as sorting by average same-month return in years 2–5) in both statistical and economic significance.

In the second additional test we show that the stocks in the model also show a remarkable amount of comovement when sorted into portfolios by their past same-month returns. In the model stocks with similar characteristics are subject to common return shocks. As a result, stocks with similar characteristics comove even when their mean returns are not exceptionally high or low. The amount of comovement in the model at the estimated parameter values is comparable to that found in the data in Table 3.

4.4 Day-of-the-week seasonality in stock returns

In our model the seasonality in stock returns emerges because past returns are informative about stock characteristics (or factor loadings) that generate the common seasonalities in expected returns. The same pattern should thus emerge at any frequency at which cross-sectional seasonalities exist. Studies such as those by Osborne (1962), French (1980), and Gibbons and Hess (1981) find that the average daily market return is significantly negative on Mondays, and that the Wednesday and Friday returns are the highest. Keim and Stambaugh (1983) show that Friday returns are particularly high for small stocks, and they reject the equality of average returns across size deciles for every weekday except Wednesdays. Lakonishok and Maberly (1990) and Chan, Leung, and Wang (2004) relate weekday seasonalities to institutional ownership and trading.

A cyclical weekly coefficient pattern should emerge in daily returns-on-returns regressions. If we take a cross-section of stocks and find those with high Friday returns, we can infer that they are more likely small stocks—or that they perhaps possess some heretofore unknown set of characteristics associated with high Friday returns. Subsequent Friday returns should thus covary positively with historical Friday returns in the cross-section because both are functions of the same characteristics. This prediction about the day-of-the-week return seasonality is a powerful test of our model.

Figure 3 plots the average coefficients from daily Fama-MacBeth regressions of day-t return against day-$t-k$ return. We include the average daily return from month-$t-12$ to month-$t-2$ as an additional explanatory variable in every regression to control for one-year return momentum. Without this control the pattern in Figure 3 is the same except that the coefficients for lags greater than $k = 20$ shift up by a small amount. The circles in the figure denote lags that are multiples of five, that is, estimates from regressions in which the historical return occurs on the same weekday. The coefficient pattern in daily returns is strikingly similar to that observed in monthly data. The difference is that here the coefficients spike at weekly lags. The estimates are negative for the first four weeks (except for the same-day spikes at lags 15 and 20) because of the one-month reversal in stock returns.[23]

The first two columns in Table 6 Panel A report the same regressions with t-values. The statistical significance in these regressions is comparable to that in the monthly regressions. In fact, the estimates for all 48 lags corresponding to the same weekday are positive from three weeks ($k = 15$) up to the one-year mark ($k = 250$; not shown), and 31 out of these 48 estimates are associated with t-values greater than 2.0. The remaining columns of Panel A show that the cyclical coefficient pattern largely disappears when these regressions control for the average historical same-weekday return over the prior one-year period. The outcome is similar to that in Table 5 for the coefficient pattern in monthly returns.

Panel B reports, separately for each weekday, estimates from univariate regressions of day-t returns against the average historical same-weekday return. The average slope coefficients are very different from one another. While the unconditional average slope estimate is 4.29 (t-value = 14.4), it is just 1.40 (t-value = 1.4) using Wednesday data but as much as 8.40 (t-value = 15.3) using Friday data. The high Monday and Friday estimates are consistent with the literature having identified significant *common* seasonalities in the weekday effect on these two days.

[23]See, for example, Lehmann (1990), Jegadeesh (1990), and Gutierrez and Kelley (2008).

4.5 Decomposing seasonalities by stock characteristics

In Heston and Sadka (2008) the seasonalities in stock returns survive one-by-one controls for size, industry, earnings announcements, dividends, and end of fiscal year. In our model the cyclical coefficient pattern can survive controls for one (or many) characteristics when there are multiple common seasonalities. If there are C independent characteristics with equal variances, a set of controls for one characteristic decreases the coefficient pattern by (approximately) $1/C$. However, as we increase the number of controls, we also increase the amount of seasonalities we can explain.

Table 7 measures how much different stock attributes explain of the return seasonalities. We begin the analysis from Figure 1 Panel A. Ignoring the momentum and long-term reversal in stock returns, the coefficients in this figure would be zero at all lags if returns exhibited no seasonality. The sum of squared deviations of the annual-lag coefficients thus measures the distance from this no-seasonalities ideal. Table 7 reports how this distance decreases as we add more controls.

We measure seasonalities at the monthly level. Given that the seasonalities related to earnings announcements, dividend declarations, and ex-dividend dates are likely to operate at a finer frequency, we use a two-stage methodology that allows us to measure seasonalities at the daily level and interpret them at the monthly level. In the first stage we estimate cross-sectional regressions of daily log returns against different stock characteristics and sum these residuals into monthly returns. In the second stage we regress these returns against lagged monthly returns. The second-stage regressions are thus comparable to the regressions in Figure 1 except that the dependent variable is now orthogonal to the set of characteristics employed in the first stage. Because the regressor in the second stage is the actual month-$t-k$ return (and not a residual), the first-stage estimation errors appear only on the left-hand side in the second-stage regressions. The second-stage regressions yield a plot similar to that in Figure 1 Panel A except that the peaks are now less pronounced *if* the characteristics used in the first stage explain some of the return seasonalities. We re-estimate the no-control regressions (Figure 1 Panel A) using log-returns to be consistent with the characteristics-controlled regressions.

Table 7 groups characteristics into five groups: dummy variables for the 48 Fama-French industries; decile dummy variables for book-to-market ratio and dividend yield, decile dummy variables for firm size; dummy variables for ex-dividend dates and dividend declaration dates; and dummy variables for earnings announcement dates. We include date dummies for the dividend and earnings announcement events to examine whether these *stock-specific* events contribute to the seasonality in stock returns.[24] The date dummy variables take values of one for the $[-1, +1]$ window around the event and are zero otherwise. Because the Compustat database has no information on earnings announcements before 1973, we analyze both the full sample without a control for earnings announcements and the post-1973 subsample with a control for earnings announcements.

We estimate the explanatory power of stock characteristics by entering them into the first-stage regressions in random order. Because the first-stage regressions can reduce the covariance between month-t and -$t-k$ returns just by reducing the variation in the dependent variable, we estimate every regression twice. The first regression uses actual stock characteristics, but the second regression randomizes the rows of the data matrix so that stocks, for example, now belong to randomized industries. The explanatory power of characteristics set A is:

$$\text{Explanatory power of characteristics set A} = 1 - \frac{\text{SSQ}_{\text{actual characteristics} \mid \text{A}}}{\text{SSQ}_{\text{randomized characteristics} \mid \text{A}}}, \tag{14}$$

where the sums of squared deviations are computed at lags $k = 12, 24, \ldots, 240$. This computation

[24] Frazzini and Lamont (2007) and Barber, De George, Lehavy, and Trueman (2012) show that stock prices rise around scheduled earnings announcements, and Hartzmark and Solomon (2012) find stocks to earn higher returns around ex-dividend days and dividend announcements.

compares the new sum of squared deviations to the decrease we would obtain using randomized characteristics. We bootstrap the time series of the second-stage Fama-MacBeth coefficient estimates to obtain standard errors for the model's overall explanatory power and the change in the explanatory power when adding more characteristics.

The first four (Panel A) or five (Panel B) columns in Table 7 show the average change in the sum of squared residuals when the characteristics enter in different order. The first row in Panel A, for example, indicates that when the regressions only control for the 48 Fama-French industries, the cyclical coefficient pattern decreases by 33% using the sum-of-squares metric. If the first-stage regression, however, already controls for all other characteristics, the *incremental* decrease in this metric is only 7% (column $c = 4$). The last column reports the average change in the explanatory power of the model, that is, when the characteristics are entered into the model in random order.

The last column in Table 7 shows that firm size (35%), BE/ME and dividend yield (22%), and industry (17%) effects generate a significant fraction of the cyclical coefficient pattern. The two *stock-specific* controls, dividends and earnings announcements, are less important. In Panel B, these variables explain only 4% and 1% of the coefficient pattern, respectively, and the point estimates are not significant even when these controls appear alone in column $c = 1$. The sum of the last column's estimates corresponds to a first-stage regression that includes all characteristics. The estimates imply that the characteristics jointly explain approximately 80% of the seasonality in stock returns. The bootstrapped standard error for this estimate is 3%. Because these regressions control only for *salient* variables correlated with common seasonalities, and because these variables enter as 10 dummies, this 80% estimate should be considered as the lower bound for the amount common factors explain of the seasonality in stock returns.

Figure 4 plots the average Fama-MacBeth coefficients from regressions of month-t log-returns against month-$t-k$ log-returns (solid line), as well as the average coefficients from regressions of *characteristics-controlled* residuals on month-$t-k$ log-returns (dashed line). The characteristics-controlled residuals are the residuals from the first-stage regressions used in Table 7 Panel A. The average coefficient estimates are not very different for the baseline and characteristics-controlled regressions for non-annual lags. For annual lags, however, the characteristics-controlled coefficients are smaller than the baseline coefficients. The difference between these two series of coefficient is particularly large for longer lags. This coefficient pattern suggests that the characteristics used in our decomposition analysis are more stable than any seasonality-generating characteristics omitted from our regressions.

4.6 Seasonalities in individual stock returns and seasonality-mimicking factors

Table 8 measures the extent to which seasonalities in well-diversified portfolios explain the seasonalities in individual stock returns. The dependent variable is the return on the equal-weighted strategy that buys the 10% of stocks with the highest 20-year average same-month returns and sells the bottom 10% of these stocks. The first column shows that the average return on this strategy is 1.11% per month with a t-value of 8.99.

The other columns regress this return against various seasonality-mimicking factors. These factors are the returns on high-minus-low strategies that buy and sell well-diversified portfolios based on their 20-year average same-month returns. Whereas the three-factor model cannot explain any of the returns earned by the individual stock seasonality strategy (see Table 1 Panel B), columns (2) through (10) show that the individual stock seasonality strategy correlates significantly with seasonalities extracted from *every set* of well-diversified portfolios.

The alpha on the seasonality strategy decreases in these regressions. The intercept is 0.49% per month ($t = 4.64$) in column (9) that regresses the returns against all seasonality-mimicking factors. In this multivariate regression the individual stock seasonality strategy loads significantly on the seasonalities in size, industry, and HML-slope portfolios. The R^2 is 46% in this regression, suggesting that these factors capture a significant amount of the variation in the returns on this strategy. Importantly, these common seasonalities can be represented by just one factor. The last column shows that the intercept is 0.41% ($t = 4.04$) when the seasonality-mimicking factor is derived from the combined set of 124 size, book-to-market, dividend-to-price, and industry portfolios.

These results are consistent with those in Table 7 and Figure 4. The common factors we consider explain the majority, but not all, of the seasonalities in individual stock returns. The estimates in Table 2 again represent lower bounds because we attempt to explain individual stock seasonalities using factors derived from a fixed set of characteristics. A broader set of portfolios could give rise to a factor that captures more of the seasonalities in individual stock returns.

5 Conclusions

Well-diversified portfolios display return seasonalities that are similar to those found in individual stock returns. The seasonalities in portfolio returns themselves share commonalities: much as in the tests of differences in average returns, the seasonalities present in book-to-market- and size-sorted portfolios span those present in E/P- and D/P-sorted portfolios. The set of factors that generates the return seasonalities appears to be close to the one—or even the same as the one—that explains differences in average stock returns. This result suggests return seasonalities and the factors present in average returns are intimately related to one another.

We develop a factor model in which factors' risk premia vary over the calendar year. Even if the seasonalities in the risk premia are small, the seasonalities in individual stock returns can be large because stocks both amplify and aggregate risk premia seasonalities. Our model offers several testable predictions. For example, we would expect the cyclical coefficient pattern to disappear when controlling for the average same-month historical return—a prediction that holds up in the data.

The similarity between the results that use monthly and daily data suggests that the amplification and aggregation mechanisms are general features of asset returns. They may also apply to frequencies finer than daily data. The literature on weekday seasonalities, for example, does not advocate the view that these seasonalities arise from predictable variation in risk aversion or quantities of risk. Rather, the supposition in this literature is that these daily seasonalities arise from differences in households' and institutions' demand and supply for equities across weekdays.[25] If seasonalities in supply and demand generate the seasonalities in daily returns, they may also generate the cross-sectional seasonalities found in intraday data by Heston, Korajczyk, and Sadka (2010).

Our results have implications for research in asset pricing. First, our tests aggregate seasonalities across all common factors and therefore have far more power than those based on individual factors. Unlike the seasonalities found in any one factor, the aggregate factor seasonalities are so large that they are difficult to dismiss as a chance finding. Second, our results suggest that one can identify risk factors by searching not only for variables that explain variation in average returns but also for those associated with return seasonalities. Both methods can be just as powerful if risk premia exhibit seasonal variation. Third, our results inform asset pricing theory. A theory that aims to model variation in expected returns should also be able to account for the seasonalities in risk premia.

[25] See, for example, Chan, Leung, and Wang (2004) and Venezia and Shapira (2007).

Appendix: Expected Regression Coefficient

This appendix computes the expected slope coefficient from a cross-sectional regression of month-t returns on month-$t-k$ returns. Equation (7) gives the regression coefficient b_k conditional on the values of $\lambda^j_{m(t)}$, $\lambda^j_{m(t-k)}$, ξ^j_t, and ξ^j_{t-k}. Taking the expectation of b_k and then conditioning inside this expectation on the values of $\lambda^j_{m(t-k)}$ and ξ^j_{t-k},

$$
\begin{aligned}
\mathrm{E}(b_k) &= \mathrm{E}\left[(1-\delta)^k \frac{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_t)(\lambda^j_{m(t-k)} + \xi^j_{t-k})}{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2 + \sigma^2_\varepsilon}\right] \\
&= (1-\delta)^k \mathrm{E}\left[\mathrm{E}\left(\frac{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_t)(\lambda^j_{m(t-k)} + \xi^j_{t-k})}{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2 + \sigma^2_\varepsilon} \,\middle|\, \{\lambda^j_{m(t-k)}, \xi^j_{t-k}\}_{j=1}^J\right)\right] \\
&= (1-\delta)^k \mathrm{E}\left[\frac{\sigma^2_\beta \sum_{j=1}^J \mathrm{E}\left(\lambda^j_{m(t)} + \xi^j_t \,\middle|\, \{\lambda^j_{m(t-k)}, \xi^j_{t-k}\}_{j=1}^J\right)(\lambda^j_{m(t-k)} + \xi^j_{t-k})}{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2 + \sigma^2_\varepsilon}\right] \qquad \text{(A-1)}
\end{aligned}
$$

Here, $\mathrm{E}(\xi^j_t \,|\{\lambda^j_{m(t-k)}, \xi^j_{t-k}\}_{j=1}^J) = 0$ and, if $m(t) \neq m(t-k)$, then $\mathrm{E}(\lambda^j_{m(t)} \,|\{\lambda^j_{m(t-k)}, \xi^j_{t-k}\}_{j=1}^J) = 0$ as well. The expected regression coefficient at a non-annual lag is thus zero. Continuing with the assumption that $m(t) = m(t-k)$:

$$
\begin{aligned}
\mathrm{E}(b_k) &= (1-\delta)^k \mathrm{E}\left[\frac{\sigma^2_\beta \sum_{j=1}^J \lambda^j_{m(t)}(\lambda^j_{m(t)} + \xi^j_{t-k})}{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2 + \sigma^2_\varepsilon}\right] \\
&= (1-\delta)^k \mathrm{E}\left[\mathrm{E}\left(\frac{\sigma^2_\beta \sum_{j=1}^J \lambda^j_{m(t)}(\lambda^j_{m(t)} + \xi^j_{t-k})}{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2 + \sigma^2_\varepsilon} \,\middle|\, \{\lambda^j_{m(t)} + \xi^j_{t-k}\}_{j=1}^J\right)\right] \\
&= (1-\delta)^k \mathrm{E}\left[\frac{\sigma^2_\beta \sum_{j=1}^J \mathrm{E}\left(\lambda^j_{m(t)} \,\middle|\, \{\lambda^j_{m(t)} + \xi^j_{t-k}\}_{j=1}^J\right)(\lambda^j_{m(t)} + \xi^j_{t-k})}{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2 + \sigma^2_\varepsilon}\right] \\
&= (1-\delta)^k \frac{\sigma^2_\lambda}{\sigma^2_\lambda + \sigma^2_\xi} \mathrm{E}\left[\frac{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2}{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2 + \sigma^2_\varepsilon}\right] \qquad \text{(A-2)} \\
&= (1-\delta)^k \frac{\sigma^2_\lambda}{\sigma^2_\lambda + \sigma^2_\xi}\left(1 - \sigma^2_\varepsilon \mathrm{E}\left[\frac{1}{\sigma^2_\beta \sum_{j=1}^J (\lambda^j_{m(t)} + \xi^j_{t-k})^2 + \sigma^2_\varepsilon}\right]\right)
\end{aligned}
$$

$$
\begin{aligned}
&= (1-\delta)^k \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\xi^2}\left(1-\sigma_\varepsilon^2 \mathrm{E}\left[\frac{1}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)\sum_{j=1}^J \epsilon_j^2+\sigma_\varepsilon^2}\right]\right) \\
&= (1-\delta)^k \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\xi^2}\left(1-\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}\mathrm{E}\left[\frac{1}{\sum_{j=1}^J \epsilon_j^2+\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}}\right]\right) \qquad \text{(A-3)} \\
&= (1-\delta)^k \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\xi^2}\left(1-\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}\mathrm{E}\left[\frac{1}{Q_J+\frac{\sigma_\varepsilon^2}{\sigma_\beta^2(\sigma_\lambda^2+\sigma_\xi^2)}}\right]\right), \qquad \text{(A-4)}
\end{aligned}
$$

where $\epsilon_j \sim N(0,1)$ and $Q_J \sim \chi_J^2$. Line (A-2) uses the normality and independence of $\lambda_{m(t)}^j$ and ξ_{t-k}^j,

$$
\mathrm{E}\left(\lambda_{m(t)}^j \,\middle|\, \{\lambda_{m(t)}^j+\xi_{t-k}^j\}_{j=1}^J\right) = \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_\xi^2}(\lambda_{m(t)}^j+\xi_{t-k}^j), \qquad \text{(A-5)}
$$

and line (A-3) standardizes $\lambda_{m(t)}^j+\xi_{t-k}^j$ to have a unit variance and writes their sum as that of standard normal variables. The expression on line (A-4) is that given in equation (8).

REFERENCES

Ang, A., J. Liu, and K. Schwarz (2010). Using stocks or portfolios in tests of factor models. Columbia Business School working paper.

Barber, B. M., E. T. De George, R. Lehavy, and B. Trueman (2012). The earnings announcement premium around the globe. *Journal of Financial Economics*. Forthcoming.

Brown, S. J. (1989). The number of factors in security returns. *Journal of Finance 44*(5), 1247–1262.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance 56*(1), 1–43.

Chan, S. H., W.-K. Leung, and K. Wang (2004). The impact of institutional investors on the Monday seasonal. *Journal of Business 77*(4), 967–986.

Choi, H.-S. (2008, December). *Three Essays on Stock Market Seasonality.* Ph. D. thesis, Georgia Institute of Technology.

Connor, G. and R. A. Korajczyk (1993). A test for the number of factors in an approximate factor model. *Journal of Finance 48*(4), 1263–1291.

Daniel, K. and S. Titman (1997). Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance 52*(1), 1–33.

Davis, J. L., E. F. Fama, and K. R. French (2000). Characteristics, covariances, and average returns: 1929 to 1997. *Journal of Finance 55*(1), 389–406.

Dyl, E. A. (1977). Capital gains taxation and year-end stock market behavior. *Journal of Finance 32*(1), 165–175.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *Journal of*

Finance 47(2), 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance 51*(1), 55–84.

Frazzini, A. and O. A. Lamont (2007). The earnings announcement premium and trading volume. AQR Capital Management working paper.

French, K. R. (1980). Stock returns and the weekend effect. *Journal of Financial Economics 8*(1), 55–69.

Garrett, I., M. J. Kamstra, and L. A. Kramer (2005). Winter blues and time variation in the price of risk. *Journal of Empirical Finance 12*, 291–316.

Gibbons, M. R. and P. Hess (1981). Day of the week effects and asset returns. *Journal of Business 54*(4), 579–596.

Grinblatt, M. S. and T. J. Moskowitz (2004). Predicting stock price movements from past returns: the role of consistency and tax-loss selling. *Journal of Financial Economics 71*(3), 541–579.

Gutierrez, Jr, R. C. and E. K. Kelley (2008). The long-lasting momentum in weekly returns. *Journal of Finance 63*(1), 415–447.

Hartzmark, S. M. and D. H. Solomon (2012). The dividend month premium. *Journal of Financial Economics*. Forthcoming.

Heston, S. L., R. A. Korajczyk, and R. Sadka (2010). Intraday patterns in the cross-section of stock returns. *Journal of Finance 65*(4), 1369–1407.

Heston, S. L. and R. Sadka (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics 87*(2), 418–445.

Heston, S. L. and R. Sadka (2010). Seasonality in the cross-section of stock returns: The international evidence. *Journal of Financial and Quantitative Analysis 45*(5), 1133–1160.

Jacobsen, B. and N. Visaltanachoti (2009). The Halloween effect in U.S. sectors. *Financial Review 44*(3), 437–459.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance 45*(3), 881–898.

Kamstra, M. J., L. A. Kramer, and M. D. Levi (2003). Winter blues: A SAD stock market cycle. *American Economic Review 93*(1), 324–343.

Kamstra, M. J., L. A. Kramer, and M. D. Levi (2012). Seasonal variation in Treasury returns. University of Toronto working paper.

Kamstra, M. J., L. A. Kramer, M. D. Levi, and T. Wang (2012). Seasonally varying preferences: Theoretical foundations for an empirical regularity. University of Toronto working paper.

Kang, J., T. Pekkala, C. Polk, and R. M. Ribeiro (2011). Stock prices under pressure: How tax and interest rates drive returns at the turn of the tax year. Harvard University working paper.

Keim, D. B. (1983). Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics 12*(1), 13–32.

Keim, D. B. and R. F. Stambaugh (1983). A further investigation of the weekend effect in stock returns. *Journal of Finance 39*(3), 819–835.

Lakonishok, J. and E. Maberly (1990). The weekend effect: Trading patterns of individual and institutional investors. *Journal of Finance 45*(1), 231–243.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics 104*(1), 1–28.

Osborne, M. (1962). Periodic structure in the Brownian motion of stock prices. *Operations Research 10*(3), 345–379.

Reinganum, M. R. (1983). The anomalous stock market behavior of small firms in January: Empirical tests for tax-loss selling effects. *Journal of Financial Economics 12*(1), 89–104.

Roll, R. and S. Ross (1980). An empirical examination of the arbitrage pricing theory. *Journal of Finance 35*(5), 1073–1103.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*(3), 341–360.

Rozeff, M. S. and W. R. Kinney (1976). Capital market seasonality: The case of stock returns. *Journal of Financial Economics 3*(4), 379–402.

Shumway, T. (1997). The delisting bias in CRSP data. *Journal of Finance 52*(1), 327–340.

Shumway, T. and V. A. Warther (1999). The delisting bias in CRSP's Nasdaq data and its implications for the size effect. *Journal of Finance 54*(6), 2361–2379.

Venezia, I. and Z. Shapira (2007). On the behavioral differences between professional and amateur investors after the weekend. *Journal of Banking and Finance 31*(5), 1417–1426.

Wachtel, S. B. (1942). Certain observations on seasonal movements in stock prices. *Journal of Business 15*(2), 184–193.

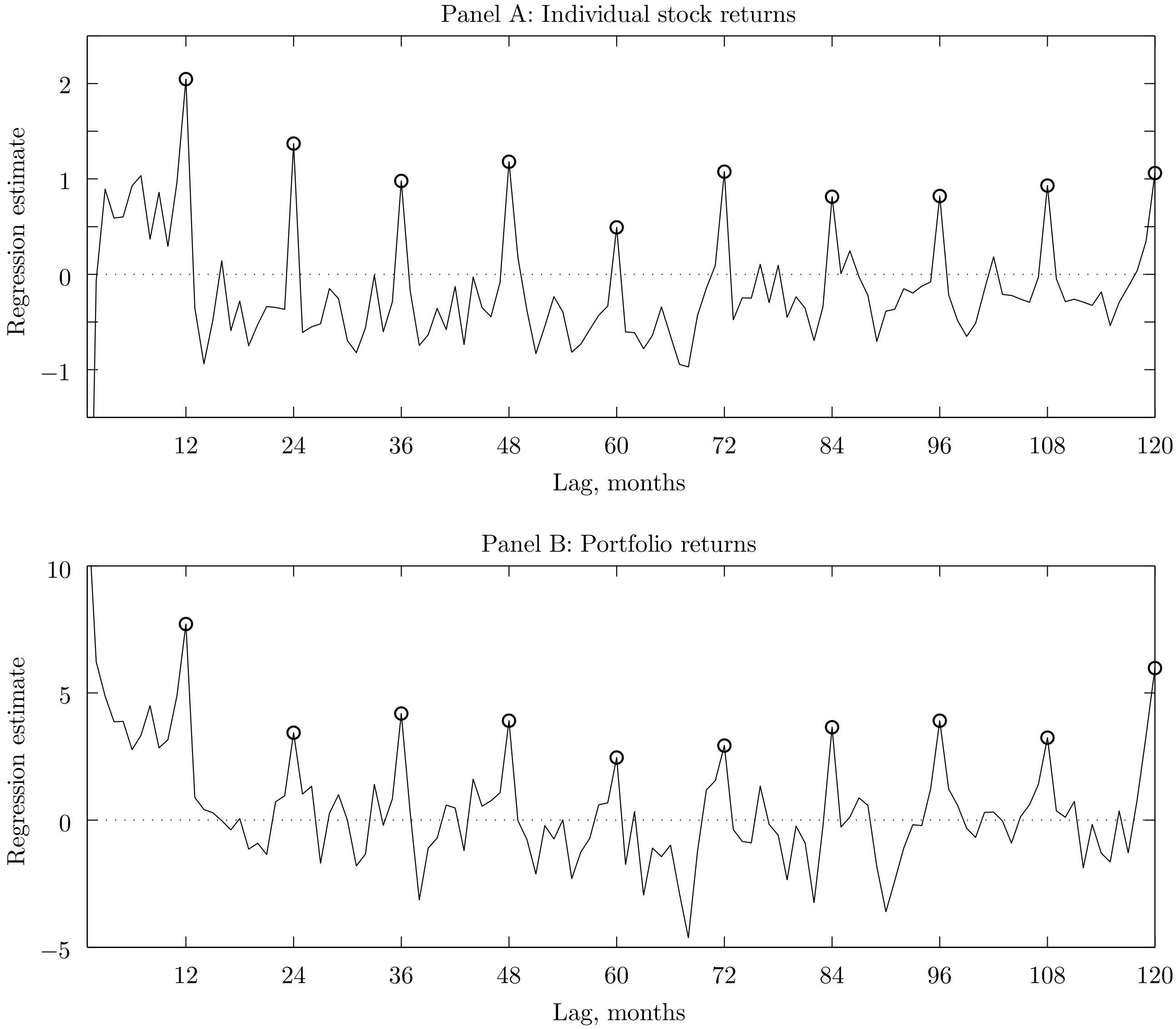

Figure 1: **Seasonalities in individual stock and portfolio returns.** Panel A uses data from January 1963 through December 2011 for NYSE, AMEX, and Nasdaq stocks to estimate univariate Fama-MacBeth regressions of month-t returns against month-$t-k$ returns, $r_{i,t} = a_t + b_t \ r_{i,t-k} + e_{i,t}$, in which the lag k ranges from one to 120 months. The regression coefficients are multiplied by 100. Panel B runs the same regressions using monthly return data on 124 portfolios: 25 portfolios sorted by market capitalization, 25 portfolios sorted by BE/ME ratios, 25+1 portfolios sorted by dividend-to-price ratio with $D = 0$ firms in a separate portfolio, and 48 Fama-French industry portfolios. The circles denote estimates at annual lags.

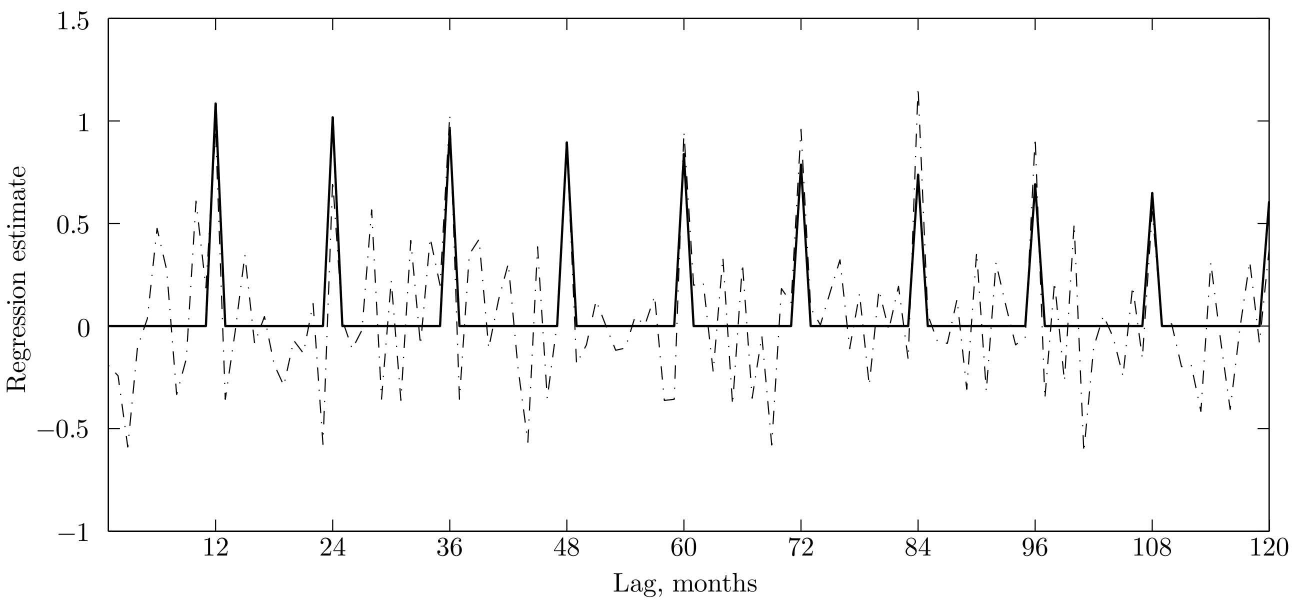

Figure 2: **Model-based seasonality in monthly returns.** The solid line in this figure plots the theoretical expected slope coefficient from a regression of month-t returns against month-$t-k$ returns, with k ranging from 1 to 120. The regression coefficients are multiplied by 100. In the model five common factors exhibit seasonal variation in their risk premia. The dashed line plots coefficient estimates from one simulation. The model is estimated using GMM (see Table 4) and the simulation is run at the estimated parameter values. The number of stocks, the length of the sample, and the survival of stocks in the simulations match the properties of actual stock data in Figure 1.

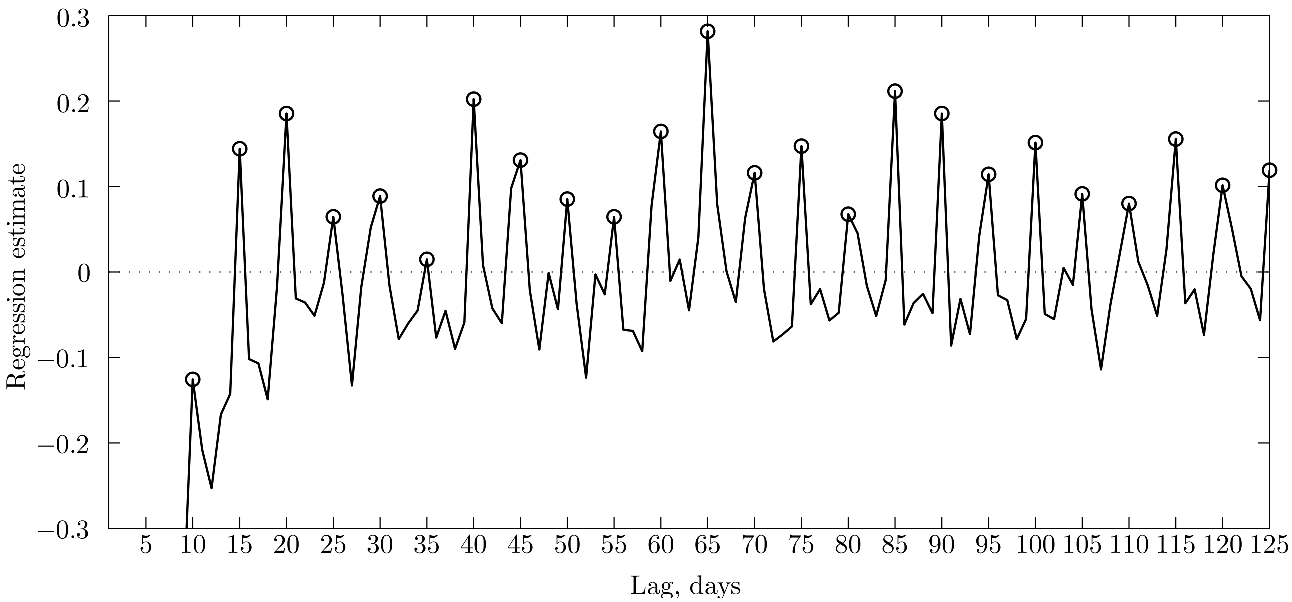

Figure 3: **Day-of-the-week seasonalities in stock returns.** This figure uses data on NYSE, AMEX, and Nasdaq stocks from January 1963 through December 2011 to estimate Fama-MacBeth regressions of day t returns against day $t - k$ returns, where k ranges from 1 to 125. The regression coefficients are multiplied by 100. These regressions also include the average daily stock return from month $t - 12$ to $t - 2$ to control for one-year return momentum (not reported). The circles denote multiples-of-five lags, that is, the same weekday from one, two, . . . , 25 weeks ago.

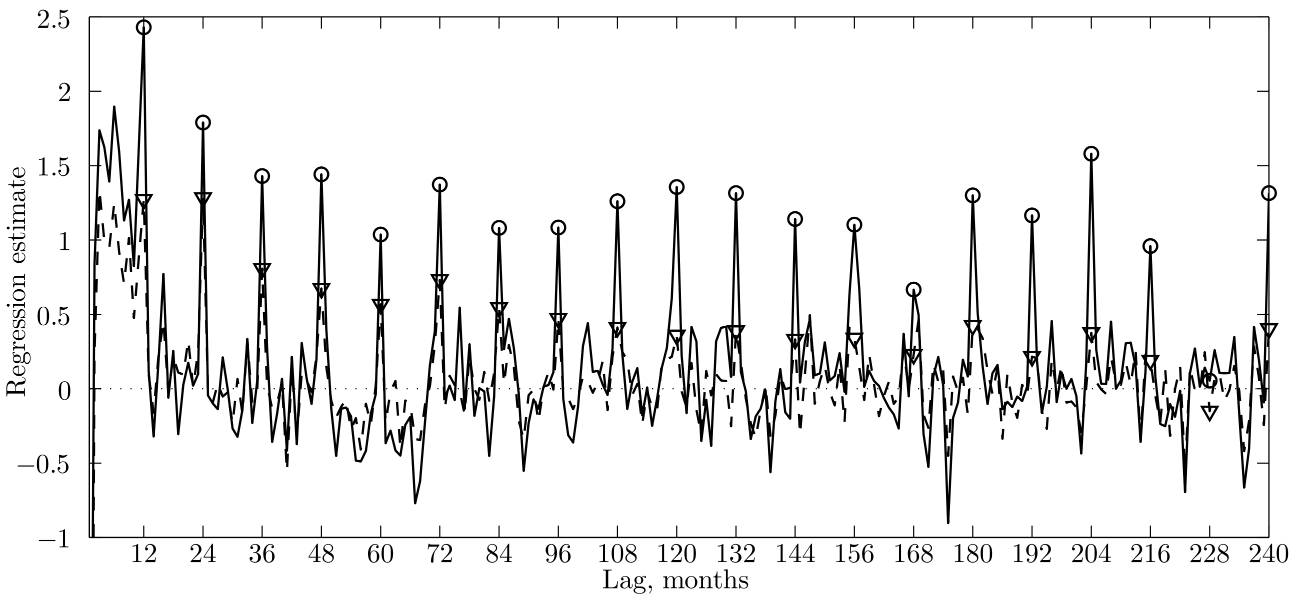

Figure 4: **Seasonality in monthly returns with controls for stock characteristics.** The solid line in this figure represents the average Fama-MacBeth coefficients from regressions of month-t log-returns against month-$t-k$ log-returns, where k ranges from 1 to 240. The regression coefficients are multiplied by 100. The circles denote estimates at annual lags. The dashed line represents coefficients from regressions of characteristics-controlled residuals on month-$t-k$ log-returns. The characteristics-controlled residuals are computed by first regressing daily returns on firm size, dividend yield, and BE/ME decile dummy variables as well as on dummy variables for the 48 Fama-French industries. The residuals from these regressions are then summed by month. Triangles denote estimates from these characteristics-controlled residual regressions at annual lags.

Table 1: Seasonalities in individual stock and portfolio returns

This table measures seasonality in individual stock and portfolio returns using cross-sectional Fama-MacBeth regressions and portfolio sorts. The Fama-MacBeth regressions in Panel A regress month-t returns on month-$t-k$ returns. Rows $k = 12, \ldots, k = 60$ show the average coefficients (with t-values in parentheses) from these regressions for the first five annual lags. Row "Average annual" reports the average coefficient over the 20 annual univariate regressions, and row "Average non-annual" is the average coefficient over the 209 non-annual univariate regressions in which k ranges from 13 to 239 and excludes all annual lags. The first column uses individual stock returns and the remaining columns use returns on portfolios of stocks. A stock is included in the first column's sample if it has at least five years of historical data available at the time of the analysis. Except for the industry portfolios, we rebalance the portfolios every June and then compute the equal-weighted returns from July of year t to June of year $t+1$. We form 25 portfolios by sorting on size (ME), book-to-market ratios (BE/ME), dividend-to-price ratios (D/P), earnings-to-price ratios (E/P), and on the estimated Fama-French three-factor model loadings. Zero dividend and negative earnings stocks are assigned into separate portfolios, so the number of D/P and E/P portfolios is 25+1. The three-factor model loadings are estimated using five years of monthly returns from July of year $t-5$ to June of year t. Column "Industry" uses the 48 Fama-French industry portfolios. Column "All" uses the returns on the 124 portfolios that result from pooling the ME, BE/ME, D/P, and industry portfolios. The same-month portfolio sorts in Panel B assign individual stocks (first column) or portfolios of stocks (other columns) into ten portfolios based on their average return in the same month over the past 20 years. "Same $-$ Other, $\hat{\alpha}_{\text{ff3}}$" is the monthly alpha from the three-factor model. In January 1963, for example, stocks and portfolios are assigned into portfolios based on the average January return in years 1943–1962. The other-month portfolio sorts assign stocks or portfolios of stocks into portfolios based on the average return in all other months. In January 1963, for example, the assignments are based on the average February through December returns in years 1943–1961. Except for the E/P portfolios, the regressions and portfolio sorts use monthly data from January 1963 through December 2011. The E/P portfolio returns begin in January 1973. The independent variables in Panel A's regressions and the sorting variables in Panel B extend to January 1943.

		Portfolios								
								Factor loadings		
	Stocks	ME	BE/ME	D/P	E/P	Industry	All	MKT	HML	SMB
Panel A: Fama-MacBeth Regressions: $r_{j,t} = a_t + b_t\, r_{j,t-k} + e_{j,t}$										
$k = 12$	2.02	12.74	7.04	4.15	3.02	6.13	7.71	5.54	6.40	8.41
	(6.81)	(5.63)	(3.84)	(1.99)	(1.40)	(4.69)	(5.88)	(2.67)	(2.84)	(3.28)
$k = 24$	1.38	9.83	3.49	2.62	4.68	1.37	3.43	3.04	2.89	6.49
	(4.90)	(4.19)	(1.76)	(1.17)	(2.04)	(1.09)	(2.63)	(1.50)	(1.25)	(2.44)
$k = 36$	0.98	4.56	7.57	4.65	6.61	2.52	4.19	4.87	7.04	7.03
	(3.73)	(1.85)	(3.73)	(2.09)	(2.73)	(1.98)	(3.21)	(2.26)	(2.76)	(2.47)
$k = 48$	1.19	7.68	6.48	0.60	6.83	1.97	3.91	4.75	7.56	5.72
	(4.38)	(3.01)	(3.09)	(0.27)	(2.68)	(1.54)	(2.96)	(2.14)	(3.02)	(1.92)
$k = 60$	0.50	2.52	3.74	0.39	2.74	0.95	2.46	−0.90	−1.07	2.08
	(2.13)	(0.99)	(1.93)	(0.19)	(1.17)	(0.73)	(1.99)	(−0.39)	(−0.48)	(0.73)
Average annual	0.97	7.08	6.08	4.02	5.90	1.92	3.94	3.21	4.67	6.86
Average non-annual	−0.26	−1.24	1.99	−0.45	0.83	−0.33	−0.25	−0.33	−0.53	−1.02
Difference, t-value	(9.48)	(6.60)	(5.81)	(5.33)	(4.55)	(6.87)	(8.05)	(4.18)	(5.44)	(4.95)
Panel B: Portfolio Sorts										
Sort by	1.11	0.88	0.91	0.42	0.78	0.91	1.02	0.46	0.62	0.79
same-month return	(8.99)	(5.95)	(7.06)	(3.90)	(5.28)	(5.19)	(7.87)	(3.45)	(5.42)	(4.18)
Sort by	−1.43	−0.55	0.76	−0.26	0.22	−0.62	−0.25	−0.31	−0.25	−0.60
other-month return	(−8.08)	(−3.44)	(4.80)	(−2.34)	(1.50)	(−3.07)	(−2.45)	(−2.19)	(−1.36)	(−3.17)
Same − Other	2.54	1.43	0.15	0.68	0.56	1.53	1.27	0.77	0.87	1.38
	(12.00)	(5.49)	(0.98)	(4.71)	(3.81)	(6.05)	(7.29)	(4.24)	(4.13)	(4.60)
Same − Other, $\hat{\alpha}_{\mathrm{ff3}}$	2.15	1.39	0.22	0.60	0.68	1.32	1.17	0.74	0.95	1.31
	(11.46)	(5.75)	(1.53)	(4.10)	(4.69)	(5.69)	(7.01)	(3.96)	(4.76)	(4.72)

Table 2: Commonalities in portfolio return seasonalities

This table measures the extent to which the return seasonalities in well-diversified portfolios of stocks share commonalities. We form 25 portfolios by sorting on size (ME), book-to-market ratios (BE/ME), dividend-to-price ratios (D/P), earnings-to-price ratios (E/P), and on the estimated Fama-French three-factor model loadings, and also use the 48 Fama-French industry portfolios. We compute returns on seasonality-mimicking portfolios by sorting each of the 25 (or 48) portfolios into 10 portfolios based on these portfolios' average same-month return over the past 20 years. In January 1963, for example, we assign portfolios into portfolios based on their average January return in years 1943–1962. This table reports the intercepts and slopes from regressions of high-minus-low seasonality portfolio returns on the returns on all other seasonal strategies; the corresponding t-values are in parentheses. The intercept, which is comparable to the "sort by same-month return" estimate in Table 1 Panel B, measures whether the seasonalities in one set of portfolios are subsumed by the seasonalities in the other portfolios. Except for the E/P portfolios, the monthly portfolio returns are from January 1963 through December 2011. Because the E/P portfolio returns begin in January 1973, the seasonalities in the E/P portfolios are not used to explain the seasonalities in other portfolios.

	Dependent variable							
Independent						Factor loadings		
variable	ME	BE/ME	D/P	E/P	Industry	MKT	HML	SMB
Intercept	0.39	0.59	−0.01	0.06	0.60	−0.06	−0.08	−0.16
	(3.97)	(5.51)	(−0.12)	(0.75)	(3.38)	(−0.59)	(−0.66)	(−1.46)
ME		0.00	0.04	0.08	0.11	0.05	0.03	0.56
		(0.08)	(1.00)	(2.48)	(1.49)	(1.05)	(0.41)	(6.67)
BE/ME	0.00		0.16	0.41	−0.24	0.05	0.37	0.07
	(0.08)		(2.29)	(12.10)	(−2.28)	(1.03)	(4.72)	(1.01)
D/P	0.06	0.20		0.03	0.12	0.22	0.10	0.15
	(0.97)	(2.62)		(0.79)	(1.37)	(4.11)	(1.70)	(2.34)
Industry	0.05	−0.11	0.04	−0.09		0.02	0.13	0.10
	(1.52)	(−2.23)	(1.29)	(−3.30)		(0.62)	(2.11)	(2.40)
Factor loadings								
MKT	0.06	0.05	0.19	0.14	0.06		0.23	0.32
	(1.02)	(1.06)	(4.55)	(3.09)	(0.62)		(4.92)	(3.48)
HML	0.03	0.40	0.09	0.09	0.34	0.23		0.14
	(0.40)	(7.78)	(1.75)	(1.68)	(2.36)	(4.58)		(1.22)
SMB	0.47	0.06	0.09	0.19	0.18	0.23	0.10	
	(15.85)	(1.02)	(2.28)	(5.79)	(2.62)	(3.59)	(1.20)	
Adj. R^2	0.47	0.34	0.34	0.69	0.21	0.45	0.49	0.59

Table 3: Comovement within average-return-sorted portfolios

This table measures excess comovement among stocks within average same-month return-sorted portfolios. Stocks are sorted into deciles every month based on their average same-month return over the prior 20-year period. We estimate ten regressions for each stock. The first regression is against the average return on decile-1 stocks, the second regression is against the average return on decile-2 stocks, and so forth. We use three years of post-sort returns, excluding same-month returns $t + 12$, $t + 24$, and $t + 36$, to estimate these regressions. These regressions also include the average return on all other stocks as a control. The stock that appears on the regression's left-hand side is not included in the right-hand side portfolios. The pth row in this table reports the average betas against portfolios $1, \ldots, 10$ for decile-p stocks. The last two columns report the difference between the own-portfolio beta and the average of the nine other-portfolio betas and the t-value associated with this difference. A stock is included in the analysis at time t if it has at least five years of historical and one year of post-sort data. This table uses data on NYSE-, AMEX-, and Nasdaq-listed stocks from 1963 through 2011.

	Beta against portfolio										Own	
Portfolio	1	2	3	4	5	6	7	8	9	10	− Other	t-value
1	0.72	0.78	0.13	−0.52	−0.74	−0.87	−0.86	−0.73	−0.23	0.44	1.01	33.57
2	0.32	0.46	0.28	0.03	−0.10	−0.19	−0.27	−0.25	−0.15	0.12	0.48	19.48
3	0.11	0.23	0.33	0.28	0.21	0.13	0.06	−0.06	−0.09	−0.02	0.24	11.01
4	−0.01	0.08	0.28	0.37	0.37	0.33	0.25	0.13	−0.03	−0.09	0.22	11.33
5	−0.07	−0.01	0.23	0.39	0.44	0.43	0.37	0.26	0.04	−0.10	0.27	16.23
6	−0.10	−0.08	0.14	0.37	0.43	0.44	0.42	0.35	0.11	−0.09	0.27	14.58
7	−0.11	−0.14	0.08	0.30	0.39	0.43	0.46	0.38	0.20	−0.03	0.30	17.33
8	−0.08	−0.13	−0.05	0.15	0.26	0.36	0.37	0.40	0.29	0.05	0.27	13.75
9	0.00	−0.09	−0.13	−0.07	0.00	0.10	0.20	0.31	0.37	0.25	0.30	12.43
10	0.26	0.08	−0.28	−0.52	−0.51	−0.43	−0.22	−0.04	0.43	0.60	0.73	24.83

Table 4: Generalized method of moments estimation: Parameter estimates and moment conditions

This table shows parameter estimates (Panel A) and data and model moments (Panel B) from Generalized Method of Moments estimation. In the model the risk premia on $J = 5$ common factors vary over the calendar year. The model parameters are: δ is the rate of mean reversion in factor loadings; σ_λ^2 is the monthly variance of risk premia seasonalities; σ_ξ^2 is the monthly variance of factor return shocks; and σ_ε^2 is the monthly firm-specific variance. The estimation uses the iterated optimal GMM weighting matrix. Panel A sets the factor loading dispersion to $\sigma_\beta = 0.5$ or 1. This parameter does not affect the fit of the model. The moment conditions are the univariate regression coefficients from regressing month-t returns on month-$t-12$, -$t-24$, ..., -$t-240$ returns, averaged into ten two-year groups and scaled to keep the cross-sectional variance of returns constant; the cross-sectional variance in stock returns; the average covariance between month-t and month-$t-k$ returns for the first five annual lags; and the average variance of residuals from the first five annual regressions of returns on historical same-month returns. The moments in Panel B are multiplied by 100. The data are the returns on NYSE, AMEX, and Nasdaq stocks from January 1963 through December 2011. A stock is included in the computation of the data moments if it has at least five years of historical data available at the time of the cross-sectional regression.

Panel A: Parameter estimates

	$\sigma_\beta = 0.5$		$\sigma_\beta = 1$	
Parameter	Estimate	SE	Estimate	SE
δ	0.0053	0.0010	0.0053	0.0010
σ_λ	0.0144	0.0010	0.0072	0.0005
σ_ξ	0.0546	0.0024	0.0273	0.0012
σ_ε	0.1313	0.0054	0.1313	0.0054
Test of overidentifying restrictions				
J	9.99		9.99	
$\Pr(\chi^2(9) > J)$	0.351		0.351	

Panel B: Data and model moments

Moment	Data	Model	t-value, Data $-$ Model
$(b_{12} + b_{24})/2$	1.369	1.051	1.60
$(b_{36} + b_{48})/2$	0.849	0.924	-0.40
$(b_{60} + b_{72})/2$	0.614	0.813	-1.21
$(b_{84} + b_{96})/2$	0.625	0.715	-0.70
$(b_{108} + b_{120})/2$	0.652	0.628	0.18
$(b_{132} + b_{144})/2$	0.642	0.552	0.96
$(b_{156} + b_{168})/2$	0.374	0.486	-1.08
$(b_{180} + b_{192})/2$	0.564	0.427	1.02
$(b_{204} + b_{216})/2$	0.426	0.375	0.57
$(b_{228} + b_{240})/2$	0.160	0.330	-1.67
$\mathrm{var}(r_{i,t})$	2.109	2.123	-0.32
$\frac{1}{5}\sum_{k=1}^{5}\mathrm{cov}(r_{i,t}, r_{i,t-12k})$	0.021	0.022	-0.38
$\frac{1}{5}\sum_{k=1}^{5}\mathrm{var}(r_{i,t} - \hat{b}_{12k} r_{i,t-12k})$	2.099	2.113	-0.30

Table 5: Seasonalities in monthly stock returns with a control for the average same-month return

This table summarizes Fama-MacBeth regressions of month-t returns against month-$t-k$ returns. Each row in columns "no control" is a separate regression with the lagged month-$t-k$ return as the only explanatory variable. Column $\hat{b}$ reports the average coefficient and the second column reports the t-value associated with this average. Regressions in columns "control for average same-month return" include the average same-month return as an additional regressor. This average return is computed using all available data over the prior 20-year period, and stock returns are demeaned in the cross-section before taking the average. A stock is included if it has at least 5 years of return data at time t. Panel A's regressions are estimated using data simulated from the model at the estimated parameter values. The averages and t-values in this panel are averages over 100 simulations. Panel B uses returns on all NYSE, AMEX, and Nasdaq stocks from January 1963 through December 2011. Panel B's last row uses the average same-month return as the only regressor.

Panel A: Model

	No control		Control for average same-month return			
Lag,	Month-$t-k$ return		Month-$t-k$ return		Average return	
months	$\hat{b}$	t-value	$\hat{b}$	t-value	$\hat{b}$	t-value
1	−0.01	−0.03	−0.02	−0.08	7.76	9.00
2	0.03	0.08	0.01	0.04	7.71	8.97
3	0.02	0.06	0.01	0.04	7.76	9.01
4	−0.03	−0.08	−0.03	−0.09	7.75	9.00
5	−0.04	−0.11	−0.03	−0.08	7.77	9.01
6	0.00	−0.01	0.00	0.00	7.78	9.02
7	−0.02	−0.08	−0.01	−0.03	7.75	9.00
8	−0.03	−0.09	−0.03	−0.08	7.72	8.95
9	−0.04	−0.11	−0.04	−0.11	7.76	9.02
10	−0.05	−0.15	−0.05	−0.17	7.76	9.02
11	0.01	0.05	0.01	0.02	7.78	9.01
12	1.16	3.58	0.39	1.24	7.44	8.81
24	1.10	3.65	0.33	1.12	7.50	8.71
36	1.02	3.59	0.22	0.79	7.57	8.68
48	0.92	3.42	0.14	0.52	7.67	8.73
60	0.90	3.52	0.09	0.34	7.73	8.70
72	0.88	3.64	0.07	0.29	8.35	8.84
84	0.80	3.47	0.00	0.00	8.97	9.03
96	0.75	3.39	−0.01	−0.07	9.42	9.13
108	0.70	3.30	−0.09	−0.41	9.81	9.21
120	0.63	3.11	−0.11	−0.53	10.14	9.25
132	0.60	3.06	−0.16	−0.80	10.49	9.32
144	0.59	3.08	−0.15	−0.78	10.68	9.28
156	0.54	2.89	−0.22	−1.20	11.06	9.42
168	0.50	2.75	−0.23	−1.24	11.13	9.33
180	0.46	2.59	−0.26	−1.42	11.32	9.35
192	0.45	2.56	−0.28	−1.57	11.43	9.32
204	0.42	2.41	−0.28	−1.56	11.44	9.20
216	0.38	2.20	−0.32	−1.81	11.64	9.28
228	0.35	2.02	−0.34	−1.89	11.72	9.23
240	0.32	1.88	−0.37	−2.06	11.77	9.20

Panel B: Data

	No control		Control for average same-month return			
Lag,	Month-$t-k$ return		Month-$t-k$ return		Average return	
months	$\hat{b}$	t-value	$\hat{b}$	t-value	$\hat{b}$	t-value
1	−5.21	−11.50	−5.22	−11.76	6.84	8.77
2	−0.06	−0.17	−0.01	−0.04	6.85	8.70
3	0.89	2.53	0.95	2.78	6.88	8.71
4	0.59	1.66	0.64	1.87	6.91	8.70
5	0.60	1.69	0.64	1.86	7.01	8.80
6	0.93	2.69	0.91	2.72	6.85	8.54
7	1.04	3.22	1.05	3.35	6.99	8.76
8	0.37	1.04	0.45	1.33	6.81	8.62
9	0.86	2.61	0.91	2.85	6.97	8.79
10	0.29	0.94	0.29	0.96	6.92	8.62
11	0.97	3.48	0.99	3.66	6.99	8.73
12	2.05	6.90	1.35	4.67	5.73	7.34
24	1.37	4.87	0.54	1.97	6.61	8.18
36	0.98	3.73	0.08	0.33	7.02	8.59
48	1.18	4.35	0.31	1.20	6.79	8.41
60	0.49	2.11	−0.45	−1.86	7.65	8.73
72	1.08	4.58	0.17	0.74	7.75	8.23
84	0.81	3.25	−0.11	−0.44	8.70	8.69
96	0.82	3.16	−0.18	−0.72	9.60	8.96
108	0.93	3.70	−0.07	−0.31	10.08	9.31
120	1.06	4.37	0.06	0.22	10.89	9.23
132	1.08	4.43	0.12	0.46	11.27	8.97
144	0.78	2.86	−0.27	−1.01	12.03	9.35
156	0.85	3.05	−0.24	−0.87	12.61	8.99
168	0.54	1.81	−0.57	−2.09	13.88	9.83
180	1.28	4.84	0.17	0.63	14.06	9.41
192	0.89	2.86	−0.20	−0.66	14.54	9.58
204	1.30	4.10	0.10	0.34	15.24	9.76
216	0.80	2.43	−0.43	−1.36	16.34	10.32
228	−0.06	−0.18	−1.34	−4.27	17.70	11.02
240	1.07	3.48	−0.16	−0.53	17.04	10.22
None					6.95	8.44

Table 6: Day-of-the-week seasonality in stock returns

This table summarizes cross-sectional Fama-MacBeth regressions of day-t returns against day-$t-k$ returns. Each row in columns "no control" represents a separate regression with the lagged day-$t-k$ return as the main explanatory variable. Column $\hat{b}$ reports the average coefficient and the second column reports the t-value associated with this average. Regressions in columns "control for average same-weekday return" include the average same-weekday return as an additional regressor. This average return is computed using the same-weekday returns over the prior one-year period, skipping weeks one through four. Both types of regressions also include the average daily stock return from month $t-12$ to $t-2$ to control for one-year return momentum (not reported). The sample consists of returns on all NYSE, AMEX, and Nasdaq stocks from January 1963 through December 2011. Panel A's last row uses the average same-weekday return as the only regressor. Standard errors are Newey-West adjusted using ten lags.

	No control		Control for average same-weekday return			
Lag,	Day-$t-k$ return		Day-$t-k$ return		Average return	
days	$\hat{b}$	t-value	$\hat{b}$	t-value	$\hat{b}$	t-value
1	−11.38	−47.15	−11.38	−47.18	4.02	14.16
2	−3.71	−50.66	−3.71	−50.95	4.11	14.24
3	−2.30	−38.34	−2.31	−38.60	4.24	14.59
4	−1.51	−24.53	−1.51	−24.71	4.24	14.47
5	−0.81	−13.91	−0.82	−14.26	4.15	14.05
10	−0.13	−2.41	−0.13	−2.61	4.22	14.31
15	0.14	2.78	0.13	2.60	4.16	13.89
20	0.19	3.74	0.17	3.57	4.27	14.62
25	0.06	1.38	−0.02	−0.43	4.25	14.32
30	0.09	1.79	−0.02	−0.37	4.33	14.46
35	0.01	0.31	−0.09	−1.81	4.28	14.14
40	0.20	4.04	0.10	1.97	4.12	13.67
45	0.13	2.75	0.02	0.47	4.28	14.34
50	0.09	1.86	−0.02	−0.49	4.30	14.14
55	0.06	1.44	−0.05	−1.09	4.48	14.70
60	0.16	3.47	0.06	1.20	4.29	14.08
65	0.28	6.32	0.18	3.97	4.21	14.02
70	0.12	2.56	0.00	0.02	4.52	14.82
75	0.15	3.17	0.04	0.79	4.39	13.99
80	0.07	1.53	−0.05	−1.04	4.49	14.14
85	0.21	4.64	0.10	2.13	4.46	14.52
90	0.19	4.06	0.08	1.69	4.41	14.11
95	0.11	2.70	0.00	−0.03	4.57	14.47
100	0.15	3.34	0.04	0.92	4.38	13.76
105	0.09	2.03	−0.01	−0.29	4.48	14.27
110	0.08	1.73	−0.03	−0.69	4.46	13.98
115	0.16	3.45	0.05	1.04	4.42	13.57
120	0.10	2.24	−0.01	−0.20	4.52	14.11
125	0.12	2.69	0.02	0.35	4.29	13.25
None					4.29	14.43

Panel B: Regressions of daily returns against average historical same-day return, by weekday

Weekday	$\hat{b}$	t-value
Monday	8.28	12.61
Tuesday	1.97	3.11
Wednesday	0.85	1.40
Thursday	2.27	3.79
Friday	8.40	15.32
Monday through Friday	4.29	14.43

Table 7: Decomposing seasonalities by stock characteristics

This table measures the extent to which different variables explain the cyclical coefficient pattern in monthly stock returns. In the first stage we regress log-daily returns on different set of characteristics, save the residuals, and sum these residuals to monthly level. In the second stage we estimate 20 Fama-MacBeth regressions of these residuals on month-$t-12$, -$t-24$, ..., -$t-240$ returns. We compute the average coefficients from the regressions and record the sum of their squared deviations from zero. This sum of squared deviations is $\text{SSQ}_{\text{actual characteristics} \mid \text{A}}$ for characteristics set A. We also compute a randomized version of this sum of squares. This version randomizes the rows of the data matrix in the first-stage regression. The sum of squared deviations based on the randomized first-stage regressions is $\text{SSQ}_{\text{randomized characteristics} \mid \text{A}}$ for characteristics set A. This table measures a model's explanatory power as Explanatory power of characteristics set A $= 1 - \text{SSQ}_{\text{actual characteristics} \mid \text{A}}/\text{SSQ}_{\text{randomized characteristics} \mid \text{A}}$. We randomize the order in which the characteristics enter the model and repeat these computations 100,000 times. The first column ($c = 1$) reports the change in the explanatory power of the model when each characteristic is included as the only control. Column c is the average change in the explanatory power of the model when the characteristic is entered as the cth control. Panel A uses the full sample period from 1963 through 2011 but does not include a control for earnings announcement dates. Panel B uses the post-1973 subsample and includes a control for earnings announcement dates. The last column reports the average change in the explanatory power, which equals the average of the first four (Panel A) or five (Panel B) columns. The bottom-right corner reports the overall explanatory power of the model that includes all controls. Standard errors (in brackets) are computed by bootstrapping the second-stage Fama-MacBeth coefficient estimates.

Panel A: Sample period 1963–2011, no control for earnings announcements

Characteristics	Change in the model's explanatory power when the control is added as the cth variable					Average change
	$c = 1$	$c = 2$	$c = 3$	$c = 4$		
48 industries	0.326	0.186	0.106	0.070		0.172
	[0.085]	[0.139]	[0.071]	[0.052]		[0.046]
BE/ME and dividend yield	0.501	0.266	0.109	0.018		0.224
	[0.068]	[0.179]	[0.111]	[0.046]		[0.057]
Firm size	0.629	0.396	0.241	0.152		0.354
	[0.060]	[0.180]	[0.114]	[0.057]		[0.057]
Dividends	0.102	0.049	0.024	0.012		0.047
	[0.122]	[0.103]	[0.061]	[0.044]		[0.044]
					Sum	0.797
						[0.032]

Panel B: Sample period 1973–2011, with a control for earnings announcements

Characteristics	Change in the model's explanatory power when the control is added as the cth variable					Average change
	$c = 1$	$c = 2$	$c = 3$	$c = 4$	$c = 5$	
48 industries	0.360	0.234	0.150	0.099	0.080	0.185
	[0.084]	[0.162]	[0.119]	[0.071]	[0.056]	[0.047]
BE/ME and dividend yield	0.500	0.314	0.175	0.076	0.013	0.215
	[0.071]	[0.204]	[0.170]	[0.114]	[0.048]	[0.060]
Firm size	0.625	0.444	0.308	0.206	0.134	0.343
	[0.064]	[0.204]	[0.170]	[0.113]	[0.057]	[0.060]
Dividends	0.096	0.053	0.029	0.016	0.014	0.042
	[0.128]	[0.131]	[0.085]	[0.060]	[0.046]	[0.043]
Earnings announcements	0.017	0.013	0.009	0.005	0.007	0.010
	[0.136]	[0.126]	[0.083]	[0.059]	[0.046]	[0.043]
					Sum	0.796
						[0.033]

Table 8: Seasonalities in individual stock returns and seasonality-mimicking factors

This table measures the extent to which the return seasonalities in well-diversified portfolios of stocks span the seasonalities in individual stock returns. We form 25 portfolios by sorting on size (ME), book-to-market ratios (BE/ME), dividend-to-price ratios (D/P), earnings-to-price ratios (E/P), and on the estimated Fama-French three-factor model loadings, and also use the 48 Fama-French industry portfolios. We compute returns on seasonality-mimicking portfolios by sorting each of the 25 (or 48) portfolios into 10 portfolios based on these portfolios' average same-month return over the past 20 years. In January 1963, for example, we assign portfolios into portfolios based on the average January return in years 1943–1962. This table reports the intercepts and slopes (t-values in parentheses) from regressions of high-minus-low seasonality portfolio returns on the returns on all other seasonal strategies. Row "All" generates the seasonality-mimicking factor using return data on 124 portfolios: 25 portfolios sorted by market capitalization, 25 portfolios sorted by BE/ME ratios, 25+1 portfolios sorted by dividend-to-price ratio with $D = 0$ firms in a separate portfolio, and 48 Fama-French industry portfolios. The intercept measures the extent to which the seasonalities in individual stock returns are explained by the seasonality-mimicking factors.

Seasonality	Model									
factor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	1.11	0.67	0.83	0.94	0.77	0.89	0.80	0.79	0.49	0.41
	(8.99)	(5.95)	(6.72)	(8.52)	(7.65)	(7.53)	(7.56)	(7.65)	(4.64)	(4.04)
ME		0.50							0.25	
		(8.54)							(5.74)	
BE/ME			0.30						0.01	
			(3.90)						(0.19)	
D/P				0.40					−0.04	
				(4.54)					(−0.70)	
Industry					0.37				0.19	
					(5.81)				(5.10)	
$\hat{\beta}_{\text{mkt}}$						0.46			0.09	
						(5.05)			(1.42)	
$\hat{\beta}_{\text{hml}}$							0.49		0.20	
							(6.51)		(2.45)	
$\hat{\beta}_{\text{smb}}$								0.40	0.08	
								(5.62)	(1.03)	
All										0.69
										(10.45)
Adj. R^2		0.29	0.08	0.11	0.23	0.21	0.25	0.29	0.46	0.41