Can stock returns be explained with macroeconomic factors ?

Diplomarbeit

zur Erlangung des Grades
eines Diplom-Volkswirtes

an der Wirtschaftswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

vorgelegt von

Emanuel Mönch

(Matrikel-Nr. 138098/98ws)

Prüfer : Prof. Harald Uhlig, Ph.D.

Berlin, 18. Juli 2002

Contents

1 Introduction

A central issue in macroeconomics is the question of how financial markets are connected to the real side of the economy. The ongoing integration of international capital markets and the repeated occurrence of large financial crises have raised the concern about this topic beyond academic circles. However, compared to the considerable amount of interest the subject attracts, as yet little is known about the interrelations between financial and macroeconomic variables.

One branch of research related to this topic deals with the question of how stock returns are linked to fluctuations of economic activity. Asset pricing theory and practice have long been dominated by market-inherent measures of risk. A classic result, the capital asset pricing model (CAPM), models an asset's expected return as a function of its past exposure to the market risk. Despite its simplicity, the CAPM in general works well in empirical tests. However, it fails to explain certain patterns in the cross-section of asset returns. For example, it has been shown that small stocks in terms of market capitalization on average earn higher returns than big stocks, even after accounting for the market risk. Also, stocks of firms whose book value is high compared to its market value on average exhibit higher returns than stocks with a low book-to-market equity ratio.

Recent results indicate that the variation of returns across assets and across time can to some extent be accounted for by macroeconomic variables. A variety of models have been proposed to explain effects like the two mentioned above, using as explanatory variables measures of economic activity or firm-specific characteristics that proxy for such. Merton (1973) shows that consumer-investors who solve an intertemporal utility maximization problem will hedge their portfolios against potential changes in the set of investment opportunities. Accordingly, they would demand more of those assets whose returns are positively correlated with changes in certain state variables that are expected to result in less consumption. Chen, Roll, and Ross (1986) show empirically that standard macroeconomic variables such as interest rate spreads, output and inflation growth indeed explain some of the cross-sectional variation of stock returns.

An important contribution to the explanation of certain particularities in cross-sectional return patterns has been made by Fama and French (1992, 1993, 1996). Constructing artificial portfolios on the basis of firm size and firms' book-to-market equity ratio, the authors show that together with the market return these portfolios go a long way explaining the variation of average returns across stocks. Fama and French state that the size and book-to-market portfolios mimic risk factors that are of concern to investors. In particular, they argue that the book-to-market portfolio proxies for an unobservable "financial distress" factor that earns significant risk premiums during recessions. However, there is as yet little empirical evidence for a close link between the Fama-French factors and individual macroeconomic variables.

[1] I thank David Hunt and Matthieu Nègre for helpful comments. I am very grateful to Kenneth French, Sydney Ludvigson, and Mark Watson for making their data accessible on the internet. I am further grateful to Martin Lettau for some helpful comments and for providing me with additional data. I especially thank my advisor, Harald Uhlig. The view expressed and all errors are my own.

More recently, the work of Lettau and Ludvigson (2001a, 2001b) shows that the log consumption-aggregate wealth ratio helps forecasting stock returns in the time and the cross-sectional dimension. Starting from a standard intertemporal budget constraint of a consumer-investor, they argue that consumption, asset wealth and labor income share a common trend, the deviations of which summarize the agent's expectations about future returns on asset wealth. Lettau and Ludvigson provide empirical evidence that an estimate of the residual of this cointegrating relationship has forecasting power for stock returns in time series regressions. They further show that this variable can also be successfully used as an instrument in conditional pricing models.

In further studies, alternative macroeconomic variables have been employed to explain the cross-sectional return patterns of stocks. These include, among others, the yield spread between BAA- and AAA-rated bonds, labor income growth (Jagannathan and Wang (1996)), returns on physical investment (Cochrane (1996)), innovations to the dividend yield, deviations of the T-bill rate from its medium-term trend (Campbell(1996)), and news related to future GDP growth (Vassalou (2002)). The results of these investigations indicate that all these variables explain some of the variation of returns across stocks, suggesting that they account for sources of risk that investors hedge against.

Given this evidence, I put forward the hypothesis that summary measures of economic activity - capturing the common variation in a large number of economic time series variables - are potentially useful candidates for factors in a pricing model. In a pioneering paper, Stock and Watson (1998) have recently proposed an estimation methodology to extract common dynamic factors from a large panel of time series variables. Based on principal components analysis, the authors construct the factors as weighted cross-sectional averages that minimize the idiosyncratic influences of individual variables. Corresponding to the use of similar averages with fixed weighting schemes in business-cycle analysis, Stock and Watson label these factors "diffusion indexes". They provide evidence that their estimates exhibit striking forecasting power for standard measures of economic activity compared to benchmark forecasting methods. Stock and Watson further suggest an extension of their method that makes it applicable to unbalanced data panels containing series with missing observations or of different frequency.

According to Stock and Watson's findings, diffusion indexes summarize different sources of systematic macroeconomic risk. It thus seems natural to presume that they can profitably be used as factors in a pricing model. Also, since risk premiums have been shown to vary over the business cycle, diffusion indexes as measures of the common variation of economic time series might prove useful as instrumental variables in conditional pricing models. In this study, I investigate whether these conjectures can be confirmed empirically. Following a common practice in empirical tests of asset pricing models, I limit my analysis to the explanation of the cross-sectional patterns of size and book-to-market sorted stock portfolios. I employ the cross-sectional regression approach of Fama and MacBeth (1973). This method is intuitively appealing and therefore widely used in cross-sectional tests of asset pricing models. However, since it exhibits some severe shortcomings, additional tests of model misspecification are warranted.

The empirical results of my study can be summarized as follows. Some lower-order diffusion indexes, each capturing only a small portion of the common variation in the data set, are shown to be significantly priced in cross-sectional tests of the model. Together with the market return, a set of three such diffusion indexes explains the variation of average returns across 100 size and book-to-market sorted stock portfolios about as well as the Fama-French three-factor model, and performs considerably better than other tested benchmark models. However, the diffusion index pricing model as well as most of the considered benchmark models, fails additional tests of model misspecification. Consequently, diffusion indexes cannot beyond all doubt be advocated as useful pricing factors. Yet, the same holds for variables that have been suggested by other authors. Similar remarks apply to the tested usefulness of diffusion indexes as scaling variables in conditional pricing models. Although they perform better than alternative macroeconomic variables, the overall model fit is unsatisfactory and the estimated relationships do not appear to be robust to tests of misspecification.

The remainder of this text is organized as follows. In the next section I briefly review some basic concepts of asset pricing theory that represent the starting point for my empirical investigations. In section 3, some standard asset pricing models that serve as benchmarks for tests of the diffusion indexes are presented. In section 4 I introduce Stock and Watson's methodology and briefly summarize their principal paper on diffusion indexes. Finally, in section 5, the employed econometric methodology is discussed and the empirical results of my investigation presented. Section 6 concludes the paper. A description of the data, some additional figures and tables, and a brief review of the most recent contributions to the literature on dynamic factor models are provided in the appendix.

2 Some Key Concepts of Asset Pricing

In this section I briefly summarize some central concepts of asset pricing that build the grounds for the theoretical and empirical derivations I will make in the remainder of the text. I mainly adopt the notation and presentational approach of Cochrane (2001).

2.1 The Basic Pricing Equation

We want to model the prices of financial assets. Consider a representative investor with utility function

$$U(c_t, c_{t+1}) = u(c_t) + \beta E_t[u(c_{t+1})],$$

where c_t denotes consumption at date t. Given this utility and an annual labor income of e_t and e_{t+1}, respectively, the investor chooses the amount to consume and to save in period t. Assume there is only one asset with price p_t offering a random payoff x_{t+1} at the beginning of period $t+1$. Then, the investor faces the following decision problem :

$$\max_{\xi} U(c_t, c_{t+1}) \quad s.t.$$

$$c_t = e_t - p_t \xi,$$

$$c_{t+1} = e_{t+1} + x_{t+1} \xi,$$

where ξ denotes the quantity of the asset the investor chooses to buy. Solving this maximization problem, we obtain the following first-order necessary condition

$$\frac{\partial U}{\partial \xi} = -p_t u'(c_t) + \beta E_t[u'(c_{t+1})x_{t+1}] = 0,$$

or

$$p_t u'(c_t) = E_t[\beta u'(c_{t+1})x_{t+1}].$$

This equation reflects the typical tradeoff between a loss of marginal utility in period t if the investor buys one additional unit of the asset and thus consumes less, and the marginal gain of utility in period $t+1$ resulting from the extra consumption the investor can afford from the (risky) payoff. The equation can be solved for p_t to get

$$p_t = E_t\left[\beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1}\right].$$

Now define the *stochastic discount factor* or *pricing kernel* m_{t+1} as

$$m_{t+1} \equiv \beta \frac{u'(c_{t+1})}{u'(c_t)}.$$

With this definition, the above equation simplifies to

$$p_t = E_t[m_{t+1}x_{t+1}], \tag{2.1}$$

which is *the* central asset pricing equation.

m_{t+1} is a measure of how the representative agent discounts uncertain future consumption. It is unknown in period t and thus correctly specified as *stochastic*. Note that the relation (2.1) is a very general statement that applies to basically all asset pricing problems since it does not refer to a particular utility function or to a particular asset. It can be shown that if the law of one price holds and if there is free portfolio formation, i.e. if investors can buy *any* linear combination of the available assets, then there is a unique stochastic discount factor m_{t+1} that prices all payoffs x_{t+1} in the payoff space correctly.[2]

If the investor pays p_t in period t to receive the payoff x_{t+1} in the following period, then the gross return R_{t+1} to her investment is given by

$$R_{t+1} = \frac{x_{t+1}}{p_t}.$$

Substituting this relation into (2.1), we obtain

$$1 = E_t[m_{t+1}R_{t+1}]. \tag{2.2}$$

Hence, gross returns can be interpreted as assets which always have an expected discounted value of 1. It follows directly that any *excess return*, $R^e_{t+1} = R^a_{t+1} - R^b_{t+1}$, has an expected discounted value of zero,

$$0 = E_t[m_{t+1}R^e_{t+1}].$$

[2] See Cochrane (2001) for two different proofs of this statement.

Since the main purpose of my investigation is to estimate the cross-section of returns on stock portfolios, equation (2.2) is the starting point for the empirical work that I have carried out. Yet, as it describes the representative agent's expectation of a discounted future return *conditional* on his time t information set, it is still unsuited for estimation. Since we do not know the agents' information sets in general, we can only estimate relations of this type unconditionally. To derive the unconditional version of (2.2), simply take the unconditional expectation of both sides,

$$E[1] = E[E_t[m_{t+1}R_{t+1}]].$$

Then, if the unconditional moment exists, the *Law of iterated expectations* tells us that this is equivalent to

$$1 = E[m_{t+1}R_{t+1}], \tag{2.3}$$

which we can in principle use for estimation since the sample analogue of a variable's unconditional expectation is given by its time series mean.

2.2 Expected Return-Beta Representations

Roughly speaking, the essence of asset pricing theory is to understand the relationships between the payoffs and the prices of assets. As we have seen in equation (2.1), they are directly related by the stochastic discount factor m. Hence, one can try to draw conclusions about m by confronting prices and payoffs which are both observable. However, equation (2.1) is not adapted to the use of linear regression techniques. We thus would like to find a representation of the equation that makes such methods applicable. Expected return-beta representations are a popular starting point for testing asset pricing models via linear regressions. Still following Cochrane (2001), I will now show how one can derive an expected return-beta representation from equation (2.3). I drop time subscripts for presentational convenience.

Consider a single asset i with return R^i. Then, equation (2.3) tells us that

$$1 = E[mR^i].$$

We know further that

$$Cov(m, R^i) = E[mR^i] - E[m]E[R^i].$$

Using this result, we obtain

$$1 = E[mR^i] = E[m]E[R^i] + Cov(m, R^i),$$

which is equivalent to

$$E[R^i] = \frac{1}{E[m]} - \frac{Cov(m, R^i)}{E[m]}.$$

Now we can divide and multiply by $Var(m)$, define

$$\gamma \equiv \frac{1}{E[m]},$$

$$\beta_{i,m} \equiv \frac{Cov(m, R^i)}{Var(m)},$$

and

$$\lambda_m \equiv -\frac{Var(m)}{E[m]},$$

to finally obtain

$$E[R^i] = \gamma + \beta_{i,m}\lambda_m. \tag{2.4}$$

Equation (2.4) is called a *beta pricing model* or *expected return-beta representation*. One can easily see that $\beta_{i,m}$ is the coefficient in a linear regression of the return R^i on the stochastic discount factor m. It can be interpreted as the *quantity of risk* associated with asset i. Accordingly, λ_m is often referred to as the *price of risk*. It is the same for all assets and rises and falls with the volatility of the discount factor.

2.3 Factor Pricing Models

While equation (2.4) has a form that makes linear regression techniques applicable, it still does not represent a testable model since we cannot directly observe the pricing kernel m. However, m is likely to depend on the overall macroeconomic conditions. We can thus try to model m in terms of observable state variables. Defining a vector f of explanatory variables and assuming a linear relationship between m and f, the most basic factor model is given by

$$m = a + b'f. \tag{2.5}$$

It is straightforward to show that equations (2.3) and (2.5) together imply an expected return-beta representation where the betas are the regression coefficients of R^i on the factors f. In fact, as we have seen above, equation (2.3) is equivalent to

$$E[R^i] = \frac{1}{E[m]} - \frac{Cov(m, R^i)}{E[m]}. \tag{2.6}$$

For computational convenience, assume that the factors are demeaned and thus $E[m] = a$. Now substitute $a = E[m]$ and $m = a + b'f$ into the above equation to obtain

$$E[R^i] = \frac{1}{a} - \frac{Cov(a + b'f, R^i)}{a} = \frac{1}{a} - \frac{b'Cov(f, R^i)}{a}. \tag{2.7}$$

According to the notation used in the previous section, define $\beta_{i,f}$ as the vector of coefficients resulting from a time series regression of R^i on the factors f. If there is a constant in the regression, $\beta_{i,f}$ is given by

$$\beta_{i,f} = \frac{Cov(f, R^i)}{Var(f)}.$$

Further define

$$\gamma \equiv \frac{1}{E[m]} = \frac{1}{a},$$

and

$$\lambda \equiv -\frac{Var(f)b}{a}.$$

Then, equation (2.7) is equivalent to

$$E[R^i] = \gamma + \lambda' \beta_{i,f}. \tag{2.8}$$

Since equation (2.8) holds for every single asset i, we are now in a position to test whether a given vector of pricing factors explains the variation of average returns across assets. Let there be k factors $f_1, f_2, \ldots, f_k$ and N assets with returns $R_1, R_2, \ldots, R_N$. Further, let $\mathbf{B}$ denote the $N \times k$ matrix of the β estimates obtained from linear regressions of each asset on the vector of pricing factors,

$$\mathbf{B} = \begin{pmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1k} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N1} & \beta_{N2} & \ldots & \beta_{Nk} \end{pmatrix},$$

where

$$\beta_i = \frac{Cov(f, R^i)}{Var(f)}.$$

Then, the cross-sectional regression model

$$E[R^i] = \gamma_i + \beta_{i1}\lambda_1 + \beta_{i2}\lambda_2 + \cdots + \alpha_i \qquad i = 1, 2, \ldots, N, \tag{2.9}$$

can equivalently be written in a more compact form as

$$E[R] = \gamma + \mathbf{B}\lambda + \alpha \tag{2.10}$$

where E denotes the time series mean, R denotes the $N \times 1$ vector of returns, γ is the $N \times 1$ vector of constants, and α is the $N \times 1$ vector of pricing errors. In section 5.1 we will see how equations (2.9) and (2.10) can be employed to test the ability of a set of pricing factors to explain the cross-section of portfolio returns.

2.4 Scaled Factor Pricing Models

If we estimate a model using unconditional moments as in (2.9) or (2.10), we can only do this because we have assumed that the parameters a and b in equation (2.5) are constant. To see this, consider the case of a factor model with time-varying parameters,

$$m_{t+1} = a_t + b_t' f_{t+1}. \tag{2.11}$$

Plugging this equation into (2.6), we obtain

$$E[R^i_{t+1}] = \frac{1}{E[a_t + b_t' f_{t+1}]} - \frac{Cov(a_t + b_t' f_{t+1}, R^i_{t+1})}{E[a_t + b_t' f_{t+1}]}.$$

Since a_t and b_t are time-varying, the above equation contains additional covariance terms compared to equation (2.7). Thus, an unconditional expected return-beta representation similar to (2.8) cannot be obtained. One way to allow for time-varying parameters in factor models when applying unconditional moments consists in introducing *conditioning information.* Assume there is a variable z_t that incorporates some information about possible shifts in the relationship between the state variables and the way investors discount risky payoffs. Then, we can model the parameters a_t and b_t as functions of this conditioning variable. The simplest way to do so is by assuming linear relationships

$$a_t = a_0 + a_1 z_t, \qquad b_t = b_0 + b_1 z_t.$$

Now, equation (2.11) can be written as

$$\begin{aligned} m_{t+1} &= a_0 + a_1 z_t + (b_0 + b_1 z_t)' f_{t+1} \\ &= a_0 + a_1 z_t + b_0' f_{t+1} + b_1' (f_{t+1} z_t), \end{aligned}$$

and we thus have obtained a linear factor pricing model with constant coefficients and additional factors z_t and $(f_{t+1} z_t)$. Since all pricing factors f_{t+1} are multiplied with z_t, we can think of z_t as a scaling variable. Accordingly, this type of conditional model is often labelled as a *scaled factor model.* After scaling the factors with z_t, we can now apply unconditional moments without generating undesired covariance terms,

$$\begin{aligned} 1 &= E[m_{t+1} R^i_{t+1}] \\ &= E[(a_0 + a_1 z_t + b_0' f_{t+1} + b_1' (f_{t+1} z_t)) R^i_{t+1}], \end{aligned}$$

which again implies an expected return-beta representation as in (2.8). Notice that including only one conditioning variable doubles the number of factors in the pricing model. Including more than one conditioning variable is possible in principle, but would result in an exponentially growing number of factors.

In section 5, I will discuss and evaluate the suitability of different variables as instruments in conditional pricing models. As mentioned above, candidates should incorporate as much information as possible on potential shifts in the risk premiums demanded by investors.

3 Benchmark Factor Pricing Models

In this section, I will present some well-known factor pricing models that shall serve as benchmarks for the assessment of the diffusion index pricing model. These are the static CAPM of Sharpe (1964), Lintner (1965) and Black (1972), a version of Merton's (1973) ICAPM with the selection of state variables following Chen, Roll, and Ross (1986), the three-factor model of Fama and French (1993), and the conditional versions of the consumption CAPM and the human capital CAPM proposed by Lettau and Ludvigson (2001b). Aside from these influential studies, there are, of course, further authors who have suggested factor pricing models to explain the cross-sectional patterns of stock returns. However, for presentational convenience and for matters of data availability, I have limited the number of comparison models to these five, basically adopting the selection of Lettau and Ludvigson (2001b). For a more extensive overview of the range of pricing models using macroeconomic variables as factors the reader is referred to Hodrick and Zhang (2000).

3.1 The Capital Asset Pricing Model (CAPM)

The capital asset pricing model is due to Sharpe (1964) and Lintner (1965). Its basic implication is that any individual asset's expected excess return over a risk-free rate, $(E[R^i] - R^f)$, is proportional to the expected excess return on the market portfolio, $(E[R^m] - R^f)$. The proportionality factor in this relationship is given by the well-known β, i.e. the coefficient in a linear time series regression of the asset's return on the return of the market portfolio,

$$\beta_{i,R^m} = \frac{Cov(R_i, R^m)}{Var(R^m)}.$$

Hence, the CAPM pricing formula is,

$$E[R^i] = R^f + \beta_{i,R^m}(E[R^m] - R^f). \tag{3.1}$$

Equation (3.1) is a direct implication of Sharpe's and Lintner's result that the market portfolio is mean-variance efficient if all investors in a market without frictions optimally hold mean-variance efficient portfolios. Let us derive the CAPM formula for the simplest case of a market with three assets, one of which guarantees the risk-free return R^f, while the two risky assets have returns R_1 and R_2 with expected values μ_1 and μ_2, respectively. Given these expected returns, an investor who wants to hold a mean-variance efficient portfolio earning the expected return μ_p, must choose portfolio weights a and b, such that

$$\begin{aligned} \{a, b\} &= \arg\min_{a,b\epsilon\mathbf{R}} \{Var(aR_1 + bR_2)\} \\ \text{s.t.} \qquad \mu_p &= a\mu_1 + b\mu_2 + (1 - a - b)R^f. \end{aligned}$$

Letting σ_1^2 and σ_2^2 denote the variances of R_1 and R_2, respectively, and σ_{12} the covariance of the two risky assets, the Lagrangian associated with the investor's minimization problem is given by,

$$L(a, b, \lambda) = a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\sigma_{12} + \lambda[\mu_p - a\mu_1 - b\mu_2 - (1 - a - b)R^f].$$

Solving for a, b, and λ obtains the first-order necessary conditions

$$\begin{aligned} &(i) && 2a\sigma_1^2 + 2b\sigma_{12} - \lambda\mu_1 + \lambda R^f &&= 0, \\ &(ii) && 2b\sigma_2^2 + 2a\sigma_{12} - \lambda\mu_2 + \lambda R^f &&= 0, \\ &(iii) && \mu_p - a\mu_1 - b\mu_2 - (1 - a - b)R^f &&= 0. \end{aligned}$$

Multiplying (i) with a and (ii) with b and adding both equations yields

$$2a^2\sigma_1^2 + 4ab\sigma_{12} + 2b^2\sigma_2^2 - \lambda a(\mu_1 - R^f) - \lambda b(\mu_2 - R^f) = 0.$$

Now let $\sigma_p^2 = Var(aR_1 + bR_2) = a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\sigma_{12}$ denote the variance of the portfolio. Then, the above equation simplifies to

$$2\sigma_p^2 - \lambda[a\mu_1 + b\mu_2 - (a + b)R^f] = 0.$$

Consider a market portfolio m that only consists of risky assets, i.e. that satisfies $a+b = 1$. Further, let $\mu_m = a\mu_1 + b\mu_2$ denote its expected return and σ_m^2 its variance. Then, it follows from the above relationship that

$$2\sigma_m^2 = \lambda(\mu_m - R^f),$$

which is equivalent to

$$\frac{1}{\lambda} = \frac{\mu_m - R^f}{2\sigma_m^2}. \tag{3.2}$$

Notice that this equation implies that the market portfolio is mean-variance efficient. In fact, since λ and R^f are constants, one cannot increase μ_m without increasing σ_m^2, and, inversely, one cannot decrease σ_m^2 without also decreasing μ_m. Resolving (i) for μ_1 yields

$$\mu_1 = R^f + \frac{1}{\lambda}(2a\sigma_1^2 + 2b\sigma_{12}).$$

Plugging (3.2) into this equation, we obtain

$$\mu_1 = R^f + \frac{\mu_m - R^f}{\sigma_m^2}(a\sigma_1^2 + b\sigma_{12}). \tag{3.3}$$

Now, given that the covariance between R_1 and the return on the market portfolio, R_m, is

$$\begin{aligned} Cov(R_1, R_m) &= E[(R_1 - \mu_1)(aR_1 + bR_2 - a\mu_1 - b\mu_2)] \\ &= E[(R_1 - \mu_1)[a(R_1 - \mu_1) + b(R_2 - \mu_2)]] \\ &= a\sigma_1^2 + b\sigma_{12}, \end{aligned}$$

equation (3.3) is obviously the CAPM formula for asset 1,

$$\mu_1 = R^f + \frac{Cov(R_1, R_m)}{Var(R_m)}(\mu_m - R^f). \tag{3.4}$$

Black (1972) derived a more general version of the CAPM which does not assume the existence of a risk-free rate. In the Black version of the CAPM,

$$E[R^i] = E[R_0] + \beta_{i,R^m}(E[R^m] - E[R_0]), \tag{3.5}$$

$E[R_0]$ denotes the average return on a zero-beta portfolio, i.e. a portfolio that is uncorrelated with the market portfolio. Obviously, equation (3.5) implies an expected return-beta representation

$$E[R^i] = \gamma_i + \beta_{i,R^m}\lambda_m. \tag{3.6}$$

Hence, employing the equivalence of $m = a + b'f$ and $1 = E[mR^i]$ with $E[R^i] = \gamma + \lambda'\beta_{i,f}$ that we have derived in section 2.2, the CAPM can be interpreted as a factor pricing model where the return on the market portfolio is the only risk factor. The expected return-beta representation (3.6) can be used to test whether the CAPM explains the cross-section of average returns. Contrary to the Sharpe-Lintner version of the model where the risk-free rate is included as a fixed regression constant, in the Black-version the expected zero-beta rate is thus simply estimated as the constant γ_i.

While first empirical tests of the capital asset pricing model based on single assets did not support its validity, the problem was soon diagnosed as being due to measurement errors in the estimated individual stock betas (see Cochrane (2001)). In fact, grouping stocks into portfolios based on sorts with respect to β, Fama and MacBeth (1973) and Black, Jensen, and Scholes (1972) found strong evidence for the linear relationship between asset returns and beta risk as predicted by the model.[3] However, despite this success, researchers soon began to notice that the model has severe problems explaining the cross-section of average returns when other sorting criteria are applied. In particular, Basu (1977) first reported that firms with low price-earnings ratios tend to have higher sample returns and firms with high price-earnings ratios on average have lower returns than their betas would suggest. Another failure of the CAPM, first documented by Banz (1981) and often referred to as the *size effect*, is the finding that small capitalization stocks on average earn returns that are too high compared to what the CAPM would predict. Further so-called *CAPM anomalies* have been found, including the evidence that stocks sorted by the firms' book-to-market equity ratios show average return patterns that cannot be explained by the CAPM. In particular, firms whose book value is high compared to their market value on average earn higher returns than their betas would suggest. This is often called the *value effect*. Further, DeBondt and Thaler (1985) report that portfolios formed on performance over the past three to five years exhibit some sort of *mean reversal*. More concretely, they find evidence that medium-term past losers outperform medium-term past winners. Jegadeesh and Titman (1993) come to a somewhat inverse result showing that short-term (up to one year) past winners earn significantly higher returns than short-term past losers. This is referred to as the *momentum effect*.

All these results show that the CAPM is a factor pricing model that exhibits some severe shortcomings. While beta certainly proxies for risk that is important for an investor's assessment of the value of an asset (namely its correlation with the market), it should not surprise us that this is not the only source of risk that investors are concerned about. However, the set of anomalies being quickly collected, it turns out that finding the risk factors that generate them is a rather delicate task. For each CAPM anomaly there is now a whole bunch of factor models trying to explain it. The models that I describe in the following sections represent some of the wide variety of attempts to parsimoniously describe the cross-section of average returns on stocks sorted by different criteria.

3.2 The Intertemporal CAPM

We have seen that the standard CAPM of Sharpe, Lintner, and Black has been derived under the assumption that investors, in order to maximize their instantaneous utility, choose a portfolio that is mean-variance efficient. However, this type of consumer-investor maximization behavior abstracts from potential changes of the set of investment opportunities over time and thus implicitly assumes that mean-variance efficiency is a time-invariant concept. Based on a critique of this restriction, Merton (1973) develops an intertemporal version of the capital asset pricing model, henceforth referred to as the ICAPM, where it is instead assumed that investors trade assets continually in time and thus take into account their expectations about the relationship between current and future returns.

In order to capture this idea, Merton makes the important assumption that investors

[3]As Campbell (2000) and Campbell, Lo, and MacKinlay (1997) point out, the slopes of the relation found by these authors were not consistent with the Sharpe-Lintner version of the model since the estimated mean return on the zero-beta portfolio was higher than the risk-free return. Yet, this could be accounted for by the Black version.

at each point in time have knowledge about the transition probabilities for returns on each asset over the next trading interval *and* in all subsequent periods. They are thus assumed to know the investment opportunity set vector *and* the stochastic process describing its changes over time. Making some further assumptions, Merton obtains an Itô process representation of the return of a single asset, the coefficients of which are equally assumed to follow Itô processes.[4] Then, he shows that the dynamics of the entire system can be described as a vector Itô process for some state variables.[5] The author further assumes the existence of a fixed number of risky assets and of one instantaneously riskless asset.[6]

Given the dynamics of the system, Merton shows that the maximization of the consumer-investors' intertemporal utility function, subject to some accumulation equation and budget constraint, obtains first-order conditions for both optimal consumption and optimal investment as functions of the state variables. From these equilibrium conditions, the author derives explicit demand functions for the different assets. One important result of his derivations follows directly from the particular form of the demand functions. In fact, these consist of two additive components the first of which is equivalent to the usual demand function of a single-period mean-variance maximizing investor as implied by the standard CAPM. Moreover, depending on the state variables, the second component has a straightforward interpretation as an investment to hedge against potential shifts in the investment opportunities set that are likely to reduce future consumption. More precisely, Merton states that this implies that investors will demand more (less) of an asset if its return is positively (negatively) correlated with changes in those state variables that are expected to result in less consumption.

Referring to Merton (1972), the author provides a two-fund theorem for the case of a time-invariant investment opportunities set. This theorem states that under the assumptions made, a unique pair of mutual funds can be established, one only containing the risky assets and the other only the riskless asset, such that investors are indifferent between choosing portfolios from these two funds or the original assets.[7] In a next step, Merton generalizes this result by providing a three-fund theorem for the case of time-varying investment opportunities sets. Assuming that there exists one of the risky assets which is perfectly negatively correlated with the risk-free rate,[8] he proves that under some additional assumptions about investors' expectations, one can establish three funds, one containing all risky assets, one containing only the risky asset that is correlated with the

[4]For matters of simplicity I dispense with a formal derivation of Merton's model. To establish a version of the ICAPM that is testable in a cross-sectional regression, it shall suffice here to highlight the baseline ideas of Merton's arguments.

[5]These are the price vector of all risky assets, as well as the drift and variance describing its stochastic change over time. Although Merton does not offer any straightforward interpretation of these state variables as measurable economic quantities at this point, later in the text he refers to the simplest form of a model with time-varying investment opportunities as one with the only state variable being a stochastically changing interest rate.

[6]Notice that "instantaneously riskless" does not mean constant. It only implies that all investors know at date t that they can earn a certain return if they hold the riskless asset over the next period. This general definition also allows for situations where the instantaneously riskless asset is not the same at each point in time.

[7]A mutual fund is defined as some financial intermediary holding all available individual assets and issuing shares of its own for purchase by the investors.

[8]The author mentions long-term bonds as an example for such an asset.

risk-free asset, and the third only containing the riskless asset, such that risk-averse investors will be indifferent between choosing portfolios from the original set of assets or these three funds. More precisely, as Merton argues, this theorem implies that if individual investors believe that the fund managers estimate the distribution of returns at least as well as they would do themselves, then all investment decisions can equivalently be made between the shares of the three funds. Since the second fund was defined to contain only the risky asset that is perfectly correlated with the instantaneously riskless rate, it can be understood as a fund that hedges against unfavorable changes in the investment opportunities set.

Finally, Merton derives an asset pricing formula equivalent to the CAPM formula, implying that the expected excess return on an individual asset over the risk-free rate depends on its correlation with the excess market return *and* on its correlation with the excess return on an asset that proxies for intertemporal changes of the investment opportunities set.[9] Hence, investors expect returns not only as a compensation for holding risk associated to the market, but also for holding risk connected to potential changes of the set of investment opportunities which, in turn, could arise due to changes of some state variables. As Merton further notes, his pricing formula implies that if a security that is perfectly negatively correlated with changes in the risk-free rate does not exist by itself, it would be created by investors since it always earns a premium over the risk-free rate.

Breeden (1979) provides an interesting simplification of Merton's ICAPM. He shows that under the same assumptions, the multi-beta pricing equation of Merton (1973) can be transformed into a single-beta representation where the beta is the asset's covariance with aggregate consumption. This version of the intertemporal CAPM is therefore often referred to as the consumption capital asset pricing model or CCAPM. Although the theoretical appeal of Breeden's model still prevails in asset pricing theory, it has as yet widely failed empirical tests (see Lettau and Ludvigson (2001b)).

According to the above derivations, the ICAPM can be written in expected return-beta language as

$$E[R^i] = \gamma + \lambda' \beta_{i,f},$$

where f is a vector of state variables and $\beta_{i,f}$ is the vector of betas of R_i with the state variables. Although Merton does not carry out empirical tests of the model himself, the ICAPM is often referred to as the theoretical groundwork for asset pricing models involving macroeconomic variables as pricing factors (Fama and French (1996), Hodrick and Zhang (2000), Lettau and Ludvigson (2001b), Brennan, Wang, and Xia (2002) etc.).

One of the first and most often cited empirical investigations of the question whether risk associated to macroeconomic variables earns premiums in asset prices has been made by Chen, Roll, and Ross (1986). Starting from the implications of Merton's ICAPM and Ross's APT,[10] they argue that in principle, any variable that systematically affects the

[9] Accordingly, some authors refer to the ICAPM as a "multi-beta CAPM".

[10] The central idea of the arbitrage pricing theory (APT) of Ross (1976) is that in equilibrium, idiosyncratic risk will not be priced since investors can diversify it away by holding portfolios. Thus, only systematic risk that is common to all assets in the market will carry risk prices. These common risk factors can, for example, be determined applying a statistical factor analysis to the covariance matrix of

economy's pricing operator, the stream of expected future dividend payments or, more generally, that helps describing the investment opportunities set, should be priced. From the simple observation that stock prices should equal expected discounted dividends,[11] Chen, Roll, and Ross identify the systematic forces that determine asset returns as those variables which influence the discount factor or future dividend yields. Precisely, they argue that unanticipated changes in the riskless interest rate represent one potential source of shifts in the discount rate. Another channel through which the discount factor will be affected is given by the risk-premium, i.e. the additional return investors demand over the risk-free rate for holding risky assets. Moreover, according to the CCAPM, Chen et al. suggest changes in real consumption as a proxy for changes in the marginal utility of investors and thus for the discount factor. The authors further identify the variables that potentially influence future cash flows or dividends as the following. First, they argue that unanticipated changes of the price-level will systematically influence asset returns since "pricing is done in real terms". Chen et al. argue that changes of the average inflation rate can also affect asset returns since relative prices may change due to movements of the inflation rate. Second, unanticipated changes of the level of production are likely to have an impact on future cash flows. Thus, the authors suppose that innovations to the time series behavior of industrial production are likewise linked to asset prices.

From these somewhat ad hoc assumptions about the interaction of individual macroeconomic variables and asset prices, Chen et al. construct a set of twelve state variables for which they investigate empirically the impact on stock portfolio returns in a cross-sectional asset pricing test. These variables comprise the monthly growth rate of industrial production, the annual growth in industrial production, the expected inflation as computed by Fama and Gibbons (1984), the unexpected inflation, the real short-term interest rate, the change in expected inflation, the risk premium as the return differential of low-grade corporate bonds and long-term government bonds, and the term structure as the return differential of long-term government bonds and the treasury bill rate, further the value-weighted and the equally-weighted NYSE index, the growth rate of real aggregate consumption, and an oil price index.

In a next step, Chen et al. test the ability of these state variables to explain the cross-sectional return variation of stock portfolios sorted by size. Applying a version of the Fama-MacBeth cross-sectional regression methodology that I will discuss more in detail in section 5.1, they obtain the following results. The monthly growth rate of industrial production, the risk premium, and the term spread appear to be highly significant in all considered subperiods. Moreover, unanticipated inflation and changes in expected inflation seem to be priced significantly during periods characterized by a high volatility of these variables. Interestingly, the considered stock indexes, although significant when tested alone, lose their predictive power in the presence of these state variables. Further, neither the change in real per capita consumption nor the oil price index show a significant impact on the size-sorted stock returns. Chen, Roll, and Ross conclude from their findings that stock returns are indeed subject to different sources of systematic risk which can be

asset returns (Connor and Korajczyk (1986)).

[11] Without taking into account the possibility of selling the asset, this is exactly the central asset pricing equation we have derived above, namely $p = E[mR]$.

proxied for by the innovations to some standard macroeconomic variables.

Based on the results of Chen, Roll, and Ross, I employ a simplified version of their pricing model as a benchmark for the diffusion index pricing model. Specifically, in addition to the market return, R^m, I include the following variables as factors: the monthly change of industrial production (ΔIP), the monthly change of the consumer price index as a measure of inflation (ΔCPI), the term spread as represented by the difference between long-term government bonds and the 1-month T-bill rate ($TSpr$), and the difference between low-grade corporate bonds and the 1-month T-bill rate as a measure for the risk premium or default spread ($DSpr$).[12] Hence, this factor pricing model has the expected return-beta representation

$$\begin{aligned} E[R^i] &= E[R_0] + \beta_{i,R^m}\lambda_{R^m} + \beta_{i,\Delta IP}\lambda_{\Delta IP} + \beta_{i,\Delta CPI}\lambda_{\Delta CPI} + \\ &\quad + \beta_{i,Tspr}\lambda_{Tspr} + \beta_{i,Dspr}\lambda_{Dspr}, \end{aligned} \tag{3.7}$$

where R_0 is the return on the zero-beta portfolio and where $\beta_{i,R^{em}}, \beta_{i,\Delta IP}, \beta_{i,\Delta CPI}, \beta_{i,Tspr}$, and $\beta_{i,Dspr}$ denote the betas of the respective factors with the i-th asset. I use equation (3.7) as the basis for cross-sectional tests of this version of the ICAPM.

3.3 The Three-Factor Model of Fama and French

In a widely recognized work, Fama and French (1993) have proposed a three-factor model that has quickly become a benchmark for all asset pricing models constructed to capture the cross-sectional variation of average stock returns. Two observations build the starting point for their reflections. First, the CAPM market beta as well as the consumption-based ICAPM beta both fail in explaining most of the observed cross-sectional patterns of stock returns. Second, several authors report good explanatory power of firm-related return variables such as size, book-to-market, leverage, and earnings-to-price for the cross-section of average returns. Based on these findings, Fama and French (1992) in a first step show that when these variables are tested jointly, size and book-to-market alone capture an important portion of the cross-sectional variation in average stock returns. In a second step, Fama and French (1993) construct a set of portfolios designed to mimic the returns on the size (price of the stock times number of shares traded) and book-to-market (ratio of the firm's book equity to its market capitalization (size)) factors. They show that together with the excess market return, these mimicking portfolios exhibit a striking good ability to explain the cross-section of average stock returns. In the following, I will briefly review the study of Fama and French (1993).

Fama and French construct their mimicking portfolios as weighted averages of the returns on six size and book-to-market (BE/ME)-sorted portfolios of all NYSE, Amex and

[12] Chan, Karceski, and Lakonishok (1997) perform an extensive empirical comparison of different state variables with regard to their ability to explain the cross-section of stock returns. They find that besides size, book-to-market, past returns and dividend yields, the term spread and the default spread are the only macroeconomic variables that are significantly priced. This results supports the use of these two variables in my version of the ICAPM. However, the observation made by Chan et al. that "industrial production growth and unanticipated inflation do not seem to be more useful than a randomly generated series of numbers" strongly questions my selection of these variables. Yet, as we will see in section 5, the empirical results I obtain do not fully confirm this view.

NASDAQ stocks. In particular, in June[13] of each year, they split all stocks into two size groups (small, S, and big, B) using the size median of all NYSE stocks as the breakpoint. Equivalently, based on the breakpoints for the top 30%, the middle 40% and the bottom 30% of the book-to-market ranked NYSE stocks, they sort all NYSE, Amex and NASDAQ stocks into three BE/ME equity categories (high, H, medium, M, and low, L).[14] From the intersections of these two size and three BE/ME categories, Fama and French construct the six portfolios S/L, S/M, S/H, B/L, B/M, and B/H,[15] for which they calculate monthly value-weighted returns from July to next years's June when the rankings are redone. It is important to note that only firms that have appeared in COMPUSTAT for at least to years are included in the portfolios to avoid survivor bias.[16]

Finally, Fama and French construct the factor mimicking portfolio SMB (small minus big), designed to capture the risk associated to firm size, as the difference each month between the average returns on the three small-stock portfolios and the average returns on the three big-stock portfolios,

$$SMB = 1/3 \cdot (S/L + S/M + S/H) - 1/3 \cdot (B/L + B/M + B/H).$$

In the same vein, they construct the portfolio HML (high minus low), designed to mimic the risk factor related to firms' book-to-market equity, as the difference each month of the average returns on the two high-BE/ME portfolios minus the average returns on the two low-BE/ME portfolios,

$$HML = 1/2 \cdot (S/H + B/H) - 1/2 \cdot (S/L + B/L).$$

Obviously, Fama and French's way of constructing SMB and HML aims at minimizing the influence of size in HML and that of BE/ME in SMB. Hence, the two portfolios should be good proxies for the risk factors they are assumed to mimic. Fama and French emphasize that since return variances are negatively related to size, their use of value-weighted returns for the construction of the factor mimicking portfolios is in line with the purpose of minimizing the variance of firm-specific factors. Finally, as a third factor designed to proxy for the market risk, Fama and French consider the excess market return, i.e. the return on all stocks in the six size and BE/ME categories minus the one-month bill rate, henceforth denoted R^{em}. As we have seen above, Black (1972) has derived a version of the CAPM that does not assume the existence of a risk-free interest rate. He instead estimates the expected zero-beta rate, i.e. the expected return on a portfolio that is uncorrelated with the market, as the regression constant. It is now a widely accepted result that the Black version of the CAPM performs better in empirical tests than the Sharpe-Lintner version. Hence, instead of employing the excess market return, in my tests of the Fama-French three-factor model I simply use the return on the market portfolio,

[13] This is to ensure that the accounting data of all firms be known when the ranking is performed.

[14] Fama and French define a firm's book equity as the COMPUSTAT book value of stockholders' equity, plus balance-sheet deferred taxes and investment tax credit, minus the book value of preferred stock (redemption, liquidation, or par value depending on availability).

[15] S/M, for example, contains all stocks that are simultaneously in the small size group and in the middle BE/ME category.

[16] However, as we will see later, there has been an interesting debate on the issue of survivor bias in the construction of the Fama-French factors.

R^m, as the third pricing factor. Note that Lettau and Ludvigson (2001b) proceed similarly in their model comparison exercise. Given these three factors, the expected return-beta representation of Fama and French's model is[17]

$$E[R^i] = E[R_0] + \beta_{i,R^m}\lambda_{R^m} + \beta_{i,SMB}\lambda_{SMB} + \beta_{i,HML}\lambda_{HML}, \tag{3.8}$$

where $E[R_0]$ is the return on the zero-beta portfolio, and β_{i,R^m}, $\beta_{i,SMB}$, and $\beta_{i,HML}$ are the time series regression coefficients of R^i on R^m, SMB, and HML, respectively. While I use equation (3.8) as the basis for cross-sectional regression tests of the Fama-French model, Fama and French (1993) adopt a different estimation methodology. Their objective is to explain average returns across stocks *and* bonds, in particular to test for overlap in the stochastic processes characterizing them. Hence, they seek for an estimation method which delivers estimates that have a clear interpretation as risk-factor sensitivities for both stocks and bonds. Given this requirement, Fama and French advocate the time series regression approach suggested by Black, Jensen, and Scholes (1972) as a suitable estimation methodology. This method consists in regressing separately the returns of all individual assets on the risk factors. Thus, one has to perform as many time series regressions as there are assets in the cross-section. Fama and French (1993) emphasize the following features of this method which make its use convenient for their purpose. First, the regression slopes have a clear interpretation as risk-factor sensitivities for both stocks and bonds. Second, and more importantly, a well specified factor model should produce regression intercepts that are indistinguishable from zero. Thus, the time series regression intercepts represent a simple summary measure of model fit.

The main results of Fama and French's analysis are the following. Although the excess return on the market portfolio has the largest explanatory power in describing the common time series variation of size and BE/ME-sorted stock portfolios, it leaves some variation - especially that of small-stock and high BE/ME portfolios - unexplained. SMB and HML show less overall predictive power, but work particularly well for these groups of portfolios. Tested jointly, all three factors, R^{em}, SMB, and HML, each capture some variation that the other factors fail to explain. Altogether, the time series variation of the 25 size and book-to-market sorted stock portfolios is very well explained by the three-factor model.[18] The empirical results of Fama and French further show that the term-structure related risk factors, when tested alone, explain a considerable portion of the time series variation of stock returns. Interestingly, they lose their predictive power when tested jointly with the three stock factors, but remain significant when the excess return on the market portfolio is excluded as a regressor. A time series regression of the excess market return on the term-structure factors shows that these have strong explanatory power. In a nutshell, there is variation of stock returns related to term-structure risk factors, but this type of risk turns out to be largely captured by the excess market return.

[17]In their paper, Fama and French also consider two term structure factors similar to those considered by Chen, Roll, and Ross (1986), and constructed to capture common risk in bond returns. They find, however, that these do not add much to the explanation of the cross-section of average stock returns when tested jointly with the three stock-market factors. Thus, in the model comparisons I perform, only the three-factor version of Fama and French's model is considered.

[18]In 21 out of the 25 regressions the R^2 is greater than 0.9.

The exceptionally good test results of Fama and French's three-factor model have induced an intense discussion about the worth of its scientific contribution. Although I cannot summarize the entire debate here, it appears necessary to briefly review the main arguments.[19] One important critique of the Fama-French model is that there is no clear economic interpretation of what sources of common risk the mimicking portfolios SMB and HML proxy for. Based on their finding that low book-to-market firms tend to have high earnings while high BE/ME firms have persistently low earnings, Fama and French (1995) argue that HML can be seen as a premium for risk related to "financial distress". They further note that both SMB and HML are correlated with firm profitability which could also suggest that they summarize variation of returns that is due to financial distress. They admit, however, that the empirical link between SMB and profitability is mainly determined by the fact that small stocks did not participate in the boom of the middle and late 1980s. As Cochrane (2001) further observes, there is only weak empirical evidence that HML is correlated with other indicators of financial distress. Another argument against the Fama-French interpretation of HML is given by Griffin and Lemmon (2002) who find that the Fama-French factors do not explain well the returns of low BE/ME stocks of firms that actually are in financial distress. Daniel and Titman (1996) state that the strong correlation of returns on stocks with high book-to-market ratios is not due to a common risk factor associated with BE/ME, but rather the result of similar firm-specific properties as industry or geographic location. Controlling for firm characteristics, they find that none of the Fama-French factors adds predictive power to the explanation of expected returns. On the other hand, Liew and Vassalou (2000) find that SMB and HML help forecasting GDP growth in different countries, supporting the view that they act as state variables in the context of Merton's (1973) ICAPM. Vassalou (2002) provides similar evidence, showing that much of the information in SMB and HML is news related to future GDP growth. Finally, Brennan, Wang, and Xia (2002) show that the Fama-French mimicking portfolios have predictive power for both real interest rates and the Sharpe ratio, which would be consistent with the existence of risk premia associated to SMB and HML.

Another line of argumentation against the Fama-French model focuses on the potential problem of survivor bias inherent to the construction of the factor mimicking portfolios. In particular, Kothari, Shanken and Sloan (1996) support this view, stating that the Fama-French factors, since based on COMPUSTAT data, exhibit a serious survivorship bias that is the major reason for the observed premium on book-to-market. However, as Daniel and Titman (1996) note, there are several authors who provide evidence against this argument. Davis (1994), for example, finds that even if the portfolios are constructed in a way such that survivorship bias can be excluded, the risk premium on HML remains significant.

Although not directly alleviating the critique that there is no clear economic interpretation of SMB and HML, Fama and French (1996) refresh the theoretical and practical justification of their three-factor model by showing that it helps to explain most of the asset pricing anomalies discussed in section 3.1. As pointed out above, one of the mo-

[19] In the introductions to their articles, Daniel and Titman (1996) and Brennan, Wang, and Xia (2002) both provide extensive but largely complementary summaries of work related to the findings of Fama and French. This may be seen as an indicator of the pervasiveness of literature on that issue.

tivations for the elaboration of their three-factor model had been the so far unexplained size effect that they show to be captured by SMB and HML. In Fama and French (1996) they investigate whether the three-factor model also helps to explain the value effect, the long-term reversal of stock returns, and the momentum effect. They find evidence that the three factors R^{em}, SMB, and HML largely capture the average return patterns associated with the value effect as well the long-term return reversal. However, the model fails in explaining the often observed momentum effect, i.e. the fact that short-term past winners tend to outperform short-term past losers.

Despite the continuing controversy about the interpretation of Fama and French's results, their three-factor model remains a cornerstone in the literature on cross-sectional asset pricing tests. As yet, there is no model that performs better in explaining the cross-section of average returns on size and book-to-market sorted portfolios. It is thus perfectly suited to serve as a benchmark for the assessment of the diffusion index pricing model.

3.4 The Conditional (C)CAPM of Lettau and Ludvigson

In a recent paper, Lettau and Ludvigson (2001b) have suggested conditional versions of several CAPM specifications that explain the cross-sectional return patterns of size and BE/ME-sorted portfolios about as well as the Fama-French three-factor model. In particular, Lettau and Ludvigson show that the simple static CAPM, the consumption CAPM, and a version of the CAPM augmented with a proxy for human capital - when scaled with a conditioning variable derived in an earlier paper (Lettau and Ludvigson (2001a)) - perform far better than the models' respective unconditional versions. The main contribution of Lettau and Ludvigson consists in the derivation of the conditioning variable $\widehat{cay}$ that I will now briefly summarize.

The theoretical starting point for Lettau and Ludvigson's (2001a) investigation is the idea that the empirically proved predictability of excess returns at business cycle frequencies "could simply reflect the rational response of agents to time-varying investment opportunities, possibly driven by cyclical variation in risk aversion". Then, macroeconomic variables that proxy for the time-variation of risk premia demanded by the investors should forecast excess stock returns. Lettau and Ludvigson thus explicitly seek for a variable that can serve as a summary measure of investors' conditional expectations of excess returns. They argue that one such variable is given by the consumption-aggregate wealth ratio, since in a wide class of forward looking models it summarizes agent's expectations of future returns on the market portfolio. To see this, start with an investor's intertemporal budget constraint

$$W_{t+1} = (1 + R_{m,t+1})(W_t - C_t),$$

where W_t denotes the investor's aggregate wealth (human capital plus asset holdings), $R_{m,t+1}$ the net return on aggregate wealth, and C_t consumption in period t. Following Campbell and Mankiw (1999), the authors show that under the assumption of a stationary consumption-aggregate wealth ratio, a first-order Taylor expansion of the budget constraint obtains[20]

$$\Delta w_{t+1} \approx k + r_{w,t+1} + (1 - 1/\rho_w)(c_t - w_t),$$

[20]Unless otherwise indicated, lowercase letters denote log variables in the remainder of this section.

where k is some constant, and where ρ_w denotes the steady-state ratio of new investment to total wealth.[21] Solving this equation forward, imposing that $\lim_{i\to\infty} \rho_w^i (c_{t+i} - w_{t+i}) = 0$, and taking unconditional expectations, yields an explicit expression for the consumption-aggregate wealth ratio as a function of expected future returns on aggregate wealth and expected future consumption growth,

$$c_t - w_t = E_t \left[\sum_{i=1}^{\infty} \rho_w^i (r_{w,t+i} - \Delta c_{t+i}) \right]. \tag{3.9}$$

Lettau and Ludvigson insist on the fact that since this expression has been derived on the basis of a very general intertemporal budget constraint, the consumption-aggregate wealth ratio summarizes the investors' expectations of future excess returns in a wide class of optimal consumption models.[22]

However, since aggregate wealth incorporates human capital that is not directly observable, the above expression of the consumption-aggregate wealth ratio is not useful for empirical applications. As a consequence, one has to find a proxy for human capital that only consists of observable variables. Lettau and Ludvigson do this by decomposing log human capital into $h_t = \kappa_t + z_t + y_t$, where κ is a constant, z_t is a mean-zero stationary component, and y_t is a non-stationary component which is assumed to be well-described by aggregate labor income. The authors justify this assumption by offering a number of different model specifications that link labor income to the stock of human capital and that all imply that the log of aggregate labor income summarizes the non-stationary component of human capital. Given this approximation, Lettau and Ludvigson express aggregate wealth, W_t, as the sum of asset wealth, A_t, and human capital, H_t, so that log aggregate wealth is given by

$$w_t \approx \omega a_t + (1-\omega) h_t,$$

where ω is the average share of asset holdings in total wealth. Accordingly, there is an approximate relationship between the log return on total wealth, $r_{w,t}$, and the log returns on asset wealth, $r_{a,t}$, and human capital, $r_{h,t}$,

$$r_{w,t} \approx \omega r_{a,t} + (1-\omega) r_{h,t}.$$

Plugging these two relations into equation (3.9) yields

$$c_t - \omega a_t - (1-\omega) h_t = E_t \left[\sum_{i=1}^{\infty} \rho_w^i \{ [\omega r_{a,t+i} + (1-\omega) r_{h,t+i}] - \Delta c_{t+i} \} \right],$$

which still contains the unobservable human capital h_t. Substituting the decomposition $h_t = \kappa_t + z_t + y_t$ for h_t, Lettau and Ludvigson finally obtain

$$c_t - \omega a_t - (1-\omega) y_t = E_t \left[\sum_{i=1}^{\infty} \rho_w^i \{ [\omega r_{a,t+i} + (1-\omega) r_{h,t+i}] - \Delta c_{t+i} \} \right] + (1-\omega) z_t. \tag{3.10}$$

[21] Strictly following the authors' derivations, unimportant linearization constants have been omitted.

[22] Obviously, this statement requires that the conditional expectation of future consumption growth be approximately constant, a fact that the authors conjecture to be confirmed by the data.

On the right-hand side of equation (3.10) figure only stationary variables. This implies that c, a, and y share a common trend whose deviation is given by $c_t - \omega a_t - (1-\omega)y_t$, henceforth denoted as cay.

Lettau and Ludvigson (2001a) estimate cay by applying a dynamic least squares (DLS) technique as suggested by Stock and Watson (1993). They report that the estimation results give strong support to the view that there is a single cointegrating vector for consumption, labor income, and asset wealth. Further, their finding that the shares of asset holdings and human capital in aggregate wealth are about one-third and two-thirds, respectively, are consistent with the capital and labor shares documented in the real business cycle literature.

Again assuming that Δc_{t+i} and $r_{h,t+i}$ are not too volatile, another implication of equation (3.10) is that cay summarizes the conditional market expectations of future asset returns. This property strongly qualifies it as a potential predictor for stock returns. In fact, Lettau and Ludvigson (2001a) show that their estimate of the cointegration residual, $\widehat{cay}$, performs very well in forecasts of the S&P 500 Index and the CRSP value-weighted index over short and intermediate horizons. Perhaps the most striking of their results is that $\widehat{cay}$ appears to be the best univariate predictor of stock returns for horizons up to one year.

With what has been said in section 2.4, the predictive power of $\widehat{cay}$ for the market return makes of it a promising candidate for a scaling variable in factor pricing models. Lettau and Ludvigson (2001b) explore this issue. They use $\widehat{cay}$ as a conditioning variable in a simple static CAPM, in a CAPM augmented with a proxy for human capital, and in a consumption CAPM. The consumption CAPM can be derived from the basic intertemporal consumption maximization framework introduced in section 2.1. In fact, we have defined the stochastic discount factor m_{t+1} as

$$m_{t+1} \equiv \beta \frac{u'(c_{t+1})}{u'(c_t)}.$$

By simply taking a first-order Taylor expansion of this equality, one can derive a linear relationship between m and Δc independently of the functional form one assumes for the utility function $u(c_t)$,

$$m_{t+1} = a_t + b_t \Delta c_{t+1}.$$

This relation can be considered as a single-factor model where Δc_{t+1} is the only pricing factor. Combined with $1 = E[m_{t+1} R^i_{t+1}]$, the Black version of this factor model has the expected return-beta representation

$$E[R^i] = E[R_0] + \beta_{i,\Delta c}\lambda, \tag{3.11}$$

where $E[R_0]$ is the expected return on a zero-beta portfolio, and where $\beta_{i,\Delta c}$ is the regression coefficient of R^i on Δc. Scaling with $\widehat{cay}$ results in a three-factor model with expected return-beta representation

$$E[R^i] = E[R_0] + \beta_{i,cay}\lambda_{cay} + \beta_{i,\Delta c}\lambda_{\Delta c} + \beta_{i,cay\Delta c}\lambda_{cay\Delta c}. \tag{3.12}$$

Following Campbell (1996) and Jagannathan and Wang (1996) who find that labor income growth as a proxy for human capital is significantly priced in cross-sectional tests of the CAPM, Lettau and Ludvigson (2001b) further test the specification

$$E[R^i] = E[R_0] + \beta_{i,cay}\lambda_{cay} + \beta_{i,R^m}\lambda_{R^m} + \beta_{i,\Delta y}\lambda_{\Delta y} + \beta_{i,R^m cay}\lambda_{R^m cay} + \beta_{i,cay\Delta y}\lambda_{cay\Delta y}. \quad (3.13)$$

Lettau and Ludvigson document strong improvements of the (C)CAPM specifications scaled by $\widehat{cay}$ over the unconditional versions. In their cross-sectional regressions using quarterly returns of 25 Fama-French size and BE/ME-sorted portfolios, the scaled consumption CAPM and the scaled human capital CAPM exhibit R^2's of 0.70 and 0.77, respectively, with the factors scaled by $\widehat{cay}$ being always significantly priced.

Despite these good overall results, in a recent paper, Brennan and Xia (2002) cast doubt on Lettau and Ludvigson's assertion that the predictive power of $\widehat{cay}$ is due to the variable's ability to proxy for investors' conditional expectations of future returns. Instead, they argue that the good forecasting performance arises as a result of the fact that the cointegrating relation between consumption, labor income, and asset wealth has been estimated in-sample and therefore incorporates a considerable look-ahead bias. To support their view, Brennan and Xia construct the variable *tay* as the residual of a similar cointegrating regression, where log consumption has been replaced with calendar time, and provide empirical evidence that *tay* forecasts excess stock returns at least as well as $\widehat{cay}$.

In a response to Brennan and Xia's critique, Lettau and Ludvigson (2002) defend their findings. Referring to results from cointegration theory, they argue that the mere fact that $\widehat{cay}$ has been estimated in-sample does not necessarily imply that its forecasting power be spurious. They further point out that - contrary to Brennan and Xia's characterization of *tay* as a mechanistic variable - this is in fact a forecasting variable which is consistent with forward-looking behavior of investors. Finally, based on recent theoretical work suggesting that in-sample tests are more reliable than out-of-sample tests, Lettau and Ludvigson defend the use of the former in their study.

It is beyond the scope of this paper to analyze the validity of the arguments put forward by Brennan and Xia and Lettau and Ludvigson, respectively. With the only exception being the paper of Hodrick and Zhang (2000), the predictive power of $\widehat{cay}$ as a conditioning variable in cross-sectional tests of asset pricing models has been widely accepted so far (Cochrane (2001), Campbell (2000), Menzly (2001), Vassalou (2002)). In my study, I will thus use Lettau and Ludvigson's conditional versions of the consumption CAPM and the human capital CAPM (henceforth denoted CCAPM and HCCAPM) as benchmarks for the diffusion index pricing model.

4 The Diffusion Indexes of Stock and Watson

In the previous section we have seen that a central issue in asset pricing theory is the search for state variables that are of special hedging concern to investors. The CAPM indicates that investors care about the covariation of an asset's return with the return of the market portfolio. However, as the evidence on CAPM anomalies such as the size effect or the value effect suggests, there are differences in average returns across groups of assets that are not captured by the market risk. Fama and French find artificially constructed factors that explain the cross-section of stocks sorted on size and book-to-market surprisingly well. Yet, there is an ongoing controversy about the question what state variables these portfolios poxy for. Lettau and Ludvigson state that investors' conditional expectations about future returns are related to the consumption-aggregate wealth ratio. Showing that this ratio can equivalently be expressed by the deviations of consumption, asset wealth and labor income from their shared common trend, they find evidence that the estimated residual of this cointegrating relationship indeed helps forecasting future excess returns. They further demonstrate that conditioning on the log consumption-aggregate wealth ratio substantially improves the predictive power of the consumption CAPM in cross-sectional regression tests.

Still other asset pricing models that are not reported in section 3 relate macroeconomic variables to cross-sectional variations of average stock returns. Cochrane (1996) from a supply-side point of view studies the asset pricing implications of firms' investment decisions. He develops a factor pricing model with returns on investment - inferred from a production function - as the pricing factors. Campbell (1996) advances a multifactor model allowing for intertemporal changes in investment opportunities. Exclusively selecting variables as factors that predict the market return in time series regressions, he considers the innovations of the dividend yield, the relative bill rate, the yield spread and labor income growth as pricing factors.[23] Jagannathan and Wang (1996) also use labor income growth as a proxy for human capital in their model. They further include the yield spread between high and low-grade bonds as a pricing factor. Finally, Vassalou (2002) considers a measure of news related to future GDP growth as an alternative risk factor.

All these models have in common that they try to identify the state variables which represent sources of risk that are priced by investors. The results are encouraging in that they provide empirical evidence for existing links between asset returns and macroeconomic variables. They also suggest that investors care about a variety of macroeconomic risk factors such as, for example, interest rates, aggregate production, consumption, labor income, and investment. However, all individual variables tested as factors in one of the above models are just what they are, individual variables. To exhaustively account for all types of risk that have been found to be priced by investors, one would have to include a lot of factors, thus leaving the path of parsimony. Yet, since all of these variables are closely linked to the business cycle, there must be a large common component in their variation. Hence, identifying the common components in the variation of the state variables could possibly provide us with a parsimonious representation of the sources of macroeconomic

[23]The relative bill rate is calculated as the difference between the one-month T-bill rate and its 1-year backward moving average.

risk that asset prices are subject to. Let us still go a bit further. Assume there are other types of macroeconomic risk that investors hedge against, but which have not yet been pinpointed by empiricism or theory. Then, if we were able to find a *priced* factor that has been constructed to capture some of the common variation of macroeconomic variables, we could use the correlation patterns of the common factor with the individual variables to identify those which contribute the most to the priced variation.

However, the first building block we need before we can pursue this idea is a suitable method to identify the common variation in macroeconomic variables. One such method has recently been proposed by Stock and Watson (1998). From a large number of variables they construct a small number of common factors using principal components analysis. They show that forecasts of output and inflation based on these factors outperform univariate autoregressions, small vector autoregressions, and leading indicator models. Given the apparent predictive power of the Stock-Watson factors for macroeconomic variables, I conjecture that they are potentially promising candidates for the summary measures of priced risk we are searching for. The central issue of this paper is to investigate whether this conjecture can be confirmed empirically. Before I document my empirical findings in section 5, I will present Stock and Watson's method in the remainder of this section.

4.1 The Model

As a motivation for the development of the diffusion index methodology, Stock and Watson point out that the information technology nowadays enables the access to a very large number of macroeconomic time series while forecasting still is mostly based on only a small number of variables. Thus, potentially valuable information is neglected. To overcome this deficiency, the authors develop a theoretical framework for forecasting single time series using a large number of predictors. They get their inspiration from a common practice of NBER business cycle analysts who construct weighted averages of many time series with fixed weighting schemes which they call *diffusion indexes*. Stock and Watson argue that since these diffusion indexes are averages of a large number of variables, they summarize the common information contained in these by cross-sectionally averaging out idiosyncratic disturbances. According to this characteristic, they propose to estimate diffusion indexes as the unobserved factors in an approximate dynamic factor model. In a dynamic factor model setting, factor loadings are allowed to vary over time. This ensures that correlation patterns between individual variables be time-varying as one observes to be the case in reality. Further, Stock and Watson permit correlation of idiosyncratic errors across series, which implies an approximate factor structure. Finally, they extend their model to be applicable to unbalanced panels, i.e. data sets that exhibit missing values or time series of different frequency. Stock and Watson show that when both the number of observations in the time series and the number of times series in the cross-section tend to infinity, then the forecasts based on the estimated dynamic factors are as efficient asymptotically as if the true factors were observed. Stock and Watson's theory builds on previous work on dynamic factor models (e.g. Sargent and Sims (1977), Geweke (1977)) as well as on approximate static factor models, mostly studied for financial applications (e.g. Chamberlain and Rothschild (1983), Connor and Korajczyk (1986, 1988, 1993)).

Consider the scalar times series variable y_t as the variable to be forecasted and X_t the N-dimensional multiple time series variable containing useful information for the prediction of y_{t+1}.[24] Further assume that X_t can be represented by the factor structure,

$$X_t = \Lambda_t F_t + e_t \tag{4.1}$$

where F_t is the $r \times 1$ vector of common factors, Λ_t is the $N \times r$ matrix of factor loadings, and e_t is the $N \times 1$ vector of idiosyncratic disturbances. We are interested in predicting y_{t+1} with the information contained in X_t. To reduce the number of predictors, we make use of the factor structure of X_t, and model y_{t+1} as,

$$y_{t+1} = \beta_t' F_t + \epsilon_{t+1} \tag{4.2}$$

where β_t is the $r \times 1$ vector of coefficients and ϵ_{t+1} is the scalar error term. We now assume that $E[\epsilon_{t+1}|X_t, y_t, \beta_t, X_{t-1}, y_{t-1}, \beta_{t-1}, \ldots] = 0$. This assumption is central to the model since it implies that $E[y_{t+1}|X_t, y_t, \beta_t, X_{t-1}, y_{t-1}, \beta_{t-1}, \ldots]$ only depends on F_t and not on X_t hence ensuring the dimension reduction from a large number of predictor time series variables (N) to a smaller number of factors (r). Further, the above assumption implies that no lags of F_t or y_t enter the forecasting equation. This implication can be relaxed easily by interpreting F_t as including lags (stacking) and y_{t+1} as a quasidifference thus also including lagged values of y_t.

Stock and Watson assume that the factor loadings, Λ_t, and their coefficients in the forecasting equation, β_t, vary over time according to,

$$\Lambda_t = \Lambda_{t-1} + H\xi_t \tag{4.3}$$

and

$$\beta_t = \beta_{t-1} + \eta_t \tag{4.4}$$

where ξ_t and η_t denote a $N \times r$ matrix and a $r \times 1$ vector of stochastic disturbances, and where H is a $N \times N$ scaling matrix. Notice that specific assumptions have to be made about H when deriving the asymptotics for the estimators. Since I will not discuss thoroughly the proofs of consistency, the reader is referred to Stock and Watson's paper for details on this.

It is worth noting at this point that the model (4.1)-(4.4) is more general than the dynamic factor models without time-variation of Geweke (1977), Sargent and Sims (1977), and others, as well as the static factor models proposed by Chamberlain and Rothschild (1983), and Connor and Korajczik (1986, 1993). Both sets of models can be obtained as special cases of (4.1)-(4.4). In particular, a static factor model representation of (4.1)-(4.4) can be derived assuming that there is no time variation in the factor loadings ($\Lambda_t = \Lambda_0 \; \forall \, t$), that e_t are serially uncorrelated, and that F_t and $\{e_{it}\}$ are i.i.d. and mutually uncorrelated. If there is no correlation of idiosyncratic disturbances across series ($e_{it} \perp e_{jt} \; \forall \, i \neq j$), the model is referred to as an exact static factor model. If the idiosyncratic disturbances are allowed to be slightly cross-correlated, then the model is called an approximate factor model.

[24] Except for some notational changes and omissions of some minor steps, I strictly follow the derivations in Stock and Watson's (1998) seminal work.

In standard dynamic factor models without time-variation, the factor loadings are similarly supposed to be constant. However, dynamics are introduced by assuming that (i) the factors evolve according to a time series process, (ii) lagged values of the factors are allowed to enter the model, and (iii) the idiosyncratic error terms can be serially correlated. Stock and Watson suggest two ways of transforming a dynamic factor model into a static one without losing too much information. The first transformation is based on the assumption that only a finite number q of lagged factors enter the model. Then, by stacking the contemporaneous factor with its first q lags, the factor loadings can be represented by a matrix of constants. Specifically, let Z_t denote a $N \times 1$ vector of time series variables that follow the dynamic factor model,

$$Z_{ti} = \alpha_i(L) f_t + v_{ti} \tag{4.5}$$

where $g_i(L)v_{ti} = \eta_{ti}$, η_{ti} $i.i.d.N(0, \phi_i^2)$, f_t is a vector of factors that are independent of $\{v_{ti}\}$, $\alpha_i(L)$ has finite order q, and $g_i(L)$ is a finite order lag polynomial with all roots outside the unit circle. Then, letting $X_t = Z_t$, stacking f_t and its q lags into a $(dim(f_t) \cdot (q+1) \times 1)$ vector $F_t = (f_t', f_{t-1}', \ldots, f_{t-q}')'$, further letting $\Lambda_0 = (\alpha_0, \alpha_1, \ldots, \alpha_q)$ and $e_t = v_t$, one can easily see that model (4.5) is equivalent to model (4.1) where $r = dim(f_t) \cdot (q+1)$, $\Lambda_t = \Lambda_0 \;\forall\; t$ and $e_t = v_t$ being serially correlated. That is, under the assumption that there is only a finite number q of factors entering the dynamic model, one can stack them to obtain a static model with constant coefficients. This is the approach adopted by Stock and Watson. The second way to rewrite the dynamic factor model in a static form consists in stacking the data *and* the factors and then again estimating the factor loadings as constant coefficients. This representation has the advantage that more information is included since also lagged values of the explanatory variables are considered. However, for their consistency proofs and forecasts, Stock and Watson only use contemporaneous values of X.

As Stock and Watson notice, dynamic factor models are usually estimated by maximum likelihood using the Kalman filter. However, they point out that the use of the Kalman filter is not appropriate for the estimation of their model since the number of parameters to estimate is large, missing data have to be handled, and nonlinear filters would be needed for computing the likelihood when both factors and factor loadings were treated as random. To avoid these problems, Stock and Watson suggest the following approach towards the estimation of the factors. They start with some restrictive assumptions. In particular, they suppose that (i) the factor loadings are constant ($\Lambda_t = \Lambda_0 \forall\; t$), (ii) the idiosyncratic disturbances $\{e_{it}\}$ are i.i.d.$N(0, \sigma_e^2)$ and cross-sectionally uncorrelated, and (iii) $\{F_t\}$ is a $T \times r$ dimensional unknown non-random parameter. Notice that the last two assumptions represent considerable restrictions compared to the assumptions made in standard dynamic factor models. However, Stock and Watson prove the consistency of the estimated factors under weaker nonparametric assumptions. It is in this sense that they classify their method as quasi-maximum likelihood.

Before we derive the factor estimates, it is convenient to introduce some additional notation. In the following, X_{ti} denotes the observation on variable i at time t, the i-th time series variable is given by the $T \times 1$ vector $X_i = (X_{1i}, X_{2i}, \ldots, X_{Ti})'$, and correspondingly, X is the $T \times N$ matrix of all observations in the data set. Equivalently, let F_{ti} be the observation on factor i at time t and $F = (F_1, F_2, \dots, F_T)'$ the $T \times r$ matrix of factors. Let F^0 denote the true value of F. Further, let $\Lambda_t = (\lambda_{1t}, \lambda_{2t}, \dots, \lambda_{Nt})'$ be the $N \times r$ matrix of factor loadings where λ_{it} is the $r \times 1$ vector of factor loadings on variable i at time t. Finally, let I_{ti} denote a nonrandom indicator function where $I_{ti} = 1$ if the i-th variable is observed at time t and $I_{ti} = 0$ otherwise. This indicator function is needed to address the problem of missing data in an unbalanced panel.

4.2 Derivation of the Estimator

Given this notation and the assumptions made above, Stock and Watson obtain the factor and factor loading estimates (F, Λ_0) by solving the nonlinear least squares problem

$$V_{NT}(F, \Lambda_0) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} I_{ti}(X_{ti} - \lambda'_{i0} F_t)^2 \tag{4.6}$$

which is implied by the maximization of the likelihood function. Solving the first order conditions of this problem for the minimizers $(\widetilde{F}, \widetilde{\Lambda}_0)$ of $V_{NT}(F, \Lambda_0)$ yields

$$\widetilde{\lambda}_{i0} = \left[\sum_{t=1}^{T} I_{ti} \widetilde{F}_t \widetilde{F}'_t \right]^{-1} \left[\sum_{t=1}^{T} I_{ti} \widetilde{F}_t X_{ti} \right], \tag{4.7}$$

and

$$\widetilde{F}_t = \left[\sum_{i=1}^{N} I_{ti} \widetilde{\lambda}_{i0} \widetilde{\lambda}'_{i0} \right]^{-1} \left[\sum_{i=1}^{N} I_{ti} \widetilde{\lambda}_{i0} X_{ti} \right]. \tag{4.8}$$

From equations (4.7) and (4.8) it is obvious that balanced and unbalanced panels require different treatment. Let us first consider the case of balanced panels, i.e. $I_{ti} = 1 \ \forall \ i \ \forall \ t$. One way of estimating the parameters consists in substituting (4.7) into (4.6) to obtain the concentrated objective function,

$$V_{NT}(F, \widetilde{\Lambda}_0) = \frac{1}{NT} \sum_{i=1}^{N} \left[X'_i X_i - X'_i P_F X_i \right], \tag{4.9}$$

where P_F denotes the projection matrix $F(F'F)^{-1}F'$.[25] Normalizing all factors F to be mutually uncorrelated and to have standard deviation of unity, $F'F/T = I_k$, it is easy to see that the minimization of (4.9) is equivalent to the maximization of $\frac{1}{N} \sum_{i=1}^{N} X'_i F F' X_i$. This expression being a scalar, one can apply the trace operator to get

$$\frac{1}{N} \sum_{i=1}^{N} X'_i F F' X_i = tr[F'(\frac{1}{N} \sum_{i=1}^{N} X_i X'_i) F] = \frac{1}{N} tr(F'XX'F),$$

the maximum of which is obtained by choosing $\widetilde{F}$ as the eigenvectors corresponding to the k largest eigenvalues of the $T \times T$ matrix XX'.[26] Obviously, the matrix XX' is the

[25] Dropping tildes for presentational convenience, this can be derived easily by plugging the matrix representation of (4.7), $\lambda_{i0} = (F'F)^{-1} F' X_i$, into (4.6), multiplying terms out, and again employing the matrix notations $F'X_i = \sum_{t=1}^{T} F_t X'_{ti}$, $X'_i F = \sum_{t=1}^{T} X_{ti} F'_t$, $F'F = \sum_{t=1}^{T} F_t F'_t$, and $X'_i X_i = \sum_{t=1}^{T} X_{ti} X'_{ti}$.

[26] This is a standard result of matrix algebra. See, for example, theorem 3.6 in Härdle and Simar (1998).

cross-sectional covariance matrix of all observations X_{ti}, i.e. its elements represent the variation and covariation across series in each period. As Stock and Watson note, a similar approach has been applied by Connor and Korajczyk (1986,1993) in their studies of the APT.

Notice that one could also maximize $V_{NT}(F, \Lambda_0)$ by substituting the first-order condition (4.8) in equation (4.6). Equivalently to the above derivations, the maximum of the concentrated objective function $V_{NT}(\widetilde{F}, \Lambda_0)$ would be obtained by choosing $\widetilde{\Lambda}_0$ as the first k principal components of X, i.e. the first k eigenvectors of the $N \times N$ variance-covariance matrix $X'X = \sum_{t=1}^{T} X_t X_t'$.

Matters are more complicated if we have an unbalanced panel. Stock and Watson note that iterating on the first-order conditions (4.7) and (4.8) subject to the normalization $F'F/T = I_k$ would in principle be possible, but computationally burdensome in the case of large N. They thus suggest to minimize $V_{NT}(F, \Lambda_0)$ by using the EM algorithm.[27] They denote X_{ti}^* as the latent value of X_{ti}. Hence, $X_{ti}^* = \lambda_{i0}' F_t + e_{ti}$ and $X_{ti} = X_{ti}^*$ if X_{ti} is observed, i.e. if $I_{ti} = 1$. The complete-data likelihood is thus given by

$$V_{NT}^*(F, \Lambda_0) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{ti}^* - \lambda_{i0}' F_t)^2.$$

Let $Q_{NT}(F^{(j)}, \Lambda_0^{(j)})$ denote the expected complete-data likelihood in iteration j,

$$Q_{NT}(F^{(j)}, \Lambda_0^{(j)}) = E\left[V_{NT}^*(F^{(j)}, \Lambda_0^{(j)}) | X, \widetilde{F}^{(j-1)}, \widetilde{\Lambda}_0^{(j-1)}\right],$$

where $F^{(j)}$ and $\Lambda_0^{(j)}$ denote the j-th iterates of F and Λ_0, respectively. Then, under the assumption that e_{ti} is i.i.d.$N(0, \sigma^2)$,

$$Q_{NT}(F^{(j)}, \Lambda_0^{(j)}) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\widehat{X}_{ti}^{*(j-1)} - (\lambda_{i0}^{(j)})' F_t^{(j)}\right)^2, \tag{4.10}$$

where terms that do not depend on F or Λ_0 have been omitted and where

$$\widehat{X}_{ti}^{*(j-1)} = E[X_{ti}^* | X, \widetilde{F}^{(j-1)}, \widetilde{\Lambda}_0^{(j-1)}] = \left\{ \begin{array}{ll} X_{ti} & \text{if } I_{ti} = 1 \\ (\widetilde{\Lambda}_{i0}^{(j-1)})' \widetilde{F}_t^{(j-1)} & \text{if } I_{ti} = 0 \end{array} \right\}.$$

Given this representation, one can again "concentrate" the likelihood by substituting the first-order condition resulting from the maximization of $Q_{NT}(F^{(j)}, \Lambda_0^{(j)})$ into (4.10). Accordingly, one obtains $\widetilde{F}^{(j)}$ as the eigenvectors of the cross-sectional covariance matrix

$$(\widehat{X}^{*(j-1)})(\widehat{X}^{*(j-1)})' = \sum_{i=1}^{N} (\widehat{X}_i^{*(j-1)})(\widehat{X}_i^{*(j-1)})',$$

[27] The expectation-maximization (EM) algorithm was first developed by Dempster, Laird, and Rubin (1977). It is an iterative method for finding the likelihood-maximizing parameters of a model by introducing some appropriate unobserved latent variable. First, the complete-data log likelihood is estimated (E-step). Second, this expected value is maximized (M-step). Both steps are repeated until convergence.

where $(\widehat{X}^{*(j-1)})$ is the $T \times N$ matrix of all $\widehat{X}_{ti}^{*(j-1)}$, i.e. the observed and the data points calculated on the basis of $\widetilde{\Lambda}_0^{(j-1)}$ and $\widetilde{F}^{(j-1)}$. Finally, the two steps (computation of the expected complete-sample likelihood and its maximization) are iterated until convergence. Notice that this algorithm can also be applied if there are series of different frequency in the sample. This makes it applicable to very large data sets and thus particularly useful for forecasts based on a large number of predictors.

4.3 Consistency Proofs

Stock and Watson exclusively provide consistency proofs for the balanced panel estimates. The authors further limit their proofs to the numerator matrix $\widehat{F}_t$ of the factor estimates given in (4.8),[28] i.e.

$$\widehat{F}_t = \frac{\sum_{i=1}^{N} I_{ti}\widetilde{\lambda}_{i0}X_{ti}}{\sum_{i=1}^{N} I_{ti}}. \tag{4.11}$$

They first show that $\widehat{F}_t$ is a uniformly consistent estimator for a linear combination of the true factors F_t^0. As they remark, the consistency is obtained by averaging over a very large number N of cross-sectional observations, relative to the number T of time series observations. Since forecasting time series variables based on the common variation of many predictors has been the main objective for Stock and Watson to develop their method, they show in a second step that forecasts of y_{t+1} based on the estimated factors $\widehat{F}_t$ are uniformly consistent for forecasts of y_{t+1} based on the true factors F_t^0. Concerning this result, two remarks are in order, however. First, Stock and Watson exclusively consider the case of time-invariant coefficients in the forecasting equation (4.2), i.e. $\beta_t = \beta \ \forall \ t$. Second, since the true number of factors r is unknown in general, they estimate the number q $(1 \leq q \leq k)$ of factors which enter the forecasting equation by using an information criterion of the form,

$$IC_q = \ln(\hat{\sigma}_\epsilon^2(q)) + g(T)q,$$

where $g(T)$ is some penalty function, and where $\hat{\sigma}_\epsilon^2(q) = SSR(q)/T$ is the squared error of forecasting y_{t+1} with q factors. Then, they show that for any $g(T)$ satisfying $g(T) \to 0$ and $\delta_{NT}g(T) \to \infty$, where δ_{NT} denotes the rate at which the estimated converge to the true factors, the information criterion estimate $\hat{r}$ of the number of factors converges in probability to the true number of factors. In a nutshell, Stock and Watson demonstrate that the efficient forecast of the forecasting variable given the true factors, $y_{t+1} = \beta' F_t^0$, can be achieved asymptotically (in L^2) by using the estimated factors, even if the number of factors has itself also been estimated.[29]

[28] Although Stock and Watson do not go into that in more detail, the reason why the denominator can be neglected seems to be the following. In matrix notation, the first-order conditions (4.7) and (4.8) are given by $\widetilde{\lambda}_{i0} = (\widetilde{F}'\widetilde{F})^{-1}\widetilde{F}'X_i$ and $\widetilde{F} = (\widetilde{\Lambda}_0'\widetilde{\Lambda}_0)^{-1}\widetilde{\Lambda}_0'X_i$. Taking into account the normalization of the factors, $F'F/T = I_k$, (4.7) reduces to $\widetilde{\lambda}_{i0} = T \cdot \widetilde{F}'X_i$. Then, $\widetilde{\lambda}_{i0}'\widetilde{\lambda}_{i0} = \frac{1}{T^2}X_i'\widetilde{F}\widetilde{F}'X_i$. Now, again using the normalization of the factors, it is easy to see that the denominator matrix in (4.8), $\widetilde{\Lambda}_0'\widetilde{\Lambda}_0 = \sum_{i=1}^{N}\widetilde{\lambda}_{i0}'\widetilde{\lambda}_{i0}$, is not stochastic and can thus be ignored for the derivation of the consistency proofs.

[29] Stock and Watson notice, however, that the convergence criterion $\delta_{NT}g(T) \to \infty$ for the penalty function deviates from the usual information criterion condition $Tg(T) \to \infty$. Since neither the Akaïke (AIC) nor the Bayesian information criterion (BIC) satisfy the condition $\delta_{NT}g(T) \to \infty$, they apply the penalty function $g(T) = \omega \ln T/\delta_{NT}$ which obviously does satisfy it. Here, ω denotes some constant and $\delta_{NT} = \min(N^{1/2}/T^{1+\epsilon}, T^{1-\epsilon})$ is the rate of convergence derived in the proof of consistency.

4.4 Monte Carlo Analysis Results

In order to verify whether their method really delivers factor estimates that are good forecasters in a finite sample setting, Stock and Watson perform a Monte Carlo analysis. Specifically, they test for three properties. First, for different values of $T, N, r,$ and k, they investigate if the estimated factors are good proxies for the true factors. Second, assuming that the true number of factors is known, they analyze whether the forecasting errors significantly rise if one uses the estimated rather than the true factors. Third, they seek to quantify the additional error that results from forecasting y_{t+1} with the estimated number $\hat{r}$ of factors compared to forecasts where the true number of factors, r, is known. Since the finite-sample performance of the SW factors in forecasting individual macroeconomic time series variables is not the focus of my paper, I omit discussing the design of Stock and Watson's Monte Carlo analysis here. Let me nonetheless briefly sketch their main results.

First, even for relatively small N and T, the estimated are quite close to the true factors, as measured by the coefficient of determination in a multivariate regression of $\widehat{F}$ on F^0, denoted $R^2_{\widehat{F},F^0}$. Increasing both the number of time series observations and the number of variables in the cross-section significantly improves the results, obtaining values of $R^2_{\widehat{F},F^0}$ close to 1 for $T \geq 100$ and $N \geq 250$. Also, they show that forecasts based on k rather than r factors do not considerably decrease the coefficient of determination. One important result is further that $R^2_{\widehat{F},F^0}$ remains high when the true model is dynamic but the factors are extracted from the stacked static model as described above. However, Stock and Watson notice that in the case of highly autocorrelated factors the Monte Carlo results deteriorate slightly. A comparison across specifications shows that the worst results are obtained in the case of time-varying factor loadings. With values of $R^2_{\widehat{F},F^0}$ in the range from 0.83 to 0.87, the estimated factors nonetheless seem to be quite good proxies for the true factors even in this case. Stock and Watson further report that forecasts based on the estimated factors in general do not differ significantly from forecasts based on the true factors. However, the results deteriorate somewhat when dynamics and time-variation in the factor loadings are introduced. Finally, forecasts based on a number $\hat{r}$ of factors determined by different information criteria are compared with forecasts based on the true number of factors. It turns out that except for the case of large ω in Stock and Watson's information criterion, all three criteria yield broadly similar results.

4.5 Forecasting Results

Having shown that the factors extracted from their algorithm are relatively good proxies for the true factors, Stock and Watson in a final step compare their forecasting power with that of standard forecasting methods. In particular, they contrast diffusion index forecasts of annual growth rates of US industrial production (IP) and consumer price index (CPI) with simple autoregressive forecasts, multivariate leading indicator forecasts, and with Phillips curve forecasts in the case of CPI. Since equation (4.2) allows for serial correlation in the error terms, Stock and Watson consider a forecasting equation that also includes lagged values of the variable to be forecasted,

$$\ln(z_{t+12}/z_t) = \beta_0 + \sum_{i=1}^{q} \beta_i \widehat{F}_{ti} + \sum_{j=1}^{p} \gamma_j \Delta \ln z_{t-j} + \epsilon_t,$$

the coefficients of which are estimated by OLS. Different specifications for the selection of the number of lags to include, p and q, are examined. Further, to simulate real-time forecasting, all three steps, i.e. factor extraction, model selection, and forecasting, are repeated at each month in the sample period so that the forecasting equation possibly changes from period to period in terms of how many factors and how many lags are included. Using the different estimation methodologies discussed below, Stock and Watson extract the factors from two sets of variables, a balanced panel containing 170 monthly macroeconomic time series for the US covering the period 1960:1 - 1997:9, and an unbalanced panel containing an additional set of 54 time series variables that have not been observed over the entire period. The variables are selected by the authors from the DRI/MacGraw Hill Basic Economics database to represent the main categories of macroeconomic times series.[30] It is important to notice that since their method is not appropriate to extract factors from integrated variables,[31] Stock and Watson preliminarily transform all non-stationary to stationary series by taking first differences or logarithms or both. Further, after this transformation, they standardize all series to have mean zero and unit variance.

Compared to multivariate leading indicator forecasts involving all or a subset of eleven variables that are known to have some forecasting power for IP, the diffusion index forecasts of industrial production exhibit considerably smaller mean squared errors. Interestingly, Stock and Watson find that the forecasts are not improved by adding lagged values of IP to the forecasting equation. That is, all information contained in past values of IP growth and potentially useful for forecasts of future IP growth, is captured by the diffusion indexes. This last observation does not hold for the inflation forecasts which the authors show to be improved by including lagged values of CPI growth. Stock and Watson report another two interesting differences concerning the results of the diffusion index forecasts of IP and CPI. First, in forecasts of IP, the first two factors capture almost all of the useful information while forecasts of CPI have minimal MSE when five to six factors are included. Second, the IP forecasts based on a number of factors selected by BIC always outperform the fixed-k forecasts, whereas in the case of CPI, the minimum MSE fixed-k forecast always performs better than the BIC-selected. Finally, the Phillips-curve model forecasts of CPI obtain results similar to the multivariate leading indicator forecasts (including all or a subset of eight leading indicators that have proved useful for forecasting inflation in the past) and are thus also outperformed by the diffusion index forecasts.

In summary, the diffusion index forecasts of US industrial production and inflation clearly outperform benchmark methods in a real-time forecasting exercise. Stock and Watson further find that the first six factors account for more than 40 % of the variation

[30] Since I use the data and programs provided by Mark Watson on his homepage, a more detailed discussion of a closely related data set is given in section 5.

[31] Since $X_t = \Lambda_0 F_t + e_t$, this is a trivial implication of the fact that the factors are normalized to be covariance-stationary, $F'F/T = I_k$, and that the idiosyncratic error terms are assumed to be normally distributed, $e_{ti} \sim i.i.d.N(0, \sigma_e^2)$.

of all series in the panel (see appendix B). These results indicate that the diffusion indexes incorporate an important amount of valuable economic information. This is in line with the way they are derived: the common dynamic factors in a large number of time series are filtered out by minimizing the influence of idiosyncratic disturbances. Stock return variations have been shown to be linked to the business cycle. Used as factors in pricing models, a number of individual macroeconomic variables have been identified which seem to account for the systematic risk underlying these variations. Since they summarize the common variation in a large number of economic time series, the Stock-Watson diffusion indexes might similarly capture sources of priced risk. Whether this is indeed the case can be analyzed by testing their usefulness as factors in a pricing model.

4.6 The Diffusion Index Pricing Model (DIPM)

I employ diffusion indexes, extracted from a large panel of economic time series using the methodology of Stock and Watson, as state variables in a factor pricing model. The information criterion suggested by Stock and Watson exclusively applies to forecasting exercises and is thus not well-suited for my empirical investigation. Consequently, I start without making any a priori assumption about the number of factors that should enter the model and compute a total of twelve factors from both the balanced and the unbalanced panel.[32] Then, I first estimate the diffusion index pricing model with the market return and the twelve indexes from the balanced and unbalanced panel, respectively, as risk factors. Letting the twelve factors denote $f_1, f_2, \ldots, f_{12}$, the diffusion index pricing model has the expected return-beta representation,

$$E[R_i] = E[R_0] + \beta_{i,R^m}\lambda_{R^m} + \beta_{i,f_1}\lambda_{f_1} + \beta_{i,f_2}\lambda_{f_2} + \ldots + \beta_{i,f_{12}}\lambda_{f_{12}}, \tag{4.12}$$

where β_{i,R^m} is the coefficient in a time series regression of R_i on the market return, R_m, and where β_{i,f_j} denotes the coefficient in a linear regression of R_i on the j-th diffusion index. I first report estimation results for models including twelve unbalanced and twelve balanced panel diffusion indexes, respectively. From both models, I then choose the factors that enter significantly, and re-estimate the diffusion index pricing model with the respective set of significant factors.

4.7 Diffusion Indexes as Conditioning Variables

Investors price assets according to their expectation of discounted future payoffs conditional on their time t information set. However, since we do not know the information sets of investors in general, we have to model their expectations unconditionally. We have seen in section 2.4 that unconditional models only have a testable expected return-beta representation if the relationship between the discount factor and the state variables is assumed to be time-invariant. Yet, there is a great deal of empirical evidence suggesting that risk prices fluctuate over the business cycle which implies that the relationship is in fact time-varying. An intuitive explanation for this finding is that investors are likely

[32] The decision to extract twelve factors has been judgmentally made. Stock and Watson show that the first twelve factors capture about 50 % of the variation in all the series in the panel, with small marginal contribution of any additional factor. I thus make the underlying assumption that a total of twelve factors suffices to summarize the information that is most important for the pricing of stocks.

to be more risk averse in "bad" times than in "good" times. We have further shown in section 2.4 that a convenient method of accounting for the time-variability of parameters in unconditional factor models consists in scaling the factors. If risk prices vary over the business cycle, we should try to capture this effect by using some business-cycle related quantity as a conditioning variable. The Stock-Watson factors are cross-sectional averages of a large number of time series and thus summarize economic information. Hence, they are closely related to the business cycle and therefore might be eligible candidates for the conditioning variable we seek. To test this hypothesis, I investigate the performance of diffusion indexes as instruments in a scaled factor pricing model. Adopting the conditional consumption CAPM setting of Lettau and Ludvigson (2001b), I compare the usefulness of other macroeconomic variables with that of diffusion indexes by running cross-sectional regressions based on the expected return-beta model,

$$E[R^i] = E[R_0] + \beta_{i,z}\lambda_z + \beta_{i,\Delta c}\lambda_{\Delta c} + \beta_{i,z\Delta c}\lambda_{z\Delta c}, \tag{4.13}$$

where $\beta_{i,z}$, $\beta_{i,\Delta c}$, and $\beta_{i,z\Delta c}$ are the estimated coefficients in time series regressions of R^i on z_{-1}, Δc, and $z_{-1} \cdot \Delta c$, respectively. z_{-1} denotes the one-period lagged value of the scaling variable. To see why we have to use z_{-1}, recall that according to the derivation of scaled factor models in section 2.4, the instrumental variable has to be included in the time t information set of investors whereas the returns and pricing factors are one-step ahead expectations. In order to compare the performance of diffusion indexes and standard macroeconomic variables as scaling variables, I carry out tests of the above model using different instruments. A description of the set of benchmark scaling variables that I have chosen for this purpose is given in section 5.9. Alternatively, controlling for a particular conditioning variable, I also test specifications where two instruments jointly enter a conditional factor pricing model. Thus, I run cross-sectional regressions of the form

$$E[R^i] = E[R_0] + \beta_{i,f_j}\lambda_{f_j} + \beta_{i,f_k}\lambda_{f_k} + \beta_{i,\Delta c}\lambda_{\Delta c} + \beta_{i,f_j\Delta c}\lambda_{f_j\Delta c} + \beta_{i,f_k\Delta c}\lambda_{f_k\Delta c}, \tag{4.14}$$

where β_{i,f_j}, β_{i,f_k}, $\beta_{i,\Delta c}$, $\beta_{i,f_j\Delta c}$, and $\beta_{i,f_k\Delta c}$ are the estimated coefficients in first-stage time series regressions of R^i on $f_{j,-1}$, $f_{k,-1}$, Δc, $f_{j,-1} \cdot \Delta c$, and $f_{k,-1} \cdot \Delta c$, respectively. Using this specification, one can draw conclusions about the relative performance of f_j and f_k as conditioning variables by comparing the significance levels of $\lambda_{f_j\Delta c}$ and $\lambda_{f_k\Delta c}$.

5 Econometric Methodology and Empirical Results

In this section I will present the results of my empirical investigation. I have tested whether diffusion indexes, extracted from a large data set using the procedure proposed by Stock and Watson (1998), can sensibly be used as factors in a pricing model and/or as instruments in a conditional model. I begin with a discussion of the econometric methodology employed and of the model comparison statistics reported. A brief description of the data follows, before I finally turn to the main findings.

5.1 The Cross-Sectional Regression Approach of Fama and MacBeth

Most of the factor pricing models discussed in section 3 have originated from the objective to describe the cross-sectional variation of average returns better than the CAPM, and thus to draw a more precise picture of the risk structure that causes these variations. Since the size effect and the value effect have been the first two CAPM anomalies documented in the literature, a great deal of empirical asset pricing research has focused on their explanation during the last two decades. Moreover, as the search for models that parsimoniously describe the return patterns associated with these two effects has not yet been accomplished, it remains a common practice to assess the performance of asset pricing models with respect to their usefulness in explaining the cross-section of size and book-to-market sorted portfolios. I follow this tradition and correspondingly investigate whether the diffusion indexes of Stock and Watson capture sources of systematic risk that affect the distribution of returns across size and book-to-market sorted assets.

There are several estimation strategies available for such purposes, and an exhaustive analysis of their respective merits is certainly beyond the scope of this paper. Chen, Roll, and Ross (1986), Fama and French (1992), Lettau and Ludvigson (2001b) and others apply the cross-sectional regression procedure of Fama and MacBeth (1973). In order to stay in line methodologically with the benchmark models, I follow these authors and likewise perform Fama-MacBeth cross-sectional regressions. Since it is intuitively appealing and easy to implement, the method is an often used tool in empirical tests of asset pricing models. However, the Fama-MacBeth algorithm exhibits some important drawbacks which I will partially examine in sections 5.4 to 5.9. I will now briefly sketch the main ideas behind the Fama-MacBeth approach.

As we have seen in section 2.2, for each factor pricing model that is represented by the two equations $m = a + b'f$ and $1 = E[mR^i]$, there exists an equivalent expected return-beta representation,

$$E[R^i] = \gamma + \lambda' \beta_{i,f},$$

where $\beta_{i,f}$ is the vector of time series regression coefficients of R^i on the factors f, measuring the exposure of R^i to risk associated with f, and where λ is a coefficient referred to as the price of risk. Our purpose is to find out whether a certain risk factor is significantly priced. Hence, we would like to know if λ is different from zero. As a method to draw statistically sensible conclusions about λ, Fama and MacBeth propose the following two-step procedure.

First, the beta risk-factors are estimated by simple OLS time series regressions,

$$R_t^i = a_i + f'\beta_{i,f} + \epsilon_t^i \qquad\qquad i = 1, 2, \ldots, N, \tag{5.1}$$

i.e. for each individual asset in the cross-section one estimates the covariation with the vector of pricing factors.[33] Then, these beta estimates are used in the cross-sectional regression,

$$R_t = \gamma_t + \hat{\mathbf{B}}\lambda_t + \alpha_t, \tag{5.2}$$

which is estimated at each time period. R_t denotes the $N \times 1$ vector of returns on all assets in the cross-section at date t, γ_t is a constant, $\hat{\mathbf{B}}$ is the $N \times k$ matrix of estimated betas, λ_t is the $k \times 1$ vector of factor risk prices to be estimated, and α_t is the $N \times 1$ vector of pricing errors. Fama and MacBeth suggest to estimate λ as the time series average of the T cross-sectional estimates $\hat{\lambda}_t$,

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_t = \frac{1}{T}\sum_{t=1}^{T}\left[(\hat{\mathbf{B}}'\hat{\mathbf{B}})^{-1}\hat{\mathbf{B}}'R_t\right],$$

and to derive sampling errors from the standard deviations of the elements of $\hat{\lambda}$. Consider the j-th element of $\hat{\lambda}$ representing the average estimated price of factor j. Under the assumption that the individual λ estimates are independent, its empirical variance is given by

$$\hat{\Sigma}_\lambda = Var(\hat{\lambda}_j) = \frac{1}{T^2}Var(\sum_{t=1}^{T}\hat{\lambda}_{tj}) = \frac{1}{T}Var(\hat{\lambda}_{tj}) = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{\lambda}_{tj} - \hat{\lambda}_j)^2,$$

and thus its standard deviation is

$$\hat{\sigma}_{\hat{\lambda}_j} = \left(\frac{1}{T^2}\sum_{t=1}^{T}(\hat{\lambda}_{tj} - \hat{\lambda}_j)^2\right)^{1/2}.$$

Given the estimate and its standard error, Fama and MacBeth test whether the j-th factor is significantly priced by computing its t-value,

$$t_{\hat{\lambda}_j} \equiv \frac{\hat{\lambda}_j}{\hat{\sigma}_{\hat{\lambda}_j}} \sim t\ (T-1). \tag{5.3}$$

Of course, the same applies to the regression constant,

$$t_{\hat{\gamma}} \equiv \frac{\hat{\gamma}}{\hat{\sigma}_{\hat{\gamma}}} \sim t\ (T-1),$$

where

$$\hat{\gamma} = \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_t,$$

[33]Fama and MacBeth estimate the betas in 5-year rolling regressions, i.e. they recompute them annually based on data covering the previous 60 months. This is to account for possible delisting of individual assets in the portfolios. However, many authors estimate the betas once over the entire data sample thus implicitly assuming them to be constant.

and

$$\hat{\sigma}_{\hat{\gamma}} = \left(\frac{1}{T^2} \sum_{t=1}^{T} (\hat{\gamma}_t - \hat{\gamma})^2 \right)^{1/2} .$$

Since the betas that are used as explanatory variables in the cross-sectional regressions have previously been estimated, it is obvious that the method of Fama and MacBeth exhibits an errors-in-variables problem. Fama and MacBeth argue that this problem is substantially mitigated by using portfolio returns instead of returns on individual assets, since this increases the precision of the beta-estimates. As they admit, however, this requires that the errors in the individual betas be largely uncorrelated. Under the assumption that given the realization of the factors, the asset returns have a conditionally joint distribution with constant covariance matrix, i.e. assuming conditional homoskedasticity, Shanken (1992) explicitly derives a correction term for the variance estimate that is designed to adjust for the bias due to the errors-in-variables problem. In particular, he shows that

$$\sqrt{T}(\hat{\lambda} - \lambda) \longrightarrow N\left(0, (1 + \hat{\lambda}'\Sigma_f^{-1}\hat{\lambda}) \cdot \hat{\Sigma}_\lambda + \Sigma_f\right),$$

where $\hat{\lambda}$ is the estimated vector of pricing errors, Σ_f denotes the covariance matrix of the pricing factors, and where $\hat{\Sigma}_\lambda$ is the estimated covariance matrix of the Fama-MacBeth risk price estimates. Since many authors report Shanken-adjusted t-statistics, I do also provide them. As we will see, in some cases the corrections affect the t-statistics of individual factors and thus the inference about their significance quite considerably. Although Shanken-adjusted t-statistics are often documented, there is still some disagreement whether this correction is necessary. For example, Jagannathan and Wang (1998) show that when the assumption of conditional homoskedasticity is violated, the standard errors from Fama and MacBeth's method do not necessarily overstate the precision of the risk price estimates. As a consequence of these conflicting results, one correctly would have to test for conditional homoskedasticity of the asset returns before running Fama-MacBeth regressions. However, this exercise being beyond the scope of my study, I follow other authors in taking into account both the Shanken-corrected and the uncorrected Fama-MacBeth standard errors when drawing conclusions about the significance of pricing factors.

The intuition behind the Fama-MacBeth procedure is to split the entire sample into as many as T subsamples, to estimate the pricing relation over the individual subsamples, and then to look at the time series averages and variances of the estimates. Consequently, a factor in a model will be reported as being significantly priced if the associated risk price λ_j on average is different from zero, taken into account its variation over time. Cochrane (2001) shows that in the absence of time variation in the cross-sectional regressors (i.e. the betas in our case), the estimates and standard errors produced by the Fama-MacBeth procedure are identical to those one would obtain when running a pooled OLS regression and adjusting the variance estimate for cross-sectional heteroskedasticity.[34] Hence, the Fama-MacBeth method can be interpreted as an estimation procedure that corrects for correlation of returns across assets, a situation one certainly faces quite often in finance applications.

[34] Pooled OLS regression here means stacking time series and cross-sections of returns and beta estimates and running a single OLS regression.

5.2 Model Comparison Statistics

Although the t-statistics that one obtains from the Fama-MacBeth regressions are potentially good measures of the significance of single factors, they are unsuited for comparing the ability of entire models to explain the cross-section of asset returns. However, since the main purpose of this study is to assess the diffusion index pricing model compared to benchmark factor pricing models, we need some summary measure of the goodness of fit. The first such measure to think of when running regressions is the coefficient of determination, R^2. Fama and MacBeth derive R^2-statistics just as the risk price estimates, i.e. for each cross-sectional regression they compute a value of R^2, then they compute the time series average and the standard deviation of these R^2's and report both.

Jagannathan and Wang (1996) adopt a slightly different approach. Instead of first performing cross-sectional regressions at each point in time and then taking time series averages of the estimated coefficients, they proceed inversely. That is, they first compute time series averages of the cross-section of returns and then run a single cross-sectional OLS regression of the vector of average returns on the betas. To see that the approach of Jagannathan and Wang delivers estimates that are identical to those provided by the Fama-MacBeth procedure, recall that the vector of risk price estimates according to Fama and MacBeth's method is given by

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_t = (\hat{\mathbf{B}}'\hat{\mathbf{B}})^{-1}\hat{\mathbf{B}}'\frac{1}{T}\sum_{t=1}^{T}R_t = (\hat{\mathbf{B}}'\hat{\mathbf{B}})^{-1}\hat{\mathbf{B}}'\bar{R},$$

where $\bar{R}$ is the $N\times 1$ vector of average returns, where $\hat{\mathbf{B}}$ is a $N\times(k+1)$ matrix containing a column of ones and k columns of beta estimates, and where λ is the $(k+1)\times 1$ vector of factor prices to be estimated including the regression constant. Jagannathan and Wang first average the returns, and then perform the one-step cross-sectional regression,

$$\bar{R} = \hat{\mathbf{B}}\lambda + \bar{\alpha},$$

where $\bar{\alpha}$ is a $N\times 1$ vector of pricing errors. Obviously, the OLS estimate of λ is given by,

$$\hat{\lambda} = (\hat{\mathbf{B}}'\hat{\mathbf{B}})^{-1}\hat{\mathbf{B}}'\bar{R},$$

and is thus identical to the Fama-MacBeth estimate. It is this regression from which Jagannathan and Wang report the coefficient of determination,

$$R^2 = 1 - \frac{Var_c(\bar{\alpha}_i)}{Var_c(\bar{R}_i)} = 1 - \frac{\sum_{i=1}^{N}(\bar{\alpha}_i-\bar{\alpha})^2}{\sum_{i=1}^{N}(\bar{R}_i-\bar{R})^2}, \tag{5.4}$$

where Var_c denotes the cross-sectional variance, variables with subscripts and bars over them denote time series averages of an individual cross-sectional unit, and variables with just bars over them are the cross-sectional averages of time series averages. Lettau and Ludvigson (2001b) also report this version of the R^2 as a summary measure of model fit. Following these authors I provide both the R^2 as given in equation (5.4) and its degrees of freedom-adjusted version, $\bar{R}^2$, for each model.

Although R^2 and $\bar{R}^2$ can be used to compare the ability of different models to explain the cross-section of average returns, they do not provide us with an indication of how good a model's predictions are in terms of returns. This information is contained in the residuals one obtains from the cross-sectional Fama-MacBeth regressions. Contrary to the regression estimates of the factor risk prices, it may be somewhat misleading to consider time series means of pricing errors, since large negative and large positive errors could cancel out, hence erroneously suggesting a good model fit. I therefore report as a third summary statistic for each model the *average absolute pricing error*,

$$\mu_{|\hat{\alpha}|} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} |\hat{\alpha}_{ti}|, \tag{5.5}$$

which is a somewhat naive measure of fit, since there is no distributional theory for it and thus cannot be used for statistical inference across samples. However, since I perform model comparisons using the same return data, $\mu_{|\hat{\alpha}|}$ gives a good indication of the explanatory power of one model with respect to another for that particular panel of returns. Moreover, it is an easy to interpret measure of how close a model comes to the true model in terms of returns.

For each period in the sample, the Fama-MacBeth cross-sectional regression method produces an $N \times 1$ vector of residuals,

$$\hat{\alpha}_t = R_t - \hat{R}_t = R_t - \hat{\mathbf{B}}\hat{\lambda}_t = \left(\mathbf{I}_N - \hat{\mathbf{B}}(\hat{\mathbf{B}}'\hat{\mathbf{B}})^{-1}\hat{\mathbf{B}}'\right) R_t.$$

An intuitively appealing way of assessing a model's performance would be to test whether the average pricing errors of all assets in the cross-section are jointly zero. Adapted to the Fama-MacBeth regression framework, Cochrane (2001) suggests a version of the J_T-test to explore this issue. He proposes to compute the sample covariance matrix of the average pricing errors,

$$Cov(\bar{\hat{\alpha}}) = \frac{1}{T^2} \sum_{t=1}^{T} (\hat{\alpha}_t - \bar{\hat{\alpha}})(\hat{\alpha}_t - \bar{\hat{\alpha}})',$$

where $\bar{\hat{\alpha}}$ denotes the $N \times 1$ vector of time series means of the estimated pricing errors,

$$\bar{\hat{\alpha}} = \frac{1}{T} \sum_{t=1}^{T} \hat{\alpha}_t.$$

Assuming that they are mutually independent and normally distributed, Cochrane suggests to test whether all pricing errors are jointly zero using the statistic,

$$J_T = \bar{\hat{\alpha}}'(Cov(\bar{\hat{\alpha}}))^{-1}\bar{\hat{\alpha}} \sim \chi^2(N-k). \tag{5.6}$$

Obviously, the errors-in-variables problem discussed above also applies to the pricing errors of the Fama-MacBeth procedure. As for the t-statistics, one could thus be induced to correct for the bias arising from the fact that the betas have been estimated in the first-stage time series regressions. According to Cochrane (2001), the Shanken-corrected version of the above test statistic, denoted J_T^c, is given by

$$J_T^c = (1 + \hat{\lambda}'\Sigma_f^{-1}\hat{\lambda})\ \bar{\hat{\alpha}}'(Cov(\bar{\hat{\alpha}}))^{-1}\bar{\hat{\alpha}}\ \sim \chi^2(N-k), \tag{5.7}$$

where $\hat{\lambda}$ is the estimated vector of risk prices, and where Σ_f denotes the variance-covariance matrix of the pricing factors. My empirical results show that Shanken's correction term in many cases exceeds unity quite substantially, thus completely changing the order of magnitude of the J_T-statistic. As has been argued above, the adjustment is only warranted if there is conditional homoskedasticity in the return data. Since I do not test whether the returns on the Fama-French portfolios are homoskedastic or not, I am reluctant to draw conclusions based on any of the two test statistics. In addition and perhaps most importantly, the use of the inverse of the variance-covariance matrix of the pricing errors in the J_T-test is a statistically questionable exercise. To see why, note that the pricing errors are of relatively small order and show high cross-sectional correlation so that their variance-covariance matrix is close to singularity in many cases. Hence, the inversion of the variance-covariance matrix of the pricing errors may produce outcomes of the J_T-test statistic that are not useful for statistical inference.[35]

5.3 The Data

At least since the path-breaking work of Fama and French (1992, 1993), it has been a common practice in empirical asset pricing research to asses factor pricing models with respect to their ability to explain the cross-section of returns on size and book-to-market sorted stock portfolios. For matters of simplicity and in order to facilitate model comparisons, many authors use exactly the same return data as Fama and French, provided by Kenneth French on his internet page.[36] I adopt the same strategy and perform model comparisons on the basis of the Fama-French data. In addition to the return data, Kenneth French also provides the average firm size and the average book-to-market ratio for each portfolio. I use these data in my regressions including firm characteristics, described in section 5.8.

The Fama-French portfolios are constructed as follows (see Fama and French (1993)). In June of each year t, all NYSE stocks are sorted independently by their market capitalization (ME) and by their BE/ME ratio, where BE is book common equity at $t-1$, and where ME is market equity at the end of December of $t-1$. From these two NYSE sorts, quintiles and deciles are formed, and the resulting breakpoints are used to allocate *all* NYSE, Amex, and (after 1972) NASDAQ stocks into the five (ten) ME and BE/ME quintiles (deciles). From the intersection of these quintiles (deciles), 25 (100) portfolios are formed, and monthly value-weighted and equally-weighted returns are calculated from July of t to June of $t+1$ when the sorting procedure is repeated based on refreshed values of ME and BE/ME. Hence, four different data sets for size and book-to-market sorted stock portfolio returns result from this procedure, all being available on Ken French's website: value-weighted returns on 25 ME and BE/ME-sorted portfolios, equally-weighted returns on 25 ME and BE/ME-sorted portfolios, value-weighted returns on 100 ME and BE/ME-sorted portfolios, and equally-weighted returns on 100 ME and BE/ME-sorted portfolios.

Although I also report results for the 25 portfolios in some cases, I use the value-weighted returns on 100 ME and BE/ME-sorted portfolios for most of the model com-

[35] Lettau and Ludvigson (2001b) make a similar point referring to studies that have provided evidence for the poor small-sample properties of estimates of the asymptotic variance-covariance matrix of pricing errors.

[36] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ for details.

parisons. Since for some of the pricing factors I only have data from 1963 until 1998, I restrict the sample to the period 1963:7 - 1998:12. Notice that in the intersections of the highest size decile and the three upper BE/ME deciles, the Fama-French data exhibit missing values in a number of subsamples. I therefore exclude these three portfolios from the data set and perform regressions on the remaining 97 portfolios. Finally, in order to demonstrate some interesting differences in the regression results for the benchmark models that arise when quarterly instead of monthly data is used, I also provide results for quarterly data. These require preliminary transformations of the Fama-French portfolio data from monthly to quarterly returns.

The data I employ for the construction of the pricing factors originate from different sources. I use the growth rate of the S&P's common composite stock price index included in the Stock-Watson data set as a proxy for the return on the wealth portfolio, denoted R^m.[37] R^m is the only factor in the standard CAPM, and the first factor in all benchmark models as well as in the different specifications of the diffusion index pricing model, except for the conditional CCAPM. For the implementation of the conditional versions of the consumption CAPM and the human capital CAPM, I use the data for consumption, asset wealth, labor income, and the derived variable $\widehat{cay}$ that Sydney Ludvigson provides on her homepage.[38]

The diffusion indexes are extracted from both an unbalanced panel containing 201 monthly series for the US economy and a balanced panel containing 142 series.[39] The series have been selected by Stock and Watson from the MacGraw Hill database to represent the major categories of economic time series. The sample covers the period 1959:1-1998:12. It has already been mentioned in section 4 that the series have been subject to some prior manipulations. As Stock and Watson (2002) notice, the transformations were the following. From most nonnegative series that were not already in rates or percentages, logarithms were taken. Moreover, most series were first differenced. Then, all series were screened for outliers. Finally, they were demeaned and standardized to have unit variance. A list of all variables in the unbalanced panel is given in appendix B. Mark Watson provides these data as well as GAUSS programs for the extraction of the factors on his homepage.[40] For matters of convenience, I use both the programs and the data

[37] I also performed regressions using the value-weighted return on all NYSE, AMEX, and NASDAQ stocks, provided by Kenneth French on his website. However, my tests of the CAPM using this time series produced values of R^2 that do not correspond to other authors' findings of almost no explanatory power of the market return for the cross-section of size and BE/ME-sorted portfolios.

[38] http://www.econ.nyu.edu/user/ludvigsons/. These are the data used in the empirical tests of the conditional CCAPM and HCCAPM in Lettau and Ludvigson (2001b). The monthly data I use have been provided by Sydney Ludvigson until recently, but now only quarterly and annual series are available from her website. We will see below that both the conditional CCAPM and HCCAPM perform much better using quarterly data. As Martin Lettau pointed out to me, for the derivation of the monthly time series of $\widehat{cay}$, some intrapolation had to be done since the asset wealth variable is only available in quarterly frequency. According to him, this intrapolation diminishes the reliability of the estimate and thus explains the relatively bad results using monthly data.

[39] There are 215 (149) series in the original unbalanced (balanced) panel used by Stock and Watson (2002). However, in order to avoid data-snooping effects, I excluded the stock market data from the data set, from which the smaller total number of series in the panels I use.

[40] See http://www.wws.princeton.edu/ mwatson/. Notice that for the programs to work, one has to increase the workspace memory in "gauss.cfg" beyond the default setting.

made available by Mark Watson for the computation of the diffusion indexes.

I also use the Stock-Watson data as the basis for the construction of the ICAPM factors. ΔIP and ΔCPI are the monthly changes of total industrial production and of the consumer price index, all items, respectively. $Tspr$ is the difference between the 10-year government bond and the federal funds rate, and $Dspr$ is the difference between the Moody's Baa corporate bonds rate and the federal funds rate.

5.4 Testing the Benchmark Models

As noted above, the monthly time series of Lettau and Ludvigson's scaling variable $\widehat{cay}$ has been constructed on the basis of intrapolated data points for the asset wealth variable a which is only available in quarterly frequency. The authors thus caution the use of the monthly time series of $\widehat{cay}$. However, since the diffusion indexes are only available in monthly frequency, I have to rely on the monthly data versions of the the benchmark models for comparisons with the diffusion index pricing model. In order to give some indication of the "true" performance of Lettau and Ludvigson's conditional CCAPM and HCCAPM, I first provide results for Fama-MacBeth regressions of all benchmark models using quarterly data in table 1.[41]

The results are largely consistent with what has been reported in the literature so far. As noted, among others, by Fama and French (1992), the standard CAPM explains literally none of the variation of average returns across stock portfolios sorted by size and book-to-market equity ratio. This is confirmed by the results of the Fama-MacBeth regressions that I obtain for the CAPM: the cross-sectional R^2-statistic only amounts to 1 percent. Figure 1 provides a nice visualization of this failure. In fact, plotting average returns of the 25 Fama-French portfolios against their CAPM predictions, one finds a completely flat relationship suggesting that there is - if at all - only a weak dependence between the portfolios' betas and their expected returns, contrary to what theory tell us. Compared to the other benchmark models, the CAPM further exhibits a considerably larger $\mu_{|\hat{\alpha}|}$ which confirms the finding that the CAPM completely fails in pricing the 25 Fama-French portfolios.

Table 1 further presents estimation results for the ICAPM, showing that in sharp contrast to the CAPM, this model explains more than 90 percent of the variation of average returns across the Fama-French portfolios. In particular, I find that while the default spread and the change in industrial production do not enter significantly, both the term spread and the rate of inflation earn significant risk premiums even after correction for the errors-in-variables bias. Although the latter finding is generally in line with the above presented results of Chen, Roll, and Ross (1986), there is an interesting difference. In fact, while Chen et al. report negative risk prices associated with the term spread, my regression results indicate that these are positive. However, limiting the sample period to the pre-1980's period I do likewise obtain negative regression estimates. This suggests that the relationship has been subject to a structural break. Obviously, negative risk prices are consistent with Merton's ICAPM theory which predicts that investors will demand more

[41] All estimation results documented in this paper have been obtained using a GAUSS program written by the author. This program is available on request.

Table 1: **Benchmark Models - 25 FF portfolios, quarterly data**

The table summarizes the results of Fama-MacBeth regressions that have been performed for the benchmark pricing models using the cross-section of the value-weighted returns on 25 Fama-French size and book-to-market sorted stock portfolios as the variable to be explained. The sample period is 1963:Q3 - 1998:Q3. All models have the form

$$E[R_i] = E[R_0] + \beta_{i,f_1}\lambda_{f_1} + \ldots + \beta_{i,f_k}\lambda_{f_k},$$

where $\beta_{i,f_1}, \ldots, \beta_{i,f_k}$ denote the beta estimates of the pricing factors, preliminarily determined by regressing R_i on the respective vector of factors. For each model, estimates of λ_{f_i}, their t-values and the associated p-value are reported. The "corrected $t-$ and p-values" take into account the errors-in-variables adjustment as suggested by Shanken (1992). R^2 is the coefficient of determination in a single cross-sectional regression of the average returns on the beta estimates, and $\bar{R}^2$ its degrees of freedom-adjusted version. $\mu_{|\hat{\alpha}|}$ is the average of the absolute values of all $T \cdot N$ pricing errors.

Model	Constant	Pricing Factors					Summary Statistics		
CAPM	$E[R_0]$	R^m					R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	3.86	0.03					0.01	-0.02	2.39
t-value	3.77	0.03							
p-value	**0.00**	0.49							
corrected-t	3.77	0.03							
corrected-p	**0.00**	0.49							
ICAPM	$E[R_0]$	R^m	TSpr	DSpr	ΔIP	ΔCPI	R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	4.08	-0.27	1.07	-0.31	0.01	-0.34	0.91	0.89	1.65
t-value	3.85	-0.23	4.25	-1.43	0.03	-1.96			
p-value	**0.00**	0.41	**0.00**	0.08	0.49	**0.03**			
corrected-t	2.31	-0.13	2.46	-0.85	0.01	-1.15			
corrected-p	**0.01**	0.45	**0.01**	0.20	0.49	0.13			
FF3F	$E[R_0]$	R^m	SMB	HML			R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	2.56	0.77	0.49	1.45			0.78	0.76	1.70
t-value	1.88	0.49	0.98	3.22					
p-value	**0.04**	0.31	0.17	**0.00**					
corrected-t	1.77	0.43	0.68	2.24					
corrected-p	**0.05**	0.34	0.25	**0.02**					
c. CCAPM	$E[R_0]$	cay	Δc	$cay \cdot \Delta c$			R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	4.22	-0.11	0.03	0.01			0.70	0.67	1.84
t-value	6.00	-0.36	0.28	3.14					
p-value	**0.00**	0.36	0.39	**0.00**					
corrected-t	4.18	-0.25	0.19	2.12					
corrected-p	**0.00**	0.40	0.43	**0.02**					
c. HCCAPM	$E[R_0]$	cay	R^m	$cay \cdot R^m$	Δy	$cay \cdot \Delta y$	R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	3.48	0.12	0.23	0.07	0.49	0.01	0.76	0.71	1.60
t-value	3.62	0.45	0.20	2.68	2.92	3.25			
p-value	**0.00**	0.33	0.42	**0.01**	**0.00**	**0.00**			
corrected-t	2.53	0.30	0.13	1.83	1.95	2.22			
corrected-p	**0.01**	0.38	0.45	**0.04**	**0.03**	**0.02**			

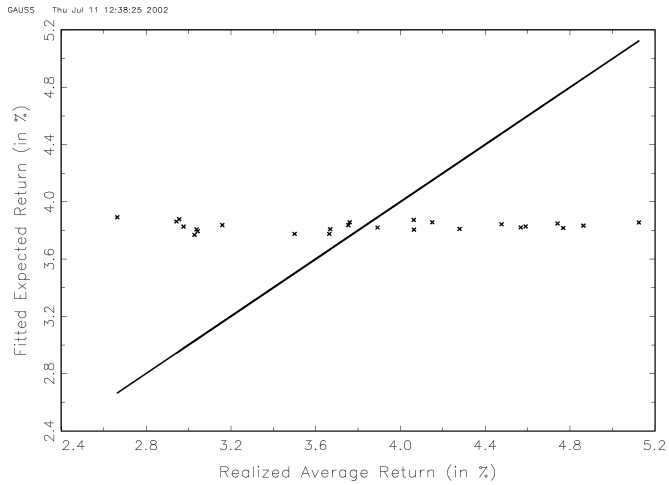


Figure 1: **CAPM - Fitted expected versus realized average quarterly returns.** Each scatter point in the graph represents one of the 25 size and book-to-market sorted Fama-French portfolios. The fitted returns have been obtained using the Fama-MacBeth cross-sectional regression methodology. The sample period is 1963:Q3-1998:Q3. The cross-sectional R^2 of this regression is 1 %.

of an asset if its return is positively correlated with changes of a state variable that are expected to result in less consumption. Indeed, a large yield spread is an indicator for high future economic activity and is thus likely to announce future consumption growth. Hence, investors who hedge against consumption risk might be induced to demand less of those assets that are positively correlated with the term spread. Chen et al. offer an alternative explanation. They argue that since the yield spread is a measure of changes in the long-term real rate of interest, its decline implies lower real returns on any form of capital. As a consequence, investors who seek to hedge against this source of risk will attribute higher values to assets that are negatively correlated with the term spread. Both interpretations are intuitively appealing, but do not seem to be confirmed by the data in the post-1970's period.[42] Although it appears difficult to find a theory-consistent explanation of this fact, one can make the following observations. First, the term spread is slightly positively correlated with the return on the market portfolio over the entire sample period. This may not be surprising since both have been shown to be positively linked to future economic activity.[43] Second, the returns on the 25 Fama-French portfolios are highly positively correlated with R^m which is also not surprising since we know that despite the observed CAPM anomalies there is a large common component in the time series variation of stock returns across assets. Hence, given these correlation patterns, one should indeed expect that the yield spread earns a positive risk premium. The question whether there has been a structural break in the relationship between the yield spread and average stock returns shall not be explored here and is thus left to future investigation.

[42] Note that Fama and French (1993) who perform time series regressions of the 25 size and book-to-market sorted portfolios on different risk factors similarly report that the term spread earns a *positive* risk premium. However, in contrast to the findings of Chen et al. and to my results, they state that the average premiums for both the default spread and the term spread "are too small to explain much variation in the cross-section of average stock returns."

[43] Estrella and Mishkin (1998), for example, provide evidence that the term spread is a powerful forecaster of recessions.

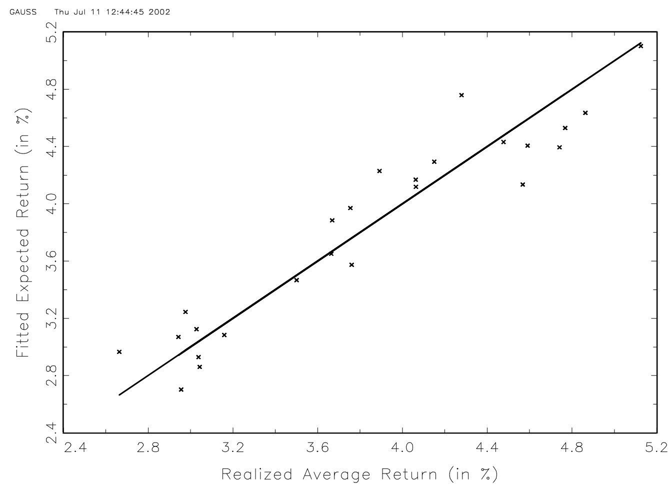


Figure 2: **ICAPM - Fitted expected versus realized average quarterly returns.** Each scatter point in the graph represents one of the 25 size and book-to-market sorted Fama-French portfolios. The fitted returns have been obtained using the Fama-MacBeth cross-sectional regression methodology. The sample period is 1963:Q3-1998:Q3. The cross-sectional R^2 of this regression is 91 %.

The estimation results of the ICAPM reported in table 1 exhibit another interesting feature: the growth rate of CPI earns a significant negative risk premium. This is in line with the results of Chen, Roll, and Ross (1986), but contradicts the finding of Chan, Karceski, and Lakonishok (1997) who argue that inflation does not help explaining returns across assets. Provided that the returns on the Fama-French portfolios are in nominal terms, one could presume that ΔCPI earns a significant risk premium since there must be some correlation between the rate of inflation and the portfolio returns. Two observations confute this interpretation, however. First, ΔCPI earns a *negative* risk premium. Hence, portfolio returns tend to be *smaller* when their inflation beta is high, i.e. when they are positively correlated with the monthly change of CPI. Second, I have performed Fama-MacBeth regressions on deflated portfolio returns in order to test whether the fact that nominal returns are used explains the significance of ΔCPI. I find that although slightly less significant, the growth rate of the consumer price index still earns a significant risk premium when deflated returns are considered. Note that the sign of the estimate suggests that investors hedge against the risk associated with high inflation. Indeed, a significantly negative $\lambda_{\Delta CPI}$ shows that assets that are inversely linked to CPI growth earn higher premiums than assets whose returns are positively correlated with inflation. All in all, the significance of TSpr and ΔCPI in the cross-sectional regression tests of the ICAPM demonstrates that the correlation of the size and BE/ME-sorted portfolios with these two variables provides a better explanation for their returns than does the mere correlation with the market portfolio. This view is also supported by the comparatively small value of $\mu_{|\hat{\alpha}|}$ indicating that the ICAPM produces considerably smaller pricing errors than the CAPM. Figure 2 visualizes this result, showing that average returns as predicted by the ICAPM are fairly close to the realized average returns.

Turning to the estimation results of the Fama-French three-factor model, we can see that the book-to-market factor HML earns a highly significant positive risk premium, while the market return and the size-related factor SMB do not enter significantly. This result is consistent with the findings of other authors who perform Fama-MacBeth regressions of the Fama-French model (e.g. Lettau and Ludvigson (2001b), Menzly (2001)).[44] The Fama-French three-factor model explains about 80 percent of the variation of average returns across the 25 ME and BE/ME-sorted portfolios. Further, its $\mu_{|\hat{\alpha}|}$ is only slightly larger than that of the ICAPM indicating that the pricing errors produced by the two models are of the same order of magnitude even though there is a slight difference in the R^2-statistic. Notice that since HML is by construction positively related to firms' book-to-market equity ratio, the positive risk price associated with this factor is compatible with the value effect, i.e. the fact that firms with high book-to-market equity ratios exhibit higher average returns than firms with low book-to-market ratios. Equivalently, the positive risk price associated with SMB provides evidence for the size effect, i.e. the fact that small firms tend to have higher returns than firms with large market capitalization.

Finally considering the results obtained from Fama-MacBeth regressions of Lettau and Ludvigson's conditional versions of the consumption CAPM and the human capital CAPM, we find that both perform about as well as the Fama-French three-factor model, explaining 70 and 76 percent of the cross-sectional variation of average returns, respectively. As in Lettau and Ludvigson (2001b), the growth rate of aggregate consumption scaled by $\widehat{cay}$ is highly significant, this even so after correction for the errors-in-variables bias. The value of $\mu_{|\hat{\alpha}|}$ obtained for the conditional CCAPM exceeds the values obtained for the competing benchmark models. This indicates that although the CCAPM explains the cross-sectional variation of average returns well, its overall model fit is slightly worse than that of the other models (except for the CAPM, of course). The results of the Fama-MacBeth regressions for the conditional human capital CAPM do also confirm Lettau and Ludvigson's findings. In particular, both aggregate labor income growth and aggregate labor income growth scaled by $\widehat{cay}$ earn significant risk premiums. Moreover, the scaled return on the market portfolio is significantly priced. This might indicate that although it does not explain the cross-sectional variation of average returns on the 25 FF portfolios unconditionally, R^m represents a source of risk that investors hedge against conditional on their expectations about future returns on the market portfolio.

I have also tested the quarterly data versions of the benchmark models running Fama-MacBeth regressions on the cross-section of 100 size and book-to-market sorted portfolios. The results of these regressions are given in appendix A. As for the 25 Fama-French portfolios, the return on the market portfolio is not a statistically significant determinant of the cross-section of average returns and consequently the CAPM has virtually no explanatory power. The R^2 statistics of the benchmark models range from 0.43 to 0.63, confirming our expectation that the overall model fit is worse when a much larger cross-section is tested. Except for aggregate consumption growth which is now significant, the significance of single factors does not appear to depend on the number of portfolios in the

[44] Fama and French (1993) only test their three-factor model on the basis of time series regressions so that their results cannot directly be compared with those presented here. Lettau and Ludvigson (2001b) show that SMB and HML are jointly significant which justifies the inclusion of the factor SMB in the model even though it is not individually significantly different from zero.

Table 2: **Benchmark Models - 100 FF portfolios, monthly data**

The table summarizes the results of Fama-MacBeth regressions that have been performed for the benchmark pricing models using the cross-section of the value-weighted returns on 100 Fama-French size and book-to-market sorted stock portfolios as the variable to be explained. The sample period is from 1963:07 to 1998:12. All models have the form

$$E[R_i] = E[R_0] + \beta_{i,f_1}\lambda_{f_1} + \ldots + \beta_{i,f_k}\lambda_{f_k},$$

where $\beta_{i,f_1}, \ldots, \beta_{i,f_k}$ denote the beta estimates of the pricing factors, preliminarily determined by regressing R_i on the respective vector of factors. For each model, estimates of λ_{f_i}, their t-values and the associated p-value are reported. The "corrected $t-$ and p-values" take into account the errors-in-variables adjustment as suggested by Shanken (1992). R^2 is the coefficient of determination in a single cross-sectional regression of the average returns on the beta estimates and $\bar{R}^2$ its degrees of freedom-adjusted version. $\mu_{|\hat{\alpha}|}$ is the average of all $T \cdot N$ pricing errors.

Model	Constant			Pricing Factors			Summary Statistics		
CAPM	$E[R_0]$	R^m					R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.50	-0.26					0.03	0.03	1.90
t-value	5.12	-0.73							
p-value	**0.00**	0.23							
corrected-t	5.11	-0.66							
corrected-p	**0.00**	0.25							
ICAPM	$E[R_0]$	R^m	TSpr	DSpr	ΔIP	ΔCPI	R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.55	-0.31	1.36	-0.11	-0.10	-0.09	0.22	0.18	1.78
t-value	5.41	-0.88	3.76	-1.36	-0.89	-2.03			
p-value	**0.00**	0.19	**0.00**	0.09	0.19	**0.02**			
corrected-t	3.93	-0.60	2.69	-0.95	-0.63	-1.43			
corrected-p	**0.00**	0.27	**0.00**	0.17	0.27	0.08			
FF3F	$E[R_0]$	R^m	SMB	HML			R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.71	-0.60	0.06	0.45			0.58	0.57	1.69
t-value	6.01	-1.71	0.43	3.34					
p-value	**0.00**	**0.05**	0.33	**0.00**					
corrected-t	5.89	-1.44	0.31	2.40					
corrected-p	**0.00**	0.08	0.38	**0.01**					
c. CCAPM	$E[R_0]$	cay	Δc	$cay \cdot \Delta c$			R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.64	-0.61	-0.20	0.00			0.28	0.27	1.97
t-value	7.64	-1.61	-1.97	1.44					
p-value	**0.00**	0.06	**0.03**	0.08					
corrected-t	5.86	-1.22	-1.50	1.08					
corrected-p	**0.00**	0.11	0.07	0.14					
c. HCCAPM	$E[R_0]$	cay	R^m	$cay \cdot R^m$	Δy	$cay \cdot \Delta y$	R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.73	-0.60	-0.35	-0.01	0.21	0.00	0.12	0.08	1.82
t-value	6.25	-2.78	-1.02	-0.99	1.51	0.15			
p-value	**0.00**	**0.00**	0.16	0.16	0.07	0.44			
corrected-t	5.52	-2.36	-0.83	-0.83	1.28	0.13			
corrected-p	**0.00**	**0.01**	0.20	0.20	0.10	0.45			

cross-section. Interestingly, both in terms of the coefficient of determination and in terms of the average absolute pricing error, the Fama-French three-factor model outperforms the other benchmark models when a cross-section of 100 size and BE/ME-sorted portfolios is considered.

Let us now turn to the results obtained from Fama-MacBeth regressions using monthly data. Since I will mainly assess the diffusion index pricing model based on regressions using a cross-section of 100 portfolios, I do also focus here on the performance of the benchmark models in explaining the returns on this larger set of Fama-French portfolios. These results are presented in table 2. Additional results obtained from regressions using monthly data and the cross-section of 25 portfolios are provided in the appendix. Except for the Fama-French three-factor model, the performance of all benchmark models in explaining the cross-sectional variation of average returns strikingly falls behind the results obtained using a cross-section of 25 portfolios and quarterly data. While the R^2 statistics of the ICAPM, the CCAPM and the HCCAPM have been in the range of 0.70 to 0.91 beforehand, all these models now explain less than 30 percent of the variation of average returns on the 100 Fama-French portfolios. As already mentioned, the monthly time series of $\widehat{cay}$ has been constructed based on intrapolations of the asset wealth variable a. This problem of data construction appears to matter a lot since none of the factors scaled by $\widehat{cay}$ is significantly priced when monthly data are being used. However, the weak performance of both the conditional CCAPM and the conditional HCCAPM cannot entirely be attributed to this problem. As indicated by the results I have obtained from regressions using quarterly data and the cross-section of 100 Fama-French portfolios (see table 9 in the appendix), the sharp decline in explanatory power is to some extent due to the use of the much larger cross-section of 100 portfolios. This view is supported by the fact that the ICAPM, although not subject to a problem of data construction, also performs considerably worse compared to the regressions using 25 Fama-French portfolios.

5.5 Testing the Diffusion Index Pricing Model

With these results at hand, we are now in a position to compare the performance of the diffusion index pricing model, henceforth denoted DIPM, with that of the benchmark models. As mentioned above, Stock and Watson (1998) have developed algorithms for the extraction of common dynamic factors from both balanced panels and unbalanced panels. In the unbalanced panel I employ there are 59 time series which are not contained in the balanced panel. The diffusion indexes extracted from the latter are thus based on considerably less economic information. In order to find out whether the missing information potentially improves the usefulness of diffusion indexes as factors in a pricing model, I therefore test two different versions of the DIPM, one using the unbalanced panel and one using the balanced panel factors.

Diffusion indexes are weighted cross-sectional averages of a large number of time series variables. Hence, in contrast to the benchmark pricing models where the factors have been chosen based on theoretical considerations, we do not have any a priori information which of the factors - if at all - is likely to earn a risk premium. I therefore start with a large-scale version of the DIPM, including a total of twelve diffusion indexes and the return on the market portfolio as pricing factors.

Table 3: **The Diffusion Index Pricing Model - 100 FF Portfolios, Monthly Data**

The table summarizes the results of Fama-MacBeth regressions that have been performed for the diffusion index pricing model using the cross-section of value-weighted returns on 100 Fama-French size and book-to-market sorted stock portfolios as the variable to be explained. The sample period is 1963:07-1998:12. The model is of the form

$$E[R_i] = E[R_0] + \beta_{i,f_1}\lambda_{f_1} + \ldots + \beta_{i,f_k}\lambda_{f_k},$$

where $\beta_{i,f_1}, \ldots, \beta_{i,f_k}$ denote the beta estimates of the pricing factors, preliminarily determined by regressing R_i on the vector of factors. See table 2 for details on the reported estimates and summary statistics.

Model	Constant	Pricing Factors													Summary Statistics		
A.1. ubp DIPM	$E[R_0]$	R^m	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.09	0.10	0.03	0.04	-0.63	0.61	0.14	-0.96	-0.37	0.44	-2.34	1.12	-3.04	0.71	0.60	0.55	1.59
t-value	3.92	0.29	0.06	0.07	-1.04	1.01	0.20	-1.56	-0.69	0.74	-3.99	2.15	-4.20	1.14			
p-value	**0.00**	0.39	0.48	0.47	0.15	0.16	0.42	0.06	0.25	0.23	**0.00**	**0.02**	**0.00**	0.13			
corrected-t	2.86	0.20	0.04	0.05	-0.73	0.71	0.14	-1.10	-0.48	0.52	-2.81	1.50	-2.99	0.80			
corrected-p	**0.00**	0.42	0.48	0.48	0.23	0.24	0.44	0.14	0.32	0.30	**0.00**	0.07	**0.00**	0.21			
A.2. ubp DIPM	$E[R_0]$	R^m									f_9	f_{10}	f_{11}		R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.57	-0.25									-2.78	1.30	-4.01		0.55	0.54	1.75
t-value	5.43	-0.70									-4.39	2.46	-3.68				
p-value	**0.00**	0.24									**0.00**	**0.01**	**0.00**				
corrected-t	3.66	-0.45									-2.88	1.60	-2.45				
corrected-p	**0.00**	0.33									**0.00**	0.06	**0.01**				
B.1. bp DIPM	$E[R_0]$	R^m	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.06	0.31	0.48	-0.45	-0.11	0.17	-0.28	-1.04	0.99	-0.13	-3.75	1.53	-0.39	-0.26	0.58	0.52	1.60
t-value	3.57	0.86	0.86	-0.74	-0.16	0.30	-0.53	-1.67	1.92	-0.22	-4.87	2.55	- 0.65	-0.39			
p-value	**0.00**	0.20	0.20	0.23	0.44	0.38	0.30	**0.05**	**0.03**	0.41	**0.00**	**0.01**	0.26	0.35			
corrected-t	2.60	0.59	0.60	-0.52	-0.11	0.21	-0.37	-1.18	1.34	-0.16	-3.47	1.79	-0.46	-0.28			
corrected-p	**0.01**	0.28	0.27	0.30	0.45	0.42	0.36	0.12	0.09	0.44	**0.00**	**0.04**	0.32	0.39			
B.2. bp DIPM	$E[R_0]$	R^m							f_7		f_9	f_{10}			R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.12	0.27							1.11		-4.39	1.34			0.56	0.55	1.75
t-value	4.02	0.72							2.09		-3.68	1.90					
p-value	**0.00**	0.24							**0.02**		**0.00**	**0.03**					
corrected-t	2.81	0.48							1.40		-2.55	1.29					
corrected-p	**0.00**	0.32							0.08		**0.01**	0.10					

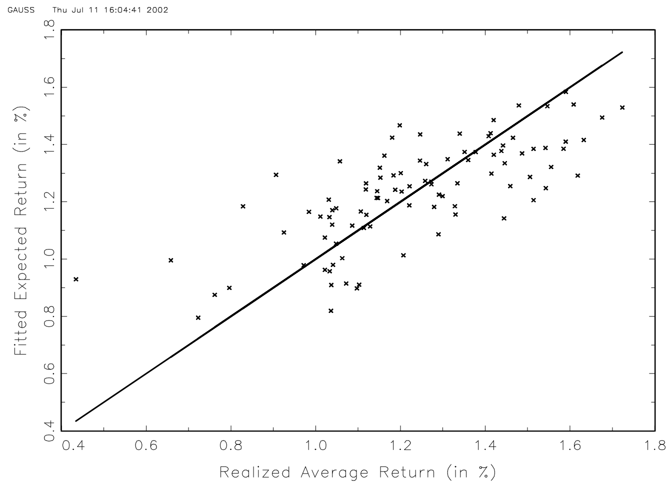


Figure 3: **ubp DIPM - Fitted expected versus realized average monthly returns.** Each scatter point in the graph represents one of the 97 size and book-to-market sorted Fama-French portfolios. The fitted returns have been obtained using the Fama-MacBeth cross-sectional regression methodology. The sample period is 1963:07-1998:12. The cross-sectional R^2 of this regression is 55 %.

Panels A.1. and B.1. of table 3 present results of Fama-MacBeth regressions performed for the large-scale versions of both the unbalanced panel and the balanced panel DIPM. One can see that with the inclusion of the diffusion indexes as pricing factors, the R^2-statistic jumps from 3 percent that I have obtained for the standard static CAPM to 60 and 58 percent for the unbalanced panel and the balanced panel DIPM, respectively. Thus, comparing this result with the results documented in table 2, I find evidence that the large-scale DIPM explains the cross-section of average monthly returns on 100 size and BE/ME-sorted stock portfolios about as well as the Fama-French three-factor model. Further, the regression results indicate that the diffusion index pricing model performs considerably better than the other benchmark models. This view is confirmed by the comparatively small values of $\mu_{|\hat{\alpha}|}$ that I obtain: in terms of absolute pricing errors, both versions of the large-scale DIPM outperform all considered benchmark models, including the Fama-French three-factor model.

Since in both the unbalanced panel and the balanced panel DIPM only three to four diffusion indexes enter significantly, I also test "trimmed" versions of the model, each including the three factors that have been the most significant in the large-scale versions. The results of Fama-MacBeth regressions for the unbalanced and balanced panel small-scale DIPM are given in panels A.2. and B.2. of table 3. They show that compared to the large-scale versions, the R^2-statistics only decrease by some percentage points. Hence, these three factors go a long way explaining the cross-section of average returns on the 100 Fama-French portfolios. All other diffusion indexes do not add much to the explanatory power of the model. Figure 3 plots the realized average returns on 100 Fama-French portfolios against their unbalanced panel three-factor DIPM predictions. In addition to the estimation results presented above, this figure allows the complementary conclusion that although this version of the DIPM seems to perform relatively well in the middle range, it has some difficulties explaining very high and very low average returns. For comparison, I provide similar graphs for four of the benchmark models in appendix A. These show that except for the Fama-French model, none of these models comes up to the ability of the DIPM to capture the cross-sectional variation of average monthly returns on 100 size and BE/ME-sorted stock portfolios.

5.6 Model Comparisons

The results presented above indicate that except for the Fama-French three-factor model, the DIPM explains the cross-section of returns on 100 size and book-to-market sorted portfolios better than all considered benchmark models. However, the measures of fit this conclusion is based on are unlikely to provide a complete picture of the relative performance of each model. I have therefore carried out further statistical tests to compare the benchmark pricing models with the DIPM. The results of these tests are briefly sketched in the following section.

It is a common practice in linear regression tests to detect the statistically more significant of several competing explanatory variables by estimating them jointly in a single regression. If the DIPM factors performed worse than the pricing factors of the benchmark models, they should in principle become insignificant in joint tests. In order to see if this is the case I perform such "horse races", testing the diffusion indexes together with the factors of each of the benchmark models.

Table 4 reports the results of cross-sectional regressions of the returns on 100 Fama-French portfolios on the beta estimates of the benchmark models, each augmented with the betas of three unbalanced panel diffusion indexes. Basically, the following conclusions can be drawn from these results. First, in none of the joint tests, the three unbalanced panel diffusion indexes significantly lose their explanatory power. Second, the only exception being the term spread which turns out to lose some of its explanatory power, all pricing factors of the benchmark models that have been significant in individual regressions remain significant when tested jointly with the diffusion indexes. Third, except for the Fama-French three-factor model, the inclusion of the diffusion indexes considerably improves the overall fit of the pricing models: compared to the regression results of the benchmark models summarized in table 2, the R^2-statistics jump from values of 12, 22, and 28 percent for the HCCAPM, the ICAPM and the CCAPM to 57, 58, and 60 percent, respectively. Notice that similar tests which I have performed using the balanced panel diffusion indexes have yielded analogous results.

It has been shown in section 5.4 that all benchmark models except for the standard static CAPM go a long way explaining the cross-section of 25 Fama-French portfolios when quarterly data is used. Since the diffusion indexes are only available in monthly frequency, a direct comparison based on results obtained from quarterly data is impossible. Yet, distributional theory allows us to contrast results of regressions based on different data sets. In fact, in section 5.2 we have defined the J_T statistic as the sum of squared standardized average pricing errors. Under the assumption that the pricing errors are mutually independent and standard normally distributed, J_T obviously follows a $\chi^2(N-k)$ distribution where N denotes the number of assets in the cross-section and k the number of pricing factors in the model. Hence, we can in principle compare J_T statistics of the

Table 4: **"Horse Races" - 100 FF portfolios, monthly data**

The table summarizes the results of Fama-MacBeth regressions that have have been performed for models that simultaneously include as pricing factors the factors of the benchmark models and three unbalanced panel diffusion indexes. The variable to be explained is the cross-section of the value-weighted returns on 100 Fama-French size and book-to-market sorted stock portfolios. The sample period is from 1963:07 to 1998:12. For each model, risk price estimates, their t-values and the associated p-values are reported. The "corrected $t-$ and p-values" take into account the errors-in-variables adjustment as suggested by Shanken (1992). R^2 is the coefficient of determination in a single cross-sectional regression of the average returns on the beta estimates.

Model	cst.	Benchmark Model Factors					DIPM Factors			
ICAPM	$E[R_0]$	R^m	Tspr	Dspr	ΔIP	ΔCPI	f_9	f_{10}	f_{11}	R^2
Estimate	1.41	-0.13	0.49	-0.08	-0.13	-0.12	-2.68	0.90	-4.35	0.58
t-value	5.07	-0.39	1.84	-1.03	-1.17	-3.04	-4.16	1.72	-4.36	
p-value	**0.00**	0.35	**0.03**	0.15	0.12	**0.00**	**0.00**	**0.04**	**0.00**	
corrected-t	3.32	-0.24	1.18	-0.65	-0.75	-1.93	-2.66	1.09	-2.82	
corrected-p	**0.00**	0.40	0.12	0.26	0.23	**0.03**	**0.00**	0.14	**0.00**	
FF3F	$E[R_0]$	R^m	SMB	HML			f_9	f_{10}	f_{11}	R^2
Estimate	1.43	-0.33	0.11	0.42			-2.01	1.60	-1.15	0.65
t-value	4.75	-0.91	0.75	3.14			-3.38	3.02	-1.91	
p-value	**0.00**	0.18	0.23	**0.00**			**0.00**	**0.00**	**0.03**	
corrected-t	4.01	-0.69	0.49	2.07			-2.72	2.41	-1.54	
corrected-p	**0.00**	0.25	0.31	**0.02**			**0.00**	**0.01**	0.06	
c. CCAPM	$E[R_0]$	cay	Δc	$cay \cdot \Delta c$			f_9	f_{10}	f_{11}	R^2
Estimate	1.36	-0.04	-0.06	0.00			-3.04	1.42	-3.19	0.60
t-value	6.44	-0.24	-0.90	1.79			-3.24	2.56	-3.27	
p-value	**0.00**	0.41	0.19	**0.04**			**0.00**	**0.01**	**0.00**	
corrected-t	4.32	-0.15	-0.59	1.17			-2.14	1.66	-2.17	
corrected-p	**0.00**	0.44	0.28	0.12			**0.02**	**0.05**	**0.02**	
c. HCCAPM	$E[R_0]$	R^m	cay	$cay \cdot R^m$	Δy	$cay \cdot \Delta y$	f_9	f_{10}	f_{11}	R^2
Estimate	1.63	-0.03	-0.20	0.00	-0.04	-0.00	-2.61	1.54	-3.71	0.57
t-value	5.85	-0.16	-0.57	0.23	-0.31	-0.51	-4.22	3.02	-3.49	
p-value	**0.00**	0.44	0.28	0.41	0.38	0.31	**0.00**	**0.00**	**0.00**	
corrected-t	3.99	-0.10	-0.37	0.15	-0.20	-0.33	-2.79	1.97	-2.35	
corrected-p	**0.00**	0.46	0.36	0.44	0.42	0.37	**0.00**	**0.03**	**0.01**	

benchmark models and the DIPM that have been obtained from Fama-MacBeth regressions using return data of different frequency. However, the aforementioned problem of consistently estimating the variance-covariance matrix of the pricing errors persists and thus inference based on the J_T statistic and especially its Shanken-corrected version J_T^c has to be cautioned.

Panels B and C of table 5 provide J_T statistics computed from average pricing errors that have been obtained from regressions using both monthly and quarterly returns of 25 size and BE/ME-sorted portfolios. The results summarized in panel B indicate that only for the Fama-French three-factor model the hypothesis that all pricing errors are jointly zero cannot be refused at the 5 percent significance level. The J_T statistics obtained for the unbalanced and the balanced panel DIPM are larger than the one obtained for the Fama-French model. However, since they are still exceeded by the values of the CAPM, the scaled CCAPM, and the scaled HCCAPM, this might imply a better model fit of the DIPM. In panel C, estimates of J_T and J_T^c which I have obtained from quarterly data are reported. As expected, the Lettau-Ludvigson scaled CCAPM and HCCAPM in this case exhibit considerably smaller J_T-statistics. Yet, the squared standardized average pricing errors remain significantly different from zero even when quarterly data are used.[45] The extremely high J_T statistic obtained for the Fama-French three-factor model can be interpreted as an example for the low power of the test. In fact, the regression results reported in table 1 have shown that the Fama-French model performs at least about as well as the other benchmark models when quarterly data are used. This suggests that its J_T statistic obtained from quarterly data is aberrant. The estimates obtained from monthly data using the larger cross-section of 100 portfolios, presented in panel D, equally confirm this view. Again, only for the Fama-French three-factor model can the Null that all pricing errors are jointly different from zero be rejected at the 5 % level.

To summarize, among the pricing models considered, the Fama-French model is the only that produces average pricing errors that are jointly statistically undistinguishable from zero in some cases. In terms of J_T test statistics obtained from monthly data, the DIPM performs about as well as the ICAPM and better than the two scaled (C)CAPM versions. However, these exhibit smaller J_T statistics when quarterly data are being used. The Shanken-corrected J_T-statistics vary quite considerably from model to model. Hence, a clear statement about the relative model fit based on these results is ruled out.

Let us finally have a brief look at the average pricing errors of all 25 individual portfolios. These are provided in panel A of table 5. Each individual portfolio is represented by a two-digit number, where the first digit refers to the size quintile (1 = small) and the second to the book-to-market category (1 = low). An "eyeballing" analysis of the signs of the pricing errors shows that all models overstate the returns on the portfolios in the highest size quintile. Further, except for the biggest stocks, the pricing errors obtained for high book-to-market portfolios in all size categories are positive across all models. This shows that the tested pricing models uniformly underestimate the returns on stocks with high book-to-market equity ratios. Hence, none of the considered models seems to be able to entirely capture the size and the value effect.[46]

[45] Lettau and Ludvigson (2001b) find slightly lower values of J_T than I do and thus reject the Null. I attribute the discrepancy between our results to slight differences in the data used and to the documented robustness problem inherent to the estimation of the J_T statistic.

[46] Plots of the sum of squared pricing errors across portfolios for each period, not provided here, show that over the entire sample there are a few months where all models equally exhibit considerable mispricing. These "outliers" are likely to influence the outcomes of Fama-MacBeth regressions quite substantially. In fact, Menzly (2001) shows that the variation of returns on Fama-French portfolios is largely determined by some influential quarters that are concentrated around recession periods. He provides evidence that pricing factors that load high on only one or two of these influential quarters earn significant risk premiums. Hence, a detailed analysis of the overall economic conditions in such influential periods could provide us with a more precise picture of the links between the stock market and macroeconomic variables.

Table 5: **Average Pricing Errors**

This table reports estimation results obtained from Fama-MacBeth regressions in terms of pricing errors. Panel A lists for each model the average monthly pricing errors (in %) for all 25 size and BE/ME-sorted portfolios in the cross-section. The individual portfolios are represented by two-digit numbers that figure in the first column. The first digit refers to the size quintile (1 = small) and the second to the book-to-market category (1 = low). Panel B reports uncorrected and Shanken-corrected J_T statistics that have been computed from the average pricing errors in panel A. An asterisk behind a number indicates that the respective statistic is *not* different from zero at the 5 % significance level. Panel C summarizes J_T statistics for the benchmark models obtained from Fama-MacBeth regressions using quarterly data. Panel D reports J_T statistics for monthly returns and a cross-section of 100 Fama-French portfolios.

	CAPM	ICAPM	Fama-French	scaled CCAPM	scaled HCCAPM	DIPM unb.panel	DIPM b.panel
A. Average Pricing Errors - 25 Fama-French Portfolios - Monthly Returns							
11	-0.4200	-0.2549	-0.3576	-0.5691	-0.4959	-0.2769	-0.2230
12	0.0437	-0.0562	-0.0477	0.0102	-0.0063	-0.1415	-0.1134
13	0.0641	-0.0634	-0.0861	-0.0094	0.0029	-0.1164	-0.0741
14	0.2673	0.2408	0.0450	0.1125	0.1739	0.0856	0.1309
15	0.4170	0.0576	0.1038	0.2827	0.2648	0.0943	0.0929
21	-0.2222	0.1605	0.0348	-0.0990	-0.0892	0.0433	0.0629
22	-0.0483	-0.0350	-0.0336	-0.1232	0.0647	0.0469	0.0165
23	0.1919	-0.0549	0.0972	0.1210	0.1385	0.0899	0.0946
24	0.2565	0.0996	0.1118	0.2497	0.3048	0.1096	0.1317
25	0.3307	0.2344	0.1228	0.1554	0.2920	0.1520	0.0416
31	-0.2027	-0.0025	0.0959	-0.1008	-0.1428	0.0252	0.0383
32	0.0150	-0.1188	0.0725	0.1146	0.0869	0.0531	0.0123
33	0.0018	-0.0249	-0.0548	0.0602	0.0080	-0.1432	-0.1758
34	0.1546	-0.0421	0.0367	0.0670	0.1272	-0.0227	0.0467
35	0.2357	0.1347	0.0625	0.2305	0.2357	0.1276	0.0244
41	-0.1819	-0.0333	0.1459	0.1179	-0.1342	0.1598	0.1662
42	-0.2588	-0.2051	-0.1236	-0.0439	-0.1415	-0.2033	-0.2427
43	-0.0482	0.1488	-0.0249	-0.0987	0.1267	0.0009	-0.0839
44	0.0690	-0.0128	0.0066	-0.0372	0.0682	0.0729	0.0753
45	0.2135	0.1757	0.1099	0.0338	0.1448	0.2206	0.2103
51	-0.2075	-0.0412	0.1436	0.0020	-0.3096	-0.0162	0.0423
52	-0.2345	-0.1309	-0.0276	-0.0409	-0.2771	0.0580	0.0106
53	-0.2435	-0.1033	-0.1255	-0.0949	-0.2277	-0.1671	-0.1629
54	-0.1255	0.0913	-0.1257	-0.1685	-0.0888	-0.1781	-0.0878
55	-0.0678	-0.1640	-0.1817	-0.1719	-0.1258	-0.0743	-0.0338
B. J_T Test Statistics - 25 FF Portfolios - Monthly Returns							
J_T	71.36	50.63	31.64*	64.76	85.85	56.35	55.72
J_T^c	71.68	332.39	33.12*	144.53	164.77	195.78	148.89
C. J_T Test Statistics - 25 FF Portfolios - Quarterly Returns							
J_T	53.31	37.38	62.73	39.78	36.52		
J_T^c	53.33	114.77	250.24	81.93	153.20		
D. J_T Test Statistics - 100 FF Portfolios - Monthly Returns							
J_T	209.74	166.52	114.68*	233.73	204.32	166.08	160.51
J_T^c	210.96	316.11	119.29	397.04	261.96	365.78	329.71

5.7 Interpreting the DIPM Factors

Having shown that the diffusion index pricing model performs comparatively well, let us now have a more detailed look at the significance patterns of the individual factors in the DIPM. Interestingly, I find that the main contribution to the explanatory power of the model cannot be attributed to those factors that correspond to the first principal components of the cross-sectional covariance matrix of all series in the panel. Instead, in both the unbalanced and the balanced panel version of the DIPM, diffusion indexes that represent *lower-order* principal axes of the cross-section of the data set prove to be significantly priced factors. This seems surprising since the lower-order principal components explain less of the total variation of all series in the panel and thus summarize less economic information than higher-order principal components. To see whether the significance of the lower-order diffusion indexes in my pricing model can sensibly be interpreted as an indication for the fact that they account for sources of risk that investors hedge against, an investigation of the individual variables on which they load is warranted. Appendix B provides a list of all series in the data set. More importantly, plots of the coefficients of determination obtained from univariate regressions of all individual series on all dynamic factors are given. Following Stock and Watson, I interpret these R^2-statistics as factor loadings of the diffusion indexes on the various time series variables in the data set. Let us first investigate the correlation patterns observed for the unbalanced panel diffusion indexes.

As we have seen, the Fama-MacBeth regressions performed for the unbalanced panel DIPM provide evidence that the factors f_9, f_{10}, and f_{11} earn significant risk premiums. That is, the returns on the 100 size and BE/ME-sorted portfolios are strongly proportional to their respective covariances with these three factors. Figure 8 in appendix B shows the loadings of the unbalanced panel factors on all time series variables in the data set. The following observations can be made: f_9 on average explains about 3 percent of the variation of each individual time series in the unbalanced panel. Being largely uncorrelated with most of the variables, it loads particularly high on the series of the category "Housing starts and sales" as well as on some variables of the two categories "Employment and hours" and "Money and credit quantity aggregates". f_{10} exhibits an average coefficient of determination of 2.4 percent in univariate regressions against the individual time series in the panel. It loads high on most of the variables of the category "Orders and unfilled orders". Further, it shows relatively strong correlations with some variables of the "Real retail, manufacturing and trade sales" category as well as with some of the consumption measures contained in the data set. Finally, f_{11} shows only weak occasional correlations with some variables of the categories "Real retail, manufacturing and trade sales", "Orders and unfilled orders", and "Money and credit quantity aggregates". It has an average R^2 of 2 percent in univariate regressions on the series in the data set.[47]

[47] Although not directly related to my problem, there is a feature of the unbalanced panel Stock-Watson factors that deserves attention: the first unbalanced panel diffusion index on average explains only 6 percent of the variation of the individual time series in the data set, whereas the second exhibits an average R^2 of 14 percent. This appears to be some particularity of the EM algorithm and has been left uncommented by the authors in their article. In fact, as one can see, the first diffusion index loads particularly high on the series No. 213 and 214 in the data set, exhibiting univariate R^2's close to unity. These two series are measures of the US trade balance and are only available from 1986 on. The first unbalanced panel diffusion

Let us next consider the factors of the balanced panel DIPM. The Fama-MacBeth regressions have identified the factors f_7, f_9, and f_{10} as being significantly priced. A quick look on the correlations with the individual time series in the data set (see figure 11 in appendix B) tells us that f_7 on average explains 2.5 percent of the variation of each variable in the panel, loading high on the category "Orders and unfilled orders", further on some retail trade and consumption variables, and on the inventory-sales ratios contained in the balanced panel. f_9 exhibits an average R^2 of 2 percent and shows the highest correlations with exchange rate variables. Further, it has small loadings on a variety of time series of the categories "Real retail, manufacturing and trade sales", "Consumption", and "Housing starts and sales". Finally, f_{10} has an average coefficient of determination of 2.2 percent most of which is due to relatively high loadings on those variables in the panel that are related to the money stock. Further, it shows some correlation with time series of the categories "Employment and hours" and "Real retail, manufacturing and trade sales".

All in all, the factor loading patterns summarized above unfortunately do not provide us with a clear picture of the set of variables that determine the significant diffusion indexes. We have seen that those Stock-Watson factors which capture the largest part of the common variation of all series in the panel turn out to be insignificant in all tested versions of the DIPM. On the other hand, those diffusion indexes that earn risk premiums explain only a small portion of the total variation in the data set and load on a variety of variables of different economic categories. More importantly, except for some consumption measures, none of the state variables that have been documented in the literature as being priced in cross-sectional regression tests, has been found to be strongly correlated with the significant diffusion indexes. In particular, it is surprising that the dynamic factors that load high on interest rates and on price level variables (factors No. 3 and No. 2 of the unbalanced and balanced panel diffusion indexes, respectively) - somewhat contrary to the findings of Chen et al. and to the results I have obtained from tests of the ICAPM - do not earn significant risk premiums.

There are broadly two possible explanations for the significance of the lower-order diffusion indexes in the cross-sectional regression tests. The first and admittedly rather optimistic one is that using pricing factors that are weighted averages of various time series variables, we have detected sources of systematic risk that have as yet been unnoticed by asset pricing theoreticians. Indeed, since financial markets and the real economy are closely entangled, both stock returns and business cycle-related variables exhibit a high degree of endogeneity. Thus, it does not seem unlikely that certain macroeconomic variables, which from a theoretical viewpoint one might not directly associate with stock returns, empirically be in fact determinants of these. For example, some of the significant

index is the eigenvector corresponding to the largest eigenvalue of the cross-sectional covariance matrix of the complete data set where missing observations have been replaced by recursively estimated values. For some unidentified reason, the recursive EM algorithm attributes a very large weight to the two trade balance series from the first period of their availability on. Figure 9 in appendix B provides a plot of the first unbalanced panel diffusion index showing that it remains almost constant from 1986 on. To complete the picture, figure 10 shows the second factor whose time series behavior is clearly cyclical. The question whether the observed feature is likely to arise when a different data set is used and whether it represents a potential drawback of Stock and Watson's method, shall not be discussed here and is thus left to future research.

diffusion indexes load on variables of the category "Orders and unfilled orders". It is obvious that these variables are closely related to the business prospects of firms, which in turn are known to affect share prices. The same might hold for other time series that have been shown to be correlated with the significant Stock-Watson factors, e.g. measures of consumption, exchange rates, inventory-sales ratios, series related to real retail, manufacturing and trade sales etc. Hence, filtering the common variation in these variables, we might have found diffusion indexes that are useful predictors for stock returns.[48]

Bernanke and Boivin (2001) show that the scope of the data set from which the Stock-Watson factors are extracted considerably affects their forecasting power. As mentioned above, I use the same set of macroeconomic variables as Stock and Watson (2002). They have selected the time series in the panel set with the intention of forecasting output and inflation. Accordingly, as Stock and Watson show, the first principal axes of the cross-sectional covariance matrix are strongly correlated with measures of current economic activity. However, an alternative selection of variables to include in the panel is likely to produce diffusion indexes that exhibit completely different time series behaviors and consequently forecast a different set of economic variables. That is, the fact that only lower-order diffusion indexes have been found to be significantly priced in cross-sectional tests of the DIPM might simply be due to the inappropriateness of the employed panel. Indeed, using basically the same data set, Watson (2001) finds only little evidence that forecasts of stock market variables based on diffusion indexes are better than alternative forecasting methods. In view of my empirical results and of the findings of Bernanke and Boivin (2001), I thus conjecture that carrying out an "informed pre-selection" of the variables to include in the panel is likely to improve the predictive power of the extracted diffusion indexes for stock returns in both the time and the cross-sectional dimension. More generally, I see considerable scope for future research concerning the design of the panel to be used in the Stock-Watson procedure.[49]

A second possible explanation of the good overall performance of the DIPM is that the model is subject to misspecification bias. As Kan and Zhang (1999) show, inference about the correctness of a pricing model based on Fama-MacBeth t-statistics is unwarranted since these are likely to overstate the significance of pricing factors. They provide evidence that "useless" factors may erroneously be detected as significant due to estimation errors of the first-stage beta estimates.[50] Kan and Zhang suggest a set of model diagnostics to avoid the problem of misspecification bias. For example, they propose to perform a subperiod joint test, i.e. to split the sample into two or more subperiods, and to run separate cross-sectional regressions on all of the subsamples. Jagannathan and Wang (1998) show that when a beta-pricing model is misspecified, the t-values for firm characteristics included as additional factors converge to infinity in probability. Hence, in well-specified models, firm

[48]However, if this were indeed the case, I would presume that the relationship found is likely to hold independently of size or book-to-market ratio.

[49]Depending on the particular objective of a certain diffusion index-based application, the selection of variables to include in the data set should ideally rely on both theoretical and statistical considerations. With respect to the latter, clustering algorithms in the frequency domain might be useful for a preliminary identification of subsets of variables with similar time series properties.

[50]Kan and Zhang (1999) consider the extreme case of "useless" factors constructed to be statistically independent of all asset returns in the cross-section.

characteristics should not be significantly priced. This result can alternatively be used to test for model misspecification bias. In order to test whether the DIPM is misspecified, I perform both a subperiod joint test and regressions including firms characteristics. The results are presented in the following section.

5.8 Testing for Model Misspecification

I start with cross-sectional regressions including firm characteristics. Jagannathan and Wang (1998) provide a formal proof that a factor that is uncorrelated with the asset returns, i.e. that has zero beta, "cannot stand up against a test with a cross-sectional variable such as firm size". Hence, joint tests of some firm-specific variable and the diffusion indexes should in principle reveal whether the latter are useless pricing factors. Lettau and Ludvigson (2001b) test for model misspecification by running Fama-MacBeth regressions using either the time series average of the log of portfolio size or the time series average of log book-to-market ratio as additional pricing factors.[51] I follow Lettau and Ludvigson in this respect and perform similar regressions.

Table 6 summarizes the results that I have obtained from Fama-MacBeth regressions of the unbalanced and the balanced panel diffusion index pricing model, each augmented with portfolio size (panel A) and the portfolio book-to-market ratio (panel B). Both firm characteristics are shown to be significantly priced when tested jointly with the diffusion indexes. In fact, with the exception of the Shanken-corrected t-values of the size factor which are not significantly different from zero at the 5 % level, the risk price estimates of the firm characteristics are statistically significant in either of the tested models. Hence, according to the result of Jagannathan and Wang, this is a clear indication that the diffusion index pricing model is subject to misspecification bias.[52]

Let us shed some more light on the issue of misspecification. Both, Kan and Zhang (1997) and Jagannathan and Wang (1998), show that a misspecified factor, i.e. a factor that is asymptotically uncorrelated with the asset returns, may have an arbitrarily large t-statistic in cross-sectional regressions. Kan and Zhang provide the following explanation for this characteristic: if the "true" betas of the assets in the cross-section with respect to a useless factor are zero, then the associated "true" risk price is undefined.[53] Consequently, the estimated risk premium must go to infinity to account for the difference in the expected returns. An analysis of the first-stage estimates of the diffusion index betas shows that these are indeed small compared to those of some of the significantly priced

[51] Unlike the standard pricing factors, these variables are directly included as explanatory variables in the second-stage estimation, i.e. without preliminarily determining beta estimates.

[52] I have carried out the same test for all benchmark models using quarterly data and the cross-section of 25 portfolios as the variable to be explained. The results indicate that the size factor indeed turns out to become insignificant when tested jointly with the beta estimates of most of the benchmark pricing factors. However, contrary to the results presented by Lettau and Ludvigson (2001b), my findings suggest that this does *not* hold for the consumption CAPM scaled by $\widehat{cay}$. Since the authors report that in their tests the factor $\widehat{cay}_t \cdot \Delta c_{t+1}$ makes SIZE insignificant, I am reluctant to place emphasis on my finding. Differences in the firm-specific data used might be one possible explanation for the contradictory results.

[53] To see this, recall that the estimated vector of risk prices is given by $\hat{\lambda} = (\hat{\mathbf{B}}'\hat{\mathbf{B}})^{-1}\hat{\mathbf{B}}'\bar{R}$ where $\hat{\mathbf{B}}$ is the matrix of beta estimates obtained from the first-stage time series regressions, and where $\bar{R}$ denotes the vector of average returns of all assets in the cross-section.

Table 6: **Fama-MacBeth Regressions Including Firm Characteristics**

This table presents results of Fama-MacBeth regressions that have been performed for the unbalanced panel and the balanced panel diffusion index pricing model, each augmented with a firm-specific cross-sectional variable as an additional pricing factor. The cross-sectional regressions are based on models of the form,

$$E[R_i] = E[R_0] + \beta_{i,f_1}\lambda_{f_1} + \ldots + \beta_{i,f_k}\lambda_{f_k} + b\Theta_i,$$

where $\beta_{i,f_1}, \ldots, \beta_{i,f_k}$ denote the beta estimates of the pricing factors, preliminarily determined by regressing R_i on the respective vector of factors. Θ_i is the time series average of the log of market capitalization of the firms in portfolio i (SIZE in Panel A) or the time series average of the log of the value-weighted book-to-market equity ratio of the firms in portfolio i (BE/ME in Panel B). The sample period is 1963:07-1998:12. The variable to be explained is the cross-section of the value-weighted returns on 100 Fama-French size and book-to-market sorted stock portfolios. For each model, estimates of the λ_{f_i}, their uncorrected and Shanken-corrected t-values, the associated p-values, and the cross-sectional R^2 are reported.

Model	Constant	Pricing Factors				Θ_i	
A.1. ubp DIPM	$E[R_0]$	R^m	f_9	f_{10}	f_{11}	SIZE	R^2
Estimate	1.78	-0.42	-2.44	0.90	-3.63	-0.30	0.59
t-value	5.45	-1.16	-4.04	1.64	-3.57	-2.16	
p-value	**0.00**	0.13	**0.00**	**0.05**	**0.00**	**0.02**	
corrected-t	3.90	-0.79	-2.80	1.13	-2.52	-1.55	
corrected-p	**0.00**	0.22	**0.00**	0.13	**0.01**	0.06	
A.2. bp DIPM	$E[R_0]$	R^m	f_7	f_9	f_{10}	SIZE	R^2
Estimate	1.35	0.07	0.79	-4.01	1.08	-0.25	0.59
t-value	4.60	0.19	1.47	-3.57	1.57	-1.94	
p-value	**0.00**	0.42	0.07	**0.00**	0.06	**0.03**	
corrected-t	3.39	0.13	1.04	-2.61	1.12	-1.43	
corrected-p	**0.00**	0.45	0.15	**0.01**	0.13	0.08	
B.1. ubp DIPM	$E[R_0]$	R^m	f_9	f_{10}	f_{11}	BE/ME	R^2
Estimate	1.30	-0.23	-1.69	1.97	-2.14	0.21	0.68
t-value	4.67	-0.64	-2.91	3.80	-2.33	3.26	
p-value	**0.00**	0.26	**0.00**	**0.00**	**0.01**	**0.00**	
corrected-t	3.75	-0.48	-2.23	2.89	-1.84	2.62	
corrected-p	**0.00**	0.32	**0.01**	**0.00**	**0.03**	**0.01**	
B.2. bp DIPM	$E[R_0]$	R^m	f_7	f_9	f_{10}	BE/ME	R^2
Estimate	1.16	-0.02	1.77	-2.38	0.30	0.20	0.67
t-value	4.19	-0.05	3.36	-2.57	0.41	3.02	
p-value	**0.00**	0.48	**0.00**	**0.01**	0.34	**0.00**	
corrected-t	3.50	-0.04	2.65	-2.10	0.33	2.53	
corrected-p	**0.00**	0.49	**0.00**	**0.02**	0.37	**0.01**	

factors in the benchmark models. According to the results of Kan and Zhang (1999) and Jagannathan and Wang (1998), this finding thus supports the view that the diffusion index pricing model is subject to misspecification bias. As a suggestion for future work in this field, the problem of misspecification of the type encountered here could possibly be avoided by preliminarily testing whether the first-stage beta estimates are statistically different from zero.

In order to explore further whether the good overall performance of the diffusion index pricing model in cross-sectional regression tests is really driven by some misspecification bias that cannot be detected by the mere use of Fama-MacBeth t-statistics and the cross-sectional R^2, I confront the DIPM with a second test for misspecification. The subperiod joint test suggested by Kan and Zhang (1999) is designed to examine whether a significant relationship between the beta risk of a pricing factor and the cross-section of returns is stable over time. Such a test can be performed very easily by splitting the entire sample into two or more subperiods and to run cross-sectional regressions over all subsamples. Kan and Zhang suggest to reject the Null hypothesis that a factor estimate is not statistically different from zero *only if* both t-statistics are significant and have the same sign. Obviously, the intuition behind this test is that if a factor proxies for some source of risk that investors hedge against, then it should earn about the same premium in every instant. Yet, this line of argumentation relies on the assumption that the betas of "true" factors are constant over time. Thus, since there is empirical evidence that betas are in fact time-varying, the subperiod joint test used alone might be unsuited for detecting model misspecification. Nonetheless, as a complement to the findings above, I report results obtained from this test in table 7.

I adopt the most basic version of the test, simply splitting the entire sample period into two equally long subperiods, 1963:07-1981:03 and 1981:04-1998:12. Then I perform separate Fama-MacBeth regressions over these two subsamples. The results of these regressions also seem to support the view that the DIPM is subject to misspecification bias: except for the balanced panel diffusion index f_{10}, none of the diffusion indexes that have been found to be significantly priced in the Fama-MacBeth regressions using the entire sample period, earns a risk premium in both subperiods. The differences in significance even turn out to be strikingly stark. The unbalanced panel diffusion index f_9, for example, is highly significant in the first subperiod, but highly insignificant in the second regression. This indicates that the correlation between average returns and pricing factors can be subject to fundamental changes. Whether these changes reveal structural breaks in otherwise valid economic relationships or rather demonstrate that the Fama-MacBeth t-statistics are susceptible to falsely suggest the presence of non-existent economic interrelations, cannot be decided at this stage of the investigation.

In order to analyze whether the poor performance of the DIPM in the subperiod joint tests is particular to this model, I have also carried out such tests for three of the benchmark models. The results of these tests are summarized in table 11 in the appendix. They show that with the only exception being the book-to-market factor HML in the Fama-French model, the pricing factors of the benchmark models do *all* fail this test. Hence, even the performance of models that are widely accepted as capturing well the return spreads across size and BE/ME-sorted stock portfolios, strongly depends on the choice of the sample period. Except for the Fama-French three-factor model, none of the tested models thus seems to represent a stable economic relationship between stock returns and pricing factors.

The results of the subperiod joint tests raise the question of how the sample period for the first-stage estimation of the betas should optimally be chosen. In their seminal study, Fama and MacBeth (1973) adopt a five-year rolling regression estimation to account for

Table 7: **Subperiod Joint Test**

This table presents results of Fama-MacBeth regressions performed for the unbalanced panel and the balanced panel diffusion index pricing model over the two subperiods 1963:07-1981:03 (panel A) and 1981:04-1998:12 (panel B). The variable to be explained is the cross-section of the value-weighted returns on 100 Fama-French size and book-to-market sorted stock portfolios. For each regression, estimates of the λ_{f_i}, their uncorrected and Shanken-corrected t-values, the associated p-values, and the cross-sectional R^2 are reported.

Model	**Constant**	**Pricing Factors**				
		A. Sample Period 1963:07 - 1981:03				
ubp DIPM	$E[R_0]$	R^m	f_9	f_{10}	f_{11}	R^2
Estimate	0.20	0.53	-4.43	5.59	2.49	0.68
t-value	0.56	1.14	-3.42	2.29	1.14	
p-value	0.29	0.13	**0.00**	**0.02**	0.13	
corrected-t	0.34	0.66	-2.05	1.39	0.69	
corrected-p	0.37	0.26	**0.03**	0.09	0.25	
bp DIPM	$E[R_0]$	R^m	f_7	f_9	f_{10}	R^2
Estimate	0.39	0.39	-0.19	4.84	4.40	0.76
t-value	1.05	0.87	-0.14	2.83	2.20	
p-value	0.15	0.20	0.44	**0.00**	**0.02**	
corrected-t	0.72	0.56	-0.10	1.93	1.51	
corrected-p	0.24	0.29	0.46	**0.03**	0.07	
		B. Sample Period 1981:04 - 1998:12				
ubp DIPM	$E[R_0]$	R^m	f_9	f_{10}	f_{11}	R^2
Estimate	3.39	-2.00	0.14	2.37	-0.11	0.68
t-value	7.56	-3.86	0.12	1.77	-0.08	
p-value	**0.00**	**0.00**	0.45	**0.04**	0.47	
corrected-t	5.81	-2.80	0.09	1.34	-0.06	
corrected-p	**0.00**	**0.00**	0.46	0.10	0.47	
bp DIPM	$E[R_0]$	R^m	f_7	f_9	f_{10}	R^2
Estimate	3.23	-1.86	2.85	-0.03	3.86	0.73
t-value	7.78	-3.52	1.68	-0.02	2.71	
p-value	**0.00**	**0.00**	**0.05**	0.49	**0.01**	
corrected-t	4.76	-2.08	1.02	-0.01	1.65	
corrected-p	**0.00**	**0.02**	0.16	0.50	0.06	

the possibility of time-varying betas. Allowing for time-varying betas is likely to have considerable effects on the performance of pricing models in general and on the significance of individual pricing factors in particular. In order to investigate these effects for the DIPM and the benchmark pricing models, I test rolling-regression versions of all models. The setup I choose for these tests is the following. For each month in the sample, the betas of all portfolios in the cross-section are re-estimated using the previous 60 months as the estimation period. Then, a cross-sectional OLS regression is performed employing these estimates as the regressors. Hence, in contrast to the cross-sectional regressions documented above, the second-stage estimation of risk prices in a rolling regression uses different beta estimates in every instant.

I provide results of these rolling regressions in appendix A. Considering the rolling regression results for the benchmark models in table 12, some interesting differences to the total sample estimations stand out. First, in terms of Fama-MacBeth t-statistics, the risk premiums associated with rolling regression beta estimates of most of the individual factors differ from their total sample counterparts. In the case of the ICAPM, this change is particularly striking. Contrary to the total sample estimation where the term spread and CPI growth have been identified as the state variables that earn a risk premium, the change in industrial production is now detected to be the only significant pricing factor. Second, employing the rolling regression approach, the factors scaled by $\widehat{cay}$ in Lettau and Ludvigson's conditional CCAPM and HCCAPM exhibit t-statistics that are significantly different from zero. This contrasts the results that I have obtained from the total sample estimations using monthly data. However, the good overall performance of the quarterly data versions of the two scaled models is still unattained in rolling regression estimations using monthly data.

The rolling regression results of the diffusion index pricing model, reported in table 13 in the appendix, do also differ considerably from their total sample estimation counterparts. While in the unbalanced panel version of the DIPM more diffusion indexes of higher order are now identified as significant, none of the balanced panel Stock-Watson factors turns out to earn a significant risk premium. However, additional tests of these models using different sample periods and different rolling regression intervals show that the performance of the DIPM remains highly susceptible to modifications of these parameters. Thus, the model does not provide reliable indications about the risk premiums associated with certain diffusion indexes.

The results obtained from the subperiod joint tests and the rolling regressions show that the choice of the sample period in the first-stage estimation of the betas can have an important influence on the outcomes of Fama-MacBeth regressions. Hence, inference based on this procedure calls for caution. Ideally, Fama-MacBeth regressions should be carried out using different sample periods or different rolling regression intervals. Only in the case of similar results, these should be interpreted in favor of the tested model. However, this is rarely done in praxis. Instead, some authors complement their presentation with results obtained from alternative estimation methods such as GMM.

What can be concluded from the results presented in the previous sections? First, the diffusion index pricing model performs well in cross-sectional regressions, but largely fails the tests of model misspecification. This shows that although intuitively appealing and widely used, the cross-sectional regression methodology suggested by Fama and MacBeth (1973) appears to be unsuited to detect certain types of model misspecification. Second, those diffusion indexes that are significantly priced in cross-sectional tests capture only a small portion of the total variation of all series in the data set. However, the correlation patterns between these factors and individual variables are broadly consistent with economic intuition. It is argued that an "informed pre-selection" of the time series to include in the data set from which the diffusion indexes are extracted is likely to result in a better performance of the DIPM. Third, the benchmark models largely exhibit similar weaknesses as the DIPM and thus cannot be considered as generally performing better. The Fama-French three-factor model is shown to be the only model which yields good results in cross-sectional regression tests *and* which is robust to tests of misspecification. Altogether, the results presented in this paper imply that there is at best mixed empirical evidence for the existence of stable linear relationships between the cross-section of average stock returns and the beta risk associated with macroeconomic factors.

5.9 Testing Diffusion Indexes as Conditioning Variables

Since diffusion indexes capture the common variation of a large number of business cycle-related time series and furthermore exhibit outstanding forecasting power for different economic variables, I have argued above that they are promising candidates for instruments in conditional factor pricing models. To test this assumption empirically, I have carried out Fama-MacBeth regressions for a conditional version of the consumption CAPM similar to the one proposed by Lettau and Ludvigson (2001b).

Instead of searching for the diffusion indexes that perform best in the Fama-MacBeth regressions, I now explicitly choose those Stock-Watson factors as conditioning variables that capture the largest amount of common variation in the time series of the data set. Appendix B shows that these are the second of the factors extracted from the unbalanced panel and the first of the factors extracted from the balanced panel. As has been reported above, the monthly time series of $\widehat{cay}$ does a much poorer job as a scaling variable than its quarterly counterpart, this being due to preliminary intrapolations of one of the variables it is derived from. Since the diffusion indexes are solely available in monthly frequency, a comparison with $\widehat{cay}$ will not be very rich in content. In my tests, I thus compare the performance of diffusion indexes as instruments in conditional pricing models with that of alternative variables. I have argued in section 4.7 that the observed variation of risk premiums over the business cycle might be well captured by using a business-cycle related time series as the scaling variable. Several standard macroeconomic time series variables present itself for such a purpose. The growth rate of industrial production, for example, is a simple indicator of the current state of the economy. Further, interest rates have been shown to be closely linked to the business cycle. For convenience I thus use ΔIP and $Tspr$ - two of the four ICAPM pricing factors - as benchmark conditioning variables in my tests of the scaled CCAPM. In addition, based on Daniel and Titman's (1997) finding that stock return patterns in January are systematically different from those in non-January months, I use as a third benchmark scaling variable a January dummy that takes the value 1 for each January and is zero otherwise.

Table 8 summarizes the results of Fama-MacBeth regressions that I have carried out to test the conditional CCAPM using different instruments. According to these results, the January dummy is the scaling variable that increases the explanatory power of aggregate consumption growth the most considerably. Somehow surprisingly, neither the growth rate of industrial production nor the term spread seem to capture well the variation of risk prices over the business cycle.[54] Although they turn out to work better than the term

[54] Lettau and Ludvigson (2001b) report that in their tests of the scaled CCAPM using quarterly data, $\widehat{cay}$ largely outperforms other instruments such as the term spread or the yield spread. However, since they do not provide detailed results, it is difficult to say whether this is in line with my findings or not. Somewhat contrary to my result, Hodrick and Zhang (2000) find that using the cyclical component of industrial production as an instrument improves the performance of several models.

Table 8: **Regression Results for the CCAPM Using Different Scaling Variables**

This table presents results from Fama-MacBeth regressions of the conditional consumption CAPM where different scaling variables have been used. All models are of the form

$$E[R^i] = E[R_0] + \beta_{i,z}\lambda_z + \beta_{i,\Delta c}\lambda_{\Delta c} + \beta_{i,z\Delta c}\lambda_{z\Delta c},$$

where $\beta_{i,z}$, $\beta_{i,\Delta c}$, and $\beta_{i,z\Delta c}$ denote the coefficients in time series regressions of R^i_{t+1} on z_t, Δc_{t+1}, and $z_t \cdot \Delta c_{t+1}$, respectively. The variable to be explained is the cross-section of the value-weighted returns on 100 Fama-French portfolios. The sample period is 1963:07-1998:12. The different instruments used are the growth rate of industrial production (panel A), the term spread (panel B), a January dummy (panel C), the second diffusion index extracted from the unbalanced data set (panel D), and the first diffusion index extracted from the balanced data set (panel E). For each regression, estimates of the risk prices of all factors, their uncorrected and Shanken-corrected t-values, and the associated p-values are reported. Further, the cross-sectional R^2, its degrees-of-freedom adjusted version and the average absolute pricing error, $\mu_{|\hat{\alpha}|}$, are provided.

A.	$E[R_{0,t}]$	$Tspr_t$	Δc_{t+1}	$Tspr_t \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.32	-0.35	-0.24	0.00	0.18	0.17	1.98
t-value	6.26	-0.69	-2.61	0.32			
p-value	**0.00**	0.25	**0.01**	0.37			
corrected-t	5.14	-0.56	-2.12	0.26			
corrected-p	**0.00**	0.29	**0.02**	0.40			
B.	$E[R_{0,t}]$	ΔIP_t	Δc_{t+1}	$\Delta IP_t \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.42	0.12	-0.23	-0.00	0.17	0.15	2.01
t-value	5.73	0.88	-2.49	-0.21			
p-value	**0.00**	0.19	**0.01**	0.42			
corrected-t	4.80	0.72	-2.06	-0.17			
corrected-p	**0.00**	0.24	**0.02**	0.43			
C.	$E[R_{0,t}]$	JAN_t	Δc_{t+1}	$\text{JAN}_t \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.23	-0.01	-0.08	-0.00	0.37	0.36	1.89
t-value	5.76	-0.28	-1.23	-2.74			
p-value	**0.00**	0.39	0.11	**0.00**			
corrected-t	4.74	-0.22	-0.99	-2.22			
corrected-p	**0.00**	0.41	0.16	**0.01**			
D.	$E[R_{0,t}]$	$f_{2,t}$	Δc_{t+1}	$f_{2,t} \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.45	-0.98	-0.23	0.01	0.25	0.23	1.92
t-value	6.91	-0.97	-2.27	1.78			
p-value	**0.00**	0.17	**0.01**	**0.04**			
corrected-t	5.52	-0.76	-1.79	1.38			
corrected-p	**0.00**	0.22	**0.04**	0.08			
E.	$E[R_{0,t}]$	$f_{1,t}$	Δc_{t+1}	$f_{1,t} \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.48	-1.30	-0.21	0.00	0.25	0.23	1.92
t-value	6.05	-1.74	-2.21	0.73			
p-value	**0.00**	**0.04**	**0.01**	0.23			
corrected-t	4.93	-1.38	-1.78	0.58			
corrected-p	**0.00**	0.09	**0.04**	0.28			

spread and the monthly change of industrial production, both diffusion indexes that I use as instruments fall behind the January dummy in terms of their t-statistics. Accordingly, among all five models considered, the CCAPM scaled by the January dummy exhibits the largest cross-sectional R^2 and the smallest average absolute pricing error. However, since none of the macroeconomic variables is closely linked to calendar time and thus able to capture the particular return characteristics that the size and book-to-market sorted stock portfolios exhibit in January, this result does not necessarily discredit the use of diffusion indexes as conditioning variables.

Whether the diffusion indexes do have predictive power as scaling variables that goes beyond the January effect can be tested by controlling for the latter. Table 14 in appendix A reports results from Fama-MacBeth regressions of the conditional CCAPM using the January dummy and a second instrument as scaling variables. There is indeed some empirical support for the view that the January dummy leaves explanatory power to the other conditioning variables, indicating that these capture some of the variation of risk prices over the business cycle. In particular, using either of the two diffusion indexes as an additional instrument raises the cross-sectional R^2 slightly from initial 37 percent to 42 percent and also slightly diminishes the average absolute pricing error. Further, the growth rate of aggregate consumption scaled by the unbalanced panel diffusion index earns a significant risk premium suggesting that investors care about consumption growth conditional on their information about the current state of the economy. Notice that I have also tested the conditional models for misspecification bias by running Fama-MacBeth regressions including firm characteristics. The results of these tests which are not reported here, provide evidence for misspecification bias in *all* tested versions of the model, including the CCAPM scaled by the January dummy. Interestingly, this variable largely captures the size effect (SIZE becomes highly insignificant when tested jointly with JAN), but is unable to attenuate the significance of the portfolio book-to-market ratio. This might indicate that the size effect is indeed a January effect. One possible explanation for such a finding would be that many money managers reorganize their portfolios at the beginning of each year with a tendency to invest in small stocks.

All in all, my results do not clearly support or reject the hypothesis that diffusion indexes can be useful instruments in conditional factor pricing models. I have employed those Stock-Watson factors as conditioning variables in a scaled version of the CCAPM that capture the largest portion of the total variation in the unbalanced and balanced panel, respectively. Although these diffusion indexes turn out to do a better job than the monthly change of industrial production and the term spread, they are outperformed by a January dummy. In terms of cross-sectional R^2's and average absolute pricing errors, none of the scaled models comes close to the Fama-French three-factor model or the diffusion index pricing model.

6 Conclusion

There is a growing body of theoretical and empirical research suggesting that the variation of stock returns across assets and across time is linked to fluctuations of macroeconomic variables. During the last two decades, a number of stylized facts have been discovered that do not match a classic result of asset pricing theory stating that the expected return of an asset is linearly related to the covariance of its return with the return of the market portfolio. For example, the observed return patterns of stocks which are sorted by size and firms' book-to-market equity ratio are not consistent with this view and thus have attracted the attention of many researchers. One of the models that have been proposed to explain these patterns is the widely recognized three-factor model of Fama and French (1993). On the basis of stock-specific information, these authors construct two portfolios the returns of which they show to capture a large portion of the cross-sectional variation of average stock returns. Fama and French argue that these portfolios mimic unobservable risk factors that are related to "financial distress" and thus to recessions. Another recent contribution to the research dealing with cross-sectional patterns of stock returns has been made by Lettau and Ludvigson (2001b). Starting from an investor's intertemporal budget constraint, they argue that expected future returns on asset wealth can be summarized by the deviations of consumption, asset wealth and labor income from their shared common trend. Based on this result, Lettau and Ludvigson show that the estimated residual from the cointegrating relationship between these three variables has predictive power for stock returns in both the time series and the cross-sectional dimension. Further studies have analyzed the links between stock returns and the business cycle. The set of macroeconomic variables that have so far been proposed as pricing factors includes, among others, interest rate spreads, the innovations to inflation, output, and investment, and news related to future GDP growth.

Since a variety of macroeconomic variables have been shown to be related to stock returns, it seems reasonable to suppose that there exist summary measures of economic activity that capture the different sources of systematic risk that investors are concerned about. With the purpose of forecasting macroeconomic variables using many predictors, Stock and Watson (1998) have recently proposed an estimation methodology for the extraction of a small number of common factors from a large panel of time series variables. Estimated on the basis of principal components analysis, these factors by construction summarize the common variation of all variables in the data set. Empirical applications of Stock and Watson's method show that the extracted factors, labelled as diffusion indexes, predict different measures of economic activity better than all considered benchmark forecasting methods. An extension of the estimation procedure allows for unbalanced panels, i.e. panels that include time series variables with missing observations or of different frequency.

In view of the way the Stock-Watson diffusion indexes are constructed, the central hypothesis of this paper is that they may act as summary measures of different sources of macroeconomic risk that are of concern to investors. Applied to the problem of explaining the cross-section of average returns on size and book-to-market sorted stock portfolios, this hypothesis has been tested empirically. The outcomes of the performed tests have been compared with estimation results for the CAPM, a version of the ICAPM, the Fama-

French three-factor model, and Lettau and Ludvigson's conditional consumption CAPM and human capital CAPM. It has further been examined whether diffusion indexes can successfully be used as instruments in conditional factor models. All tests have been carried out using the cross-sectional regression methodology of Fama and MacBeth (1973). Several drawbacks of this procedure have been diagnosed and discussed in the text. Consequently, the estimated models have been confronted with additional tests for model misspecification.

The use of diffusion indexes as factors in a pricing model is novel to the literature and therefore my investigation provides some new insights concerning two issues: the range of potential uses of the Stock-Watson procedure, and the usefulness of summary measures of economic activity as factors or scaling variables in pricing models. Since my empirical results are mixed, I am reluctant to either reject the hypothesis that diffusion indexes are useful factors or to strongly advocate this view. Standard test statistics and measures of model fit related to the Fama-MacBeth methodology indicate that the diffusion index pricing model performs well compared to the considered benchmark models. However, tests for model misspecification display that the relationship found is at best unstable over time. Yet, since the benchmark models also mainly fail these tests, a ready conclusion about the relative performance of the diffusion indexes compared to standard pricing factors is unwarranted. The tests of the usefulness of the Stock-Watson factors as conditioning variables in scaled factor models similarly yield conflicting results. Although the considered diffusion indexes perform better than other standard macroeconomic variables, the overall fit of the tested conditional models is unsatisfactory.

The inconsistency of my empirical results gives ground for further investigation in this field of study. Several potential lines of research present themselves. First, in order to rule out erroneous inference due to methodological deficiencies, the investigation that has been carried out in this paper could be repeated employing a different estimation methodology such as the time series regression approach of Black, Jensen, and Scholes (1972) or GMM. Second, as yet standard static principal components analysis has mostly been used to extract common factors in the sense of Ross'(1976) arbitrage pricing theory (APT) from a large number of asset returns. Since Stock and Watson's method represents a dynamic generalization of standard static factor models, it would be interesting to apply this procedure to a panel of asset returns and to compare the outcomes with those obtained from the classic static analysis. As an alternative to the Stock-Watson methodology, one could further employ the dynamic factor model approach of Forni et al. (2001a). Third, an interesting way of studying the relationships between macroeconomic and financial variables consists in the construction of economic "tracking" portfolios. Such tracking portfolios are the linear projections of individual macroeconomic time series variables or their innovations on a set of return data. Thus, by construction they capture all the common variation in the return series and the macroeconomic variables. This approach has first been advocated by Breeden, Gibbons, and Litzenberger (1989) and has recently been readopted by Lamont (1999) and Vassalou (2002) to study the variation of risk premia across assets and across time. Adopting this approach, one possible extension of the investigation carried out in this paper would be to construct tracking portfolios for diffusion indexes and to analyze their time series behavior and correlation patterns in order to identify those variables that capture the largest part of the variation in asset returns. Fourth, money managers often select assets either from a top-down or from a bottom-up perspective. While the former consists in choosing stocks from a particular industry or country, the latter aims at selecting individual stocks on the basis of firm-specific information. Since it is based on aggregate data, the diffusion index pricing approach thus might be more promising when used to uncover potential relationships between returns and economic variables that arise due to top-down-based trading. Hence, it would be interesting to extract dynamic factors from a data set containing series from different industries or countries, and to study whether these diffusion indexes are able to explain the variation of returns across the same industries or countries. More generally, since the choice of the data set appears to influence the outcome of the Stock-Watson procedure substantially, a method of suitably pre-selecting the series to include in the panel should be developed.

The above suggestions are only a few examples of a wide range of potential applications of dynamic factor models in asset pricing theory and praxis. Since financial markets and the real side of the economy are tightly interwoven, it is unlikely that economists be able to boil down all the interactions of financial and macroeconomic variables to a well-defined set of easily interpretable relationships. Instead, they will have to base their analyses on measures of common variation. Currently, dynamic factor models represent one of the most promising approaches to the derivation of such measures.

References

[1] Bernanke, Ben S., and Jean Boivin (2001) : "Monetary Policy in a Data-Rich Environment", *NBER Working Paper* No. 8379

[2] Black, Fischer (1972) : "Capital Market Equilibrium with Restricted Borrowing", *Journal of Business* 45, 444-455

[3] Black, Fischer, Michael C. Jensen, and Myron Scholes (1972) : "The Capital Asset Pricing Model: Some Emprical Tests", in Michael Jensen (ed.), *Studies in the Theory of Capital Markets*, Praeger, New York

[4] Breeden, Douglas T. (1979) : "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities", *Journal of Financial Economics* 7, 265-296

[5] Breeden, Douglas T., Michael R.Gibbons, and Robert Litzenberger (1989) : "Empirical Tests of the Consumption-Oriented CAPM", *Journal of Finance* 44, 231-262

[6] Brennan, Michael J., Ashley W. Wang, and Yihong Xia (2002) : "A Simple Model of Intertemporal Capital Asset Pricing and Its Implications for the Fama-French Three-Factor Model", *unpublished working paper*

[7] Brennan, Michael J., and Yihong Xia (2002) : "tay's as good as cay", *unpublished working paper*

[8] Campbell, John Y. (1996) : "Understanding Risk and Return", *Journal of Political Economy* 104, 298 - 345

[9] Campbell, John Y. (2000) : "Asset Pricing at the Millenium", *NBER Working Paper* No. 7589

[10] Campbell, John Y., Andrew W. Lo, and Craig A. MacKinlay (1997) : *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ

[11] Campbell, John Y., and Gregory Mankiw (1989) : "Consumption, Income, and Interest Rates: Reinterpreting the time series Evidence", in Olivier J. Blanchard and Stanley Fischer (eds.), *NBER Macroeconomics Annual*, MIT Press, Cambridge, MA

[12] Chamberlain, Gary, and Michael Rothschild (1983) : "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets", *Econometrica* 51(5), 1281-1323

[13] Chan, Louis K.C., Jason Karceski, and Josef Lakonishok(1997) : "The Risk and Return from Factors", *NBER Working Paper* No. 6098

[14] Chen, Nai-Fu, Richard Roll, and Stephen A. Ross (1986) : "Economic Forces and the Stock Market", *Journal of Business* 59(3), 383-403

[15] Cochrane, John H. (1996) : "A Cross-Sectional Test of an Investment-Based Asset Pricing Model", *Journal of Political Economy* 104(3), 572-621

[16] Cochrane, John H. (2001) : *Asset Pricing*, Princeton University Press, Princeton, NJ

[17] Connor, Gregory, and Robert A. Korajczyk (1986) : "Performance Measurement with the Arbitrage Pricing Theory", *Journal of Financial Economics* 15, 373-394

[18] Connor, Gregory, and Robert A. Korajczyk (1988) : "Risk and Return in an Equilibrium APT: Application of a New Test Methodology", *Journal of Financial Economics* 21, 255-289

[19] Connor, Gregory, and Robert A. Korajczyk (1993) : "A Test for the Number of Factors in an Approximate Factor Model", *Journal of Finance* 48(4), 1263-1291

[20] Cuthbertson, Keith (1996) : *Quantitative Financial Economics*, John Wiley, Chichester

[21] Daniel, Kent, and Sheridan Titman (1996) : "Evidence on the Characteristics of Cross Sectional Variation in Stock Returns", *NBER Working Paper* No. 5604

[22] Davis, James L. (1994) : "The Cross-Section of Realized Stock Returns: The Pre-COMPUSTAT Evidence", *Journal of Finance* 50, 1579-1593

[23] DeBondt, Werner F. M., and Richard H. Thaler (1985) : "Does the Stock Market Overreact ?", *Journal of Finance* 40, 793-805

[24] Dempster, A.P., N.M. Laird, and D.B. Rubin (1977) : "Maximum Likelihood from Incomplete Data via the EM Algorithm", *Journal of the Royal Statistical Society* Serie B, 39, 1-38

[25] Estrella, A. and Frederic S. Mishkin (1998) : "Predicting US Recessions: Financial Variables as Leading Indicators", *The Review of Economics and Statistics* 80(1), 45-61

[26] Fama, Eugene F., and Kenneth R. French (1992) : "The Cross-Section of Expected Stock Returns", *Journal of Finance* 47, 427-465

[27] Fama, Eugene F., and Kenneth R. French (1993) : "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics* 33, 3-56

[28] Fama, Eugene F., and Kenneth R. French (1995) : "Size and Book-to-Market Factors in Earnings and Returns", *Journal of Finance* 50, 131-155

[29] Fama, Eugene F., and Kenneth R. French (1996) : "Multifactor Explanations of Asset Pricing Anomalies", *Journal of Finance* 51, 55-84

[30] Fama, Eugene F., and Michael Gibbons (1984) : "A Comparison of Inflation Forecasts", *Journal of Monetary Economics* 13, 327-348

[31] Fama, Eugene F., and James D. MacBeth (1973) : "Risk, Return, and Equilibrium: Empirical Tests", *Journal of Political Economy* 81, 607-636

[32] Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin (2001a) : "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting", *unpublished working paper*, ECARES, Université Libre de Bruxelles

[33] Geweke, John (1977) : "The Dynamic Factor Analysis of Economic Time Series", in D.J. Aigner and A.S. Goldberger (eds.), *Latent Variables in Socio-Economic Models*, North-Holland, Amsterdam

[34] Griffin, John M., and Michael L. Lemmon (2002) : "Does Book-to-Market Equity Proxy for Distress Risk or Mispricing?", *Journal of Finance*, forthcoming

[35] Härdle, Wolfgang and Léopold Simar (1998) : "Applied Multivariate Statistical Analysis", *unpublished manuscript*

[36] Hodrick, Robert J., and Xiaoyan Zhang (2000) : "Evaluating the Specification Errors of Asset Pricing Models", *NBER Working Paper* No. 7661

[37] Jagannathan, Ravi, and Zhenyu Wang (1996) : "The Conditional CAPM and the Cross-Section of Expected Returns", *Journal of Finance* 51(1), 3-53

[38] Jagannathan, Ravi, and Zhenyu Wang (1998) : "An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression", *Journal of Finance* 53(4), 1285-1309

[39] Jegadeesh, Narasimham, and Sheridan Titman (1993) : "Returns on Buying Winners and Selling Losers: Implications for Stock Market Efficiency", *Journal of Finance* 48, 65-91

[40] Kan, Raymond, and Chu Zhang (1999) : "Two-Pass Tests of Asset Pricing Models with Useless Factors", *Journal of Finance* 54(1), 203-255

[41] Kothari, S.P., Jay Shanken, and Richard Sloan (1995) : "Another Look at the Cross-Section of Expected Returns, *Journal of Finance* 50, 185-224

[42] Lamont, Owen (1999) : "Economic Tracking Portfolios", *NBER Working Paper* No. 7055

[43] Lettau, Martin, and Sydney Ludvigson (2001a) : "Consumption, Aggregate Wealth, and Expected Stock Returns", *Journal of Finance* 56(3), 815-849

[44] Lettau, Martin, and Sydney Ludvigson (2001b) : "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying", *Journal of Political Economy* 109(6), 1238-1287

[45] Lettau, Martin, and Sydney Ludvigson (2002) : "tay's as good as cay : Reply", *mimeo*, New York University

[46] Liew, Jimmy, and Maria Vassalou (2000) : "Can Book-to-Market, Size and Momentum Be Risk Factors that Predict Economic Growth?", *Journal of Financial Economics* 57, 221-245

[47] Lintner, John (1965) : "Security Prices, Risk, and Maximal Gains from Diversification", *Journal of Finance* 20, 587-615

[48] Menzly, Lior (2001) : "Influential Observations in Cross-Sectional Asset Pricing Tests", *Dissertation Thesis*, University of Chicago

[49] Merton, Robert C. (1972) : "An Analytic Derivation of the Efficient Portfolio Frontier", *Journal of Financial and Quantitative Analysis* 7, 1851-1872

[50] Merton, Robert C. (1973) : "An Intertemporal Capital Asset Pricing Model", *Econometrica* 41, 867-887

[51] Ross, Stephen A. (1976) : "The Arbitrage Theory of Capital Asset Pricing", *Journal of Economic Theory* 13, 341-360

[52] Sargent, Thomas J., and Christopher A. Sims (1977) : "Business Cycle Modeling without Pretending to have Too Much a-priori Economic Theory", in C. Sims et al. (eds.), *New Methods in Business Cycle Research*, Minneapolis: Federal Reserve Bank of Minneapolis

[53] Shanken, Jay (1992) : "On the Estimation of Beta-Pricing Models", *Review of Financial Studies* 5, 1-33

[54] Sharpe, William F. (1964) : "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance* 19, 425-442

[55] Stock, James H., and Mark W. Watson (1993) : "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems", *Econometrica* 6, 783-820

[56] Stock, James H., and Mark W. Watson (1998) : "Diffusion Indexes", *NBER Working Paper* No. 6702

[57] Stock, James H., and Mark W. Watson (2002) : "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics* 20(2), 147-162

[58] Vassalou, Maria (2002) : "News Related to Future GDP Growth as a Risk Factor in Equity Returns", *Journal of Financial Economics*, forthcoming

Appendix

A Additional Tables and Figures

Table 9: **Benchmark Models - 100 FF portfolios, quarterly data**

The table summarizes the results of the Fama-MacBeth regressions that have been performed for the benchmark pricing models using the cross-section of the value-weighted returns on 100 Fama-French size and book-to-market sorted stock portfolios as the variable to be explained. The sample period is from 1963:Q3 to 1998:Q3. For a description of the estimates and statistics reported see, for instance, table 2 in the text.

Model	Constant	Pricing Factors					Summary Statistics		
CAPM	$E[R_0]$	R^m					R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	4.02	-0.05					0.01	-0.01	3.51
t-value	3.95	-0.05							
p-value	**0.00**	0.48							
corrected-t	3.95	-0.04							
corrected-p	**0.00**	0.48							
ICAPM	$E[R_0]$	R^m	TSpr	DSpr	ΔIP	ΔCPI	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	3.97	-0.06	0.87	-0.19	0.12	-0.34	0.54	0.52	3.11
t-value	4.19	-0.06	3.95	-1.53	0.52	-3.01			
p-value	**0.00**	0.48	**0.00**	0.06	0.30	**0.00**			
corrected-t	3.02	-0.04	2.68	-1.05	0.35	-2.00			
corrected-p	**0.00**	0.49	**0.00**	0.15	0.36	**0.02**			
FF3F	$E[R_0]$	R^m	SMB	HML			R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	4.20	-0.79	0.45	1.40			0.63	0.63	2.99
t-value	4.91	-0.71	0.87	3.08					
p-value	**0.00**	0.24	0.19	**0.00**					
corrected-t	4.71	-0.58	0.61	2.18					
corrected-p	**0.00**	0.28	0.27	**0.02**					
c. CCAPM	$E[R_0]$	cay	Δc	$cay \cdot \Delta c$			R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	3.29	-0.27	0.17	0.00			0.43	0.42	3.31
t-value	4.46	-0.95	1.94	1.74					
p-value	**0.00**	0.17	**0.03**	**0.04**					
corrected-t	3.36	-0.69	1.38	1.27					
corrected-p	**0.00**	0.25	0.08	0.10					
c. HCCAPM	$E[R_0]$	cay	R^m	$cay \cdot R^m$	Δy	$cay \cdot \Delta y$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	3.81	0.05	0.09	0.04	0.42	0.01	0.50	0.48	3.09
t-value	3.93	0.27	0.08	2.60	2.77	2.78			
p-value	**0.00**	0.39	0.47	**0.01**	**0.00**	**0.00**			
corrected-t	3.17	0.20	0.06	1.93	2.07	2.13			
corrected-p	**0.00**	0.42	0.48	**0.03**	**0.02**	**0.02**			

Table 10: **Benchmark Models - 25 FF portfolios, monthly data**

The table summarizes the results of Fama-MacBeth regressions that have been performed for the benchmark pricing models using the cross-section of the value-weighted returns on 25 Fama-French size and book-to-market sorted stock portfolios as the variable to be explained. The sample period is 1963:07 - 1998:12. All models have the form

$$E[R_i] = E[R_0] + \beta_{i,f_1}\lambda_{f_1} + \ldots + \beta_{i,f_k}\lambda_{f_k},$$

where $\beta_{i,f_1}, \ldots, \beta_{i,f_k}$ denote the beta estimates of the pricing factors, preliminarily determined by regressing R_i on the respective vector of factors. For each model, estimates of λ_{f_i}, their t-values and the associated p-value are reported. The "corrected $t-$ and p-values" take into account the errors-in-variables adjustment as suggested by Shanken (1992). R^2 is the cross-sectional coefficient of determination and $\bar{R}^2$ its degrees of freedom-adjusted version. $\mu_{|\hat{\alpha}|}$ is the average of the absolute values of all $T \cdot N$ pricing errors.

Model	Constant	Pricing Factors					Summary Statistics		
CAPM	$E[R_0]$	R^m					R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.45	-0.22					0.03	0.03	1.26
t-value	5.02	-0.62							
p-value	**0.00**	0.27							
corrected-t	5.01	-0.56							
corrected-p	**0.00**	0.29							
ICAPM	$E[R_0]$	R^m	TSpr	DSpr	ΔIP	ΔCPI	R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.40	-0.26	3.30	-0.13	-0.47	-0.17	0.64	0.57	0.99
t-value	4.60	-0.70	4.92	-0.84	-2.49	-2.60			
p-value	**0.00**	0.25	**0.00**	0.21	**0.01**	**0.01**			
corrected-t	1.80	-0.27	1.92	-0.33	-0.97	-1.01			
corrected-p	**0.04**	0.40	**0.03**	0.37	0.17	0.16			
FF3F	$E[R_0]$	R^m	SMB	HML			R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.77	-0.68	0.08	0.46			0.70	0.68	0.97
t-value	4.88	-1.62	0.56	3.48					
p-value	**0.00**	**0.06**	0.29	**0.00**					
corrected-t	4.77	-1.42	0.40	2.48					
corrected-p	**0.00**	0.08	0.35	**0.01**					
c. CCAPM	$E[R_0]$	cay	Δc	$cay \cdot \Delta c$			R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	1.75	-0.65	-0.33	0.00			0.38	0.33	1.17
t-value	5.50	-0.96	-2.20	0.72					
p-value	**0.00**	0.17	**0.02**	0.24					
corrected-t	3.68	-0.64	-1.47	0.48					
corrected-p	**0.00**	0.26	0.08	0.32					
c. HCCAPM	$E[R_0]$	cay	R^m	$cay \cdot R^m$	Δy	$cay \cdot \Delta y$	R^2	$\bar{R}^2$	$\mu_{\|\hat{\alpha}\|}$
Estimate	2.28	-1.58	-0.82	-0.05	0.08	-0.01	0.29	0.15	1.05
t-value	6.50	-3.00	-2.01	-2.85	0.33	-1.87			
p-value	**0.00**	**0.00**	**0.03**	**0.00**	0.37	**0.04**			
corrected-t	3.62	-1.67	-1.09	-1.58	0.18	-1.03			
corrected-p	**0.00**	**0.05**	0.14	0.06	0.43	0.16			

Table 11: **Subperiod Joint Test - Benchmark Models**

This table presents results of Fama-MacBeth regressions performed for three benchmark models using quarterly data. The two subperiods are 1963:Q3-1980:Q4 (panel A) and 1981:Q1-1998:Q3 (panel B). The variable to be explained is the cross-section of the value-weighted returns on 25 size and book-to-market sorted stock portfolios. For each regression, risk price estimates, their uncorrected and Shanken-corrected t-values, the associated p-values,and the cross-sectional R^2 are reported.

Model	**Constant**	**Pricing Factors**					
		A. Sample Period 1963:Q3 - 1980:Q4					
FF3F	$E[R_0]$	R^m	SMB	HML			R^2
Estimate	1.24	1.27	1.41	1.16			0.90
t-value	0.92	0.73	1.69	1.83			
p-value	0.18	0.24	**0.05**	**0.04**			
corrected-t	0.87	0.60	1.18	1.28			
corrected-p	0.20	0.28	0.13	0.11			
c. CCAPM	$E[R_0]$	cay	Δc	$cay \cdot \Delta c$			R^2
Estimate	1.70	0.35	0.11	0.01			0.84
t-value	1.58	0.82	0.73	2.73			
p-value	0.06	0.21	0.24	**0.01**			
corrected-t	1.14	0.58	0.51	1.91			
corrected-p	0.13	0.28	0.31	**0.03**			
ICAPM	$E[R_0]$	R^m	TSpr	DSpr	ΔIP	ΔCPI	R^2
Estimate	2.55	-3.23	0.53	-0.11	1.00	-0.24	0.84
t-value	2.30	-1.37	1.09	-0.55	0.99	-1.28	
p-value	**0.02**	0.09	0.14	0.29	0.17	0.11	
corrected-t	1.75	-1.00	0.81	-0.40	0.74	-0.89	
corrected-p	**0.05**	0.16	0.21	0.35	0.23	0.19	
		B. Sample Period 1981:Q1 - 1998:Q3					
FF3F	$E[R_0]$	R^m	SMB	HML			R^2
Estimate	9.87	-5.75	-0.60	1.96			0.63
t-value	7.88	-3.70	-1.05	3.11			
p-value	**0.00**	**0.00**	0.15	**0.00**			
corrected-t	6.17	-2.64	-0.66	1.93			
corrected-p	**0.00**	**0.01**	0.26	**0.03**			
c. CCAPM	$E[R_0]$	cay	Δc	$cay \cdot \Delta c$			R^2
Estimate	5.63	-0.48	0.11	0.00			0.71
t-value	7.28	-1.55	0.84	0.27			
p-value	**0.00**	0.07	0.20	0.39			
corrected-t	5.54	-1.13	0.62	0.20			
corrected-p	**0.00**	0.14	0.27	0.42			
ICAPM	$E[R_0]$	R^m	TSpr	DSpr	ΔIP	ΔCPI	R^2
Estimate	9.64	-0.33	0.57	-0.18	-0.26	0.71	0.69
t-value	6.42	-0.12	2.22	-0.87	-0.85	2.97	
p-value	**0.00**	0.45	**0.02**	0.20	0.20	**0.00**	
corrected-t	3.13	-0.06	1.05	-0.42	-0.41	1.44	
corrected-p	**0.00**	0.48	0.15	0.34	0.34	0.08	

Table 12: **Benchmark Models - 100 FF portfolios, 5-Year Rolling Regressions**

The table summarizes the results of 5-year rolling Fama-MacBeth regressions that have been performed for the benchmark pricing models using the cross-section of the value-weighted returns on 100 Fama-French size and book-to-market sorted stock portfolios as the variable to be explained. The sample period is 1968:07 - 1998:12. In every period, the betas are estimated by time series regressions of each portfolio in the cross-section on the vector of pricing factors using the previous 60 months as the sample period. For a description of the estimates and statistics reported see table 10 above.

Model	Constant	Pricing Factors					Summary Statistics		
CAPM	$E[R_0]$	R^m					R^2	$\bar{R}^2$	$\mu_{\vert\hat{\alpha}\vert}$
Estimate	1.48	-0.35					0.35	0.35	1.94
t-value	6.29	-1.31							
p-value	**0.00**	0.10							
corrected-t	6.25	-1.09							
corrected-p	**0.00**	0.14							
ICAPM	$E[R_0]$	R^m	TSpr	DSpr	ΔIP	ΔCPI	R^2	$\bar{R}^2$	$\mu_{\vert\hat{\alpha}\vert}$
Estimate	1.45	-0.27	0.07	-0.05	-0.15	0.00	0.52	0.50	1.80
t-value	6.30	-1.15	0.55	-1.36	-3.03	0.11			
p-value	**0.00**	0.13	0.29	0.09	**0.00**	0.45			
corrected-t	6.09	-0.90	0.43	-0.94	-2.26	0.09			
corrected-p	**0.00**	0.19	0.33	0.17	**0.01**	0.47			
FF3F	$E[R_0]$	R^m	SMB	HML			R^2	$\bar{R}^2$	$\mu_{\vert\hat{\alpha}\vert}$
Estimate	1.42	-0.31	-0.07	0.40			0.60	0.59	1.72
t-value	6.56	-1.50	-0.50	3.02					
p-value	**0.00**	0.07	0.31	**0.00**					
corrected-t	6.49	-0.99	-0.34	2.09					
corrected-p	**0.00**	0.16	0.37	**0.02**					
c. CCAPM	$E[R_0]$	cay	Δc	$cay \cdot \Delta c$			R^2	$\bar{R}^2$	$\mu_{\vert\hat{\alpha}\vert}$
Estimate	1.20	0.23	-0.01	0.02			0.36	0.35	1.93
t-value	4.92	0.24	-0.38	3.06					
p-value	**0.00**	0.40	0.35	**0.00**					
corrected-t	4.74	0.19	-0.32	2.55					
corrected-p	**0.00**	0.43	0.37	**0.01**					
c. HCCAPM	$E[R_0]$	cay	R^m	$cay \cdot R^m$	Δy	$cay \cdot \Delta y$	R^2	$\bar{R}^2$	$\mu_{\vert\hat{\alpha}\vert}$
Estimate	1.39	-0.00	-0.29	-0.00	0.07	-2.42	0.55	0.53	1.83
t-value	6.01	-0.00	-1.17	-0.61	1.12	-2.37			
p-value	**0.00**	0.50	0.12	0.27	0.13	**0.01**			
corrected-t	5.91	-0.00	-0.94	-0.47	0.86	-1.87			
corrected-p	**0.00**	0.50	0.17	0.32	0.20	**0.03**			

Table 13: **DIPM - 100 FF portfolios, 5-Year Rolling Regressions**

The table summarizes the results of 5-year rolling Fama-MacBeth regressions that have been performed for the unbalanced panel and the balanced panel diffusion index pricing model using the cross-section of the value-weighted returns on 100 Fama-French size and book-to-market sorted stock portfolios as the variable to be explained. The sample period is from 1968:07 to 1998:12. In every period, the betas are estimated by time series regressions of each portfolio in the cross-section on the vector of pricing factors using the previous 60 months as the sample period. For each regression, risk price estimates, their uncorrected and Shanken-corrected t-values, the associated p-values,and the cross-sectional R^2 are reported. data

Model	Constant	Pricing Factors													Summary Statistics		
A.1. ubp DIPM	$E[R_0]$	R^m	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.35	-0.18	0.31	-0.45	0.04	0.05	-1.13	0.86	-0.43	-0.07	0.47	-0.72	0.94	0.02	0.68	0.64	1.64
t-value	5.97	-0.89	0.88	-1.20	0.16	0.12	-2.95	2.12	-1.20	-0.19	1.37	-1.92	2.27	0.05			
p-value	**0.00**	0.19	0.19	0.12	0.44	0.45	**0.00**	**0.02**	0.12	0.42	0.09	**0.03**	**0.01**	0.48			
corrected-t	5.46	-0.63	0.70	-0.95	0.10	0.10	-2.35	1.69	-0.94	-0.15	1.05	-1.52	1.83	0.04			
corrected-p	**0.00**	0.27	0.24	0.17	0.46	0.46	**0.01**	**0.05**	0.17	0.44	0.15	0.07	**0.04**	0.49			
A.2. ubp DIPM	$E[R_0]$	R^m					f_5	f_6				f_{10}	f_{11}		R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.31	-0.35					-1.14	1.37				-0.58	0.74		0.54	0.52	1.79
t-value	5.58	-1.50					-2.46	2.76				-1.45	1.53				
p-value	**0.00**	0.07					**0.01**	**0.00**				0.08	0.06				
corrected-t	5.08	-1.12					-2.04	2.29				-1.17	1.27				
corrected-p	**0.00**	0.13					**0.02**	**0.01**				0.12	0.10				
B.1. bp DIPM	$E[R_0]$	R^m	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.41	-0.25	0.17	0.10	-0.09	0.50	0.60	0.20	-0.15	-0.24	-0.52	-0.08	0.51	-0.03	0.68	0.63	1.65
t-value	6.12	-1.23	0.48	0.29	-0.26	1.28	1.59	0.54	-0.39	-0.61	-1.27	-0.20	1.30	-0.08			
p-value	**0.00**	0.11	0.32	0.39	0.40	0.10	0.06	0.29	0.35	0.27	0.10	0.42	0.10	0.47			
corrected-t	5.91	-0.91	0.39	0.23	-0.21	1.05	1.34	0.45	-0.32	-0.51	-1.06	-0.16	1.08	-0.07			
corrected-p	**0.00**	0.18	0.35	0.41	0.42	0.15	0.09	0.33	0.37	0.31	0.15	0.44	0.14	0.47			

Table 14: **Controlling for the January Effect**

This table presents results from Fama-MacBeth regressions of the conditional consumption CAPM with two instruments, the January dummy and an alternative scaling variables. Thus, all models are of the form

$$E[R^i] = E[R_0] + \beta_{i,JAN}\lambda_{JAN} + \beta_{i,\Delta c}\lambda_{\Delta c} + \beta_{i,JAN\Delta c}\lambda_{JAN\Delta c} + \beta_{i,z}\lambda_z + \beta_{i,z\Delta c}\lambda_{z\Delta c},$$

where $\beta_{i,JAN}$, $\beta_{i,\Delta c}$, $\beta_{i,JAN\Delta c}$, $\beta_{i,z}$, and $\beta_{i,z\Delta c}$ denote the coefficients in time series regressions of R^i_{t+1} on JAN_t, Δc_{t+1}, $\text{JAN}_t \cdot \Delta c_{t+1}$, z_t, and $z_t \cdot \Delta c_{t+1}$, respectively. The variable to be explained is the cross-section of the value-weighted returns on 100 Fama-French portfolios. The sample period is 1963:07-1998:12. Panel A reports results using the growth rate of industrial production as the second scaling variable. The other panels refer to the following additional instruments: the term spread (panel B), the second diffusion index of the unbalanced data set (panel C), and the first diffusion index extracted from the balanced data set (panel D). For each regression, estimates of the risk prices of all factors, their uncorrected and Shanken-corrected t-values, and the associated p-values are reported. Further, the cross-sectional R^2, its degrees-of-freedom adjusted version and the average absolute pricing error, $\mu_{|\hat{\alpha}|}$, are provided.

A.	$E[R_{0,t}]$	JAN_t	Δc_{t+1}	$\text{JAN}_t \cdot \Delta c_{t+1}$	ΔIP_t	$\Delta IP_t \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.23	-0.01	-0.08	-0.00	0.09	-0.00	0.39	0.36	1.84
t-value	6.22	-0.23	-1.20	-2.67	0.60	-1.38			
p-value	**0.00**	0.41	0.12	**0.00**	0.28	0.09			
corrected-t	4.92	-0.18	-0.93	-2.09	0.46	-1.06			
corrected-p	**0.00**	0.43	0.18	**0.02**	0.32	0.15			
B.	$E[R_{0,t}]$	JAN_t	Δc_{t+1}	$\text{JAN}_t \cdot \Delta c_{t+1}$	$Tspr_t$	$Tspr_t \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.26	-0.00	-0.09	-0.00	-0.36	0.00	0.39	0.36	1.84
t-value	5.82	-0.12	-1.17	-2.68	-0.87	0.47			
p-value	**0.00**	0.45	0.12	**0.00**	0.19	0.32			
corrected-t	4.74	-0.09	-0.94	-2.15	-0.70	0.37			
corrected-p	**0.00**	0.46	0.18	**0.02**	0.24	0.35			
C.	$E[R_{0,t}]$	JAN_t	Δc_{t+1}	$\text{JAN}_t \cdot \Delta c_{t+1}$	$f_{2,t}$	$f_{2,t} \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.31	-0.00	-0.07	-0.00	-0.82	0.01	0.42	0.40	1.76
t-value	6.51	-0.03	-1.06	-2.94	-0.86	1.76			
p-value	**0.00**	0.49	0.15	**0.00**	0.20	**0.04**			
corrected-t	5.29	-0.02	-0.83	-2.34	-0.69	1.38			
corrected-p	**0.00**	0.49	0.20	**0.01**	0.25	0.09			
D.	$E[R_{0,t}]$	JAN_t	Δc_{t+1}	$\text{JAN}_t \cdot \Delta c_{t+1}$	$f_{1,t}$	$f_{1,t} \cdot \Delta c_{t+1}$	R^2	$\bar{R}^2$	$\mu_{\lvert\hat{\alpha}\rvert}$
Estimate	1.30	-0.00	-0.06	-0.00	-0.81	0.00	0.42	0.40	1.77
t-value	6.60	-0.04	-1.03	-2.96	-1.01	1.36			
p-value	**0.00**	0.49	0.15	**0.00**	0.16	0.09			
corrected-t	5.41	-0.03	-0.82	-2.38	-0.81	1.09			
corrected-p	**0.00**	0.49	0.21	**0.01**	0.21	0.14			

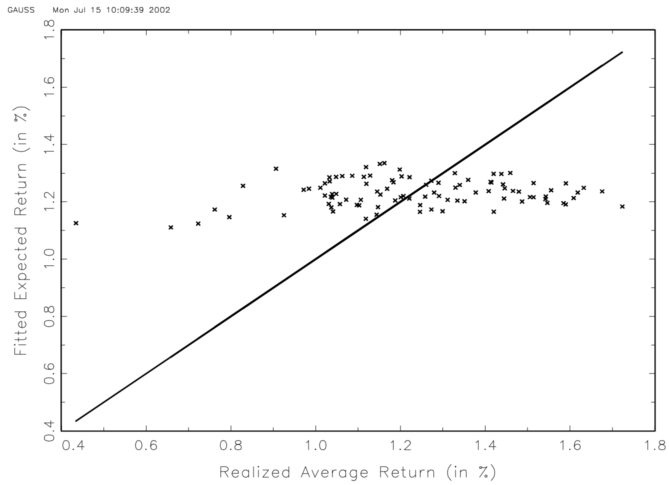


Figure 4: **CAPM - Fitted expected versus realized average monthly returns.** Each scatter point in the graph represents one of the 97 size and book-to-market sorted Fama-French portfolios. The fitted returns have been obtained using the Fama-MacBeth cross-sectional regression methodology. The sample period is 1963:07-1998:12. Visibly, the CAPM explains almost none of the variation of average monthly returns across the 97 Fama-French portfolios. The cross-sectional R^2 of this regression is 3 percent.

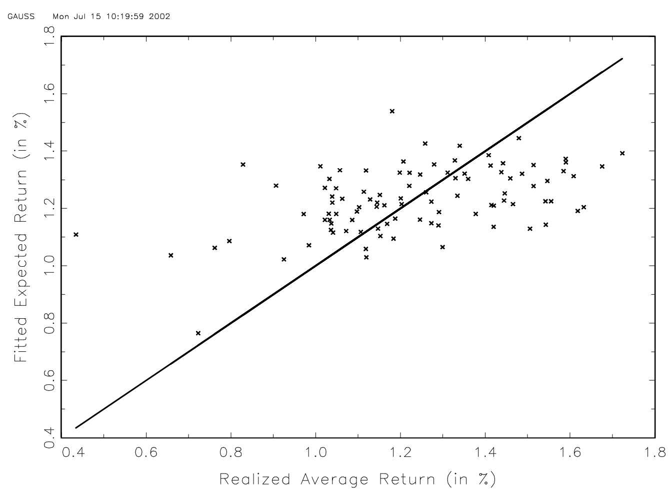


Figure 5: **ICAPM - Fitted expected versus realized average monthly returns.** Each scatter point in the graph represents one of the 97 size and book-to-market sorted Fama-French portfolios. The fitted returns have been obtained using the Fama-MacBeth cross-sectional regression methodology. The sample period is 1963:07-1998:12. Although less markedly than the CAPM, the ICAPM also largely fails in explaining the cross-section of returns on 97 Fama-French portfolios.The cross-sectional R^2 of this regression is 22 percent.

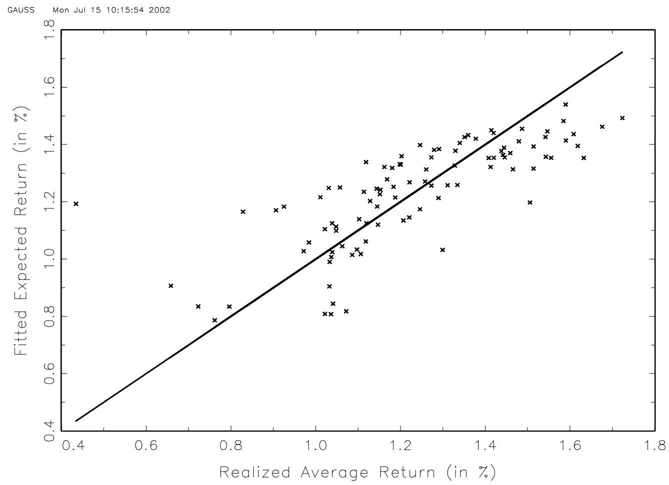


Figure 6: **FF3F - Fitted expected versus realized average monthly returns.** Each scatter point in the graph represents one of the 97 size and book-to-market sorted Fama-French portfolios. The fitted returns have been obtained using the Fama-MacBeth cross-sectional regression methodology. The sample period is 1963:07-1998:12. The Fama-French three-factor model explains the cross-section of average monthly returns on 97 ME and BE/ME-sorted stock portfolios visibly better than the other benchmark models, but still has some difficulties predicting very high and very low average returns. The cross-sectional R^2 of this regression is 58 percent.

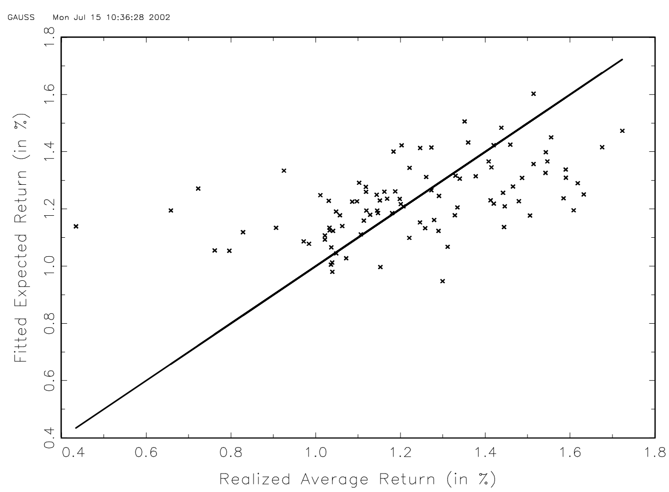


Figure 7: **c. CCAPM - Fitted expected versus realized average monthly returns.** Each scatter point in the graph represents one of the 97 size and book-to-market sorted Fama-French portfolios. The fitted returns have been obtained using the Fama-MacBeth cross-sectional regression methodology. The sample period is 1963:07-1998:12. As reported in the text, the conditional CCAPM of Lettau and Ludvigson largely fails in explaining the cross-section of average monthly returns on 97 ME and BE/ME-sorted stock portfolios. This figure displays that the model in particular has severe problems explaining low returns. The cross-sectional R^2 of this regression is 28 percent.

B The Stock-Watson Data Set

This appendix reports the data from which the diffusion indexes have been extracted. As noted earlier, they are provided on Mark Watson's homepage and were originally taken from the DBI-McGraw Hill Basic Economics database. Some of the series have been produced by the authors' own calculations. In the first column, the mnemonic of each series is given. The second column indicates the period over which the respective series is available. An asterisk after the date denotes a series that is included in the unbalanced panel but not in the balanced panel. The third column reports a transformation code indicating the manner in which a series has been manipulated preliminarily. The transformation codes are: 1 = no transformation; 2 = first difference; 4 = logarithm; 5 = first difference of logarithms; 6 = second difference of logarithms. The forth column gives a short description of the series. For the extraction of the factors used in the DIPM, I have excluded all series related to the stock market (series No. 118 - 131) in order to avoid data-snooping effects. After elimination of the stock market data, the unbalanced panel contains a total of 201 time series out of which 142 form the balanced panel.

Real output and income

1. IP	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: TOTAL INDEX (1992=100,SA)
2. IPP	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: PRODUCTS, TOTAL (1992=100,SA)
3. IPF	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: FINAL PRODUCTS (1992=100,SA)
4. IPC	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: CONSUMER GOODS (1992=100,SA)
5. IPCD	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: DURABLE CONSUMER GOODS (1992=100,SA)
6. IPCN	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: NONDURABLE CONSUMER GOODS (1992=100,SA)
7. IPE	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: BUSINESS EQUIPMENT (1992=100,SA)
8. IPI	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: INTERMEDIATE PRODUCTS (1992=100,SA)
9. IPM	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: MATERIALS (1992=100,SA)
10. IPMD	1959:01-1998:12*	5	INDUSTRIAL PRODUCTION: DURABLE GOODS MATERIALS (1992=100,SA)
11. IPMND	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: NONDURABLE GOODS MATERIALS (1992=100,SA)
12. IPMFG	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: MANUFACTURING (1992=100,SA)
13. IPD	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: DURABLE MANUFACTURING (1992=100,SA)
14. IPN	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: NONDURABLE MANUFACTURING (1992=100,SA)
15. IPMIN	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: MINING (1992=100,SA)
16. IPUT	1959:01-1998:12	5	INDUSTRIAL PRODUCTION: UTILITIES (1992-=100,SA)
17. IPX	1967:01-1998:12*	1	CAPACITY UTIL RATE: TOTAL INDUSTRY (% OF CAPACITY,SA)(FRB)
18. IPXMCA	1959:01-1998:12	1	CAPACITY UTIL RATE: MANUFACTURING,TOTAL(% OF CAPACITY,SA)(FRB)
19. IPXDCA	1967:01-1998:12*	1	CAPACITY UTIL RATE: DURABLE MFG (% OF CAPACITY,SA)(FRB)
20. IPXNCA	1967:01-1998:12*	1	CAPACITY UTIL RATE: NONDURABLE MFG (% OF CAPACITY,SA)(FRB)
21. IPXMIN	1967:01-1998:12*	1	CAPACITY UTIL RATE: MINING (% OF CAPACITY,SA)(FRB)
22. IPXUT	1967:01-1998:12*	1	CAPACITY UTIL RATE: UTILITIES (% OF CAPACITY,SA)(FRB)
23. PMI	1959:01-1998:12	1	PURCHASING MANAGERS' INDEX (SA)
24. PMP	1959:01-1998:12	1	NAPM PRODUCTION INDEX (PERCENT)
25. GMPYQ	1959:01-1998:12*	5	PERSONAL INCOME (CHAINED) (SERIES #52) (BIL 92$,SAAR)
26. GMYXPQ	1959:01-1998:12	5	PERSONAL INCOME LESS TRANSFER PAYMENTS (CHAINED) (#51) (BIL 92$,SAAR)

Employment and hours

27. LHEL	1959:01-1998:12	5	INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA)
28. LHELX	1959:01-1998:12	4	EMPLOYMENT: RATIO; HELP-WANTED ADS:NO. UNEMPLOYED CLF
29. LHEM	1959:01-1998:12	5	CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)
30. LHNAG	1959:01-1998:12	5	CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)
31. LHUR	1959:01-1998:12	1	UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%,SA)
32. LHU680	1959:01-1998:12	1	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)
33. LHU5	1959:01-1998:12	1	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)
34. LHU14	1959:01-1998:12	1	UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)
35. LHU15	1959:01-1998:12	1	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)
36. LHU26	1959:01-1998:12	1	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)
37. LPNAG	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: TOTAL (THOUS.,SA)
38. LP	1959:01-1998:12	5	EMPLOYEES ON NONAG PAYROLLS: TOTAL, PRIVATE (THOUS,SA)
39. LPGD	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: GOODS-PRODUCING (THOUS.,SA)
40. LPMI	1959:01-1998:12*	5	EMPLOYEES ON NONAG. PAYROLLS: MINING (THOUS.,SA)
41. LPCC	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: CONTRACT CONSTRUCTION (THOUS.,SA)
42. LPEM	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: MANUFACTURING (THOUS.,SA)
43. LPED	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: DURABLE GOODS (THOUS.,SA)
44. LPEN	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: NONDURABLE GOODS (THOUS.,SA)
45. LPSP	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: SERVICE-PRODUCING (THOUS.,SA)
46. LPTU	1959:01-1998:12*	5	EMPLOYEES ON NONAG. PAYROLLS: TRANS. & PUBLIC UTILITIES (THOUS.,SA)
47. LPT	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: WHOLESALE & RETAIL TRADE (THOUS.,SA)
48. LPFR	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: FINANCE,INSUR. & REAL ESTATE (THOUS.,SA
49. LPS	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: SERVICES (THOUS.,SA)
50. LPGOV	1959:01-1998:12	5	EMPLOYEES ON NONAG. PAYROLLS: GOVERNMENT (THOUS.,SA)

51. LW	1964:01-1998:12*	2	AVG. WEEKLY HRS. OF PROD. WKRS.: TOTAL PRIVATE (SA)
52. LPHRM	1959:01-1998:12	1	AVG. WEEKLY HRS. OF PRODUCTION WKRS.: MANUFACTURING (SA)
53. LPMOSA	1959:01-1998:12	1	AVG. WEEKLY HRS. OF PROD. WKRS.: MFG.,OVERTIME HRS. (SA)
54. PMEMP	1959:01-1998:12	1	NAPM EMPLOYMENT INDEX (PERCENT)

Real retail, manufacturing and trade sales

55. MSMTQ	1959:01-1998:12	5	MANUFACTURING & TRADE: TOTAL (MIL OF CHAINED 1992 $)(SA)
56. MSMQ	1959:01-1998:12	5	MANUFACTURING & TRADE:MANUFACTURING;TOTAL(MIL OF CHAINED 1992 $)(SA)
57. MSDQ	1959:01-1998:12	5	MANUFACTURING & TRADE:MFG; DURABLE GOODS (MIL OF CHAINED 1992 $)(SA)
58. MSNQ	1959:01-1998:12	5	MANUFACT. & TRADE:MFG;NONDURABLE GOODS (MIL OF CHAINED 1992 $)(SA)
59. WTQ	1959:01-1998:12	5	MERCHANT WHOLESALERS: TOTAL (MIL OF CHAINED 1992 $)(SA)
60. WTDQ	1959:01-1998:12	5	MERCHANT WHOLESALERS:DURABLE GOODS TOTAL (MIL OF CHAINED 1992 $)(SA)
61. WTNQ	1959:01-1998:12	5	MERCHANT WHOLESALERS:NONDURABLE GOODS (MIL OF CHAINED 1992 $)(SA)
62. RTQ	1959:01-1998:12	5	RETAIL TRADE: TOTAL (MIL OF CHAINED 1992 $)(SA)
63. RTNQ	1959:01-1998:12	5	RETAIL TRADE:NONDURABLE GOODS (MIL OF 1992 $)(SA)

Consumption

64. GMCQ	1959:01-1998:12	5	PERSONAL CONSUMPTION EXPEND (CHAINED)-TOTAL (BIL 92$,SAAR)
65. GMCDQ	1959:01-1998:12	5	PERSONAL CONSUMPTION EXPEND (CHAINED)-TOTAL DURABLES (BIL 92$,SAAR)
66. GMCNQ	1959:01-1998:12	5	PERSONAL CONSUMPTION EXPEND (CHAINED)-NONDURABLES (BIL 92$,SAAR)
67. GMCSQ	1959:01-1998:12	5	PERSONAL CONSUMPTION EXPEND (CHAINED)-SERVICES (BIL 92$,SAAR)
68. GMCANQ	1959:01-1998:12	5	PERSONAL CONS EXPEND (CHAINED)-NEW CARS (BIL 92$,SAAR)

Housing starts and sales

69. HSFR	1959:01-1998:12	4	HOUSING STARTS:NONFARM(1947-58);TOTAL FARM&NONFARM(1959-)(THOUS.,SA
70. HSNE	1959:01-1998:12	4	HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.
71. HSMW	1959:01-1998:12	4	HOUSING STARTS:MIDWEST(THOUS.U.)S.A.
72. HSSOU	1959:01-1998:12	4	HOUSING STARTS:SOUTH (THOUS.U.)S.A.
73. HSWST	1959:01-1998:12	4	HOUSING STARTS:WEST (THOUS.U.)S.A.
74. HSBR	1959:01-1998:12	4	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)
75. HSBNE	1960:01-1998:12*	4	HOUSES AUTHORIZED BY BUILD. PERMITS:NORTHEAST(THOU.U.)S.A
76. HSBMW	1960:01-1998:12*	4	HOUSES AUTHORIZED BY BUILD. PERMITS:MIDWEST(THOU.U.)S.A.
77. HSBSOU	1960:01-1998:12*	4	HOUSES AUTHORIZED BY BUILD. PERMITS:SOUTH(THOU.U.)S.A.
78. HSBWST	1960:01-1998:12*	4	HOUSES AUTHORIZED BY BUILD. PERMITS:WEST(THOU.U.)S.A.
79. HNS	1963:01-1998:12*	4	NEW 1-FAMILY HOUSES SOLD DURING MONTH (THOUS,SAAR)
80. HNSNE	1973:01-1998:12*	4	ONE-FAMILY HOUSES SOLD:NORTHEAST(THOU.U.,S.A.)
81. HNSMW	1973:01-1998:12*	4	ONE-FAMILY HOUSES SOLD:MIDWEST(THOU.U.,S.A.)
82. HNSSOU	1973:01-1998:12*	4	ONE-FAMILY HOUSES SOLD:SOUTH(THOU.U.,S.A.)
83. HNSWST	1973:01-1998:12*	4	ONE-FAMILY HOUSES SOLD:WEST(THOU.U.,S.A.)
84. HNR	1963:01-1998:12*	4	NEW 1-FAMILY HOUSES, MONTH'S SUPPLY @ CURRENT SALES RATE(RATIO)
85. HNIV	1963:01-1998:12*	4	NEW 1-FAMILY HOUSES FOR SALE AT END OF MONTH (THOUS,SA)
86. HMOB	1959:01-1998:12	4	MOBILE HOMES: MANUFACTURERS' SHIPMENTS (THOUS.OF UNITS,SAAR)
87. CONTC	1964:01-1998:12*	4	CONSTRUCT.PUT IN PLACE:TOTAL PRIV & PUBLIC 1987$(MIL$,SAAR)
88. CONPC	1964:01-1998:12*	4	CONSTRUCT.PUT IN PLACE:TOTAL PRIVATE 1987$(MIL$,SAAR)
89. CONQC	1964:01-1998:12*	4	CONSTRUCT.PUT IN PLACE:PUBLIC CONSTRUCTION 87$(MIL$,SAAR)
90. CONDO9	1959:01-1998:10*	4	CONSTRUCT.CONTRACTS: COMM'L & INDUS.BLDGS(MIL.SQ.FT.FLOOR SP.;SA)

Real inventories and inventory-sales ratios

91. IVMTQ	1959:01-1998:12	5	MANUFACTURING & TRADE INVENTORIES:TOTAL (MIL OF CHAINED 1992)(SA)
92. IVMFGQ	1959:01-1998:12	5	INVENTORIES, BUSINESS, MFG (MIL OF CHAINED 1992 DOLLARS, SA)
93. IVMFDQ	1959:01-1998:12	5	INVENTORIES, BUSINESS DURABLES (MIL OF CHAINED 1992 DOLLARS, SA)
94. IVMFNQ	1959:01-1998:12	5	INVENTORIES, BUSINESS, NONDURABLES (MIL OF CHAINED 1992 DOLLARS, SA)
95. IVWRQ	1959:01-1998:12	5	MANUFACTURING & TRADE INV:MERCHANT WHOLESALERS (MIL OF CHAINED 1992)(SA)
96. IVRRQ	1959:01-1998:12	5	MANUFACTURING & TRADE INV:RETAIL TRADE (MIL OF CHAINED 1992 DOLLARS)(SA)
97. IVSRQ	1959:01-1998:12	2	RATIO FOR MFG & TRADE: INVENTORY/SALES (CHAINED 1992 DOLLARS, SA)
98. IVSRMQ	1959:01-1998:12	2	RATIO FOR MFG & TRADE:MFG;INVENTORY/SALES (87$)(S.A.)
99. IVSRWQ	1959:01-1998:12	2	RATIO FOR MFG & TRADE:WHOLESALER;INVENTORY/SALES(87$)(S.A.)
100. IVSRRQ	1959:01-1998:12	2	RATIO FOR MFG & TRADE:RETAIL TRADE;INVENTORY/SALES(87$)(S.A.)
101. PMNV	1959:01-1998:12	1	NAPM INVENTORIES INDEX (PERCENT)

Orders and unfilled orders

102. PMNO	1959:01-1998:12	1	NAPM NEW ORDERS INDEX (PERCENT)
103. PMDEL	1959:01-1998:12	1	NAPM VENDOR DELIVERIES INDEX (PERCENT)
104. MOCMQ	1959:01-1998:12	5	NEW ORDERS (NET)-CONSUMER GOODS & MATERIALS, 1992 DOLLARS (BCI)
105. MDOQ	1959:01-1998:12	5	NEW ORDERS, DURABLE GOODS INDUSTRIES, 1992 DOLLARS (BCI)
106. MSONDQ	1959:01-1998:12	5	NEW ORDERS, NONDEFENSE CAPITAL GOODS, IN 1992 DOLLARS (BCI)
107. MO	1959:01-1998:12	5	MFG NEW ORDERS: ALL MANUFACTURING INDUSTRIES, TOTAL (MIL$,SA)
108. MOWU	1959:01-1998:12	5	MFG NEW ORDERS: MFG INDUSTRIES WITH UNFILLED ORDERS(MIL$,SA)
109. MDO	1959:01-1998:12	5	MFG NEW ORDERS: DURABLE GOODS INDUSTRIES, TOTAL (MIL$,SA)

110.	MDUWU	1959:01-1998:12	5	MFG NEW ORDERS:DURABLE GOODS INDUST WITH UNFILLED ORDERS(MIL$,SA)
111.	MNO	1959:01-1998:12	5	MFG NEW ORDERS: NONDURABLE GOODS INDUSTRIES, TOTAL (MIL$,SA)
112.	MNOU	1959:01-1998:12	5	MFG NEW ORDERS: NONDURABLE GDS IND.WITH UNFILLED ORDERS(MIL$,SA)
113.	MU	1959:01-1998:12	5	MFG UNFILLED ORDERS: ALL MANUFACTURING INDUSTRIES, TOTAL (MIL$,SA)
114.	MDU	1959:01-1998:12	5	MFG UNFILLED ORDERS: DURABLE GOODS INDUSTRIES, TOTAL (MIL$,SA)
115.	MNU	1959:01-1998:12	5	MFG UNFILLED ORDERS: NONDURABLE GOODS INDUSTRIES, TOTAL (MIL$,SA)
116.	MPCON	1959:01-1998:12	5	CONTRACTS & ORDERS FOR PLANT & EQUIPMENT (BIL$,SA)
117.	MPCONQ	1959:01-1998:12	5	CONTRACTS & ORDERS FOR PLANT & EQUIPMENT IN 1992 DOLLARS (BCI)

Stock prices (SPr)

118.	FSNCOM	1959:01-1998:12	5	NYSE COMMON STOCK PRICE INDEX: COMPOSITE (12/31/65=50)
119.	FSNIN	1966:01-1998:12*	5	NYSE COMMON STOCK PRICE INDEX: INDUSTRIAL (12/31/65=50)
120.	FSNTR	1966:01-1998:12*	5	NYSE COMMON STOCK PRICE INDEX: TRANSPORTATION (12/31/65=50)
121.	FSNUT	1966:01-1998:12*	5	NYSE COMMON STOCK PRICE INDEX: UTILITY (12/31/65=50)
122.	FSNFI	1966:01-1998:12*	5	NYSE COMMON STOCK PRICE INDEX: FINANCE (12/31/65=50)
123.	FSPCOM	1959:01-1998:12	5	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)
124.	FSPIN	1959:01-1998:12	5	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)
125.	FSPCAP	1959:01-1998:12	5	S&P'S COMMON STOCK PRICE INDEX: CAPITAL GOODS (1941-43=10)
126.	FSPTR	1970:01-1998:12*	5	S&P'S COMMON STOCK PRICE INDEX: TRANSPORTATION (1970=10)
127.	FSPUT	1959:01-1998:12	5	S&P'S COMMON STOCK PRICE INDEX: UTILITIES (1941-43=10)
128.	FSPFI	1970:01-1998:12*	5	S&P'S COMMON STOCK PRICE INDEX: FINANCIAL (1970=10)
129.	FSDXP	1959:01-1998:12	1	S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM)
130.	FSPXE	1959:01-1998:12	1	S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%,NSA)
131.	FSNVV3	1974:01-1997:07*	5	NYSE MKT COMPOSITION:REPTD SHARE VOL BY SIZE,5000+ SHRS,%

Exchange rates

132.	EXRUS	1959:01-1998:12	5	UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)
133.	EXRGER	1959:01-1998:12	5	FOREIGN EXCHANGE RATE: GERMANY (DEUTSCHE MARK PER U.S.$)
134.	EXRSW	1959:01-1998:12	5	FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.$)
135.	EXRJAN	1959:01-1998:12	5	FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.$)
136.	EXRUK	1959:01-1998:12*	5	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)
137.	EXRCAN	1959:01-1998:12	5	FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$)

Interest rates

138.	FYFF	1959:01-1998:12*	2	INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)
139.	FYCP90	1959:01-1998:12*	2	INTEREST RATE: 90 DAY COMMERCIAL PAPER, (AC) (% PER ANN,NSA)
140.	FYGM3	1959:01-1998:12*	2	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)
141.	FYGM6	1959:01-1998:12*	2	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)
142.	FYGT1	1959:01-1998:12*	2	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)
143.	FYGT5	1959:01-1998:12	2	INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)
144.	FYGT10	1959:01-1998:12	2	INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)
145.	FYAAAC	1959:01-1998:12	2	BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)
146.	FYBAAC	1959:01-1998:12	2	BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)
147.	FWAFIT	1973:01-1994:04*	1	WEIGHTED AVG FOREIGN INTEREST RATE(%,SA)
148.	FYFHA	1959:01-1998:12	2	SECONDARY MARKET YIELDS ON FHA MORTGAGES (% PER ANNUM)
149.	SFYCP	1959:01-1998:12	1	Spread FYCP - FYFF
150.	SFYGM3	1959:01-1998:12	1	Spread FYGM3 - FYFF
151.	SFYGM6	1959:01-1998:12	1	Spread FYGM6 - FYFF
152.	SFYGT1	1959:01-1998:12	1	Spread FYGT1 - FYFF
153.	SFYGT5	1959:01-1998:12	1	Spread FYGT5 - FYFF
154.	SFYGT10	1959:01-1998:12	1	Spread FYGT10 - FYFF
155.	SFYAAAC	1959:01-1998:12	1	Spread FYAAAC - FYFF
156.	SFYBAAC	1959:01-1998:12	1	Spread FYBAAC - FYFF
157.	SFYFHA	1959:01-1998:12	1	Spread FYFHA - FYFF

Money and credit quantity aggregates

158.	FM1	1959:01-1998:12	6	MONEY STOCK: M1(CURR,TRAV.CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL$,SA)
159.	FM2	1959:01-1998:12	6	MONEY STOCK:M2(M1+O'NITE RPS,EURO$,G/P&B/D MMMFS&SAV&SM TIME DEP(BIL$,
160.	FM3	1959:01-1998:12	6	MONEY STOCK: M3(M2+LG TIME DEP,TERM RP'S&INST ONLY MMMFS)(BIL$,SA)
161.	FML	1959:01-1998:09*	6	MONEY STOCK:L(M3 + OTHER LIQUID ASSETS) (BIL$,SA)
162.	FM2DQ	1959:01-1998:12	5	MONEY SUPPLY-M2 IN 1992 DOLLARS (BCI)
163.	FMFBA	1959:01-1998:12	6	MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL$,SA)
164.	FMRRA	1959:01-1998:12	6	DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL$,SA)
165.	FMRNBC	1959:01-1998:12	6	DEPOSITORY INST RESERVES:NONBORROW+EXT CR,ADJ RES REQ CGS(MIL$,SA)
166.	FCLS	1973:01-1998:12*	5	LOANS & SEC @ ALL COML BANKS: TOTAL (BILS,SA)
167.	FCSGV	1973:01-1998:12*	5	LOANS & SEC @ ALL COML BANKS: U.S.GOVT SECURITIES (BIL$,SA)
168.	FCLRE	1973:01-1998:12*	5	LOANS & SEC @ ALL COML BANKS: REAL ESTATE LOANS (BIL$,SA)
169.	FCLIN	1973:01-1998:12*	5	LOANS & SEC @ ALL COML BANKS: LOANS TO INDIVIDUALS (BIL$,SA)
170.	FCLNBF	1973:01-1994:01*	5	LOANS & SEC @ ALL COML BANKS: LOANS TO NONBANK FIN INST(BIL$,SA)

171. FCLNQ	1959:01-1998:12*	5	COMMERCIAL & INDUSTRIAL LOANS OUSTANDING IN 1992 DOLLARS (BCI)
172. FCLBMC	1959:01-1998:12*	1	WKLY RP LG COM'L BANKS:NET CHANGE COM'L & INDUS LOANS(BIL$,SAAR)
173. CCI30M	1959:01-1995:09*	1	CONSUMER INSTAL.LOANS: DELINQUENCY RATE,30 DAYS & OVER, (%,SA)
174. CCINT	1975:01-1995:09*	1	NET CHANGE IN CONSUMER INSTAL CR: TOTAL (MIL$,SA)
175. CCINV	1975:01-1995:09*	1	NET CHANGE IN CONSUMER INSTAL CR: AUTOMOBILE (MIL$,SA)
176. CCINRV	1980:01-1995:09*	1	NET CHANGE IN CONSUMER INSTAL CR: REVOLVING(MIL$,SA)

Price indexes

177. PMCP	1959:01-1998:12	1	NAPM COMMODITY PRICES INDEX (PERCENT)
178. PWFSA	1959:01-1998:12	6	PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)
179. PWFCSA	1959:01-1998:12	6	PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)
180. PWIMSA	1959:01-1998:12*	6	PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA)
181. PWCMSA	1959:01-1998:12*	6	PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)
182. PWFXSA	1967:01-1998:12*	6	PRODUCER PRICE INDEX: FINISHED GOODS,EXCL. FOODS (82=100,SA)
183. PW160A	1974:01-1998:12*	6	PRODUCER PRICE INDEX: CRUDE MATERIALS LESS ENERGY (82=100,SA)
184. PW150A	1974:01-1998:12*	6	PRODUCER PRICE INDEX: CRUDE NONFOOD MAT LESS ENERGY (82=100,SA)
185. PSM99Q	1959:01-1998:12	6	INDEX OF SENSITIVE MATERIALS PRICES (1990=100)(BCI-99A)
186. PUNEW	1959:01-1998:12	6	CPI-U: ALL ITEMS (82-84=100,SA)
187. PU81	1967:01-1998:12*	6	CPI-U: FOOD & BEVERAGES (82-84=100,SA)
188. PUH	1967:01-1998:12*	6	CPI-U: HOUSING (82-84=100,SA)
189. PU83	1959:01-1998:12	6	CPI-U: APPAREL & upkeep (82-84=100,SA)
190. PU84	1959:01-1998:12	6	CPI-U: TRANSPORTATION (82-84=100,SA)
191. PU85	1959:01-1998:12	6	CPI-U: MEDICAL CARE (82-84=100,SA)
192. PUC	1959:01-1998:12	6	CPI-U: COMMODITIES (82-84=100,SA)
193. PUCD	1959:01-1998:12	6	CPI-U: DURABLES (82-84=100,SA)
194. PUS	1959:01-1998:12	6	CPI-U: SERVICES (82-84=100,SA)
195. PUXF	1959:01-1998:12	6	CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA)
196. PUXHS	1959:01-1998:12	6	CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA)
197. PUXM	1959:01-1998:12	6	CPI-U: ALL ITEMS LESS MIDICAL CARE (82-84=100,SA)
198. PCGOLD	1975:01-1998:12*	6	COMMODITIES PRICE:GOLD,LONDON NOON FIX,AVG OF DAILY RATE,$ PER OZ
199. GMDC	1959:01-1998:12	6	PCE,IMPL PR DEFL:PCE (1987=100)
200. GMDCD	1959:01-1998:12	6	PCE,IMPL PR DEFL:PCE; DURABLES (1987=100)
201. GMDCN	1959:01-1998:12	6	PCE,IMPL PR DEFL:PCE; NONDURABLES (1987=100)
202. GMDCS	1959:01-1998:12	6	PCE,IMPL PR DEFL:PCE; SERVICES (1987=100)

Average hourly earnings

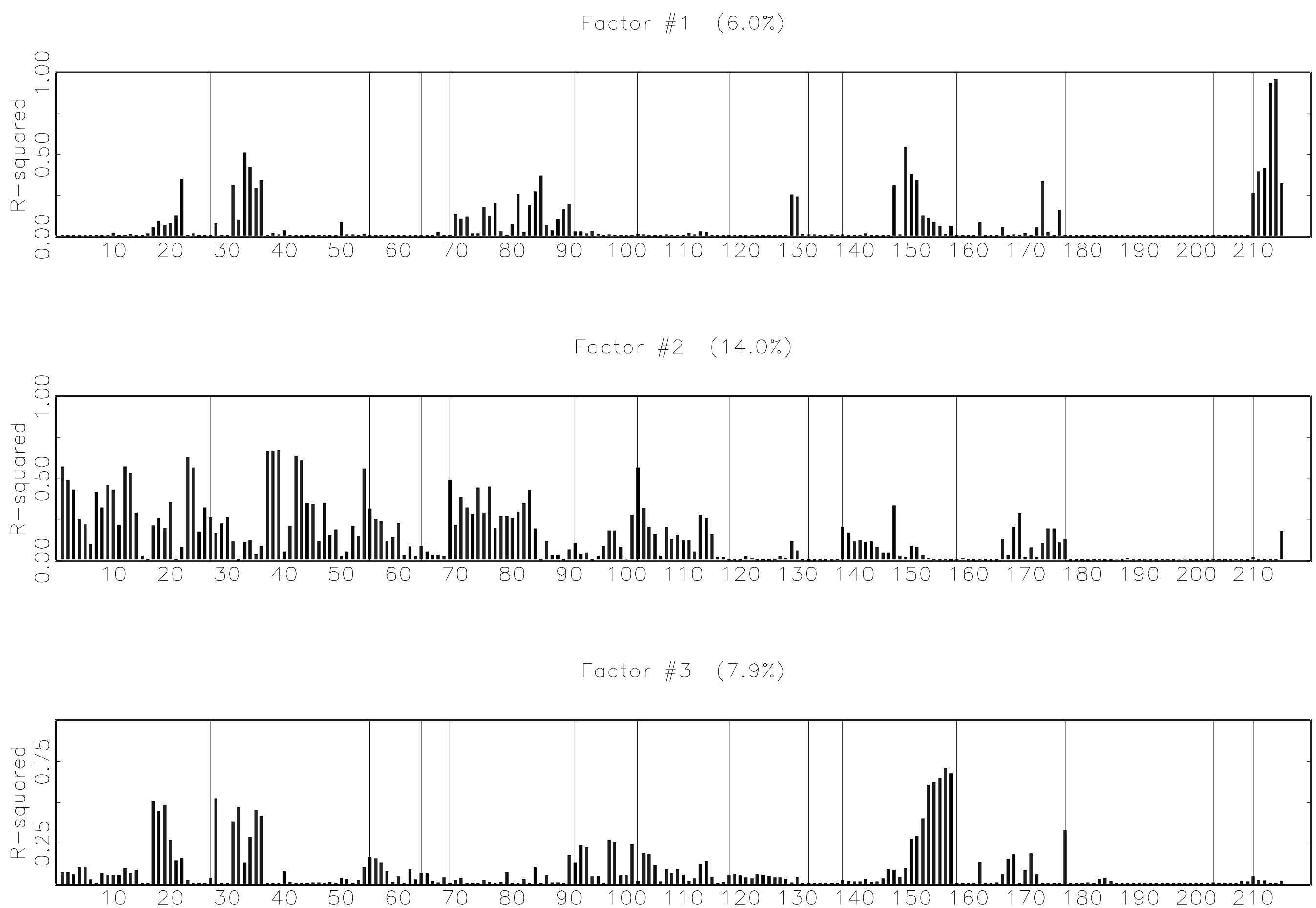
203. LEH	1964:01-1998:12*	6	AVG HR EARNINGS OF PROD WKRS: TOTAL PRIVATE NONAGRIC ($,SA)
204. LEHCC	1959:01-1998:12	6	AVG HR EARNINGS OF CONSTR WKRS: CONSTRUCTION ($,SA)
205. LEHM	1959:01-1998:12	6	AVG HR EARNINGS OF PROD WKRS: MANUFACTURING ($,SA)
206. LEHTU	1964:01-1998:12*	6	AVG HR EARNINGS OF NONSUPV WKRS: TRANS & PUBLIC UTIL($,SA)
207. LEHTT	1964:01-1998:12*	6	AVG HR EARNINGS OF PROD WKRS:WHOLESALE & RETAIL TRADE(SA)
208. LEHFR	1964:01-1998:12*	6	AVG HR EARNINGS OF NONSUPV WKRS: FINANCE,INSUR,REAL EST($,SA)
209. LEHS	1964:01-1998:12*	6	AVG HR EARNINGS OF NONSUPV WKRS: SERVICES ($,SA)

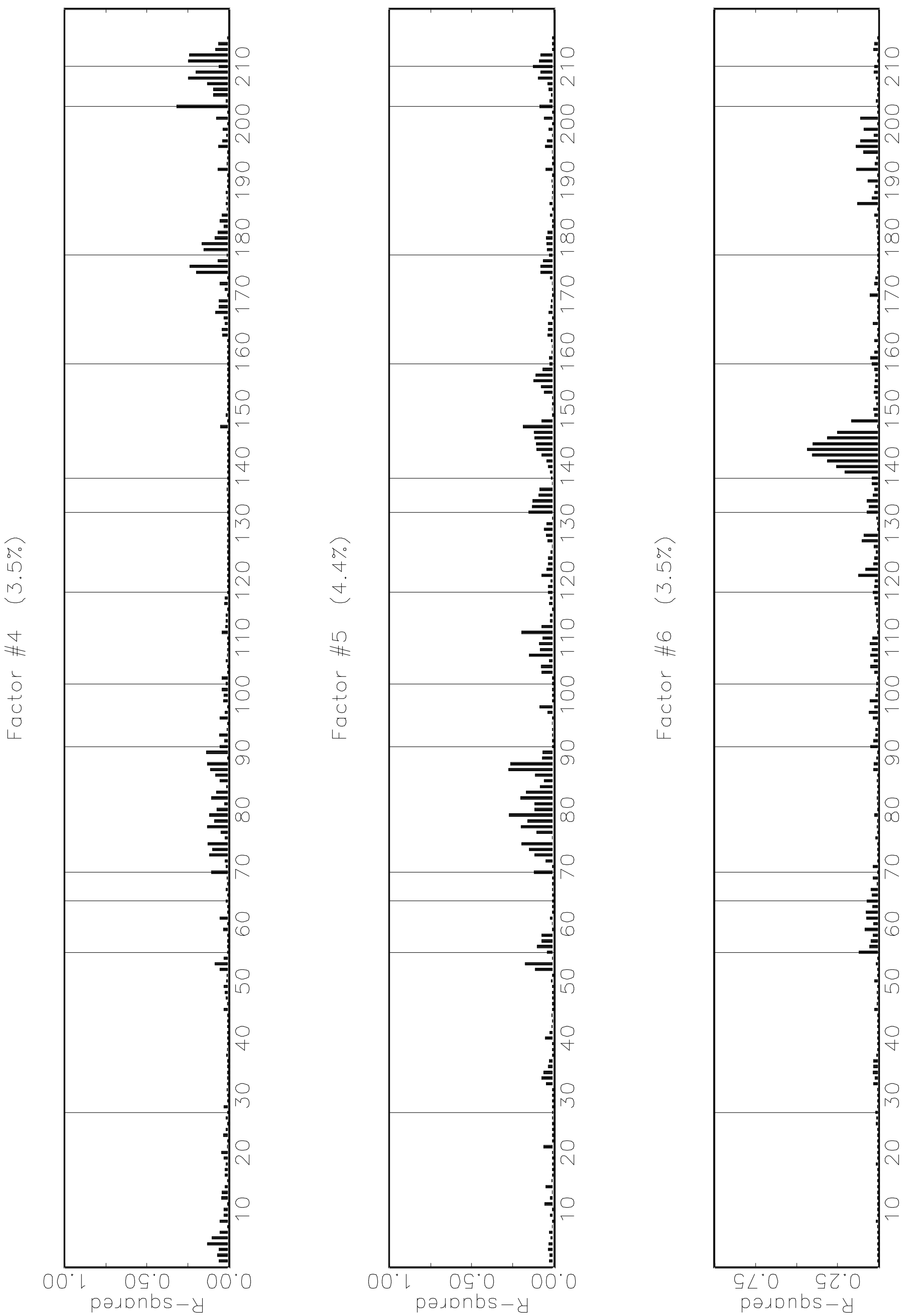
Miscellaneous

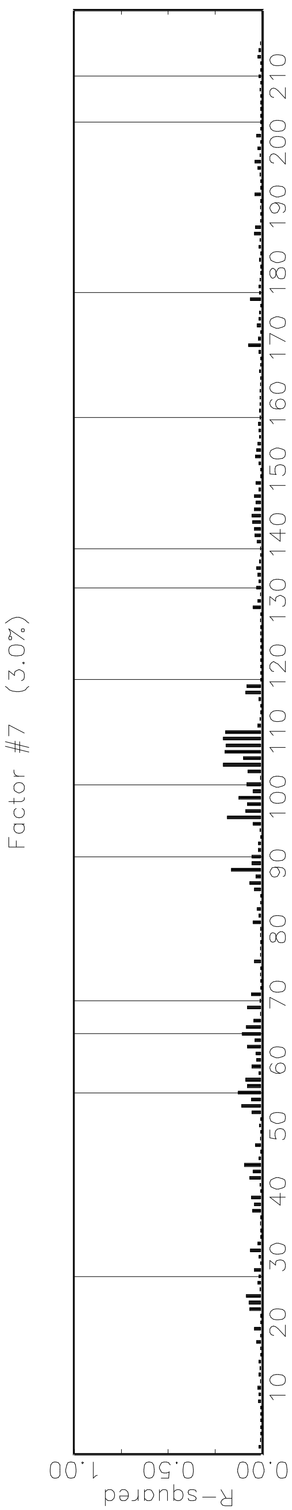
210. FSTE	1986:01-1998:12*	5	U.S.MDSE EXPORTS: TOTAL EXPORTS(F.A.S. VALUE)(MIL.$,S.A.)
211. FSTM	1986:01-1998:12*	5	U.S.MDSE IMPORTS: GENERAL IMPORTS(C.I.F. VALUE)(MIL.$,S.A.)
212. FTMD	1986:01-1998:12*	5	U.S.MDSE IMPORTS: GENERAL IMPORTS (CUSTOMS VALUE)(MIL$,S.A.)
213. FSTB	1986:01-1998:12*	2	U.S.MDSE TRADE BALANCE:EXPORTS LESS IMPORTS(FAS/CIF)(MIL$,S.A.)
214. FTB	1986:01-1998:12*	2	U.S.MDSE TRADE BALANCE:EXP.(FAS) LESS IMP.(CUSTOM)(MIL$,S.A.)
215. HHSNTN	1959:01-1998:12	1	U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)

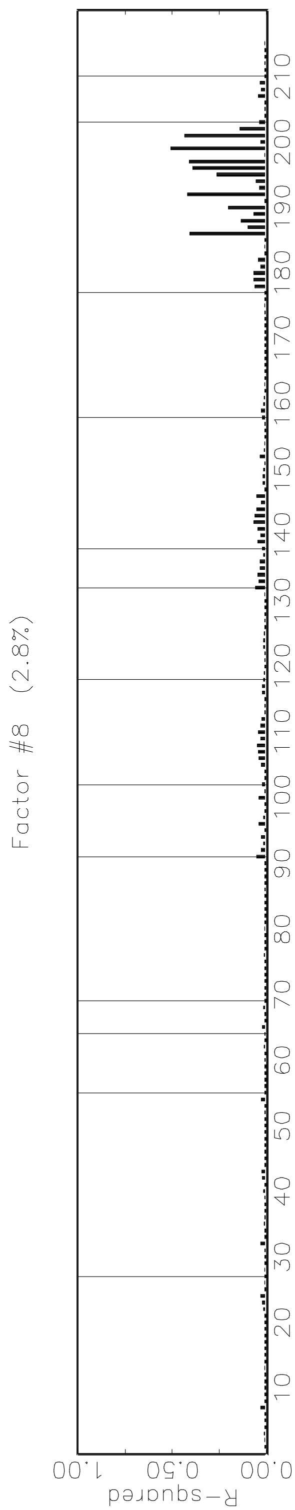
Figure 8: R^2**'s of univariate regressions of each of the individual time series on the unbalanced panel factors.**

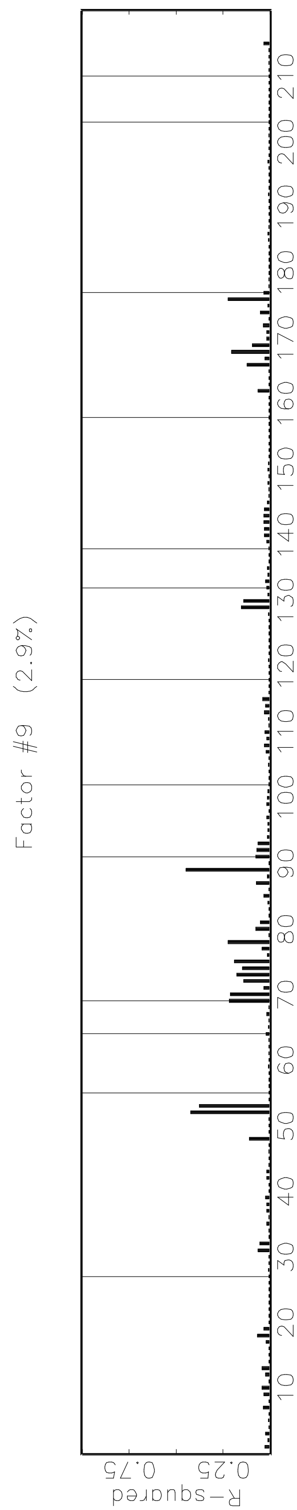
The variables are numbered as in the list of data series above, i.e. the categories are real output and income (1-26), employment and hours (27-54), real retail, manufacturing, and trade sales (55-63), consumption (64-68), housing starts and sales (69-90), real inventories and inventory-sales ratios (91-101), orders and unfilled orders (102-117), stock prices (118-131), exchange rates (132-137), interest rates (138-157), money and credit quantity aggregates (158-176), price indexes (177-202), average hourly earnings (203-209), miscellaneous (210-215). Note that although regression R^2's for the stock price series (118-131) are reported, these series have been excluded from the panel before computing the factors. The percentage values behind each factor correspond to the mean of all related univariate R^2 statistics.

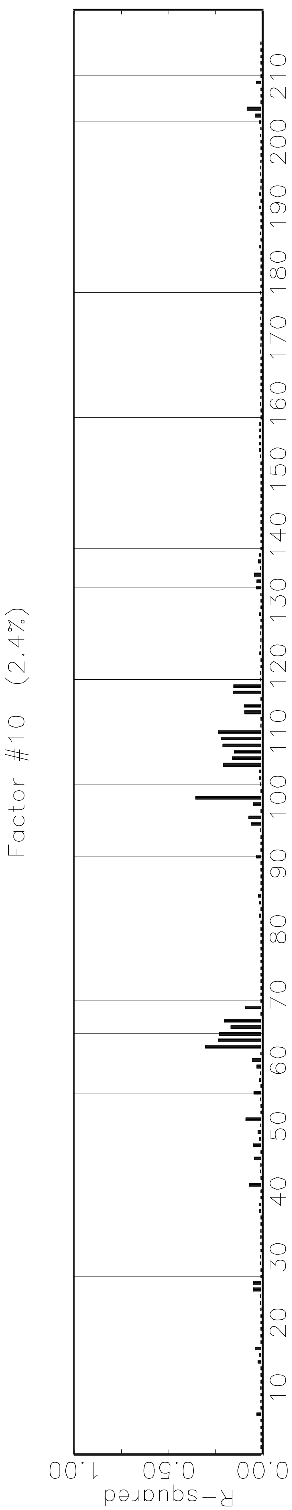


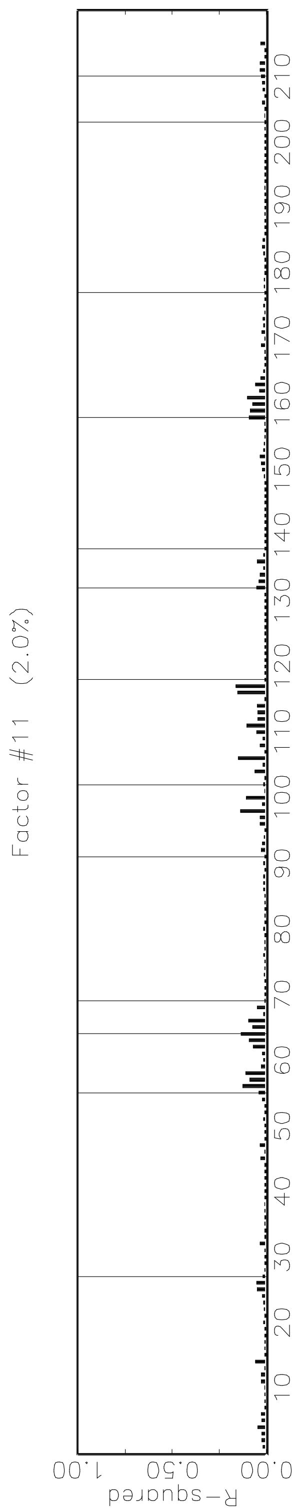

Factor #4 (3.5%)
Factor #5 (4.4%)
Factor #6 (3.5%)
R−squared
0.00 0.50 1.00
0.25 0.75
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210

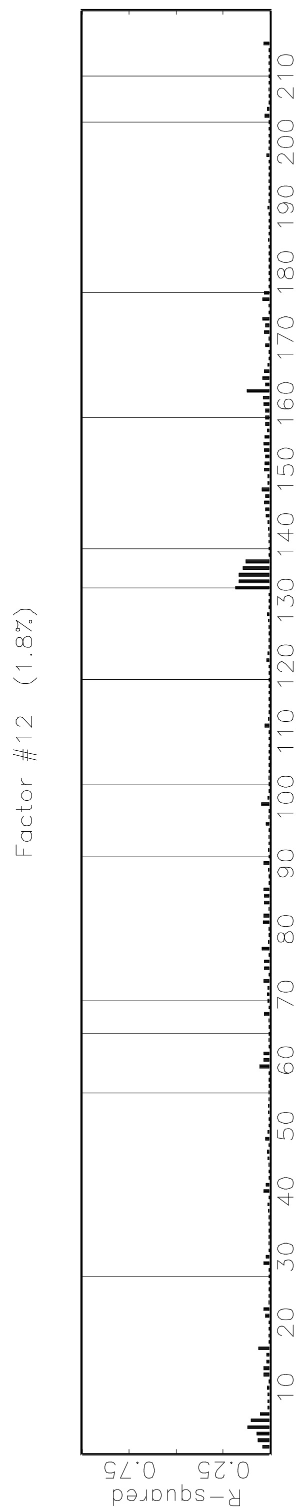

Factor #7 (3.0%)
R−squared
0.00 0.50 1.00
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210


Factor #8 (2.8%)
R−squared
0.00 0.50 1.00
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210


Factor #9 (2.9%)
R−squared
0.25 0.75
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210


Factor #10 (2.4%)
R-squared
0.00 0.50 1.00
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210


Factor #11 (2.0%)
R-squared
0.00 0.50 1.00
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210


Factor #12 (1.8%)
R-squared
0.25 0.75
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210

Plots of the first two factors extracted from the unbalanced panel

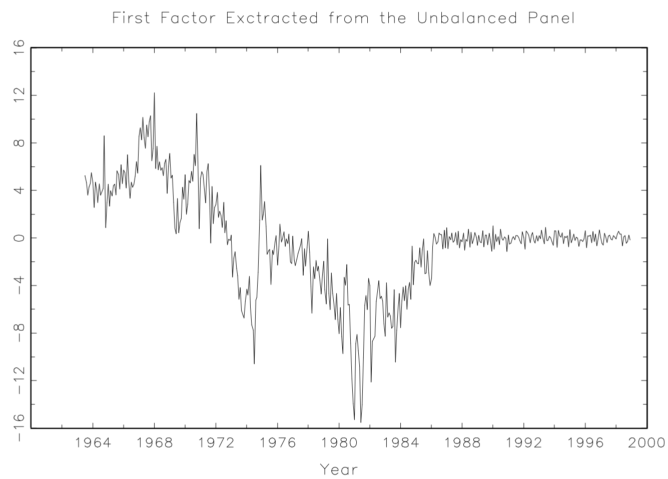


Figure 9: This figure plots the first dynamic factor extracted from the unbalanced panel using Stock and Watson's method. As one can see, the first unbalanced panel diffusion index remains almost constant from 1986 on when the series No. 210-214 become available.

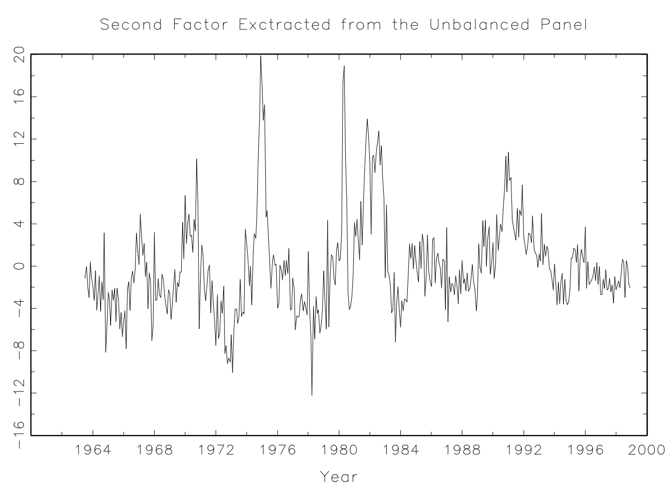
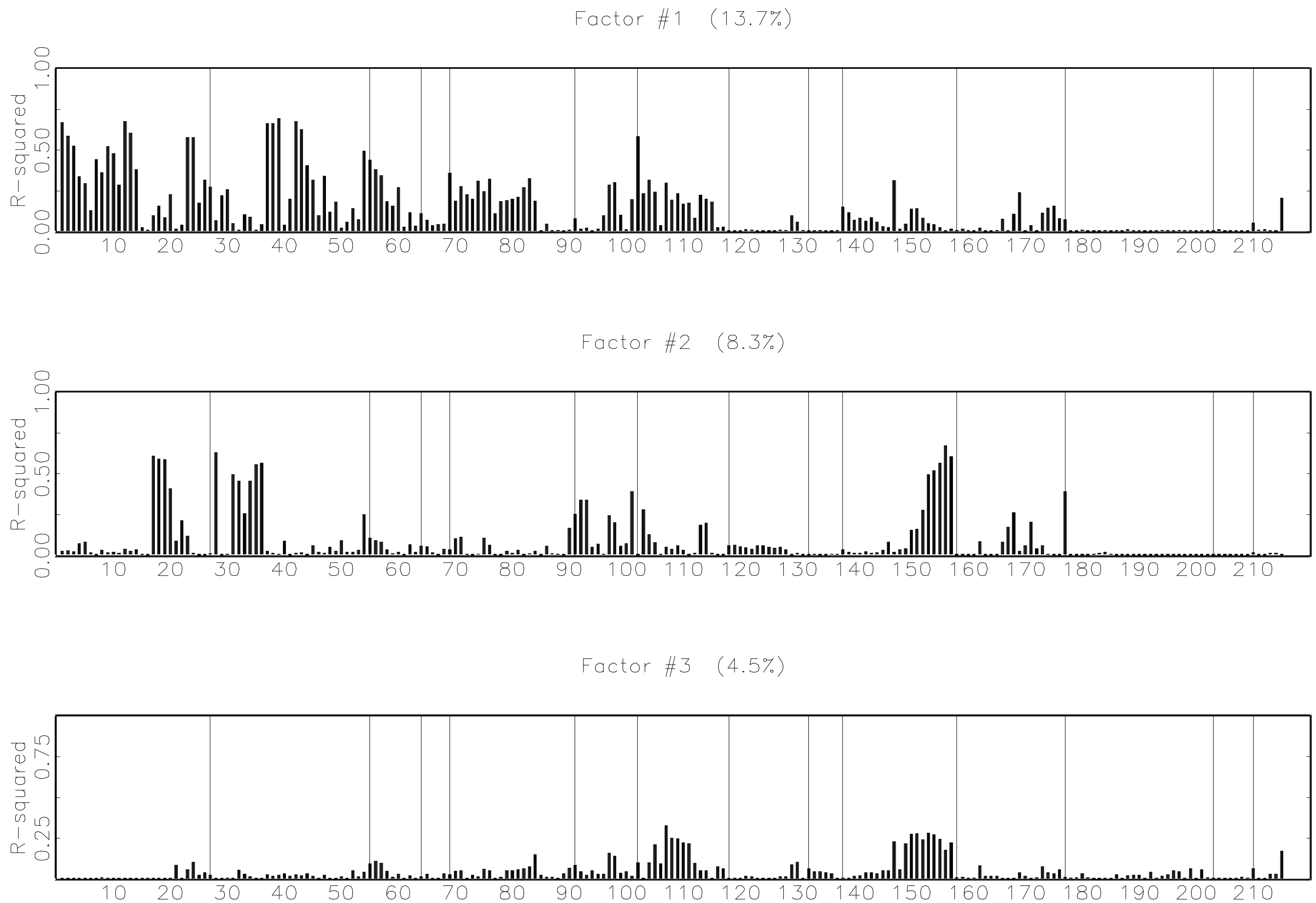
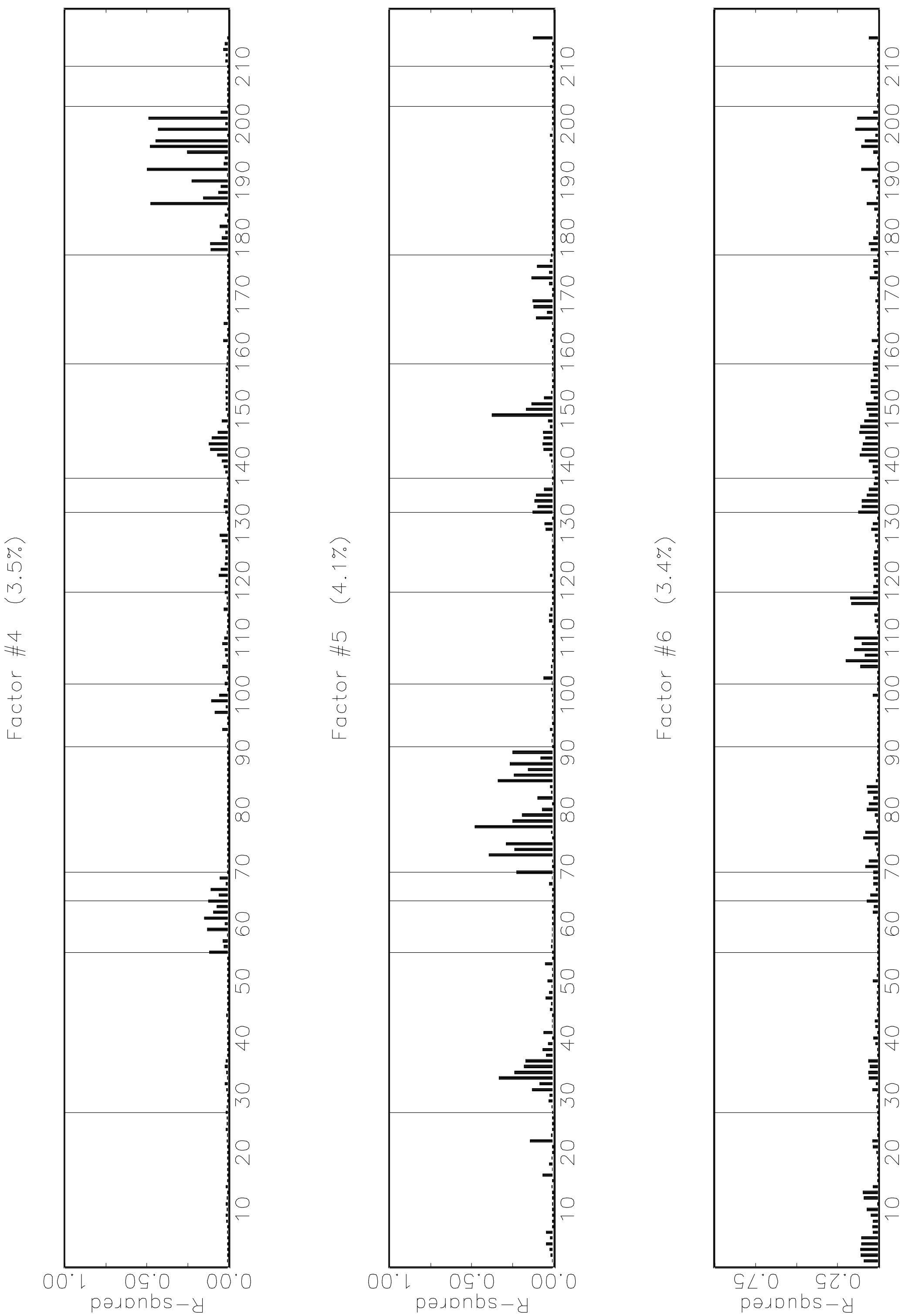


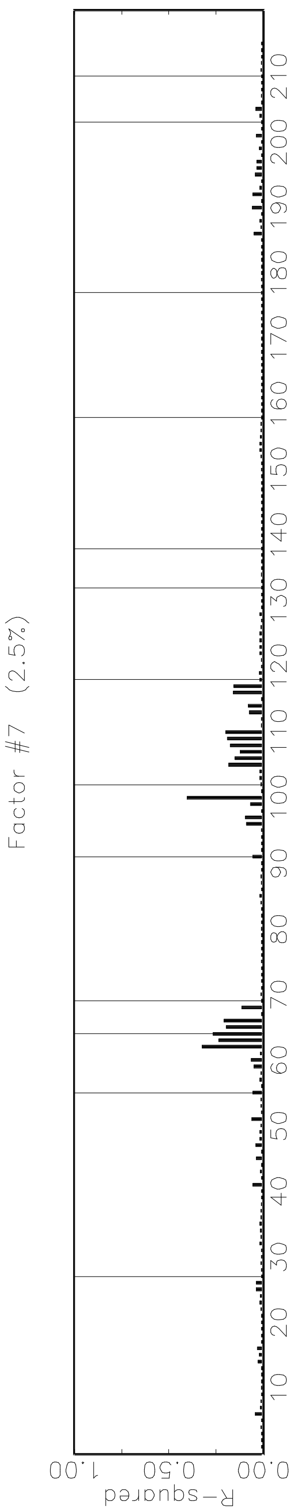
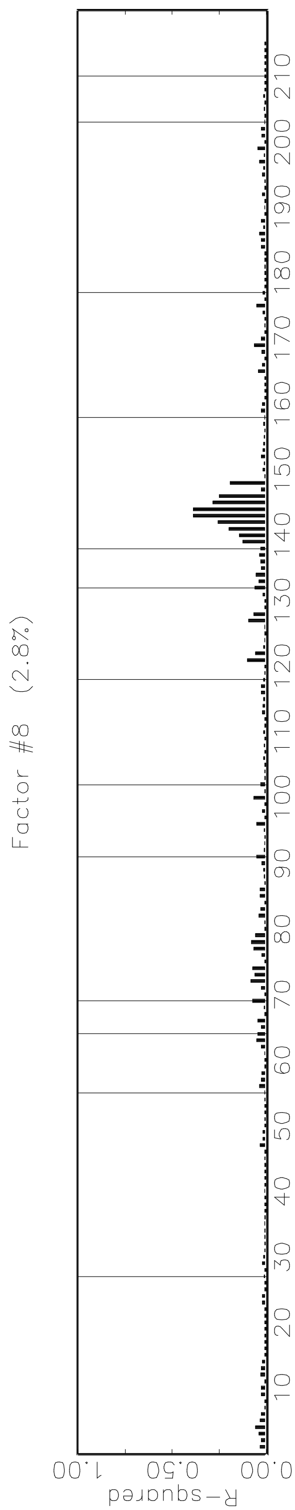
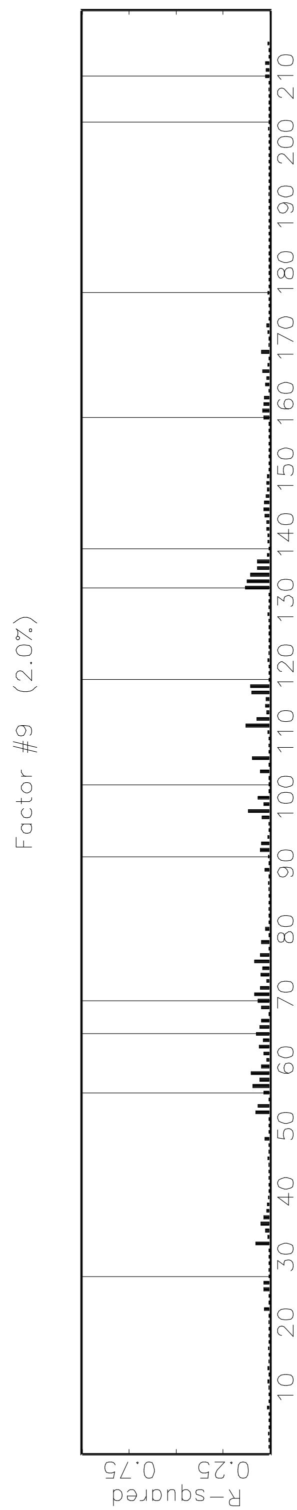
Figure 10: This figure plots the second dynamic factor extracted from the unbalanced panel using Stock and Watson's method. Contrary to the first factor, the second diffusion index shows a "normal" time series behavior.

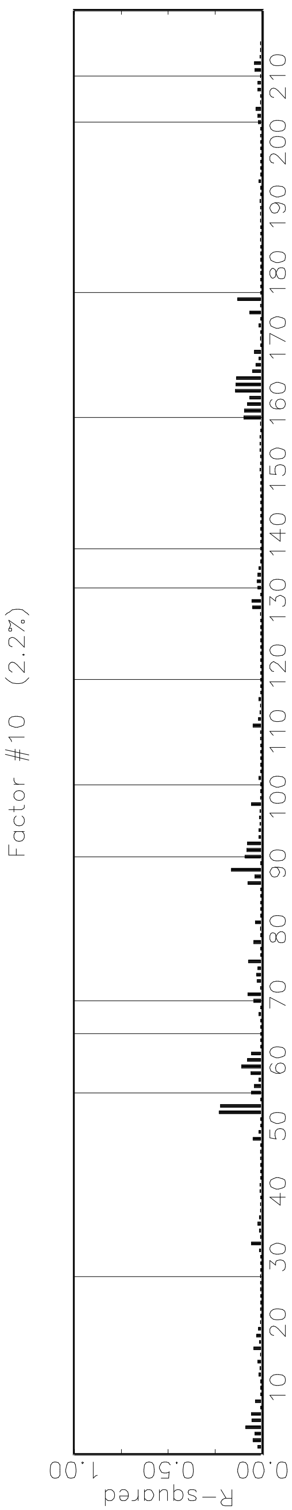
Figure 11: R^2**'s of univariate regressions of each of the individual time series on the balanced panel factors.**

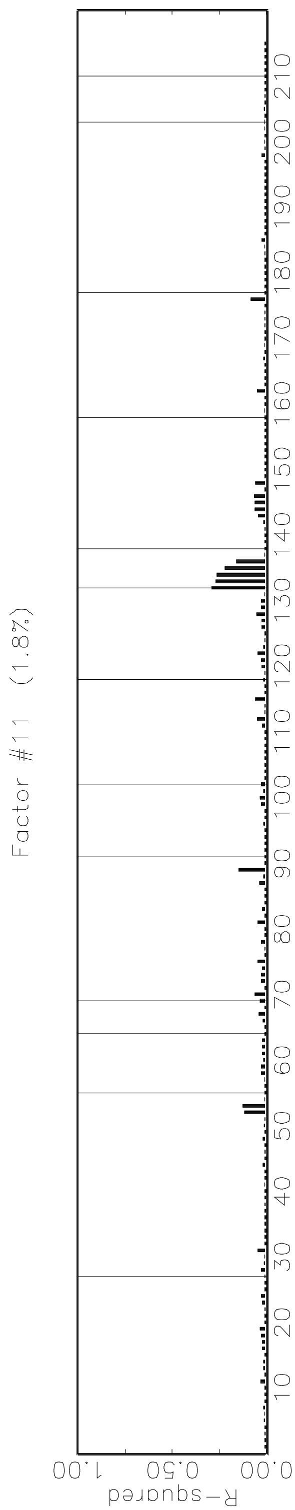
The variables are numbered as in the list of data series above, i.e. the categories are real output and income (1-26), employment and hours (27-54), real retail, manufacturing, and trade sales (55-63), consumption (64-68), housing starts and sales (69-90), real inventories and inventory-sales ratios (91-101), orders and unfilled orders (102-117), stock prices (118-131), exchange rates (132-137), interest rates (138-157), money and credit quantity aggregates (158-176), price indexes (177-202), average hourly earnings (203-209), miscellaneous (210-215). Note that although regression R^2's for the stock price series (118-131) are reported, these series have been excluded from the panel before computing the factors. The percentage values behind each factor correspond to the mean of all related univariate R^2 statistics.

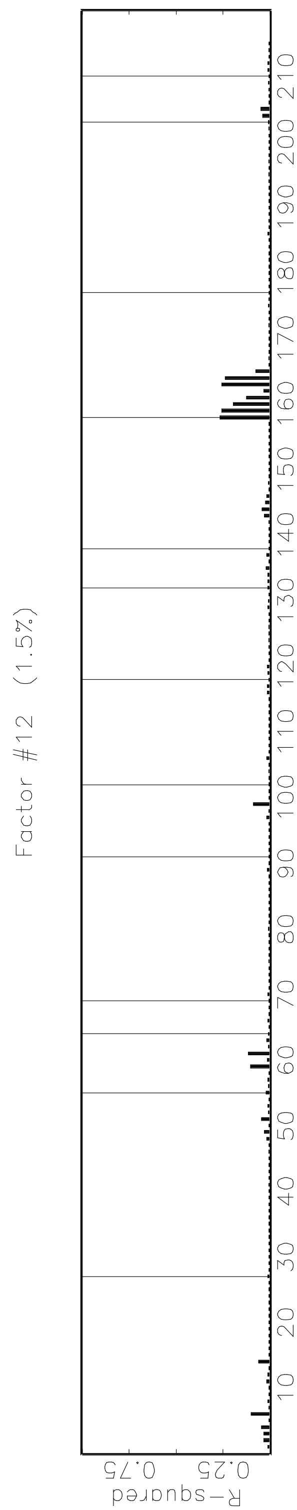



Factor #4 (3.5%)
Factor #5 (4.1%)
Factor #6 (3.4%)
R−squared


Factor #7 (2.5%)
R-squared

Factor #8 (2.8%)
R-squared

Factor #9 (2.0%)
R-squared


Factor #10 (2.2%)
R-squared
0.00 0.50 1.00
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210


Factor #11 (1.8%)
R-squared
0.00 0.50 1.00
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210


Factor #12 (1.5%)
R-squared
0.25 0.75
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210

C Recent Papers on Diffusion Indexes : A Brief Overview

Since the appearance of Stock and Watson's pioneering paper, there has been a growing body of theoretical and empirical work on diffusion indexes in macroeconomic research. In the following, I will briefly review the major recent contributions in this field.

Stock and Watson (2002) extend the forecasting exercise in Stock and Watson (1998) to a set of eight major macroeconomic variables for the US, including measures of industrial production, real personal income, manufacturing and trade sales, employment, and different price indexes. Performing forecasts at horizons of 6, 12, and 24 months, they show that with the exception of employment, diffusion indexes strongly outperform benchmark AR, VAR, and leading indicator forecasts of the real variables at all horizons. Their results are not as supportive for nominal variables, however. They report, for example, that diffusion index forecasts of inflation are subject to substantial improvement when lagged values of inflation are included. Perhaps most importantly, restricting the number of factors to one, the unbalanced panel factor model augmented with lagged values of inflation performs better than all alternatively tested versions. Hence, the authors conclude that the gains to forecasting inflation come from a single factor. As will be discussed below, Watson (2001) and Angelini et al. (2001b) obtain similar results.

Drawing on the model comparisons in Stock and Watson (2002), Watson (2001) surveys the respective merits and shortcomings of different small-scale and a large-scale diffusion index-based forecasting method. He finds evidence that in linear regressions using a very large number of explanatory variables, an important portion of the regressors have small but non-zero coefficients. From this observation, Watson concludes that since many variables each have little explanatory power rather than few variables each have much, a large-scale approach to macroeconomic forecasting is warranted. He then repeats the forecasting comparison exercise of Stock and Watson (1998, 2002) for all series in his data set, providing some additional model comparison statistics. Not surprisingly, his finding is that the factor model forecasts largely outperform standard VAR and leading indicator forecasts. However, comparing the forecasting power of all models across different sets of dependent variables, he shows that factor model forecasts only negligibly improve predictions of financial variables such as exchange rates or stock indexes. Similar to the result in Stock and Watson (2002), an examination of the links between the estimated factors and categories of variables further shows that the predictability of nominal variables is due to only one factor.

Marcellino, Stock, and Watson (2001) compare different methods of forecasting macroeconomic variables for the Euro area using country-specific as well as aggregate data. They find that while diffusion index forecasts outperform VAR forecasts based on either aggregate or country-specific data, the best results are in general obtained by pooling country-specific univariate autoregressive forecasts. Marcellino et al. offer two possible explanations for this finding. First, they suggest that their data sample 1982-1997 covers a period of great economic change in Europe implying that multivariate relations could have been particularly unstable during this time. The second explanation is that the small sample size counteracts forecasting models with a relatively large number of parameters as is the case in the diffusion index model.

A similar exercise has been performed in two studies of Angelini, Henry, and Mestre (2001a, 2001b). In a first step, they try to uncover an implicit measure of Euro area trend inflation by applying the Stock-Watson method to a set of (first-differenced) price variables of all Euro area countries. They find that the first factor, explaining almost 60 % of the variation of all series in the sample, can be viewed as representing a common trend inherent to the different inflation measures of the Euro area. Since almost all of the price variables the authors consider are non-stationary even after taking first differences, the first factor, accounting for the largest portion of common variation in these series, is likewise integrated. This, of course, raises the question whether the Stock-Watson methodology is appropriate for the data set used, since it is designed to accommodate only stationary variables. Angelini et al., without discussing the distributional implications of adopting the Stock-Watson method to non-stationary series, investigate the time series behavior of the factors and find some evidence that the first factor is indeed cointegrated with other measures of aggregate Euro area inflation. The dispersion of inflation across countries seems to be captured by the subsequent factors, although no direct link can be found between a certain set of countries and one or more of the factors. In a companion paper, Angelini, Henry, and Mestre (2001b) extend their analysis to three different data sets containing exclusively nominal, exclusively real and both, nominal and real variables for all of the Euro area countries. Investigating the time series behaviors and the leading-indicator properties of the factors extracted from these data sets, they find that (i) price variables are mainly driven by one single factor which is almost completely uncorrelated with the other factors, (ii) real variables are represented by a more complex set of nominal and non-nominal factors, and (iii) both, nominal and real factors have predictive power for aggregate measures of Euro area inflation.

Camacho and Sancho (2001) apply the Stock-Watson factor extraction procedure to a data set including a large number of Spanish economic time series that has only recently been made accessible by Spain's central bank. They report two major results. First, the time series behavior of the first factor exhibits a striking similarity to an activity indicator regularly published by the Spanish Ministry of Economy. Second, diffusion index forecasts of output and prices outperform standard AR and VAR forecasting methods both in an in-sample and simulated out-of-sample setting.

Bernanke and Boivin (2001) discuss the usefulness of the Stock-Watson diffusion index forecasts for both assessing and possibly improving the monetary policy conducted by the Federal Reserve. They first show that while there is practically no difference in using "real-time" or fully revised data, the scope of the data set from which the factors are extracted, i.e. the number and variety of series included, considerably affects the forecasting power of the diffusion indexes. Performing an ex-post comparison between Federal Reserve forecasts and diffusion index forecasts, they find that the former perform slightly better, but that averaging the two with equal weights obtains better results at almost all forecasting horizons.[55] Bernanke and Boivin further estimate a forward-looking policy

[55] Referring to Romer and Romer (2000) who find that the Federal Reserve "Greenbook" forecasts clearly outperform any for-profit private-sector forecasts, the authors argue that this result is particularly encouraging.

reaction function for the Fed assuming that interest rates are set in response to the state of the economy as measured by the Stock-Watson factors. More precisely, they estimate a potential policy response to the economic information contained in the diffusion indexes by simply regressing interest rates on the dynamic factors. Then, they use these estimated optimal interest rates which they label *excess policy response* as an additional regressor in a forward-looking Taylor-rule type policy reaction function with forecasts of output and inflation as the other explanatory variables. This specification allows them to test whether the excess policy response, i.e. the estimated reaction of the Fed to the information contained in the common factors, adds significantly to the explanation of interest rates through forecasts of output and inflation. Bernanke and Boivin report that this is indeed the case in most of the regressions they have conducted using different data sets and different sample periods. Finally, the authors discuss the possibility of developing a monetary "expert system" on the basis of factor-extracting procedures similar to the one of Stock and Watson. They argue that such a tool could accommodate the most recent economic information in updated diffusion indexes. Based on these, policy reaction functions of the type discussed above could then provide useful guidance to monetary policy authorities.

In a recent article, Bai and Ng (2002) make an important contribution to the theory of large-scale factor models of the type proposed by Stock and Watson. They criticize the information criterion suggested by Stock and Watson (1998) to determine the optimal number of factors as being too restrictive since it requires N >> T. Without requiring any relative asymptotic behavior, simply assuming that both N and T diverge, Bai and Ng develop a set of information criteria for finding the optimal number of factors in a large-scale approximate factor model. According to the framework studied by Stock and Watson, they allow for cross-sectional and serial dependence in the error terms as well as for some weak correlation of factors and idiosyncratic disturbances. In addition, they permit heteroskedasticity in the error terms in both the time and the cross-sectional dimension. Bai and Ng first provide a proof of convergence of the estimated to the true factors that relies on less restrictive assumptions than Stock and Watson's proof. Second, they propose a set of information criteria based on penalty functions that take into account the number of time series observations and the number of cross-sectional units, and show that these information criteria consistently estimate the true number of factors. They also perform some sort of Monte Carlo analysis and find that with a sufficiently large sample size ($\min\{N, T\} > 40$), their criteria precisely estimate the true number of factors. Finally applying their method to a data set containing a very large number of individual US stocks, Bai and Ng find evidence that only two factors summarize the common variation in these. However, since they do not analyze the properties of the estimated factors, their findings do not allow to draw conclusions about the potential usefulness of dynamic factor models in finance applications.

A somewhat separated although closely related strand of recent literature deals with dynamic factor models akin to the one studied by Stock and Watson, but estimated using principal components in the frequency domain. This research has been initiated by Forni and Reichlin (1996, 1998) who develop a methodology for identifying common shocks in a large cross-section of time series variables. Forni and Reichlin (1998) show that the number of common factors can be determined by applying principal components analysis to the spectral density matrix of a vector of cross-sectional averages of all variables in the panel. They then estimate the common shocks using a standard structural VAR technique. Finally, they show that the parameters in the factor model can be consistently estimated by applying OLS equation by equation.

In Forni, Hallin, Lippi, and Reichlin (2000), an alternative methodology is proposed where the estimation of the common factors is based on principal components analysis. Their model allows for infinite dynamics in the common factors and also permits some weak cross-correlation of the idiosyncratic error terms.[56] The authors thus label it a *generalized dynamic factor model*. Forni et al. first show that if the data take on a factor model with q factors, then the first q eigenvalues of the spectral density matrix of the data diverge while the subsequent eigenvalues are uniformly bounded. Second, they show that the common factors can be identified by projecting the data onto the space spanned by the first q dynamic principal components of the population spectral density matrix, and that this projection converges in mean square to the true common components when the number of observations in the cross-section goes to infinity. Third, they suggest to estimate the common components correspondingly, i.e. they propose to estimate the sample spectral density by using a Bartlett window estimator, to compute its first q eigenvectors, then to project the data onto these eigenvectors, and to obtain two-sided filter estimates as the inverse Fourier transforms of these projections. Finally, they show that applying these filters to the data set yields estimates of the common components which converge to the true factors in probability for some large N and T.

Since unboundedness of the eigenvalues cannot be examined in an empirical finite sample setting, Forni et al. argue that a formal method for finding the true number of factors in empirical applications is impossible to derive. However, they provide some heuristic observations that they assert to be good indicators of the true number of factors. The authors use their method to determine a coincident business cycle indicator for the Euro area by extracting the major common component from a large data set containing several macroeconomic variables for each member country.[57] Moreover, in a number of recent articles, further applications of the method developed in Forni et al. (2000) are provided. These comprise, among others, the identification of common shocks used to evaluate the policy response of the Fed (Giannone, Reichlin, and Sala (2002)), the role of regional, national and over-national components in business cycle fluctuations of the US and the Euro area (Forni and Reichlin (2001), the development of a core inflation index for the Euro area (Cristadoro, Forni, Reichlin, and Veronese (2001)), and the role of financial variables in forecasting Euro area industrial production and inflation (Forni et al. (2001b)).

[56]Notice that these assumptions are more general than the requirements in Stock and Watson's model that the factors admit a finite-order lag structure and that the idiosyncratic components be mutually orthogonal. However, the factor loadings in the model of Forni et al. are assumed to be constant whereas Stock and Watson explicitly model the factor loadings to be time-varying. Moreover, the methodology of Forni et al. does not apply to unbalanced panels as is the case for Stock and Watson's method. Hence, neither of two model types can be thought of as a special case of the other.

[57]Since January 2002, the Centre for Economic Policy Research (CEPR) publishes a monthly update of this indicator called EuroCOIN. It is calculated on the basis of more than 1000 time series for the six major economies in the Euro area.

The methodology of Forni et al. (2000) exhibits an important drawback. Since the filters they obtain are two-sided, the common factors that are derived from these filters are infeasible for forecasting purposes. As a solution to this problem, Forni, Hallin, Lippi, and Reichlin (2001a) suggest a forecasting method which is based on their generalized dynamic factor model and which makes use of only one-sided filters. Their procedure consists of two steps. First, from the Fourier transforms of the estimated spectral density matrix of the data set, they derive the cross-covariances for common and idiosyncratic components. More precisely, they propose to estimate the population spectral density matrix using some Periodogram-smoothing or Bartlett window. Then, again exploiting the unboundedness property of eigenvalues discussed above, they suggest to recover estimates of the spectral density matrices of the common and idiosyncratic components from the respective eigenvectors of the empirical spectral density matrix of the data set. Finally, they obtain cross-covariances of both common and idiosyncratic components as the Fourier transforms of these estimated spectral density matrices. In the second step, Forni et al. obtain one-sided filters as the solution to a generalized eigenvalue problem resulting from the maximization of the common-to-idiosyncratic variance ratio. They show that forecasts based on these filters are consistent for large T. In a Monte Carlo analysis, the authors demonstrate that their method provides good forecasting results, this even so in comparison to the Stock-Watson diffusion index forecasts.

References

[1] Angelini, Elena, Jérôme Henry, and Ricardo Mestre (2001a) : "A Multi-Country Trend Indicator for the Euro Area: Computation and Properties", *mimeo*, European Central Bank

[2] Angelini, Elena, Jérôme Henry, and Ricardo Mestre (2001b) : "Diffusion Index-Based Inflation Forecasts for the Euro Area", *mimeo*, European Central Bank

[3] Bai, Jushan, and Serena Ng (2002) : " Determining the Number of Factors in Approximate Factor Models", *Econometrica*, forthcoming

[4] Bernanke, Ben S., and Jean Boivin (2001) : "Monetary Policy in a Data-Rich Environment", *NBER Working Paper* No. 8379

[5] Camacho, Maximo, and Israel Sancho (2001) : "Spanish Diffusion Indexes", *unpublished working paper*, University of Murcia

[6] Cristadoro, Riccardo, Mario Forni, Lucrezia Reichlin, and Giovanni Veronese (2001): "A Core Inflation Index for the Euro Area", *Bank of Italy working paper*

[7] Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin (2000) : "The Generalized Dynamic Factor Model: Identification and Estimation", *Review of Economics and Statistics* 82, 540-554

[8] Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin (2001a) : "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting", *unpublished working paper*, ECARES, Université Libre de Bruxelles

[9] Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin (2001b) : "Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area?", *unpublished working paper*, ECARES, Université Libre de Bruxelles

[10] Forni, Mario, and Lucrezia Reichlin (1996) : "Dynamic Common Factors in Large Cross-Sections", *Empirical Economics* 21, 27-42

[11] Forni, Mario, and Lucrezia Reichlin (1998) : "Let's Get Real: A Factor Analytic Approach to Disaggregated Business Cycle Dynamics", *Review of Economic Studies* 65, 453-473

[12] Forni, Mario, and Lucrezia Reichlin (2001) : "Federal Policies and Local Economies: Europe and the US", *European Economic Review* 45, 109-134

[13] Giannone, Domenico, Lucrezia Reichlin, and Luca Sala (2002) : "Tracking Greenspan: Systematic and Unsystematic Monetary Policy Revisited", *unpublished working paper*, ECARES, Université Libre de Bruxelles

[14] Marcellino, Massimiliano, James H. Stock, and Mark W. Watson (2001) : "Macroeconomic Forecasting in the Euro Area: Country Specific Versus Area-Wide Information", *unpublished working paper*

[15] Watson, Mark W. (2001) : "Macroeconomic Forecasting Using Many Predictors", *unpublished working paper*, Princeton University