Testing the Duo-factor-model of Return and Volume

K.J. Martijn Cremers and Jianping Mei[1]

First Version: December 2001

Abstract

Recent theoretical work by Lo and Wang (2000) shows that a multi-factor asset-pricing model not only imposes factor restrictions on stock returns but on trading volume as well. We explicitly test their theoretical result using individual stock return and turnover data from NYSE and AMEX from 1962 to 1996. We introduce a recently developed consistent statistic by Bai and Ng (2001a) to determine the number of factors in a duo approximate multifactor model for return and turnover. While we find that the duo-factor model captures a great deal of common variation of trading volume, the data rejects a model restriction that excess return and turnover should have the same number of systematic factors. Using the duo-factor-model, we decompose excess return and turnover into systematic and idiosyncratic components. Our empirical work discovers a significant increase in the variation of idiosyncratic turnover through time, analogous to the discovery of a noticeable increase in firm level volatility by Campbell, Lettau, Malkiel and Xu (2001). We also find significant co-movement between volatility and turnover at the systematic levels. Our findings support the view that trading volume is not purely random but driven by trading activities associated with macroeconomic and firm news.

[1] Doctoral student and associate professor of finance at New York University. Please send inquiries to Jianping Mei, Department of International Business, Stern School, 44 West 4th Street, New York, NY 10012-1126. Tel: 212-998-0354, FAX: 212-998-4221. Email: jmei@stern.nyu.edu. We like to thank Burton Malkiel, Joel Hasbrouck, Yexiao Xu and seminar participants at NYU for helpful discussions.

Testing the Duo-factor-model of Return and Volume

Abstract

Recent theoretical work by Lo and Wang (2000) shows that a multi-factor asset-pricing model not only imposes factor restrictions on stock returns but on trading volume as well. We explicitly test their theoretical result using individual stock return and turnover data from NYSE and AMEX from 1962 to 1996. We introduce a recently developed consistent statistic by Bai and Ng (2001a) to determine the number of factors in a duo approximate multifactor model for return and turnover. While we find that the duo-factor model captures a great deal of common variation of trading volume, the data rejects a model restriction that excess return and turnover should have the same number of systematic factors. Using the duo-factor-model, we decompose excess return and turnover into systematic and idiosyncratic components. Our empirical work discovers a significant increase in the variation of idiosyncratic turnover through time, analogous to the discovery of a noticeable increase in firm level volatility by Campbell, Lettau, Malkiel and Xu (2001). We also find significant co-movement between volatility and turnover at the systematic levels. Our findings support the view that trading volume is not purely random but driven by trading activities associated with macroeconomic and firm news.

Introduction

While multifactor models, such as the intertemporal capital asset pricing model (ICAPM) and the arbitrage pricing model, have long been cornerstones of asset pricing, they have played minor roles in studies of trading volume. In a seminal paper, Lo and Wang (2000, LW henceforth) extend the mutual-fund separation theorem to trading volume. Their insight is that the popular multi-factor asset pricing models not only have strong implications for the cross section of expected returns, but for the cross section of trading volume as well. In contrast to much of the existing volume literature, which relies on specialized models to examine the relationship between volume-price/volatility, they derive an approximate K-factor structure for trading volumes, parallel to the classic K-factor model for asset returns in the presence of a riskless asset. As a result, we now have a duo factor model for returns and trading volume. This paper examines the implications of the duo factor model for the behavior of equity return and trading volume. We hope to add to the literature in several ways.

First, a central issue in both the theoretical and empirical content of LW is the correct identification of the number of factors. Until now, this crucial parameter is often assumed rather than determined by the data.[2] A small number of papers in the asset pricing literature have considered the problem of determining the number of factors in a multifactor model, but the present study differs from them in important ways. Roll and Ross (1980), for example, employ a likelihood ratio test using an exact factor model with normality assumptions. (See also Lehmann and Modest (1988), who test the APT for 5, 10 and 15 factors.) Connor and Korajczyk (1993) develop a test for the number of factors in asset returns under sequential limit asymptotics, i.e., N converges to infinity with a fixed T and then T converges to infinity. Mei (1993) proposes a semi-autoregressive approach to determine the number of factors, but his approach could not obtain factor estimates for some periods due to the use of autoregressors. While Jones (2001) provides a new approach to the extraction of factors, he did not provide an estimate on the number

---

[2] Brown and Weinstein (1983) emphasized the importance of obtaining the correct estimates on number of factors. They pointed out that the common practice of using a over estimate would cause spurious rejection of asset pricing models. They note "…the rejection of the five and seven factor versions is to be expected if the three factor version is correct."

of factors. This paper introduces a formal statistical procedure that can consistently estimate the number of factors from observed data. This procedure is developed by Bai and Ng (2001a, BN thereafter) under the assumption that both N and T converge to infinity.[3] This extension is of empirical relevance because it fully exploits the advantage of a large panel data set. In addition, our empirical study employs an approximate factor structure for both returns and trading volume. Our results hold under heteroskedasticity in both the time and the cross-section dimensions. This renders it more general than Connor and Korajczyk (1993) who assume homoskedasticity over time.[4] Our results also hold under weak serial and cross-sectional dependence.

Second, we explicitly test the duo factor model using monthly turnover data for NYSE and AMEX securities from 1962 to 1996. Unlike Lo and Wang, our empirical study uses data from individual stocks rather than beta-sorted portfolios. By exploiting the advantage of a large cross-section of individual stocks, we get around the nonstationarity issue in turnover. Our results are robust to the presence of either a trend or a unit root in the systematic component of turnover. Berk (2000) has shown a significant drop in statistical power in asset pricing tests using firm characteristics sorted portfolios.[5] As our own empirical work shows, the number of factors in the duo factor model of return and turnover changes dramatically when individual stocks are used instead of beta-sorted portfolios. In addition, we use an EM algorithm to handle the problem of missing values for individual stocks (i.e. unbalanced panels) so that our study is less subject to a survivorship bias that is associated with balanced panels used in other studies.

Third, using the duo-factor-model, we decompose turnover into systematic and idiosyncratic turnover. Likewise, we also decompose individual stock volatility into systematic risk and idiosyncratic risk. These decompositions allow us to examine the relationship between return and turnover factors as well as the relationship between return and turnover betas. Such studies give us a deeper understanding on the relationship between different components of stock returns, risk and trading volume. Recently, there

[3] Xu (2001) has also developed a Maximum Explanatory Components analysis to extract factors from security returns, which is similar in spirit to Bai and Ng. However, he did not study the duo factor model.
[4] Recently, Jones (2001) also provides a new factor estimate for an *approximate* factor model with heteroscedasticity. He does not, however, provide a test on the number of factors in the model.
[5] Brennan, Chordia, and Subrahmanyam (1998) also discover that "…inferences are extremely sensitive to the sorting criteria used for portfolio formation, so that results based on regressions using portfolio returns should be interpreted with caution."

is an increasing interest in the study of the dynamics of stock returns, risk and trading volume.[6] These studies shed important light on the different motives of trading and their impact on asset pricing. However, most of these studies use total return and turnover rather than their components. It is conceivable that macroeconomic shocks may induce more systematic trading across stocks for portfolio rebalancing while firm specific news may affect firm-specific turnover more due to information arbitrage. By decomposing return and volume into systematic and firm-specific components, we provide a new framework for studying liquidity, asymmetric information and their impact on asset pricing.

Fourth, using a balanced panel of excess return and turnover data, our empirical study finds that there are as many as three systematic factors in excess returns and five factors in turnover. While the duo-factor model of Lo and Wang captures a great deal of time variation of trading volume, the data rejects the model restriction that excess return and turnover have the same number of factors. In addition, we document significant co-movements of volatility and turnover at the firm-level as well as at the systematic level. These results are consistent with the view that portfolio rebalancing as a result of macroeconomic shocks drives systematic turnover while firm specific news drives "abnormal" trading at the firm level. Overall, we find that the duo-factor model of Lo and Wang provides a parsimonious description of return and turnover data. Furthermore, There is stronger presence of commonality in turnover in the monthly data.

Our study complements recent studies in the market microstructure literature on the common variation in liquidity or trading volume.[7] Chordia, Roll, and Subrahmanyam (2000) explore cross-sectional interactions in liquidity measures using quote data. They use the market portfolio to analyze the commonality in liquidity. Hasbrouck and Seppi (2001) use a multi-factor model to characterize relationships involving returns and order flows by using the thirty actively traded Dow Industrial firms. The above studies all use high frequency data rather than the monthly data used in our study.

[6] See for example, Amihud (2000), Brennan, Chordia and Subrahmanyam (1998), Chordia, Subrahmanyam and Anshuman (2001), Chordia, Roll and Subrahmanyam (2000), Hasbrouck and Seppi (2001).
[7] The issue of common factor in liquidity was highlighted during the LTCM debacle, when there appeared to be a world wide "flight-to quality" and significant drop in trading volume across many assets.

The paper is organized as follows. We begin in section I by introducing the approximate multifactor model for both excess return and turnover. We provide the main theoretical results of LW that if mutual fund separation holds for stock returns then turnover satisfies an approximate linear K-factor structure as well. Section II discusses a recently developed consistent statistic by Bai and Ng (2001a) to determine the number of factors in the duo-factor-model for return and turnover. It also discusses our empirical methodology and provides a description of the data set. In section III, we use the principle component approach to extract systematic factors from the return and turnover data and provide an explicit test of the hypothesis that the numbers of factors are the same for excess returns and turnover. Using the duo-factor-model, section IV provides the empirical results of decomposing turnover and individual stock volatility into systematic and idiosyncratic components. We then study the relationship between stock returns, risk and turnover. Given the large body of empirical literature on asset pricing, our study will focus on the time and cross-sectional variation of turnover. Section V summarizes our results and concludes.

I. The Duo Multifactor Model For Return And Turnover

Following Lo and Wang, our analysis begins by denoting I investors indexed by $i = 1, ..., I$ and stocks indexed by $j = 1, ..., N$. Assume that asset returns are generated by the following approximate K-factor model:

$$R_{jt} = E_t(R_{jt}) + f_{1t}\beta_{j1} + ... + f_{Kt}\beta_{jK} + e_{jt} \qquad j = 1,...,N; \quad t = 1,...,T. \qquad (1)$$

where $f_t\,' = (f_{1t},...,f_{Kt})$ is a vector of unobservable pervasive shocks, $(\beta_{j1},..., \beta_{jK})$ is a vector of factor loadings which are constant over the sample period, and e_{jt} represents an idiosyncratic risk specific to asset j at time t. We also assume that e_{jt} has mean zero and is orthogonal to f_{kt}. To derive the consistency result of the BN statistic for the number of factors in the above model, some additional regularity conditions are imposed, which are provided in the appendix.

As discussed in Chamberlain (1983), the above economy implies the following linear pricing relationship if there exist K well-diversified portfolios[8]:

$$E_t(R_{jt}) = r_{ft} + \lambda_{1t}\beta_{j1} + \ldots + \lambda_{Kt}\beta_{jK}, \tag{2}$$

where $(\lambda_{1t},\ldots, \lambda_{Kt})$ is a vector of risk premiums corresponding to the pervasive shocks $(f_{1t},\ldots,f_{Kt})$, and r_{ft} is the return on a riskless asset. Denoting $F_t' = (f_{1t},\ldots,f_{Kt}) + (\lambda_{1t},\ldots, \lambda_{Kt})$ and $B_i' = (\beta_{i1},\ldots, \beta_{iK})$, we obtain,

$$r_{j,t} = R_{j,t} - r_{ft} = F_t' B_j + e_{j,t}. \qquad j=1,\ldots, N;\ t=1,\ldots,T. \tag{3}$$

where $r_{j,t}$ is excess return for asset j at time t. For simplicity, we will now call F_t factors and f_t systematic shocks in the paper. We stack the J time series of excess returns in the T×N matrix $\mathbf{r}$.

To establish the link between asset returns and trading volume, we note that under the presence of K well-diversified portfolios, Chamberlain (1983) shows that the above asset pricing model also satisfies K-fund separation. To derive a parallel K-fund separation theorem for trading volume, we begin by denoting Q_{jt} as the total number of shares outstanding for each stock j. Without loss of generality, we assume that the total number of shares outstanding for each stock is constant over time. For each investor i, let S^i_{jt} denote the number of shares of stock j he holds at date t. Finally, denote X_{jt} to be the total number of shares of security j traded at time t, that is, share volume, hence

$$X_{jt} = \frac{1}{2}\sum_{i=1}^{I}\left|S^i_{jt} - S^i_{jt-1}\right| ,$$

where the coefficient 1/2 corrects for the double counting when sum the shares traded over all investors. The turnover τ_{jt} of stock j at time t is defined as X_{jt}/Q_j, where X_{jt} is the

[8] Connor (1984) derived a same result under the condition that the supplies of the assets are well diversified.

share trading volume of security j at time t and Q_j is the total number of shares outstanding of stock j. We stack the N×T time series of turnover in the matrix τ.

Under the assumptions that the separating stock funds are constant over time and the amount of trading in the separating portfolios is small for all investors, LW derive the proposition that the turnover of each stock has an approximate K'-factor structure. More formally, we have:

$$\tau_{jt} = \tau_j + \delta_{j1} g_{1t} + \ldots + \delta_{jK'} g_{K't} + \xi_{jt} \tag{4}$$

Here, δ_{jk} is the exposure of firm j to economy–wide liquidity shocks g_{kt}. g_{kt} could be functions of f_{kt} but it is not specified in the model and τ_j is a constant. Using similar term from asset pricing, we will call δ_{jk} turnover betas. ξ_{jt} has mean zero and it is assumed to be orthogonal to g_{kt}. In addition, we assume that ξ_{jt} also satisfy the regularity condition given in the appendix. For simplicity, we will call the multi-factor models of (3) and (4) the duo-factor model for return and volume.

Moreover, LW derive a easily testable hypothesis about the duo-factor-model of (3) and (4) that the two models should have exactly the same number of factors, i.e. K = K'. This test allows us to gain important insights on the number of pervasive factors determine asset return and trading volume. Moreover, using the duo-factor-model, we decompose turnover into systematic and idiosyncratic turnover. Likewise, we also decompose individual stock volatility into systematic and idiosyncratic risk. With this decomposition, we can examine the relationship between return and turnover factors and betas. Such analysis gives us a deeper understanding of the relationship between different components of stock returns, risk and trading volume.

II. Estimation Procedure

A. A Partial Solution to Nonstationarity in Turnover Data

To extract factors from the return and turnover data, LW apply principal components approach to the variance-covariance matrix of return and turnover *among ten* *portfolios*. As shown in LW and numerous other studies, aggregate turnover appears to be nonstationary, exhibiting a time trend and time-varying volatilities. Thus, the variance-covariance matrix for turnover, var(τ_i ,τ_j), is not well defined when τ_{it} or τ_{jt} is nonstationary. As a result, conventional statistical inference may not apply.

This paper provide a partial solution to the nonstationary problem by taking advantage of a large cross-section of individual stocks. Rather than using the variance-covariance matrix of turnover *among ten portfolios,* we will rely on the variance-covariance matrix of turnover *over different time periods*. In other words, we will apply principal component approach to var(τ_t ,τ_s), where

$$\mathrm{Var}(\tau_t, \tau_s) = N^{-1}\sum_{i=1}^{N}(\tau_{it} - \bar{\tau}_t)(\tau_{is} - \bar{\tau}_s), \text{ and } \bar{\tau}_s = \sum_{i=1}^{N}(\tau_{it} - \bar{\tau}_t).$$

As we can see from the above equation, var(τ_t ,τ_s) is well defined for give time period t and s, as long as the cross-sectional mean and variance for turnover exist. Intuitive speaking, var(τ_t ,τ_s) depends on N-consistency rather than T-consistency, which require stationarity.

B. The Bai and Ng (2001a) Statistic

We will begin by estimating the common factors in (3) using the asymptotic principal component method of Connor and Korajczyk (1988). Since the true number of factors K is unknown, we start with an arbitrary number k_{max} (k_{max} < min (N, T)). Denoting B^k and F^k are the estimates of k factors and factor loadings, respectively that solve the following optimization problem:

$$V(k) = \min_{B^k, F^k} T^{-1} N^{-1} \sum_{t=1}^{T}\sum_{j=1}^{N}(r_{jt} - B_j^k F_t^k)^2 \qquad (5)$$

To determine the number of factors, BN propose the following statistic based on information criteria (IC):

$$\hat{K} = argmin_{0<k<kmax} PC_1(k), \tag{6}$$

where $PC_1(k)$ equals a measure of the goodness-of-fit V(k) in (5) plus a second term that serves as an adjustment for the increased "degree of freedom" as a result of increasing k:

$$PC_1(k) = V\left(k, \hat{F}^k\right) + k\hat{\sigma}^2\left(\frac{N+T}{NT}\right) ln\left(\frac{NT}{N+T}\right). \tag{7}$$

where $\hat{\sigma}^2$ is the mean variance for idiosyncratic risk under k_{max} ($\hat{\sigma}^2 = V(k_{max})$). BN show that $\hat{K}$ is a consistent estimate for the true number of factors in the factor model.[9] Intuitively, the estimation procedure treats the determination of factors as a model selection problem. As a result, the selection criterion depends on the usual trade-off between goodness-of-fit and parsimony. The difference here is that we not only take the sample size in both the cross section and the time series dimensions in consideration, but also the fact that the factors are not observed. There are three distinctive advantage of the BN approach comparing to the method of Connor and Korajczyk (1993). First, BN do not impose any restrictions between N and T, allowing for both large N and large T. Second, the results hold under heteroskedasticity in *both* the time and the cross-section dimensions. Third, the results also hold under *both* weak serial dependence and cross-section dependence. In addition, the model selection procedure is easy to implement. The conditions under which the consistency of $\hat{K}$ holds are given in the appendix. Bai and Ng (2001b) further point out that the consistency of $\hat{K}$ holds in the presence of trend or unit root in F_t.

---

[9] Bai and Ng also proposed two other asymptotically equivalent statistics as follows. Our empirical study has found that the PCs gave almost identical results in our balanced panel.

$$PC_2(k) = V\left(k, \hat{F}^k\right) + k\hat{\sigma}^2\left(\frac{N+T}{NT}\right) ln\, C^2_{NT},$$

$$PC_3(k) = V\left(k, \hat{F}^k\right) + k\hat{\sigma}^2\left(\frac{N+T}{NT}\right),$$

Here, C_{NT} is defined as min($\sqrt{N}, \sqrt{T}$). In addition, BN also proposed three other asymptotically equivalent statistics called ICs. Our own simulation studies has show that, while the ICs have the advantage of not having to specify kmax and estimating $\hat{\sigma}^2 = V(k\max)$, their estimates tend to be biased towards finding smaller number of factors in small samples. These results are available upon request.

C. Data Description

Following Lo and Wang, we use the CRSP Monthly Master File to construct monthly turnover series for individual NYSE and AMEX securities from July 1962 to December 1996.[10] The choice of the monthly horizon makes our results comparable to earlier asset pricing studies and is a compromise between maximizing sample size while minimizing the day-to-day volume and return fluctuations that have less direct economic relevance. Since our focus is the implications of portfolio theory for volume behavior, we limit our attention to ordinary common shares on the NYSE and AMEX (CRSP share codes 10 and 11 only), omitting ADRs, REITs, closed-end funds, and others whose turnover may be difficult to interpret. We also omit NASDAQ stocks because of differences in market structure between the NASDAQ and the NYSE/AMEX exchanges. In addition to turnover, we also collect data on firm market price, capitalization, trading volume, and returns.

Like LW, we throw away firms that have no or problematic turnover data. In particular, we remove firms that have turnover with no, zero or extremely large standard deviation of turnover (respectively data errors 1, 2 and 3 in the table). As LW argue, such large standard deviations probably indicate data errors. This removes about 5% of the firms in the unbalanced panel. Table 1 presents some summary statistics about our sample section. These include the number of securities in each sample, the percentage of securities with missing observation in returns and turnover. In addition, we also report the number of firms that were excluded from the sample for three different reasons: The first error indicates firms that have less than 25% of the data available, so have more than 45 missing entries in either return or turnover over the 60 month period. The second error indicates firms that have constant turnover in the time period. The third error indicates of firms that have likely data entry problems as evidenced by an unusual large standard deviation (ten times the average standard deviation, see also the discussion on the Z-flag in Lo and Wang (2000)).

Panel B provides summary statistics of the excess return and turnover of the value-weighted portfolio of all NYSE and AMEX ordinary common shares from July

[10] Lo and Wang graciously provided us with MiniCRSP data manual.

1962 to December 1996. We report the annualized mean, standard deviation and autocorrelation for each sample period. Not surprisingly, we observe a particularly high market volatility during the 1987-1991 time period, which is largely due to the October 1987 market crash. We also document a corresponding increase in the variation of turnover during the same time period. Based on the mean turnover of the seven time periods, there seems to be a significant increase in trading volume over time. While the autocorrelation of the returns varies over time and is generally quite small, the autocorrelation for turnover is quite large and positive over time, suggesting that changes in turnover are quite persistent over time. These results are consistent with LW.

III. Analysis of the Duo Multifactor Model

A. Test On Number Of Factors in a Balanced Panel

The common factors in F are estimated non-parametrically by the method of asymptotic principal components, such that we select the eigenvectors corresponding to the k_{max}-largest eigenvalues of the T×T matrix $\mathbf{r} \cdot \mathbf{r}'$ for returns, and $\tau \cdot \tau'$ for turnover. Regressing the return and turnover data on their respective factors (eigenvectors) gives the beta's. Finally, we compute the model selection criteria PC_1 for both returns and turnover separately, for models including 1 to k_{max} = 10 factors.

Table 2 provides the results of the test of the number of factors in excess return and turnover. We report the incremental proportion of explained variation (R^2) from the k-th factor of the return and turnover data of the NYSE and AMEX common shares for seven subperiods from July 1962 to December 1996. Note that in the case of the balanced panel, the incremental R^2 from the k-th factor equals the k-th largest eigenvalue θ_k , k = 1, …,10, of the covariance matrix of returns and turnover, respectively, where the eigenvalues are normalized to sum to 100%. The first principal component of returns typically explains between 11% and 36% in the variation of the normalized excess returns while the first principal component of turnovers typically explains between 11% and 24% in the variation of the normalized turnover. This is quite different from LW, who use returns from broadly diversified portfolios and find their first principal component typically explains over 70% (sometimes as high as 90%) of the variation in both returns and turnovers. Further examination of our results suggests that the second and third component still explains a fair amount of variation in excess returns and turnovers. For example, the third component still explains 6.13% of variation in turnover for the 1962-1966 period. Similar to LW, we also find that the second, third and fourth principal components seem to be more important for the 1992-1996 time period. These results seem to indicate that we have more than one systematic factor in both returns and turnovers.

To determine the number of factors in excess returns and turnovers, we compute the "goodness-of-fit" statistic, PC_1 of BN conditional on a wide range of included numbers of factors. Table 2 reports the number of factors corresponding to the minimum PC_1 statistic. For example, comparing $PC_1(k)$ for $k = 1, 2, \ldots, 10$ indicates that $k = 2$ provides the minimum $PC_1(k)$ for turnover for the 1962-1966 sample period. This indicates that there are two systematic factors for turnover during the first sample period. It is reassuring to see that the number of factors identified by the PC statistic closely corresponds with the eigenvalues of the principal components. The eigenvalues typically exceed 3% for those principal components identified as factors. In summary, the "goodness-of-fit" statistic suggests that there were two or three systematic factors in excess returns and there were four or five systematic factors in turnover during the various sample periods. The difference in the number of factors between return and turnover seems to reject the restriction of the duo-factor model of LW. We will provide a more rigorous test of the restriction that the number of factors in returns and turnover is equal in section III B.

Our result of four or five factors in turnover is different from the results reported in Lo and Wang, who find one or two factors in turnover. This difference could be due to two reasons. First, LW use weekly data while our study is based on monthly data. Second, LW use beta-sorted portfolios while we use individual stocks. As a result, due to diversification their covariance matrix contains much less cross-section variation in excess return and turnover than our matrix. This will lead their principal components to explain more cross-section variation in excess returns and turnover. As pointed out by Shukla and Trzcinka (1990), because beta-sorted portfolios tend to mask some cross-section differences in exposure to other sources of systematic risk, the principal components approach based on beta-sorted portfolios is biased towards finding small number of factors. While our test does not specifically identify what exactly the factors are, they do provide some guidance for theorists in their equilibrium model construction. Our results suggests that, while the two-factor model of Lo and Wang (2001) provides a best prediction of future market returns, they still leave out a few systematic factors in their model. This may help explain why their model does not fully capture the cross-section of expected returns.

Table 2 also reports the average R^2 of regressing individual stock excess returns and turnovers on their respective systematic factors for each sample periods. For the 1962-1966 period, a two-factor model explains on average about 28.9% of variation in excess returns of individual stocks, with a standard deviation of 13.3%. During the same time period, a four-factor model explains on average about 33.7% of variation in turnover of individual stocks, with a standard deviation of 16.6%. The average R^2 for returns and turnovers over the whole sample period are 33.8% and 36.4%, respectively. Thus, turnover factors are just as important as return factors in explaining the time variation of turnover across individual stocks. Comparing to empirical results about trading volume found in market microstructure studies by Hasbrouck and Seppi (2001), we have found a stronger presence of commonality in liquidity.[11]

Since trading volume determines the transaction costs in the stock market, our results imply that trading volume may have a systematic impact on after-cost returns. This implies that liquidity risk could be a systematic risk that should be priced. As a result, our results is consistent with the empirical results of Amihud (2001) and Pastor and Stambaugh (2001) that liquidity is an important risk factor in financial markets.

Table 2 also shows a significant drop in average R^2 for excess returns for the last sample period, suggesting a significant increase in contribution of idiosyncratic risk to total return variation. This result is consistent with the result of Campbell, Lettau, Malkiel and Xu (2001, CLMX thereafter), who find a noticeable increase in firm level volatility

[11] Hasbrouck and Seppi (2001) use order flow data from a sample of 30 Dow stocks during 1994 to study the common factors in stock prices and liquidity. They find the first three common factors explain about 20% of the variation in order flows. They do not provide an explicit test for the number of factors in the factor model. Chordia, Roll and Subrahmanyam (2000) also use transaction data from a sample of 1,169 stock in 1992. They examine the common movement in market depth using value- and equal- weight indices. They find the mean $\mathbf{R^2}$ to be less than 2%.

relative to market volatility in recent years. However, our results for excess returns suggest that CLMX may under-estimate the importance of systematic factors in returns, since the R^2 obtained through their market model appears to be substantially below the average R^2 found in our study under a multi-factor model.[12]

The most intriguing result of Table 2 is an apparent positive correlation (78.2%) between the average R^2 of returns and the average R^2 of turnover across different time periods. This means that when return factors explain a larger proportion of the variation in returns, turnover factors also tend to drive more turnover variation for individual stocks. This suggests a positive relationship between systematic return factors and systematic turnover factors.

To further study the relationship between return and turnover factors, Panel B of Table 1 also decomposes monthly value-weighted portfolio of excess returns and turnovers into systematic and firm specific components, using return and turnover factors determined in Table 2. There appears to be a close relationship between volatility and turnover at the systematic level. Their correlation is 32%. With the exception of the last sample period, there also seems to be a rising trend in systematic volatility and turnover. We formally study the relationship between the various components of volatility and turnover in the next section.

B. Monte Carlo Simulation and Test of Same Number of Factors in Excess Return and Volume

In this section, we first provide a simulation study to demonstrate that the PC estimates have good small sample properties.[13] We then provide a formal test of the same number of factors in equation (3) and (4). Because a realistic model of how returns and volumes are added and deleted from the sample is not obvious, we restrict our attention in this section to the cases in which both return and turnover have no missing observations.

---

[12] The contribution of the market to total volatility was 13.4% during the 1988-1997 in CLMX while the average R^2 found in our study was 39.6% for the 1987-1991 period and 16.9% for the 1992-1996 period. However, difference in time period and weighting (CLMX used value-weighting while we use equal weighting in Table 2) may account for some of the difference in results.

[13] While Bai and Ng did provide a simulation study on the small sample properties of the PC estimator, they used a general data generating processes (DGP) that is not calibrated for stock return and turnover.

The data generating processes (DGP) used in the simulations follows Jones (2001) and is designed to mimic the actual data as closely as possible. Rather than simulating factors under some arbitrary assumptions, bootstrap samples of factor estimates extracted from the actual data are used as the true factors in the simulations. Given estimates of the $T \times K$ matrix **F** of factor realizations, we sample (with replacement) T rows of **F** to use as the true factors in the simulations. Let $\mathbf{F}_i$ denote the *ith* bootstrap draw of the factor matrix. The factor betas assumed in the DGP are bootstrap samples of the least squares estimates of the betas from the actual data and we assume then to be constant over time. Denoting **B** to be the $N \times K$ matrix of OLS estimates of the factor betas from real data, we follow Jones (2001) by drawing with replacement N rows of the **B** matrix to use as the true betas in the simulations. We then draw the corresponding elements of the $N \times N$ diagonal matrix Ω, whose *(j, j)* element is the unconditional sample variance of the residual of stock j. We denote $\mathbf{B}_i$ to be the *ith* bootstrap draw of the beta matrix and Ω_i the corresponding draw of Ω. As a result, the $N \times T$ matrix of simulated excess returns $\mathbf{R}_i$ will then be generated by the equation

$$\mathbf{R}_i = \mathbf{B}_i \mathbf{F}_i + \Psi_i * \mathbf{E}_i \qquad (9)$$

where Ψ_i is the Cholesky factor of Ω_i and $\mathbf{E}_i$ is an $N \times T$ matrix of independent standard normals. Here, we assume all alphas to be zero.

Similarly, given estimates of the $T \times K'$ matrix **G** of factor realizations for normalized turnover, we sample (with replacement) T rows of **G** to use as the true factors in the simulations. Let $\mathbf{G}_i$ denote the *ith* bootstrap draw of the factor matrix. The factor betas assumed in the DGP are the bootstrap samples of the least squares estimates of the turnover betas from the actual data, which are assumed to be constant over time. Denoting **D** to be the $N \times K$ matrix of OLS estimates of the turnover betas from real data, we draw with replacement N rows of the **D** matrix that we use as the true betas in the simulations. We then draw the corresponding elements of the $N \times N$ diagonal matrix Σ, whose *(j, j)* element is the unconditional sample variance of the residual turnover of stock j.

To maintain the correlation found in the data between residual excess return and residual turnover, we simulate residual turnover by the following equation,

$$\xi_{jt} = \gamma_i e_{j,t} + \mu_{jt}, \tag{10}$$

where γ_i is a scaling coefficient to make the correlation between ξ_{jt} and $e_{j,t}$ to be ρ_j and μ_{jt} is independent standard normal. Here, ρ_j is the sample correlation between residual excess return and residual turnover for stock j. We then further scale ξ_{jt} so that its variance equal to the *jth* diagonal element of Σ. As a result, the $N \times T$ matrix of simulated turnover Γ_i will then be generated by the equation

$$\Gamma_i = \mathbf{D}_i \mathbf{G}_i + H_i \tag{11}$$

where H_i is the *ith* draw of the NxT matrix whose elements are ξ_{jt}.

Table 3 presents the frequency on the number of factors that minimizes the PC_1 criterion for return and turnover data over 100 simulations. The value of k_{max} is again set to equal to 10. Conditional on the number of factors found in Table 2, each simulation involves the draw of a set of $N \times T$ individual return and turnover data for the corresponding sample period. For example, each simulation draws 1441 × 60 individual returns and turnovers for the 1992 - 1996 period, using equation (9)-(11).

As the first row of the top panel shows, if the true number of factors is two, the PC criterion finds the right number of factors in 94% of the simulations using parameters calibrated to resemble the data in the 1962 - 1966 sample period. The mean of the estimated number of factors equal to 1.94 shows a slight downward bias compared to the true number of factors. The worst performance for the PC estimates is for the 1972 - 1976 period, when the mean estimates of the number if included factors for returns is 2.64 compared to the true number of factors of three. The PC criterion shows a similar degree of accuracy in estimating the number of factors in turnover. As the first row of the bottom panel shows, if the true number of factors is four, the PC criterion has a 89% chance of finding the right number of factors in the simulation using parameters calibrated to resemble data during the 1962 - 1966 sample period. It worth noting that the accuracy of the PC approach depends upon N and T: as we increase the number of companies used in the sample or the length of the sample period, accuracy tends to improve.

After examining the accuracy of the PC estimates, we now turn to a formal test of the hypothesis of Lo and Wang that the number of return factors equals the number of turnover factors in the duo factor model. While Table 2 documents some apparent differences in the number of return and turnover factors over the sample periods, one cannot be sure that these differences are statistically significant. To formally address this issue, Table 4 presents the Type I and Type II error estimates for the test of the difference between the numbers of return and turnover factors. The error estimates are based on 100 simulations for each time period and each simulation involves the draw of a set of $N \times T$ individual return and turnover data. For type I error estimates, we assume that the true numbers of return and turnover factors are three. We choose the number three because that is the highest number of return factors found in the data. For type II error estimates, we assume that the true numbers of return and turnover factors are the same as those found in the data. Thus, the true difference is K-K', which is based on the difference in the numbers of factors found in Table 2.

As the first row of the panel A shows, if the true number of factors are the same for return and turnover, then the probability that the PC criterion finds a difference of two during the 1962 - 1966 sample period (period 1) is only 1%. The only period the LW hypothesis is not rejected is for 1972 - 1976, when the significance level is 11%. While our test has statistical power in rejecting the hypothesis, our test seems to have poor power against the hypothesis that return has one less factor then turnover for the time periods of four of the seven sample periods considered, 1962 - 1966, 1967 - 1971, 1987 - 1991, and 1992 - 1996.[14] However, it is reassuring that the PC criterion has fairly small Type II errors conditional on the actual number of factors found in the data. The probability of accepting the null of same factors while it is not true never exceeds 5% for all sample periods. In summary, our simulation study indicates a strong rejection of the

[14] This may be expected, however, since the actual number of return factors found in the data is two while these simulations force it to be three. As a result, the third factor drawn in the simulation is white noise, which may bias the PC criterion to find only two factors.

null hypothesis that there are same numbers of systematic factors in the duo factor model of return and turnover.

The rejection of the "same number of factors" restriction should not be surprising, since the turnover factor model was derived based on k-fund separation, implying common mimicking factor portfolios held by all investors. To the extent that investors use private information to speculate on small stocks, this could lead to a violation of k-fund separation and thus the violation of the turnover factor model. One possible explanation for the difference in the number of factors between returns and turnovers in the balanced panel could be the presence of private information. For example, Llorente, Michaely, Saar, and Wang (2001) find that small firms tend to have high trading volume associated with private information. Another explanation could be a sample selection bias. Since our sample exclude bonds and Nasdaq stock, our sample may not be able to reflect all systematic risks in the economy. For example, Fama and French (1993) find that with stocks only three factors are necessary but five factors are needed when bonds are included in asset pricing studies. To the extent that changing in rising sector demand (such as high technology) and interest rates may have a disproportionate impact on the return of excluded assets, investors may need to rebalance their position on all assets. As a result, we may observe systematic changes in turnover but fail to detect significant return impact on our sample.

C. Test On Number Of Factors in an Unbalanced Panel

So far, our discussion has focused on balanced panels, which require the firms in panel to having no missing return and turnover data during the sample period. Obviously, this requirement will lead to a survivorship bias. Fortunately, as discussed in Stock and Watson (1998) and BN, the problem can be solved easily by using an iterative EM algorithm to fill missing data with estimated values. The idea is to replace return or turnover by their value as predicted by the parameters obtained from the last iteration when they are not observed. Using returns as an example, if $B_j(m)$ and $F_t(m)$ are estimated values of B_j and F_t from the m^{th} iteration, let $r^*_{jt}(m-1) = r_{jt}$ if r_{jt} is observed, and $r^*_{jt}(m-1) = B'_j\,(m-1)F_t\,(m-1)$ otherwise. We then minimize $V^*(k)$ with respect to $F(m)$ and $B(m)$, where $V^*(k) = (NT)^{-1} \sum^T_{j=1} \sum^T_{t=1} (r^*_{jt}(m-1) - B^k_j(m)\, F^k_t\,(m))^2$. This is equivalent to compute the T x T matrix $r^*(m-1)\, r^*(m-1)'$ with projected and observed data and then extract their eigenvalues (and associated eigenvectors). We then iterate this process until convergence.

We apply this algorithm to return and turnover data for the unbalanced panel that includes missing observations. We set the starting values of these missing observations equal to the mean of that firm's nonmising observations. Table 5 reports the number of factors corresponding to the minimum PC_1 statistic. For example, the $PC_1(k)$ test indicates that k = 3 and k = 9 provide the minimum of $PC_1(k)$ for return and turnover, respectively, for the 1962-1966 sample period. This suggests that there are three systematic factors for returns and nine factors for turnover during the sample period. In contrast to our results reported in Table 2, we can see that there is a sizable increase in the number of factors for both returns and turnover. It is worth noting that the difference between the unbalanced and the balanced panel is that the latter consists of firms that by construction survived at least five years. The unbalanced sample thus includes much more younger and less mature firms or delisted firms due to merger or bankruptcy. The difference in the selected number of factors suggests that the returns and turnover of mature and young firms are not driven by the same factors.

Table 5 also reports the average R^2 of regressing individual stock excess returns and turnovers on their respective systematic factors for each sample period. For the 1962-1966 period, a three-factor model explains on average about 40.5% of variation in excess returns of individual stocks, with a standard deviation of 23.9%. During the same time period, a nine-factor model explains on average about 60.2% of variation in turnover of individual stocks, with a standard deviation of 24.9%. In contrast to the results reported in Table 2, we find a large increase in the average R^2 (as well as its standard deviation) for the turnover model. We like to note, however, that one of the main results of Table 2, namely the positive correlation between average R^2 of returns and average R^2 of turnovers across different time periods, remains unchanged.

In summary, the results of Tables 2, 3, 4 and 5 indicate that, contrary to LW, a one-factor model for turnover cannot capture the commonality for the time-series and cross-sectional variation in turnover. This calls into question the practice of estimating "abnormal" volume by using an event-study style "market model", for example, Bamber

(1986), Brennan, Chordia, and Subrahmanyam (1997), Jain and Joh (1988), Lakonishok and Smidt (1986), Richardson, Sefcik, Thompson (1986), Stickel and Verrecchia (1994), Tkac (1996), and Llorente, Michaely, Saar, and Wang (2001). We believe that a multi-factor model similar to those of the three-factor model of Fama and French (1993) is needed in estimating "abnormal" volume for individual stocks.

IV. The Determinants of Turnover

A. A Graphic Presentation

In order to obtain a better understanding of the duo-factor model, Figure 1a-1d provide a graphic depiction of turnover for a value-weight portfolio and its respective systematic and idiosyncratic components using the number of factors determined in Table 2.[15] Here, turnover is decomposed into systematic and idiosyncratic components. We annualize turnover in percentages. Figure 1a presents turnover for the value-weighted portfolio. One distinguishing feature of turnover for the value-weighted portfolio is that it increases dramatically over the whole sample period, rising from 10% per annum in the 1962-1966 period to over 70% in the 1992-1996 period.

Next, we examine the systematic turnover in Figure 1b. Not surprisingly, systematic turnover has similar patterns as total turnover in Figure 1a. The most intriguing results here are about idiosyncratic turnover, which are presented in Figure 1c and shows a dramatic increase in the variation of idiosyncratic turnover during the later sample periods. This is analogous to the result of CLMX, who find a noticeable increase in firm-level volatility over the sample period. This suggests a possibly close link between the increase in firm-level volatility and the increase in firm-level trading volume over time. To remove the short-term fluctuations in idiosyncratic turnover, we take their absolute values and then plot their twelve-month moving average in Figure 1d. We see a resemblance of this chart to Figure 4 from CLMX for firm-specific risks, which also display a upward trend. Moreover, there also appears to be a close relationship between

[15] To save space, we only provide the turnover charts. We also plotted similar charts for return volatility and their components. They are available upon request.

volatility and turnover at systematic level. We formally study these relationships between various components of volatility and turnover in the next section.

B. The Determinants of Turnover Factors over Time

In order to understand the main drivers of turnover factors, Table 6 provides OLS coefficients as well as R^2 of regressing return factors as well as turnover factors on Fama and French factors (henceforth FF-factors). Not surprisingly, the FF-factors have high explanatory power for the systematic factors extracted from the return data. What is interesting is that the FF-factors have fairly high explanatory power over systematic factors extracted from the turnover data as well.[16] The fact that the FF-factors are significant in explaining turnover is consistent with Lo and Wang (1998). They develop a formal dynamic equilibrium asset-market model in which volume, prices, and other state variables evolve through time together in an economically consistent way. They explicitly model the motives to trade as a function of preferences, endowments and economic conditions and demonstrate that trading volume satisfies an approximate three-factor structure that includes the market factor.[17]

If return factors have a significant impact on stock turnover, then it is natural to assume that stock volatility also impacts trading volume. Possibly, macroeconomic news associated with systematic risk has a different impact on trading volume compared to company-specific news associated with idiosyncratic risk. Thus, we decompose stock volatility into systematic and idiosyncratic risk and examine their impact on turnover separately. The results are reported in Table 7. Given the presence of a trend term in turnover and idiosyncratic risk found in previous studies, we have de-trended all variables in Table 7.[18] To simplify our presentation, we use the turnover of the value-weighted portfolio and decompose total turnover into systematic and firm-specific

[16] Here, we only report regression results for the first factor. The results for other factors are quite similar. We have also included innovations in some economic variables in the regression. However, these economic variables are found to have little additional explanatory power over the variation of turnover factors and thus dropped from the regressions. They are available upon request.
[17] To further understand the relationship between return factors and turnover factors, we also conducted regression of return factors on turnover factors and vice versa for each sample period. Not surprisingly, we find that there is significant co-movement between return factors and turnover factors, suggesting return factors and trading volume are highly related. However, return factors and turnover factors do not span each other. These results are available upon request.
[18] Our results are quite similar using raw data.

turnover. Table 7 presents the results of regressing these turnover components on systematic and idiosyncratic risk. We also include volatility lags for predictive analysis. The results show that systematic turnover is affected mostly by contemporaneous systematic risk. We find that the higher the systematic volatility, the larger the systematic turnover. This is certainly consistent with the view that changes in trading volumes over time are driven by portfolio re-balancing needs as a result of systematic changes in risk-return trade-off. The table also shows that idiosyncratic turnover on the value-weighted portfolio is affected by contemporaneous systematic risk as well. Somewhat surprisingly, changes in idiosyncratic risk do not seem to have a significant impact on the idiosyncratic turnover of the value-weighted portfolio. We conjecture that this could be the result of averaging over a large number of stocks, which could mask the impact of firm risks on idiosyncratic turnover. We will further study the issue using individual stock data in section C.

C. The Determinants of Turnover in the Cross-section

To develop a sense for cross-sectional difference in turnover, Figure 2 provides a graphic depiction of turnover for value-weight decile portfolios. For simplicity, we only report those for the first, fourth, seventh and tenth decile portfolios. Figure 2 is similar to Figure 3a of LW, which provides a graphic representation of turnover for decile portfolios. There are several interesting patterns. First, turnover for the 10th decile portfolio, which consists of the largest 10 percentile of stocks, rises sharply during the mid-1960s, then falls suddenly in the late of 1960s and remained relatively low in the remaining sample periods. The dramatic rise and fall in turnover for large stocks reflect the "nifty-fifty" craze for large-cap growth stocks in the mid-1960s, when stocks like IBM were traded much like internet stocks in the late 1990s. Second, there is a dramatic increase in turnover for decile 1 and 4 portfolio for small stocks over time, especially after 1975 when fixed commissions were abolished.

It is clear from Figure 2 that turnover varies across stocks. LW examine the cross-sectional relationship between turnover and a set of firm variables, including Jensen's α, market beta, idiosyncratic risk using the market model, dividend yield and four other firm-specific variables. This paper further examines the impact of risks on turnover by using the multi-factor model (3) for measuring systematic and idiosyncratic risk. Our motivation is motivated by the idea that realized returns often generate portfolio rebalancing needs such that the various components of return volatility should be positively related to turnover.[19]

Table 8 contains the coefficients as well as R^2 of the cross-sectional regression model of mean turnover. We estimate two regression models for each sample period. The first model only includes turnover betas as regressors. The second model includes turnover betas as well as return betas. As one can see from the first regression, the multi-factor model of equation (4) provides a fair explanation of the cross-sectional variation of turnover. Specifically, all turnover betas are statistically significant. The explanatory power of these cross-sectional regressions ranges from 1.37% (1992-1996) to 16.5% (1972-1976), comparable to the R^2's of typical cross-sectional return regressions. A close inspection of data and Figure 3 reveals that while there is great variation in firm turnover, ranging from 0.26% to 21% a month, the variation in firm turnover betas is much smaller.[20] As a result, the differences in firm exposure to systematic changes in turnover only explain a small proportion of the cross-sectional variation in firm turnover.

The second regression of Table 8 reveals that the systematic risk $B_{r,j}$ has a significant impact on turnover in all sample periods. The inclusion of return betas significantly increases the explanatory power of these cross-sectional regressions in some time periods – as measured by R^2- from 7.32% to 35.9 % in the 1982 – 1986 period. And the regression coefficients of almost all return betas are statistically significant. Thus, the firm's exposure to systematic risk has an important and statistically significant impact on mean trading volume. In addition to the time series evidence found in LW and in Table 6 and 7 of this paper, these R^2's at the cross-sectional level provide some confidence that variations in mean turnover are not purely random but are related to economic factors.

[19] Turnover may also relate to systematic risk indirectly through expected excess return. Amihud (2000), Brennan, Chordia and Subrahmanyam (1998), Chordia, Subrahmanyam and Anshuman (2001), Chordia, Roll and Subrahmanyam (2000), Hasbrouck and Seppi (2001), and Hu (1997) have shown that expected excess return may contain a premium associated with liquidity. Wang (1994) and He and Wang (1995) have also shown that heterogeneous information are associated with expected excess return and trading volume.

[20] Here, in order to make a clearer presentation on the variation of turnover betas, we have shifted $\beta 1$, $\beta 2$, and $\beta 3$ up by adding 30, 20, and 10 to their respective values. The data is sorted by turnover.

Note that the cross-sectional relationship between residual turnover in equation (4) and idiosyncratic risk is particularly strong. Table 9 reports the coefficients as well as R^2 of the cross-sectional regression of residual turnover on idiosyncratic risk. The highly significant coefficients provide clear evidence that firm-specific turnover is strongly related to news about the firm cash flow and risks. The explanatory power of these cross-sectional regressions ranges from 6.12% (1987-1991) to 59.5% (1967-1971). These results are consistent with Chordia, Roll, Subrahmanyam (2000) who find that recent market volatility exerts a strong influence on stock trading activity. Our contribution here is the decomposition of market volatility into systematic risk and residual risk and the focus on the impact of volatility on trading activity at the firm level. Here, our results reaffirm our earlier discovery in Figure 2 that there is strong co-movement between volatility on trading activity at the firm level.

V. Conclusion

Trading activity is fundamental to a deeper understanding of interactions between stock returns and economic news. In this article we provide a formal test of the duo-factor model developed by Lo and Wang (2000) on return and trading volume. We make two methodological contributions to the literature. First, we introduce a recently developed consistent statistic by Bai and Ng (2001) to determine the number of factors in a duo approximate multifactor model. The approach allows for correlation and heteroskedasticity at both time and cross-section dimension. Second, our empirical study uses data from individual stocks rather than from beta-sorted portfolios. By exploiting the advantage of a large cross-section of individual stocks, we get around the nonstationarity issue in turnover. Our results are robust to the presence of either a trend or a unit root in the systematic component of turnover. Moreover, we are able to detect more cross-sectional variation in turnover and relate them to volatility at the firm level.

Based on a balanced panel of return and turnover data from NYSE and Amex stocks, we find the following results: First, turnover factor models are quite useful in explaining the variation of turnover for large panel data set. We find that there are four or five systematic factors driving firm turnover and that a significant portion of variation in firm trading volume is determined by the common factors in the market. Second, there is a significant increase in the variation of idiosyncratic turnover during the 1962 - 1996 sample period, parallel to the discovery of a noticeable increase in firm level volatility by Campbell, Lettau, Malkiel and Xu (2001) over the same time period. Third, there is a significant co-movement between volatility and turnover at the systematic level. These findings suggest that trading volume are not purely random but are driven by trading activities associated with macroeconomic and firm-specific news. However, we reject the restriction of Lo and Wang that excess return and turnover should have the same number factors in the duo-factor model.

There are many issues that remain to be examined. If the duo-factor model has provided a parsimonious description of monthly data, it is interesting to know whether the model works equally well on high frequency data. Hasbrouck and Seppi (2001) have taken a step in that direction, though they do not explicitly test the model constraints and their sample is limited to the thirty Dow Industrial stocks. Moreover, if the duo-factor model fits the US data reasonably well, it could also help us understand stock price and trading behavior in foreign markets. Furthermore, if firm news and asymmetric information drive trading volume, then by using the return and turnover decomposition developed in this article we may obtain a better measure of "abnormal " trading volume and gain additional insights in trading behavior.[21] In addition, while we find that the empirical results of unbalanced panel is quite different from those of balanced panel, we have not explored the return and trading behavior of the firms with missing observations. We leave these for future research.

[21] See Llorente, Michaely, Saar, and Wang (2001).

Appendix

To derive the consistency result of the statistic for the number of factors in the APT model of (3), Bai and Ng introduced the following assumptions:

Assumption A: Factors

$E\|F_t\|^4 < \infty$ and $T^{-1}\sum_{t-1}^{T} F_t\, F_t' \to \sum_F$ as $T \to \infty$ for some positive definite matrix $\sum_F$.

Assumption B: Factor Loadings

$\|\beta_i\| \le \bar{\beta} < \infty$, and $\|B' B / N - D\| \to 0$ as $N \to \infty$ for some KxK positive definite matrix D.

Assumption C: Time and Cross Section Dependence and heteroskedasticity

There exists a positive constant $M < \infty$, such that for all *N* and *T,*

1. $E(e_{it}) = 0\, E|e_{it}|^8 \le M$;
2. $E(e'_s\, e_t / N) = E\left(N^{-1}\sum_{i-1}^{N} e_{is} e_{it}\right) = \gamma_N(s,t), |\gamma_N(s,s)| \le M$ for all s, and

$$T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T} |\gamma_N(s,t),| \le M \, ;$$

3. $E(e_{it} e_{jt}) = \tau_{ij,t}$ with $|\tau_{ij,t}| \le |\tau_{ij}|$ for some τ_{ij} and for all t. In addition,

$$N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N} |\tau_{ij}| \le M \, ;$$

4. $\underline{E(e_{it} e_{js}) = \tau_{ij,ts}}$ and $(NT)^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{T=1}^{T}\sum_{s=1}^{T} \|\tau_{ij,ts}| \le M$;
5. For every (*t,s*), $E\left|N^{-1/2}\sum_{t=1}^{N}\left[e_{is}e_{it} - E(e_{is}e_{it})\right]\right|^4 \le M$.

Assumption D: Weak dependence between factors and idiosyncratic errors

$$E\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}F_t\ e_{it}\right\|^2\right)\leq M\ .$$

Assumption A and B are fairly standard for factor models and their ensure that each factor would have a bounded and non-trivial contribution to the variance of asset returns (or turnover). While we only consider non-random factor loadings here, the results still hold when B is random, provided they are independent of the factors and idiosyncratic errors. Assumption C allows for limited time series and cross section dependence in the idiosyncratic risks. Heteroskedasticity in both the time and cross section dimensions are also allowed. Therefore, our model is more general than a strict factor model of Ross (1976) that assumes no correlation across e_{it}. BN has shown that the above assumption C is consistent with the approximate factor model of Chamberlain and Chamberlain and Rothchild (1983) in the sense that it ensures that the largest eigenvalue of the N x N covariance matrix for the idiosyncratic risks must be bounded. While Chamberlain and Rothchild did not make any explicit assumption about the time series behavior of the factor, BN allows for serial correlation and heteroskedasticity. They have shown that Assumption C3 maintains the condition that the largest eigenvalue of the covariance matrix for the idiosyncratic risks will be bounded, thus their results is consistent with the approximate factor pricing model of Chamberlain and Rothchild. Here our discussion focus on the return factor model of (3), but the same assumptions A-D should also apply to the turnover factor model of (4) for estimating the number of factors.

References

Amihud, Y. and H, Mendelson, 1986, Asset Pricing and the Bid-Ask Spread;' *Journal of Financial Economics,* 17, 223-249.

Amihud, Yakov. "Illiquidity And Stock Returns: Cross-Section and Time-Series Effects." NYU WP, August 2000.

Andersen, T., 1996, "Return Volatility and Trading Volume: An Information Interpretation;' *Journal of Finance,* 51, 169-204.

Bai, J. and S. Ng, 2001a, Determining The Number Of Factors In Approximate Factor Models, Econometrica, forthcoming.

Bai, J. and S. Ng, 2001b, Determining A PANIC Approach to Unit Root in Panel Data, Working Paper, Boston College.

Bamber, L., 1986, "The Information Content of Annual Earnings Releases: A Trading Volume Approach," *Journal of Accounting Research,* 24, 40-56.

Berk, J. 2000, Sorting Out Sorts, *Journal of Finance,* 55, 407-427.

Brennan, M., T. Chordia, and A. Subrahmanyam, 1997, "Alternative Factor Specifications, Security Characteristics and the Cross-Section of Expected Stock Returns," *Journal of Financial Economics*.

Brown, S., 1989, The Number of Factors in Security Returns, *Journal of Finance,* 44, 1247-1262

Chordia, T., R. Roll and A. Subrahmanyam, 2000, Commonality in Liquidity, *Journal of Financial Economics* 56, 3-28.

Chordia, Tarun, Avanidhar Subrahmanyam, and Ravi Anshuman, 2001, The Volatility of Trading Activity and Expected Stock Returns, *Journal of Financial Economics,* 59

Chordia, T., R. Roll and A. Subrahmanyam, 2001, Market Liquidity and Trading Activity, *Journal of Finance,* 56, 501-530.

Chamberlain, G., 1983, "Funds, Factors, and Diversification in Arbitrage Pricing Models," *Econometrica,* 51, 1305-1323.

Chamberlain, G., and M. Rothschild, 1983, "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets," *Econometrica,* 51, 1281-1304.

Connor, G. and R. Korajczyk, 1993, A Test for the Number of Factors in an Approximate Factor Model, *Journal of Finance*, Vol. 48, No. 4., 1263-1291.

Fama, F. and K. French, 1996, "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance*.

Gallant, R., P. Rossi, and G. Tauchen, 1992, "Stock Prices and Volume," *Review of Financial Studies,* 5, 199-242.

Gervais, S., R. Kaniel, D. Mingelgrin, 2001, The High Volume Return Premium, *Journal of Finance*, 56, 877-920.

HASBROUCK, J., D Seppi, 2001, Common Factors in Prices, Order Flows and Liquidity, *Journal of Financial Economics*, 59, 2, 383-411.

Hiernstra, C., and *J.* Jones, 1994, "Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation," *Journal of Finance,* 49, 1639-1664.

He, H., and *J.* Wang, 1995, "Differential Information and Dynamic Behavior of Stock Trading Volume," *Review of Financial Studies,* 8, 919-972.

Hu, *S.,* 1997, "Trading Turnover and Expected Stock Returns: Does It Matter and Why?," Working paper, National Taiwan University.

Jain, P., and G. Ioh, 1988, "The Dependence between Hourly Prices and Trading Volume," *Journal Financial and Quantitative Analysis,* 23, 269-282.

James, C., and R. Edmister, 1983, "The Relation Between Common Stock Returns, Trading Activity, and Market Value," *Journal of Finance,* 38, 1075-1086.

Christopher S. Jones, 2001, Extracting Factors from Heteroskedastic Asset Returns, *Journal of Financial Economics* , forthcoming.

Karpoff, *J.,* 1987, "The Relation between Price Changes and Trading Volume: A Survey," *Journal Financial and Quantitative Analysis,* 22, 109-126.

Lamont, O., 1998, Earnings and Expected Returns, *Journal of Finance* 53, 1563-1587.

Lamoureux, C., and W. Lastrapes, 1990, "Heteroskedasticity in Stock Return Data: Volume vs. GARCH Effects;' *Journal of Finance,* 45, 487-498.

Lehmann, B. N. and Modest, D. (1988), The Empirical Foundations of the Arbitrage Pricing Theory, *Journal of Financial Economics* 21, 213–254.

Llorente, G., R. Michaely, G. Saar, and J. Wang, 2001, "Dynamic Volume-Return Relation of Individual Stocks" forthcoming in the *Review of Financial Studies*.

Lo, A., and J. Wang, 2001, "Trading Volume: Implications of an Intertemporal Capital Asset-Pricing Model," working paper, MIT.

Lo, A., and J. Wang, 2000, Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory, *Review of Financial Studies* 13, 257-300.

Mei, J., 1993, "A Semi-autoregression Approach to the Arbitrage Pricing Theory", *Journal of Finance*, 48, 599-620.

Michaely, R., and J. Vila, 1996, "Trading Volume with Private Valuation: Evidence from the Ex-Dividend Day," *Review of Financial Studies,* 9, 471-509.

Morse, D., 1980, "Asymmetric Information in Securities Markets and Trading Volume," *Journal of Financial and Quantitative Analysis,* 15, 1129-1148.

Roll, R. and S. Ross, "An empirical investigation of the APT" *Journal of Finance*, 35, 1073-1103.

Ross, *S.,* 1978, "Mutual Fund Separation in Financial Theory-The Separating Distributions" *Journal of Economic Theory,* 17,254-286.

Smidt, *S.,* 1990, "Long-Run Trends in Equity Turnover;"' *Journal of Portfolio Management,* Fall, 66-73.

Stickel, *S.,* 1991, "The Ex-Dividend Day Behavior of Nonconvertible Preferred Stock Returns and Trading Volume;' *Journal of Finance and Quantitative Analysis,* 26, 45-61.

Tauchen, G., and M. Pitts, 1983, "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica,* 51, 485-506.

Tkac, P., 1996, "A Trading Volume Benchmark: Theory and Evidence," working paper, University of Notre Dame.

Wang, *J.,* 1994, "A Model of Competitive Stock Trading Volume;' *Journal of Political Economy,* 102, 127~168.

Table 1: Summary Statistics

Panel A: Number of Securities in Each Sample, Percentage of securities with missing observation in returns, Percentage of securities with missing observation in turnover. Common shares are from CRSP share codes 10 and 11, excluding stocks containing zeros in reported volume.

Period	Dates	Number of firms in unbalanced panel	Percent of firms missing returns	Percent of firms missing turnover	Number of firms in balanced panel	Number of error 1 firms	Number of error 2 firms	Number of error 3 firms
1	1962-1966	2407	17.41%	17.21%	1595	209	3	67
2	1967-1971	2819	24.59%	24.39%	1562	395	3	115
3	1972-1976	2815	12.61%	12.56%	1998	174	9	25
4	1977-1981	2762	18.50%	18.26%	1788	318	2	64
5	1982-1986	2580	21.34%	21.03%	1513	347	17	93
6	1987-1991	2519	25.56%	25.22%	1353	382	73	94
7	1992-1996	2788	25.66%	25.41%	1441	472	39	148

Panel B: Summary statistics for monthly value-weighted excess return and turnover of NYSE and AMEX ordinary common shares for July 1962 to December 1996. Turnover and returns are measured in percentages (annualized). We report mean, SD, and autocorrelation for the whole sample.

	Value-weighted excess return					Value-weighted turnover				
Period	mean	Total stand. dev.	R1	Systematic stand. dev	Idiosyncratic Stand. dev	Mean	Total stand. dev.	R1	Systematic stand. dev	Idiosyncratic stand. dev
1	9.18%	9.90%	2.33%	10.04%	2.53%	11.45%	0.66%	59.16%	0.60%	0.28%
2	5.85%	15.36%	9.25%	15.00%	3.60%	18.18%	0.70%	30.04%	0.67%	0.25%
3	1.19%	17.45%	4.28%	17.20%	4.35%	18.90%	1.19%	58.04%	1.17%	0.32%
4	6.08%	15.23%	-2.14%	14.39%	4.55%	30.53%	2.26%	74.45%	2.23%	0.46%
5	14.65%	14.56%	-10.90%	14.15%	3.50%	60.96%	3.41%	61.73%	3.33%	0.93%
6	10.75%	18.74%	17.83%	18.53%	3.74%	63.52%	3.98%	47.43%	3.77%	1.18%
7	11.41%	8.27%	-14.11%	7.61%	3.17%	61.94%	2.24%	51.89%	2.12%	0.74%

Table 2: Test of number of factors in the excess return and turnover models for balanced panels

Incremental R^2, θk , k = 1 ; :: :; 10 of the covariance matrix of weekly turnover and returns of NYSE and AMEX ordinary common shares in percentages for subperiods of the sample period from July 1962 to December 1996. We also report IC1 and average R^2 for each sample periods.

Returns													
Period	$\theta 1$	$\theta 2$	$\theta 3$	$\theta 4$	$\theta 5$	$\theta 6$	$\theta 7$	$\theta 8$	$\theta 9$	$\theta 10$	# factors	average R^2	stdev R^2
1	24.07%	4.86%	2.76%	2.67%	2.40%	2.24%	2.13%	2.10%	2.02%	1.96%	2	28.93%	13.26%
2	34.14%	4.26%	3.12%	2.14%	1.98%	1.91%	1.81%	1.70%	1.63%	1.54%	2	38.39%	11.47%
3	35.66%	4.98%	3.36%	2.50%	2.25%	2.03%	1.83%	1.78%	1.66%	1.64%	3	43.99%	13.59%
4	32.47%	4.83%	3.43%	3.07%	2.23%	1.86%	1.77%	1.63%	1.61%	1.53%	3	40.73%	13.73%
5	23.22%	5.25%	3.45%	2.47%	2.39%	2.09%	1.93%	1.90%	1.81%	1.79%	2	28.47%	14.37%
6	34.53%	5.10%	2.81%	2.57%	2.29%	1.97%	1.80%	1.73%	1.64%	1.56%	2	39.63%	17.00%
7	10.87%	6.02%	3.11%	2.87%	2.58%	2.36%	2.22%	2.13%	2.09%	2.01%	2	16.89%	12.05%

Turnover													
Period	$\theta 1$	$\theta 2$	$\theta 3$	$\theta 4$	$\theta 5$	$\theta 6$	$\theta 7$	$\theta 8$	$\theta 9$	$\theta 10$	# factors	average R^2	stdev R^2
1	17.02%	6.47%	6.13%	4.08%	3.66%	3.21%	2.98%	2.92%	2.68%	2.52%	4	33.71%	16.63%
2	17.95%	10.06%	7.58%	3.91%	3.23%	2.84%	2.64%	2.38%	2.14%	2.05%	5	42.73%	16.84%
3	24.35%	11.70%	3.92%	3.77%	3.04%	2.80%	2.52%	2.34%	2.17%	1.99%	5	46.78%	18.32%
4	18.09%	7.73%	5.83%	4.00%	3.33%	2.85%	2.71%	2.56%	2.31%	2.21%	4	35.66%	16.99%
5	16.23%	10.59%	4.66%	4.11%	3.09%	2.77%	2.41%	2.27%	2.23%	2.04%	4	35.58%	15.50%
6	17.78%	7.44%	4.84%	3.66%	3.25%	2.87%	2.55%	2.37%	2.24%	2.15%	4	33.72%	16.72%
7	11.15%	6.31%	5.04%	4.62%	3.71%	3.13%	2.87%	2.52%	2.40%	2.22%	4	27.12%	15.40%

Table 3: Simulation Test for the number of factors extracted for return and turnover Using PC Criterion

The table presents the frequency on the number of factors extracted from return and turnover data over 100 simulations. The kmax is set to be 10. Each simulation involves the draw of a set of *JxT* individual return and turnover data.

Return		Frequency found in 100 simulation studies							
Time period	True k	1	2	3	4	5	6	Mean k	Std k
1	2	6%	**94%**	0%	0%	0%	0%	1.94	0.24
2	2	0%	**100%**	0%	0%	0%	0%	2.00	0.00
3	3	0%	36%	**64%**	0%	0%	0%	2.64	0.48
4	3	0%	14%	**86%**	0%	0%	0%	2.86	0.35
5	2	0%	**100%**	0%	0%	0%	0%	2.00	0.00
6	2	0%	**100%**	0%	0%	0%	0%	2.00	0.00
7	2	1%	**99%**	0%	0%	0%	0%	1.99	0.10

Turnover		Frequency found in 100 simulation studies							
Time period	True k	1	2	3	4	5	6	mean k	Std k
1	4	0%	0%	11%	**89%**	0%	0%	3.89	0.31
2	5	0%	0%	0%	33%	**67%**	0%	4.67	0.47
3	5	0%	0%	0%	23%	**77%**	0%	4.77	0.42
4	4	0%	0%	5%	**95%**	0%	0%	3.95	0.22
5	4	0%	0%	3%	**97%**	0%	0%	3.97	0.17
6	4	0%	0%	8%	**92%**	0%	0%	3.92	0.27
7	4	0%	0%	13%	**87%**	0%	0%	3.87	0.34

Table 4: Simulation Results on the Factor Number Difference Using PC Criterion

The table presents the Type I and Type II Error Estimates for test on the difference between the number of return factors and the number of turnover factors based on 100 simulations for each time period. Each simulation involves the draw of a set of *JxT* individual return and turnover data. Here, k_{max} is set to be 10. For type I error estimates, we assume that the true numbers of return and turnover factors are three.

Panel A: Type I Error Estimates based on 100 Simulation for Each Time Period

	Difference			Frequency	Found		
Time period	Found	-3	-2	-1	0	1	2
1	-2	0%	**1%**	67%	32%	0%	0%
2	-3	**0%**	0%	25%	75%	0%	0%
3	-2	0%	**0%**	16%	82%	2%	0%
4	-1	0%	0%	**11%**	88%	1%	0%
5	-2	0%	**0%**	20%	80%	0%	0%
6	-2	0%	**0%**	41%	58%	1%	0%
7	-2	0%	**0%**	50%	50%	0%	0%

Panel B: Type II Error Estimates based on 100 Simulation for Each Time Period

				Frequency	Found		
Time period	True K-K’	-4	-3	-2	-1	0	1
1	-2	0%	5%	85%	10%	**0%**	0%
2	-3	0%	67%	33%	0%	**0%**	0%
3	-2	0%	28%	57%	15%	**0%**	0%
4	-1	0%	0%	14%	81%	**5%**	0%
5	-2	0%	0%	97%	3%	**0%**	0%
6	-2	0%	0%	92%	8%	**0%**	0%
7	-2	0%	1%	86%	13%	**0%**	0%

Table 5: Test of number of factors in the excess return and turnover models for unbalanced panels

Incremental R^2 θκ , k = 1 ; :: :; 10 of the covariance matrix of weekly turnover and returns of NYSE and AMEX ordinary common stocks for subperiods of the sample period from July 1962 to December 1996. We also report IC1 and average R^2 for each sample periods.

Returns													
Period	θ1	θ2	θ3	θ4	θ5	θ6	θ7	θ8	θ9	θ10	# factors	average R^2	stdev R^2
1	28.31%	6.94%	5.26%	3.22%	3.31%	2.63%	2.47%	2.10%	1.96%	1.81%	3	40.51%	23.92%
2	42.37%	5.50%	3.89%	3.34%	2.14%	2.10%	2.39%	1.73%	1.95%	1.34%	7	61.72%	21.57%
3	37.67%	6.66%	3.71%	3.09%	2.05%	2.45%	2.09%	1.95%	1.68%	1.64%	6	55.63%	19.71%
4	36.83%	7.38%	4.47%	2.78%	2.66%	2.10%	2.29%	1.78%	2.03%	1.38%	6	56.21%	22.33%
5	28.44%	6.26%	5.41%	3.40%	3.07%	3.21%	2.56%	2.68%	1.87%	1.72%	4	43.51%	25.32%
6	39.92%	8.26%	4.05%	3.70%	3.19%	1.83%	2.57%	2.04%	1.70%	1.34%	8	65.55%	23.48%
7	18.94%	8.32%	5.86%	5.51%	3.96%	2.68%	2.52%	2.64%	1.98%	1.88%	5	42.58%	30.66%
Turnover													
Period	θ1	θ2	θ3	θ4	θ5	θ6	θ7	θ8	θ9	θ10	# factors	average R^2	stdev R^2
1	20.47%	8.62%	6.20%	5.71%	6.24%	3.14%	4.26%	3.05%	2.49%	2.28%	9	60.17%	24.92%
2	24.59%	13.86%	8.65%	5.93%	3.33%	3.66%	2.73%	2.17%	2.48%	1.38%	9	67.40%	24.09%
3	26.83%	12.34%	5.97%	5.04%	3.50%	2.95%	2.95%	2.22%	2.27%	1.77%	7	59.58%	23.41%
4	18.69%	10.67%	10.24%	4.53%	5.76%	2.97%	2.34%	2.67%	2.26%	2.07%	8	57.87%	26.12%
5	16.72%	14.50%	8.07%	5.69%	6.36%	3.87%	3.30%	2.10%	1.92%	1.53%	7	58.52%	26.66%
6	19.40%	8.55%	14.01%	4.05%	6.45%	3.31%	3.02%	2.02%	2.02%	1.56%	9	62.82%	26.51%
7	18.35%	12.62%	6.55%	8.02%	3.98%	4.52%	2.56%	2.57%	2.04%	1.53%	8	59.18%	29.59%

Table 6: The determinants of turnover factors

Report OLS coefficients (and their t-statistics below these between parenthesis) as well as R^2 of regressing return and turnover factors on Fama French factors for each sample period. The number below R^2 gives the P-value of the F-test of joint significance for the coefficients.

	Return factor 1					Turnover factor 1				
Period	Constant	Mkt-RF	SMB	HML	R2	Constant	Mkt-RF	SMB	HML	R2
1	-0.268	0.259	0.162	0.099	98.49%	0.026	-0.104	0.226	0.022	27.64%
	-14.915	44.056	21.187	9.474	0.00%	0.209	-2.551	4.284	0.311	0.10%
2	-0.135	0.167	0.113	0.040	99.26%	-0.102	-0.035	0.175	0.069	25.31%
	-11.730	54.299	27.362	8.089	0.00%	-0.889	-1.122	4.208	1.377	0.09%
3	0.016	-0.145	-0.117	-0.054	99.10%	-0.025	0.052	0.002	0.028	7.22%
	1.174	-53.216	-29.859	-12.337	0.00%	-0.189	1.889	0.040	0.631	23.73%
4	0.184	-0.183	-0.128	-0.057	98.16%	-0.055	-0.030	0.050	0.036	3.77%
	9.454	-32.799	-15.551	-7.039	0.00%	-0.390	-0.730	0.846	0.615	53.80%
5	0.199	-0.229	-0.107	-0.014	98.25%	0.008	0.066	0.055	-0.074	22.16%
	9.736	-43.931	-11.431	-1.549	0.00%	0.061	1.909	0.878	-1.211	0.27%
6	-0.094	0.179	0.086	0.026	98.76%	-0.005	0.013	0.048	-0.029	3.41%
	-6.294	52.565	13.894	2.991	0.00%	-0.035	0.435	0.871	-0.383	58.12%
7	-0.422	0.405	0.135	0.110	97.08%	-0.047	-0.021	0.040	0.090	5.41%
	-16.517	41.466	14.360	10.770	0.00%	-0.324	-0.376	0.744	1.553	36.98%

Table 7: The time series relationship between components of volatility and turnover

Time series regression of systematic volatility and idiosyncratic volatility on systematic turnover and idiosyncratic turnover (value weight) and vice versa. The number below the R^2 is the P value for the joint F-test that the regression coefficients are all zero. Include lags for predictive analysis. The variables are de-trended.

	Constant	Sys. Vol	Sys. Vol -1	Sys. Vol -2	Idio. Vol	Idio. Vol -1	Idio. Vol -2	R^2
Total Turnover	2.43	435.47	10.58	51.98	-1090	-1394	-1223	28.8%
	10.93	12.70	0.30	1.49	-1.44	-1.86	-1.63	0.0%
Sys. Turnover	2.48	360.6	23.16	76.17	-1046	-1266	-1127	22.8%
	11.32	10.66	0.68	2.22	-1.40	-1.71	-1.53	0.0%
Idio. Turnover	-0.01	3.37	0.15	2.12	-18.21	-29.57	-5.23	5.9%
	-1.91	4.18	0.19	2.59	-1.02	-1.68	-0.30	0.0%

Table 8: The impact of asset risk on turnover (contemporaneous)

Cross-sectional regressions of median monthly turnover of NYSE and AMEX ordinary common shares for sub-periods from July 1962 to December 1996. Report OLS coefficients as well as R2 of regressing median turnover on turnover betas, return betas, and idiosyncratic risks. Report OLS coefficients as well as R2 of regressing median turnover on turnover betas, lagged return betas, and lagged idiosyncratic risks. Same sets of regressions again plus a set of firm specific variables. (Four regressions for each time period.). The number below the R^2 is the P value for the joint F-test that the regression coefficients are all zero.

Period	Constant	TO beta 1	TO beta 2	TO beta 3	TO beta 4	TO beta 5	Ret beta 1	Ret beta 2	Ret beta 3	R2
1	0.007	0.001	0.001	-0.003	-0.003					4.18%
	19.632	1.892	2.600	-5.273	-5.713					0.00%
	0.005	0.001	0.001	-0.002	-0.003		0.049	0.127		6.89%
	7.476	4.047	2.532	-3.863	-5.816		2.534	6.785		0.00%
2	0.008	-0.001	-0.001	0.001	-0.003	-0.001				9.54%
	19.327	-6.580	-3.833	2.187	-7.554	-1.342				0.00%
	0.004	-0.001	-0.001	0.001	-0.003	-0.001	0.070	0.094		11.90%
	4.496	-4.660	-3.222	3.335	-8.099	-2.738	4.620	5.602		0.00%
3	0.001	-0.003	-0.005	-0.002	0.001	0.000				16.56%
	2.278	-9.298	-10.758	-3.867	1.260	-0.002				0.00%
	0.011	0.001	-0.005	-0.001	0.001	0.000	0.174	0.079	-0.020	37.78%
	16.739	2.379	-12.870	-1.693	2.206	0.241	18.265	8.515	-1.736	0.00%

4	0.009	-0.002	0.002	0.000	-0.001					3.61%
	24.123	-5.000	5.998	0.691	-2.309					0.00%
	0.007	0.000	0.001	-0.002	0.000		-0.055	-0.120	-0.112	11.67%
	8.934	-1.190	3.273	-3.949	-1.020		-3.724	-7.174	-6.922	0.00%
5	0.010	0.001	-0.004	-0.002	0.000					7.32%
	15.157	2.983	-9.256	-2.819	-0.079					0.00%
	0.012	0.002	0.000	-0.003	0.001		0.030	-0.447		35.98%
	13.767	4.532	-1.012	-5.699	1.887		1.476	-24.780		0.00%
6	0.003	0.002	0.001	-0.001	0.002					3.31%
	3.761	4.333	2.037	-1.242	3.477					0.00%
	0.008	0.001	0.002	0.001	0.002		-0.097	-0.301		18.69%
	6.317	1.004	4.336	1.544	3.768		-4.583	-13.196		0.00%
7	0.008	-0.001	0.001	0.000	0.001					1.37%
	18.640	-3.146	1.778	-0.942	1.219					0.05%
	0.007	-0.001	0.001	-0.001	0.000		0.069	-0.075		2.68%
	9.749	-2.894	1.751	-1.111	0.851		2.777	-4.114		0.00%

Table 9: Relationship between idiosyncratic volatilities and turnover

Regression of the idiosyncratic volatility on idiosyncratic turnover volatility. Reported are the OLS coefficients (and their t-statistics below), the R^2. The number below the R^2 is the P value for the joint F-test that the regression coefficients are all zero.

Period	Constant	idio vol	R^2
1	-0.467	5.729	40.28%
	-9.925	32.787	0.00%
2	-1.163	8.379	59.54%
	-22.795	47.929	0.00%
3	-0.119	3.720	27.91%
	-2.930	27.806	0.00%
4	-0.158	4.879	35.60%
	-3.583	31.428	0.00%
5	0.488	3.228	15.98%
	8.781	16.958	0.00%
6	0.975	1.556	6.12%
	19.266	9.389	0.00%
7	0.380	3.802	26.94%
	7.994	23.041	0.00%

Figure 1a: VW total turnover

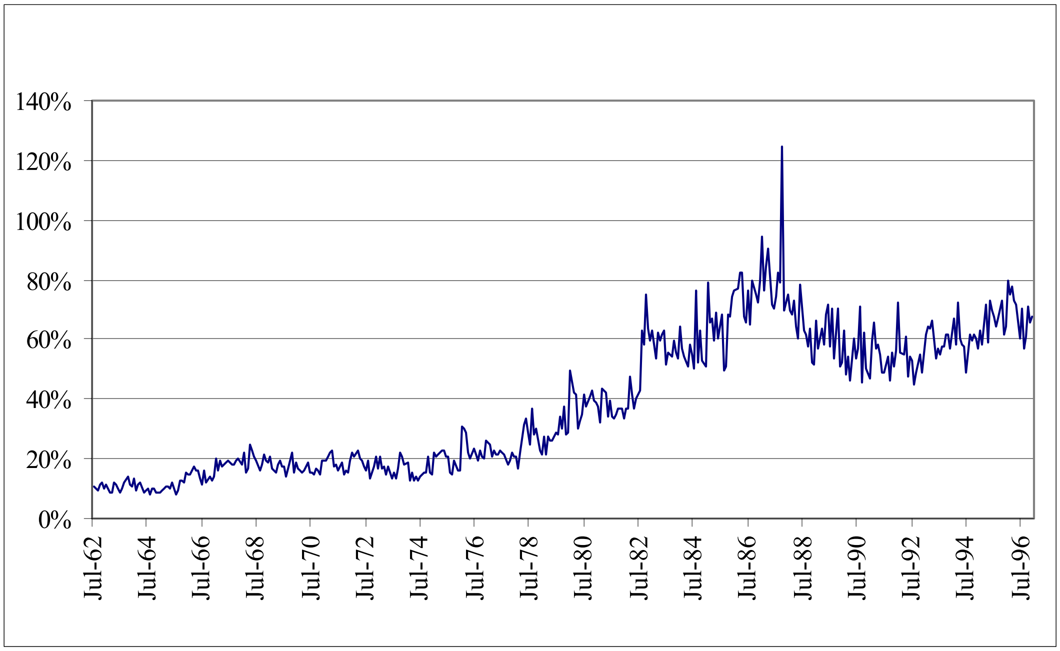

Figure 1b: VW systematic turnover (annualized)

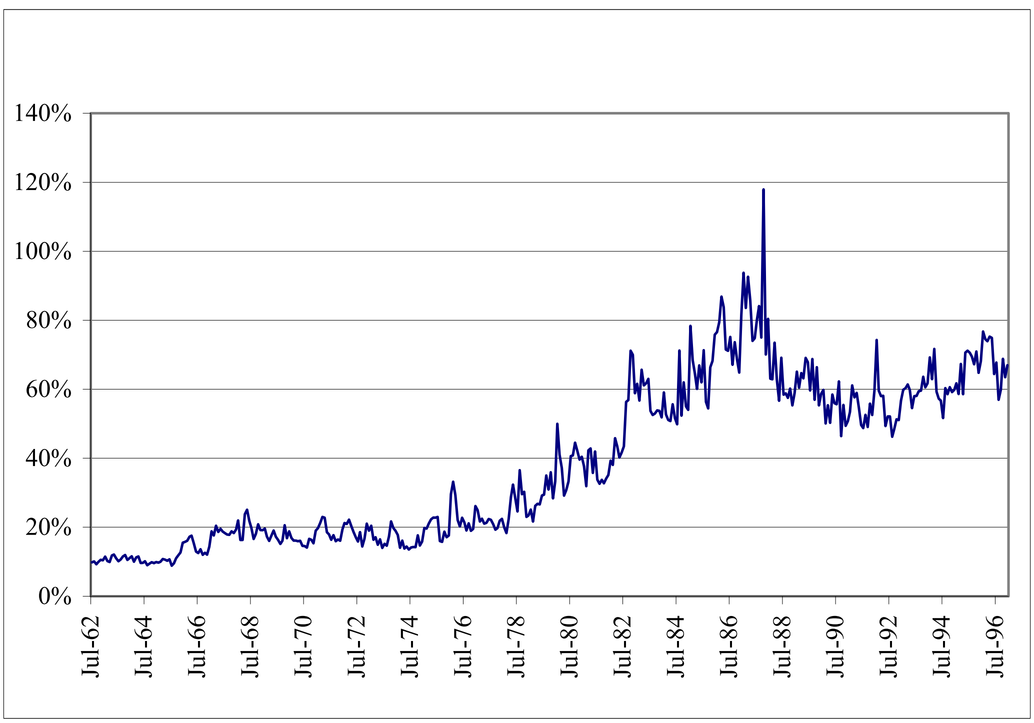

Figure 1c: VW idiosyncratic turnover(annualized)

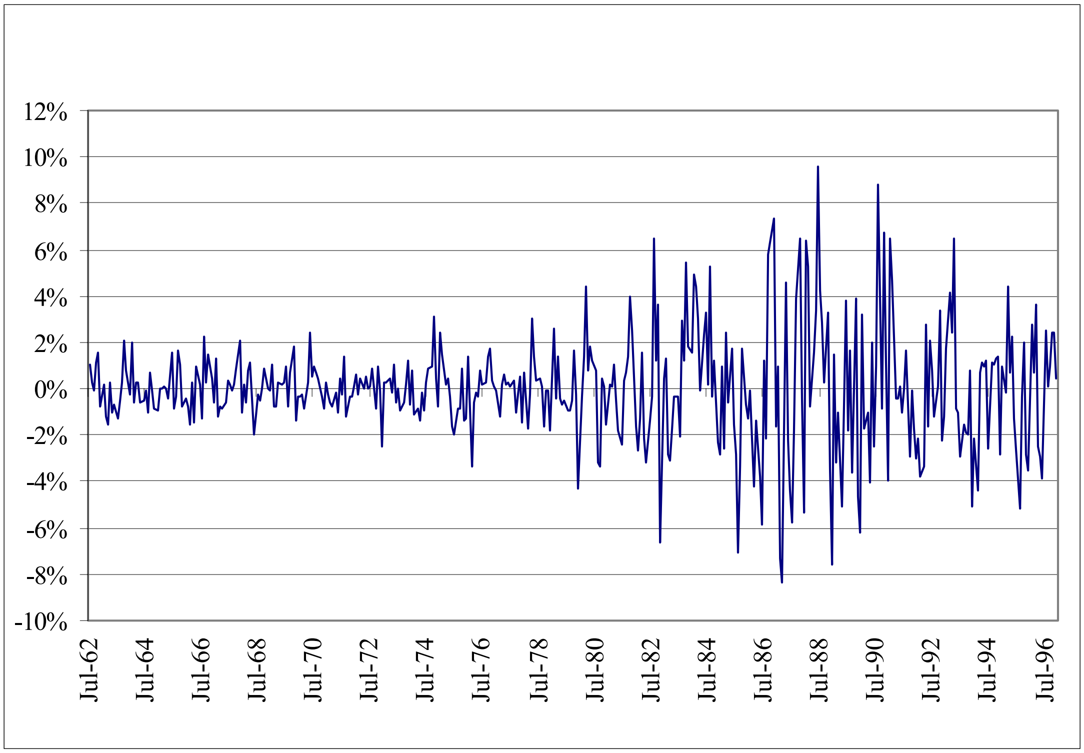

Figure 1d: VW idiosyncratic turnover (Absolute Value, 12-month Moving Average)

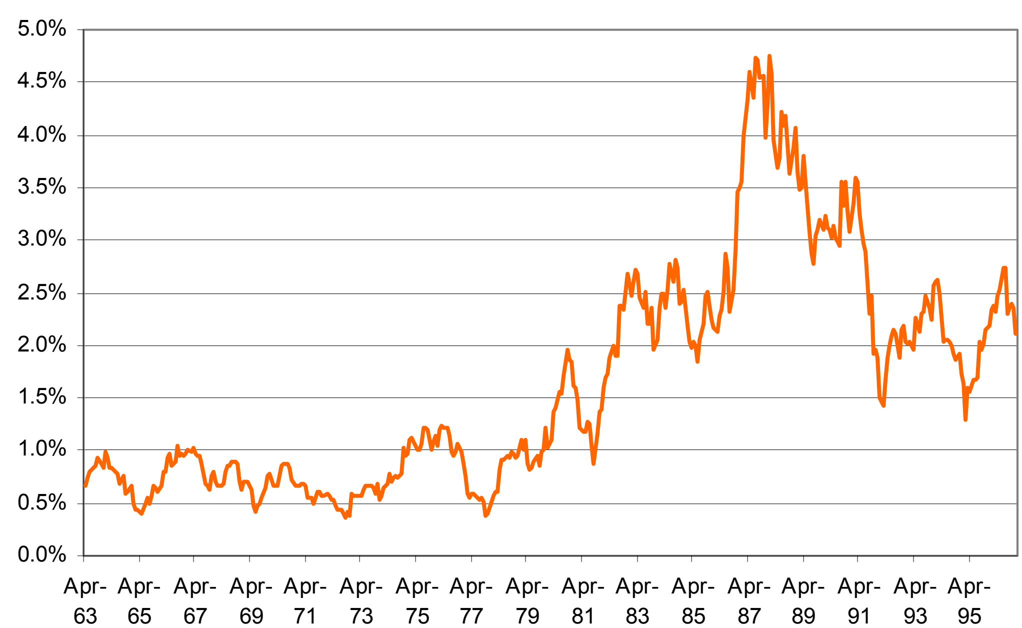

Figure 2: Total Turnover for Four Decile Portfolios (1/4/7/10, VW)

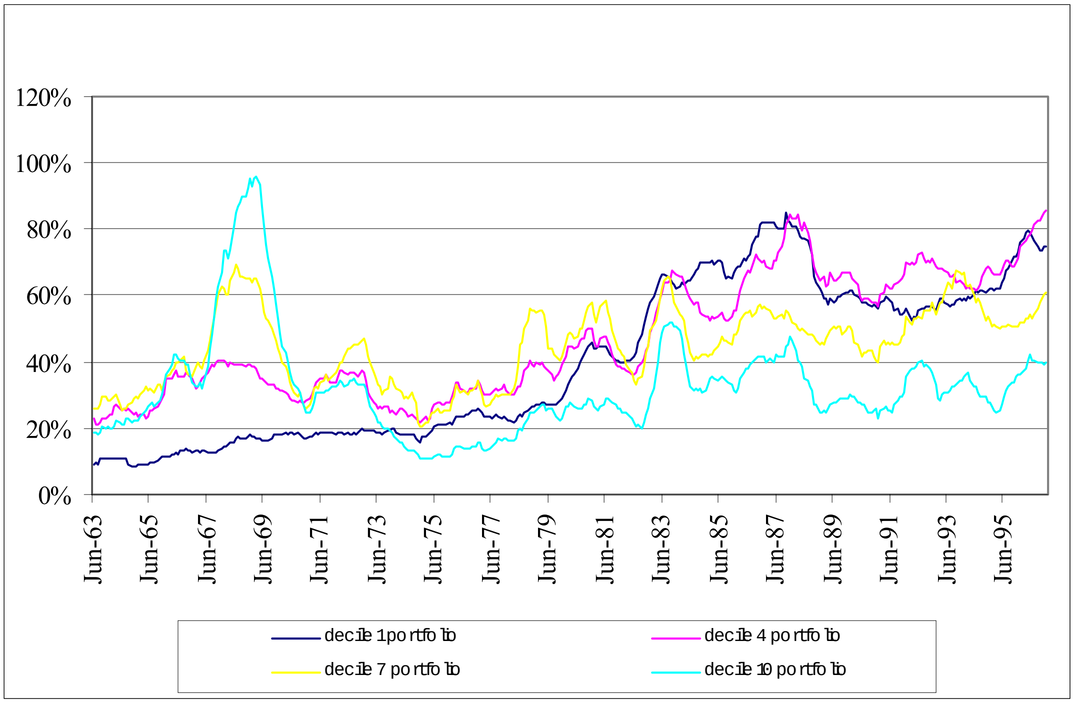

Figure 3: Cross Sectional Variation of Mean Turnover (M TO) and Turnover Betas (b1, b2, b3, b4)