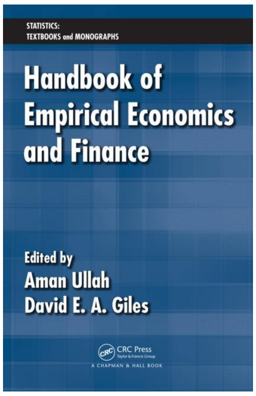


Handbook of Empirical Economics and Finance

Ullah Aman, E. A. Giles David

A Factor Analysis of Bond Risk Premia

Publication details
https://test.routledgehandbooks.com/doi/10.1201/b10440-13
C. Ludvigson Sydney, Ng Serena
Published online on: 20 Dec 2010

12

A Factor Analysis of Bond Risk Premia

Sydney C. Ludvigson and Serena Ng

CONTENTS

12.1 Introduction 313
12.2 Predictive Regressions 316
12.3 Estimation of Latent Factors 317
 12.3.1 Static Factors 318
 12.3.2 Dynamic Factors 319
 12.3.3 Comparison of $\hat{f}_t$ and $\hat{g}_t$ 330
12.4 Predictive Regressions 331
 12.4.1 Two-Year Returns 334
 12.4.2 Longer Maturity Returns and Overview 343
12.5 Inference Issues 346
 12.5.1 Asymptotic Bias 346
 12.5.2 Bias When the Predictors Are Functions of $\hat{f}_t$ 348
 12.5.3 Bootstrap Inference 349
 12.5.4 Posterior Inference 354
12.6 Countercyclical Yield Risk Premia 356
12.7 Conclusion 363
12.8 Acknowledgment 363
Data Appendix 363
References 369

12.1 Introduction

The expectations theory of the term structure posits that variables in the information set at time t should have no predictive power for excess bond returns. Consider the predictive regression

$$r_{t+h} = a + b' Z_t + e_{th}$$

where r_{t+h} is excess returns for holding period h, and Z_t is a set of predictors. Conventional tests often reject the null hypothesis that the parameter vector b is zero. Some suggest that over-rejections may arise if r is stationary and the variables Z are highly persistent, making inference highly distorted in finite samples. For this reason, researchers often use finite sample corrections or the bootstrap to conduct inference. However, it is often the case that robust inference still points to a rejection of the null hypothesis.

For a long time, the Zs found to have predictive power are often financial variables such as default premium, term premium, dividend price ratio, and measures of stock market variability and liquidity. Cochrane and Piazzesi (2005) find that a linear combination of five forward spreads explains between 30% and 35% of the variation in next year's excess returns on bonds with maturities ranging from 2 to 5 years. Yet theory suggests that predictive power for excess bond returns should come from macroeconomic variables. Campbell (1999) and Wachter (2006) suggest that bond and equity risk premia should covary with a slow-moving habit driven by shocks to aggregate consumption. Brandt and Wang (2003) argue that risk premia are driven by shocks to inflation as well as aggregate consumption; notably, both are macroeconomic shocks.

In an effort to reconcile theory and evidence, recent work has sought to establish and better understand the relation between excess returns and macroeconomic variables. Piazzesi and Swanson (2004) find that the growth of nonfarm payroll employment is a strong predictor of excess returns on federal funds futures contracts. Ang and Piazzesi (2003) use a no-arbitrage factor model of the term structure of interest rates that also allows for time-varying risk premia and finds that the pricing kernel is driven by a few observed macroeconomic variables and unobserved yield factors. Kozicki and Tinsley (2005) use affine models to link the term structure to perceptions of monetary policy. Duffie (2008) finds that an "expectations" factor unrelated to the level and the slope has strong predictive power for short-term interest rates and excess returns, and that this expectations factor has a strong inverse relation with industrial production. Notably, these studies have focused on the relation between expected excess bond returns, risk premia, and a few selected macroeconomic variables. The evidence falls short of documenting a direct relation between expected excess bond returns (bond risk premia) and the macroeconomy.

In Ludvigson and Ng (2007), we used a new approach. We used a small number of estimated (static) factors instead of a handful of observed predictors in the predictive regressions, where the factors are estimated from a large panel of macroeconomic data using the method of asymptotic principal components (PCA). Such a predictive regression is a special case of what is known as a "factor augmented regression" (FAR).[1] The factors enable us to substantially reduce the dimension of the predictor set while still being able to use the information underlying the variables in the panel. Furthermore, our latent factors are estimated without imposing a no-arbitrage condition or any parametric structure. Thus, our testing framework is nonstructural, both from an economic and a statistical point of view. We find that latent factors associated

[1] See Bai and Ng (2008) for a survey on this literature.

with real economic activity have significant predictive power for excess bond returns even in the presence of financial predictors such as forward rates and yield spreads. Furthermore, we find that bond returns and yield risk premia are more countercyclical when these risk premia are constructed to exploit information in the factors.

This chapter investigates the robustness of our earlier findings with special attention paid to how the factors are estimated. We first reestimate the FAR on a panel of 131 series over a longer sample. As in our previous work, these (static) factors, denoted $\hat{f}_t$, are estimated by PCA. We then consider an alternative set of factor estimates, denoted $\hat{g}_t$, that differ from the PCA estimates in two important ways. First, we use a priori information to organize the 131 series into 8 blocks. Second, we estimate a dynamic factor model for each of the eight blocks using a Bayesian procedure.

Compared with our previous work, we now use information in the large macroeconomic panel in a different way, and we estimate dynamic factors using a Bayesian method. It is thus useful to explain the motivation for doing so. The factors estimated from large panels of data are often criticized for being difficult to interpret, and organizing the data in blocks (such as output and price) provides a natural way to name the factors estimated from a block of data. At this point, we could have used PCA to estimate one static factor for each block. We could also have estimated dynamic factors using dynamic principal components, which is frequency-domain based. Whichever principal components estimator we choose, the estimates will not be precise as the number of series in each block is no longer 131 but a much smaller number. Bayesian estimation is more appropriate for the newly organized panels of data and Bayesian estimation yields a direct assessment of sampling variability. Using an estimator that is not principal components based also allows us to more thoroughly assess whether the FAR estimates are sensitive to how the factors are estimated. This issue, to our knowledge, has not been investigated in the literature. Notably, the factors that explain most of the variation in the large macroeconomic panel of data need not be the same as the factors most important for predicting excess bond returns. Thus for each of the two sets of factor estimates, namely, $\hat{f}_t$ and $\hat{g}_t$, we consider a systematic search of the relevant predictors, including an out-of-sample criterion to guard against overfitting the predictive regression with too many factors. We also assess the stability of the relation between excess bond returns and the factors over the sample.

An appeal of FAR is that when N and T are large and $\sqrt{T}/N$ tends to zero, the estimated factors in the FAR can be treated as though they are the true but latent factors. There is no need to account for sampling error incurred when the factors are estimated. Numerous papers have studied the properties of the (static and dynamic) principal components estimators in a forecasting context.[2] To date, little is known about the properties of the FAR estimates when $\sqrt{T}/N$ is not negligible. We show that principal components estimation may induce a

[2] See, for example, Boivin and Ng (2005).

bias in the parameter estimates of the predictive regression and suggest how a bias correction can be constructed. For our application, this bias is very small.

Our main finding is that macro factors have strong predictive power for excess bond returns and that this result holds up regardless of which method is used to estimate the factors. The reason is that both methods are capable of isolating the factor for real activity, which contributes significantly to variations in excess bond returns. However, the prior information that permits us to easily give names to the factors also constrains how information in the large panel is used. Thus, as far as predictability is concerned, the factors estimated from the large panel tend to be better predictors than the factors estimated from the eight blocks of data, for the same total number of series used in estimation. Recursive estimation of the predictive regressions finds that the macroeconomic factors are statistically significant throughout the entire sample, even though the degree of predictability varies over the 45 years considered. While the estimated bond and yield risk premia without the macro factors are acyclical, these premia are countercyclical when the estimated factors are used to forecast excess returns. This implies that investors must be compensated for risks associated with recessions.

Our empirical work is based on a macroeconomic panel that extends the one used in Stock and Watson (2005), which has since been used in a number of factor analyses.[3] The original data set consists of monthly observations for 132 macroeconomic time series from 1959:1 to 2003:12. We extend their data to 2007:12 and our panel consists of 131 series. Our empirical work uses data from 1964:1 to 2007:12.

12.2 Predictive Regressions

For $t = 1, \ldots T$, let $rx_{t+1}^{(n)}$ denote the continuously compounded (log) excess return on an n-year discount bond in period $t+1$. Excess returns are defined as $rx_{t+1}^{(n)} \equiv r_{t+1}^{(n)} - y_t^{(1)}$, where $r_{t+1}^{(n)}$ is the log holding period return from buying an n-year bond at time t and selling it as an $n-1$ year bond at time $t+1$, and $y_t^{(1)}$ is the log yield on the one-year bond. That is, if $p_t^{(n)}$ is log price of n-year discount bond at time t, then the log yield is $y_t^{(n)} \equiv -(1/n)p_t^{(n)}$.

A standard approach to assessing whether excess bond returns are predictable is to select a set of K predetermined conditioning variables at time t, given by the $K \times 1$ vector Z_t, and then estimate

$$rx_{t+1}^{(n)} = \beta' Z_t + \epsilon_{t+1} \tag{12.1}$$

by least squares. For example, Z_t could include the individual forward rates studied in Fama and Bliss (1987), the single forward factor studied in Cochrane and Piazzesi (2005), or other predictor variables based on a few macroeconomic series. Such a procedure may be restrictive when the number of eligible

[3] See, for example, Bai and Ng (2006b) and DeMol, Giannone, and Reichlin (2006).

predictors is quite large. In particular, suppose we observe a $T \times N$ panel of macroeconomic data with elements $x_t = (x_{1t}, x_{2t}, \ldots x_{Nt})'$, $t = 1, \ldots, T$, where the cross-sectional dimension, N, is large, and possibly larger than the number of time periods, T. The set of eligible predictors consists of the union of x_t and Z_t. With standard econometric tools, it is not obvious how a researcher could use the information contained in the panel because unless we have a way of ordering the importance of the N series in forming conditional expectations (as in an autoregression), there are potentially 2^N possible combinations to consider. The regression

$$rx_{t+1}^{(n)} = \gamma' x_t + \beta' Z_t + \epsilon_{t+1} \tag{12.2}$$

quickly runs into degrees-of-freedom problems as the dimension of x_t increases, and estimation is not even feasible when $N + K > T$.

The approach we consider is to posit that x_{it} has a factor structure so that if these factors were observed, we would have replaced Equation 12.2 by the following (infeasible) "factor augmented regression"

$$rx_{t+1}^{(n)} = \alpha' F_t + \beta' Z_t + \epsilon_{t+1}, \tag{12.3}$$

where F_t is a set of k factors whose dimension is much smaller than that of x_t but has good predictive power for $rx_{t+1}^{(n)}$. Equation 12.1 is nested within the factor-augmented regression, making Equation 12.3 a convenient framework to assess the importance of x_{it} via F_t, even in the presence of Z_t. The Z_t that we will use as benchmark is the forward rate factor used in Cochrane and Piazzesi (2005). This variable, hereafter referred to as *CP*, is a simple average of the one-year yield and four forward rates. These authors find that the predictive power of forward rates, yield spreads, and yield factors are subsumed in CP_t. To implement the regression given by Equation 12.3, we need to resolve two problems. First, F_t is latent and we must estimate it from data. Second, we need to isolate those factors with predictive power for our variable of interest, $rx_{t+1}^{(n)}$.

12.3 Estimation of Latent Factors

The first problem is dealt with by replacing F_t with an estimated value $\widehat{F}_t$ that is close to F_t in some well-defined sense, and this involves making precise a model from which F_t can be estimated. We will estimate two factor models, one static and one dynamic, using data retrieved from the Global Insight database and the Conference Board. The data are collected to incorporate as many series as that used in Stock and Watson (2005). However, one series (ao048) is no longer available on a monthly basis after 2003. Accordingly, our new data set consists of 131 series from 1959:1 to 2007:12, though our empirical analysis starts in 1964:1 because of availability of the bond yield data. As in the original Stock and Watson data, some series need to be transformed to be stationary. In general, real variables are expressed in growth rates, first differences are used for nominal interest rates, and second log differences are used for prices. The data description is given in appendix. This data can be downloaded from our Web site http://www.econ.nyu.edu/user/ludvigsons/Data&ReplicationFiles.zip.

12.3.1 Static Factors

Let N be the number of cross-section units and T be the number of time series observations. For $i = 1, \dots N, \quad t = 1, \dots T$, a *static* factor model is defined as

$$x_{it} = \lambda_i' f_t + e_{it}. \tag{12.4}$$

In factor analysis, e_{it} is referred to as the idiosyncratic error and λ_i are the factor loadings. This is a vector of weights that unit i put on the corresponding r (static) common factors f_t. In finance, x_{it} is the return for asset i in period t, f_t is a vector of systematic risk, λ_i is the exposure to the risk factors, and e_{it} is the idiosyncratic returns. Although the model specifies a static relationship between x_{it} and f_t, f_t itself can be a dynamic vector process that evolves according to

$$A(L) f_t = u_t,$$

where $A(L)$ is a polynomial (possibly of infinite order) in the lag operator. The idiosyncratic error e_{it} can also be a dynamic process, and e_{it} can also be cross-sectionally correlated.

We estimate f_t using the method of asymptotic principal components (PCA) originally developed by Connor and Korajzcyk (1986) for a small T large N environment. Letting "hats" denote estimated values, the $T \times r$ matrix $\hat{f}$ is $\sqrt{T}$ times the r eigenvectors corresponding to the r largest eigenvalues of the $T \times T$ matrix $xx'/(TN)$ in decreasing order with $\hat{f}'\hat{f} = I_r$. The normalization is necessary as the matrix of factor loadings Λ and f are not separately identifiable. The normalization also yields $\hat{\Lambda} = x'\hat{f}/T$. Intuitively, for each t, $\hat{f}_t$ is a linear combinations of each element of the $N \times 1$ vector $x_t = (x_{1t}, ..., x_{Nt})'$, where the linear combination is chosen optimally to minimize the sum of squared residuals $x_t - \Lambda f_t$. Bai and Ng (2002) and Stock and Watson (2002a) showed that the space spanned by f_t can be consistently estimated by $\hat{f}_t$ defined as above when $N, T \to \infty$. The number of static factors in x_t can be determined by the panel information criteria developed in Bai and Ng (2002). For the panel of 131 series under investigation, the IC_2 criterion finds eight factors over the full sample of 576 observations (with the maximum number of factors set to 20).

A common criticism of the method of principal components estimator is that the factors can be difficult to interpret. Our interpretation of the factors is based on the marginal R^2s, obtained by regressing each of the 131 series on the eight factors, one at a time. Because the factors are mutually uncorrelated, the marginal R^2 is also the explanatory power of the factor in question holding other factors fixed. Extending the sample to include three more years of data did not change our interpretation of the factors. Figures 12.1 through 12.8 show the marginal R-square statistics from regressing the series number given on the x-axis onto the estimated factor named in the heading. As in Ludvigson and Ng (2007), $\hat{f}_1$ is a real activity factor that loads heavily on employment and output data. The second factor loads heavily on interest rate spreads, while the third and fourth factors load on prices. Factor 5 loads on interest rates (much more strongly than the interest rate spreads). Factor 6 loads predominantly on the housing variables while factor 7 loads on measures of the money supply. Factor 8 loads on variables relating to the stock market. Thus, loosely speaking, factors 5–8 are more strongly related to money, credit, and finance.

While knowing that there are eight factors in the macroeconomic panel is useful information in its own right, of interest here are not the N variables $x_t = (x_{1t}, \ldots, x_{Nt})'$, but the scalar variable $rx^{(n)}_{t+1}$ which is not in x_t. Factors that are pervasive for the large panel of data need not be important for predicting $rx^{(n)}_{t+1}$. For this reason, we make a distinction between $F_t \subset f_t$ and f_t. The predictive regression of interest is

$$rx^{(n)}_{t+1} = \alpha'_F \hat{F}_t + \beta'_F Z_t + \epsilon_{t+1}, \tag{12.5}$$

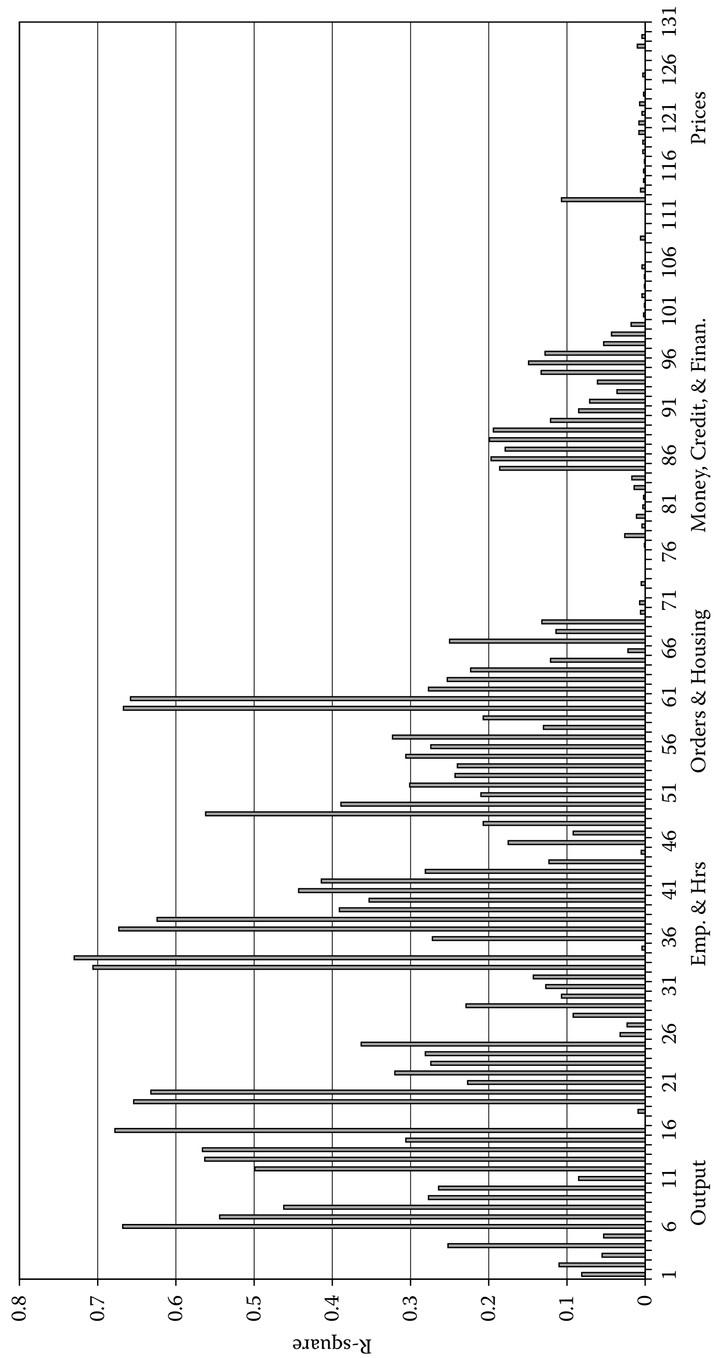
which has a vector of generated regressors, $\hat{F}_t$.

Consistency of $\hat{\alpha}_F$ follows from the fact that the difference between $\hat{f}_t$ and the space spanned by f_t vanishes at rate $\min[N, T]$, a result established in Bai and Ng (2002).[4] Bai and Ng (2006a) showed that if $\sqrt{T}/N \to 0$ as $N, T \to \infty$, the sampling uncertainty from first-step estimation is negligible. The practical implication is that standard errors can be computed for the estimates of α_F as though the true F_t were used in the regression. This is in contrast to the case when $\hat{F}_t$ is estimated from a first-step regression with a finite number of predictors. As shown in Pagan (1984), the standard errors for $\hat{\alpha}_F$ in such a case are incorrect unless they are adjusted for the estimation error incurred in the first step of F_t.

12.3.2 Dynamic Factors

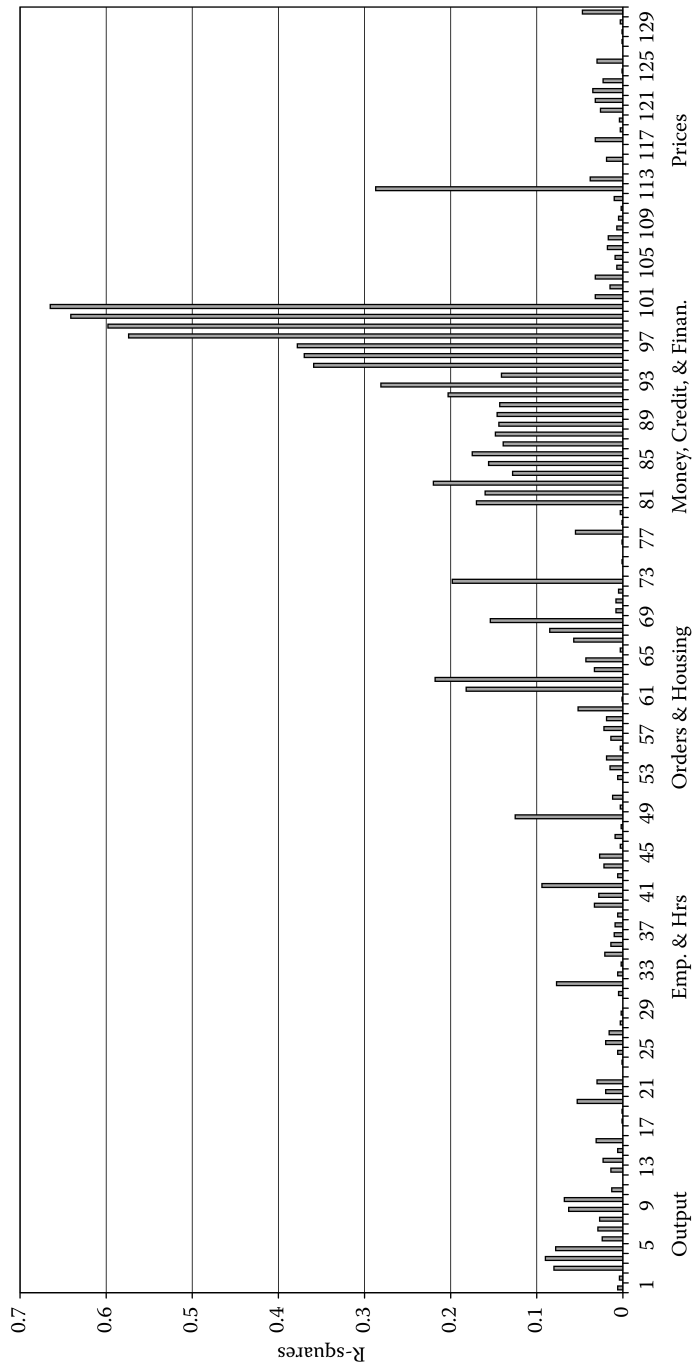
An advantage of the method of principal components is that it can handle a large panel of data at little computation cost, one reason being that little structure is imposed on the estimation. To be convinced that factor augmented regressions are useful in analyzing economic issues of interest, we need to show that estimates of the FAR are robust to the choice of the estimator *and* to the specification of the factor model. To this end, we consider an alternative way of estimating the factors with two fundamental differences.

[4] It is useful to remark that the convergence rate established in Stock and Watson (2002a) is too slow to permit consistent estimation of the parameters in Equation 12.5.



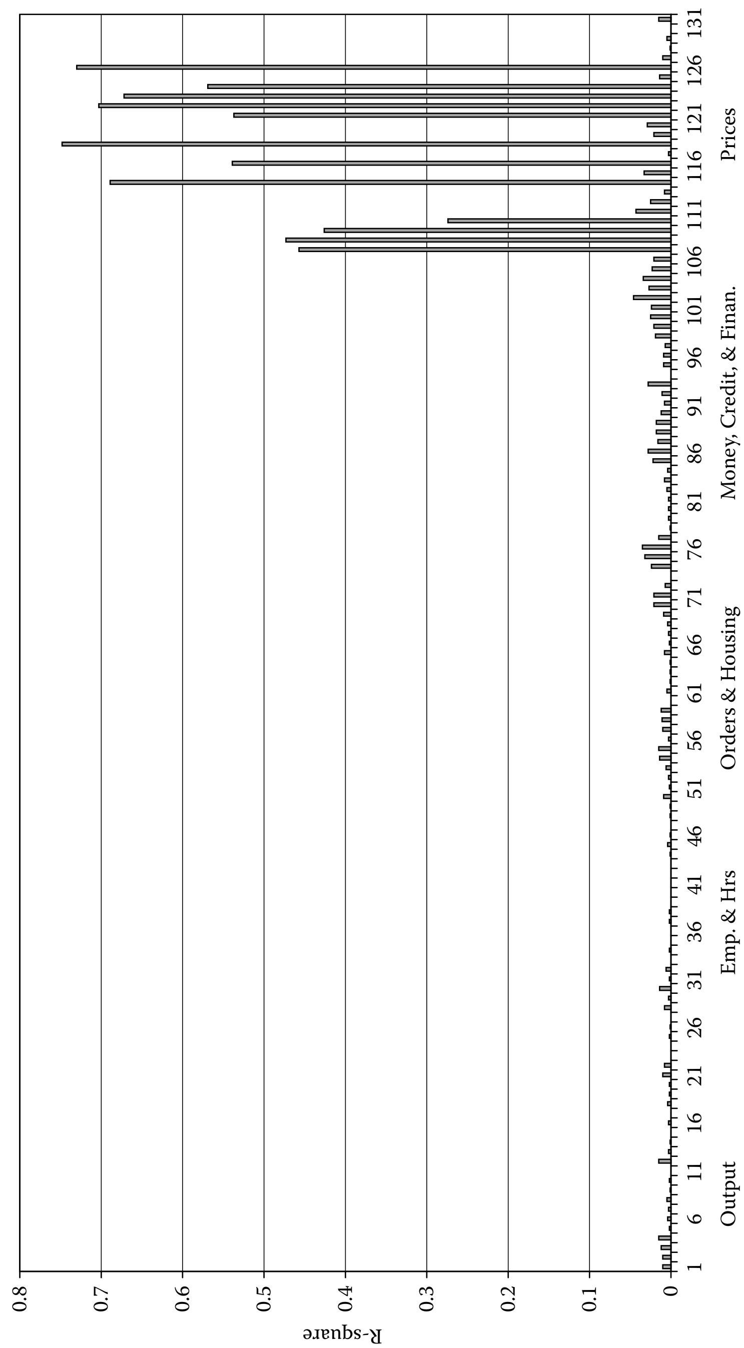
Notes: Chart shows the R-square from regressing the series number given on the *x*-axis onto the estimated factor named in the heading. See the appendix for a description of the numbered series. The factors are estimated using data from 1964:1-2007:12.

FIGURE 12.1
Marginal R-squares for F_1.



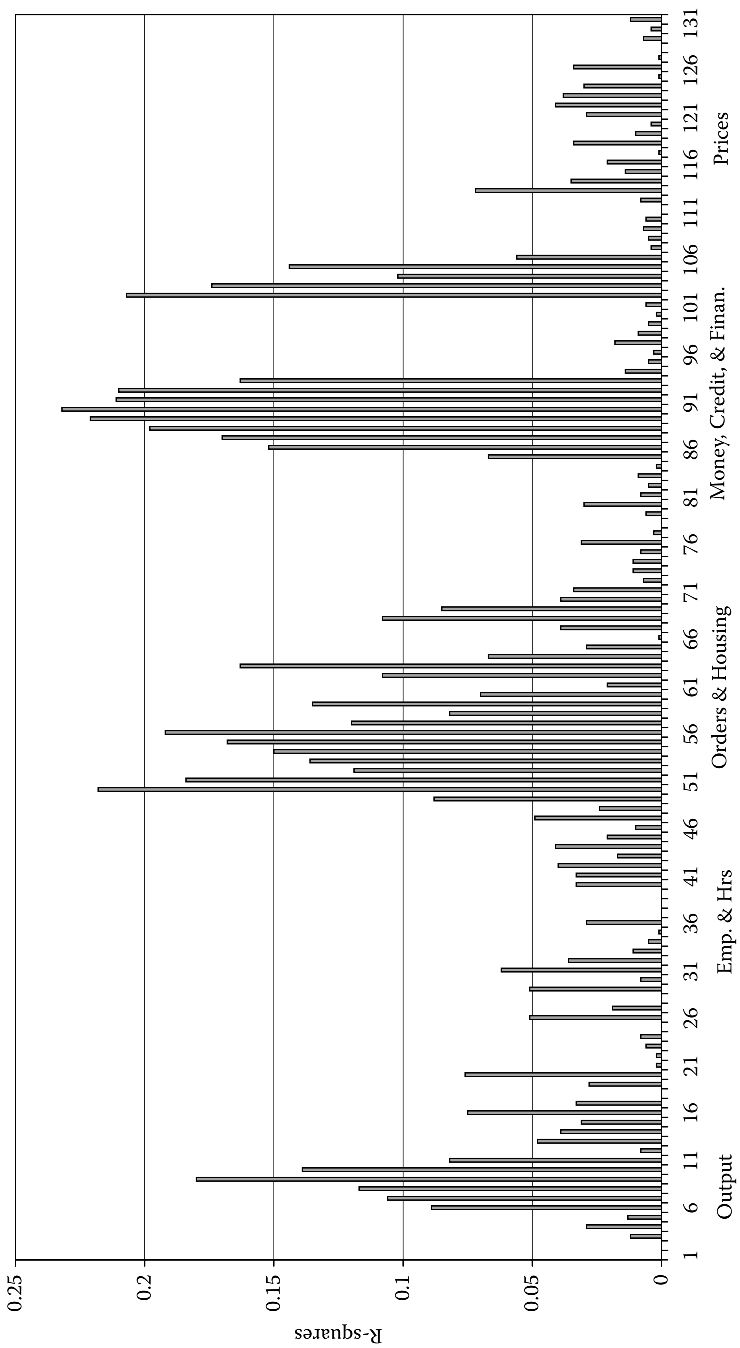
Notes: See Figure 12.1.

FIGURE 12.2
Marginal R-squares for F_2.



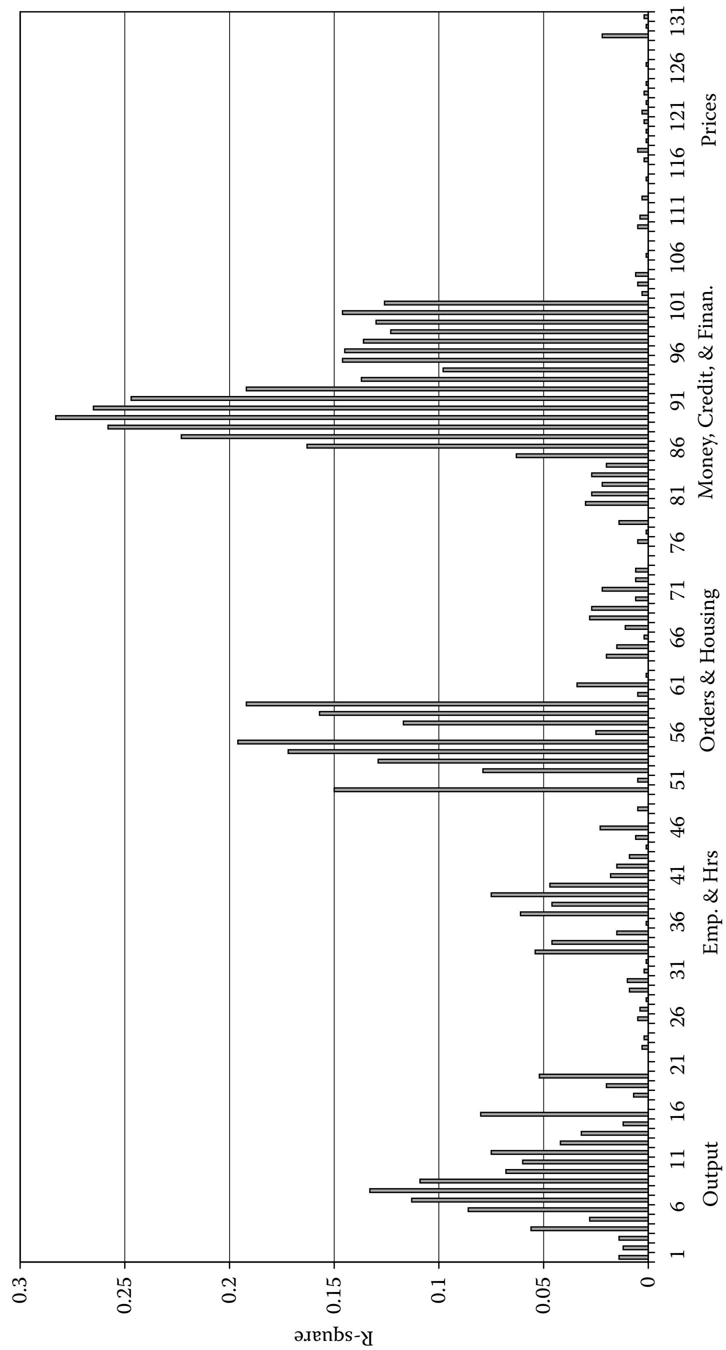
Notes: See Figure 12.1.

FIGURE 12.3
Marginal R-squares for F_3.



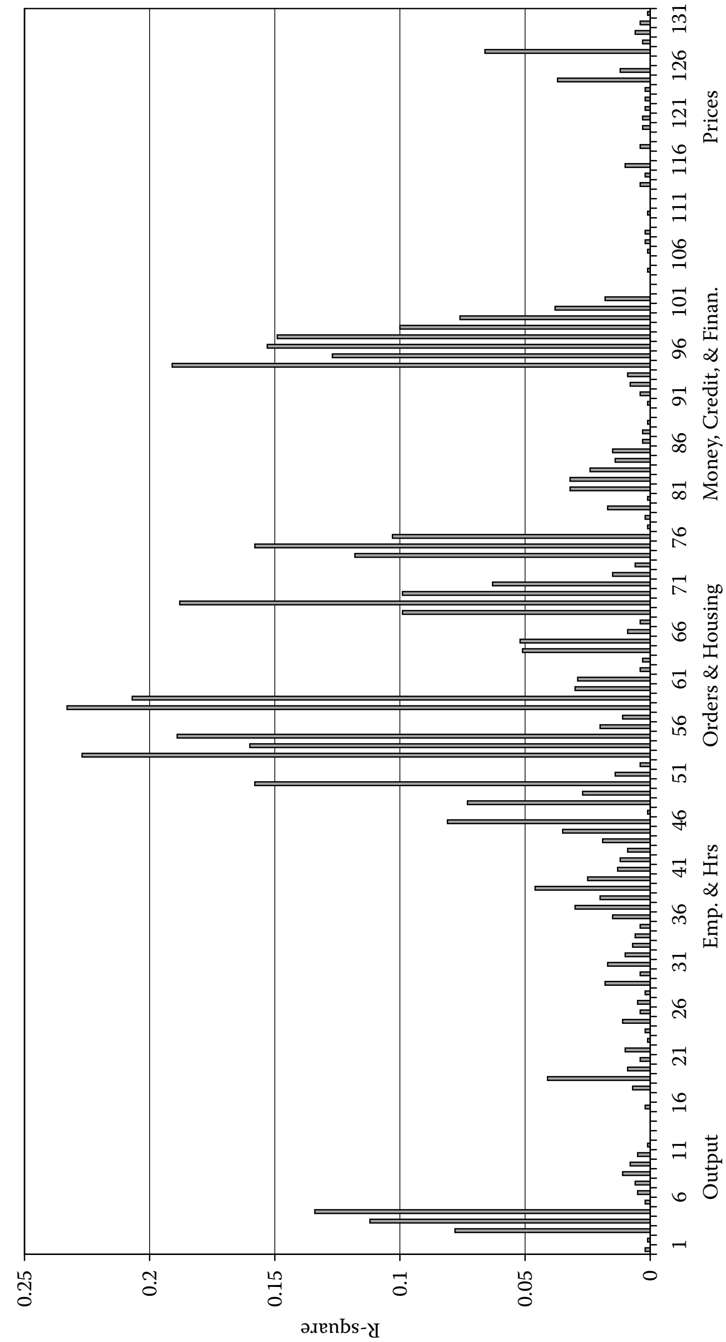
Notes: See Figure 12.1.

FIGURE 12.4
Marginal R-squares for F_4.



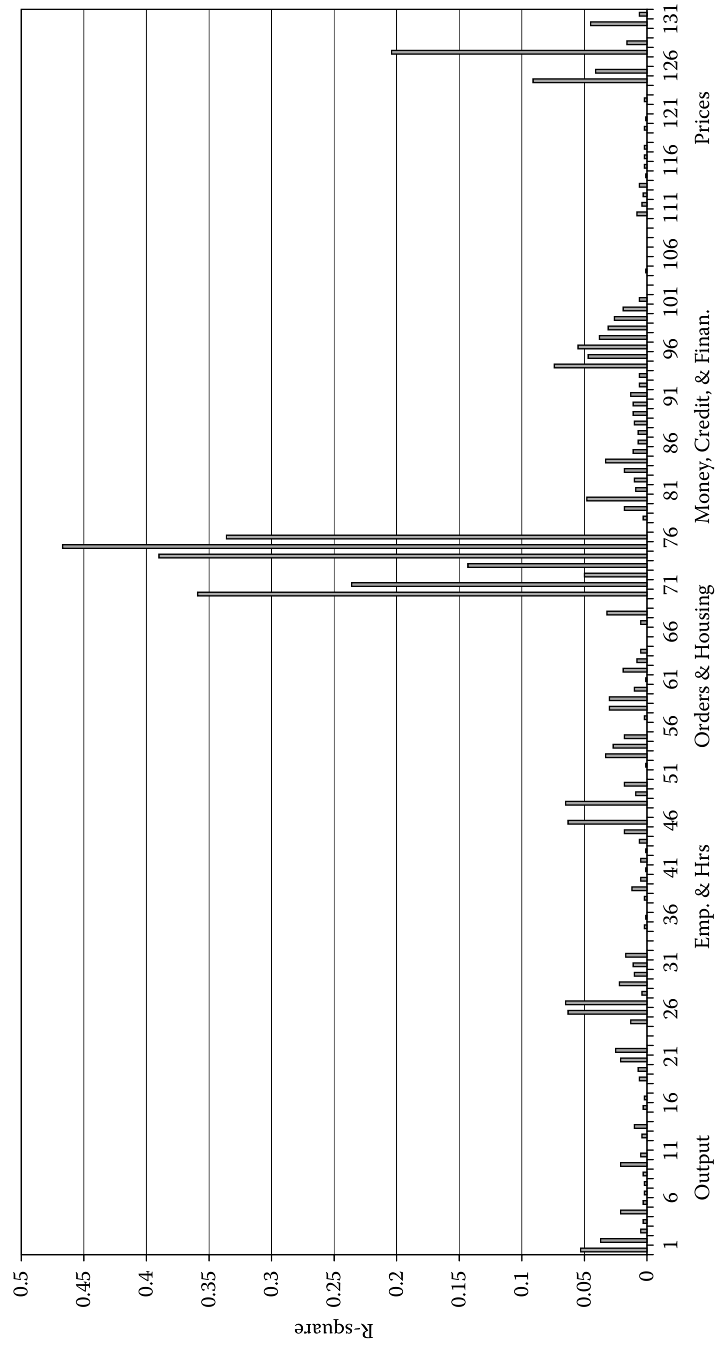
Notes: See Figure 12.1.

FIGURE 12.5
Marginal R-squares for F_5.



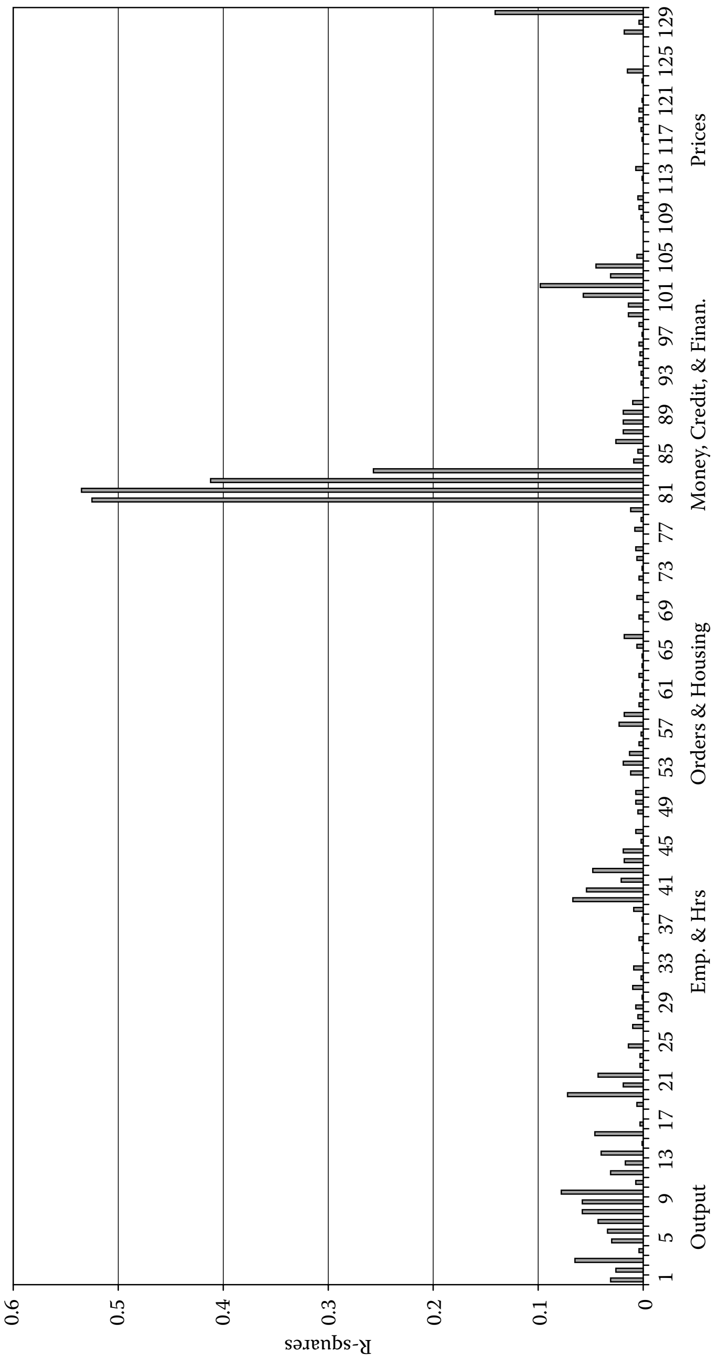
Notes: See Figure 12.1.

FIGURE 12.6
Marginal R-squares for F_6.



Notes: See Figure 12.1.

FIGURE 12.7
Marginal R-squares for F_7.



Notes: See Figure 12.1.

FIGURE 12.8
Marginal R-squares for F_8.

First, we use prior information to organize the data into eight blocks. These are (1) output, (2) labor market, (3) housing sector, (4) orders and inventories, (5) money and credit, (6) bond and forex, (7) prices, and (8) stock market. The largest block is the labor market which has 30 series, while the smallest group is the stock market block, which only has four series. The advantage of estimating the factors (which will now be denoted g_t) from blocks of data is that the factor estimates are easy to interpret.

Second, we estimate a dynamic factor model specified as

$$x_{it} = \beta_i'(L)g_t + e_{xit}, \tag{12.6}$$

where $\beta_i(L) = (1 - \lambda_{i1}L - \ldots - \lambda_{is}L^s)$ is a vector of dynamic factor loadings of order s and g_t is a vector of q "dynamic factors" evolving as

$$\psi_g(L)g_t = \epsilon_{gt},$$

where $\psi_g(L)$ is a polynomial in L of order p_G, ϵ_{gt} are i.i.d. errors. Furthermore, the idiosyncratic component e_{xit} is an autoregressive process of order p_X so that

$$\psi_x(L)e_{xit} = \epsilon_{xit}.$$

This is the factor framework used in Stock and Watson (1989) to estimate the coincident indicator with $N = 4$ variables. Here, our N can be as large as 30.

The dimension of g_t, (which also equals the dimension of ϵ_t), is referred to as the number of dynamic factors. The main distinction between the static and the dynamic model is best understood using a simple example. The model $x_{it} = \beta_{i0}g_t + \beta_{i1}g_{t-1} + e_{it}$ is the same as $x_{it} = \lambda_{i1}f_{1t} + \lambda_{i2}f_{2t}$ with $f_{1t} = g_t$ and $f_{2t} = g_{t-1}$. Here, the number of factors in the static model is two but there is only one factor in the dynamic model. Essentially, the static model does not take into account that f_t and f_{t-1} are dynamically linked. Forni et al. (2005) showed that when N and T are both large, the space spanned by g_t can also be consistently estimated using the method of dynamic principal components originally developed in Brillinger (1981). Boivin and Ng (2005) find that static and dynamic principal components have similar forecast precision, but that static principal components are much easier to compute. It is an open question whether to use the static or the dynamic factors in predictive regressions though the majority of factor augmented regressions use the static factor estimates. Our results will shed some light on this issue.

We estimate a dynamic factor model for each of the eight blocks. Given the definition of the blocks, it is natural to refer to g_{1t} as an output factor, g_{7t} as a price factor, and so on. However, as some blocks have a small number of series, the (static or dynamic) principal components estimator which assumes that N and T are both large will give imprecise estimates. We therefore use the Bayesian method of Monte Carlo Markov Chain (MCMC). MCMC samples a chain that has the posterior density of the parameters as its stationary distribution. The posterior mean computed from draws of the chain are then unbiased for g_t. For factor models, Kose, Otrok, and Whiteman (2003) use an algorithm that involves inversion of N matrices that are of dimension $T \times T$, which can be computationally demanding. The algorithms used in Aguilar and West (2000), Geweke and Zhou (1996), and Lopes and West (2004) are extensions of the MCMC method developed in Carter and Kohn (1994) and Fruhwirth-Schnatter (1994). Our method is similar and follows the implementation in Kim and Nelson (2000) of the Stock–Watson coincident indicator closely. Specifically, we first put the dynamic factor model into a state-space framework. We assume $p_X = p_G = 1$ and $s_g = 2$ for every block. For $i = 1, \ldots N_b$ (the number of series in block b), let x_{ibt} be the observation for unit i of block b at time t. Given that $p_X = 1$, the measurement equation is

$$(1 - \psi_{bi} L)x_{bit} = (1 - \psi_{bi} L)(\beta_{bi0} + \beta_{bi1} L + \beta_{bi2} L^2)g_{bt} + \epsilon_{Xbit}$$

or more compactly,

$$x^*_{bit} = \beta^*_i(L)g_{bt} + \epsilon_{Xbit}.$$

Given that $p_G = 1$, the transition equation is

$$g_{bt} = \psi_{gb} g_{bt-1} + \epsilon_{gbt}.$$

We assume $\epsilon_{Xbit} \sim N(0, \sigma^2_{Xbi})$ and $\epsilon_{gb} \sim N(0, \sigma^2_{gb})$. We use principal components to initialize g_{bt}. The parameters $\beta_b = (\beta_{b1}, \ldots, \beta_{b,Nb})$, $\psi_{Xb} = \psi_{Xb1}, \ldots, \psi_{Xb,Nb}$ are initialized to zero. Furthermore, $\sigma_{Xb} = (\sigma_{Xb1}, \ldots, \sigma_{Xb,N_b})$, ψ_{gb}, and σ^2_{gb} are initialized to random draws from the uniform distribution. For $b = 1, \ldots, 8$ blocks, Gibbs sampling can now be implemented by successive iteration of the following steps:

1. Draw $g_b = (g_{b1}, \ldots g_{bT})'$ conditional on β_b, ψ_{Xb}, σ_{Xb} and the $T \times N_b$ data matrix x_b.
2. Draw ψ_{gb} and σ^2_{gb} conditional on g_b.
3. For each $i = 1, \ldots N_b$, draw β_{bi}, ψ_{Xbi} and σ^2_{Xbi} conditional on g_b and x_b.

We assume normal priors for $\beta_{bi} = (\beta_{i0}, \beta_{i1}, \beta_{i2})$, ψ_{Xbi} and ψ_{gb}. Given conjugacy, β_{bi}, ψ_{Xbi}, ψ_{gb}, are simply draws from the normal distributions whose posterior means and variances are straightforward to compute. Similarly, σ^2_{gb} and σ^2_{Xbi} are draws from the inverse chi-square distribution. Because the model is linear and Gaussian, we can run the Kalman filter forward to obtain the conditional mean $g_{bT|T}$ and conditional variance $P_{bT|T}$. We then draw g_{bT} from its conditional distribution, which is normal, and proceed backwards to generate draws $g_{bt|T}$ for $t = T-1, \ldots, 1$ using the Kalman filter. For identification, the loading on the first series in each block is set to 1. We take 12,000 draws and discard the first 2000. The posterior means are computed from every 10th draw after the burn-in period. The $\hat{g}_t$s used in subsequent analysis are the means of these 1000 draws.

As in the case of static factors, not every g_{bt} need to have predictive power for excess bond returns. Let $G_t \subset g_t = (g_{1t}, \ldots g_{8t})$ be those that do. The analog to Equation 12.5 using dynamic factors is

$$rx^{(n)}_{t+1} = \alpha'_G \hat{G}_t + \beta'_G Z_t + \epsilon_{t+1}, \quad (12.7)$$

TABLE 12.1

First Order Autocorrelation Coefficients

	$\hat{f}_t$	t	$\hat{g}_t$	t
1	0.767	20.589	−0.361	−6.298
2	0.748	18.085	0.823	22.157
3	−0.239	−2.852	0.877	32.267
4	0.456	7.594	0.660	14.385
5	0.362	6.819	−0.344	−1.635
6	0.422	4.232	0.448	4.552
7	−0.112	−0.672	0.050	0.609
8	0.225	4.526	0.157	2.794

We have now obtained two sets of factor estimates using two distinct methodologies. We can turn to an assessment of whether the estimates of the predictive regression are sensitive to how the factors are estimated.

12.3.3 Comparison of $\hat{f}_t$ and $\hat{g}_t$

Table 12.1 reports the first order autocorrelation coefficients for f_t and g_t. Both sets of factors exhibit persistence, with $\hat{f}_{1t}$ being the most correlated of the eight $\hat{f}_t$, and $\hat{g}_{3t}$ being the most serially correlated amongst the $\hat{g}_t$. Table 12.2 reports the contemporaneous correlations between $\hat{f}$ and $\hat{g}$. The real activity factor $\hat{f}_1$ is highly correlated with the $\hat{g}_t$ estimated from output, labor, and manufacturing blocks. $\hat{f}_2$, $\hat{f}_4$, and $\hat{f}_5$ are correlated with many of the $\hat{g}$, but the correlations with the bond/exchange rate seem strongest. $\hat{f}_3$ is predominantly a price factor, while $\hat{f}_8$ is a stock market factor. $\hat{f}_7$ is most correlated with $\hat{g}_5$, which is a money market factor. $\hat{f}_8$ is highly correlated with $\hat{g}_8$, which is estimated from stock market data.

The contemporaneous correlations reported in Table 12.2 do not give a full picture of the correlation between $\hat{f}_t$ and $\hat{g}_t$ for two reasons. First, the $\hat{g}_t$ are not mutually uncorrelated, and second, they do not account for correlations that might occur at lags. To provide a sense of the dynamic correlation between $\hat{f}$

TABLE 12.2

Correlation between $\hat{f}_t$ and g_t

	$\hat{g}_1$ Output	$\hat{g}_2$ Labor	$\hat{g}_3$ Housing	$\hat{g}_4$ Mfg.	$\hat{g}_5$ Money	$\hat{g}_6$ Finance	$\hat{g}_7$ Prices	$\hat{g}_8$ Stocks
$\hat{f}_1$	0.601	0.903	0.551	0.766	−0.067	0.489	0.126	−0.092
$\hat{f}_2$	0.181	−0.120	0.376	0.269	0.095	−0.462	−0.227	0.449
$\hat{f}_3$	0.037	0.027	−0.150	−0.010	−0.148	0.144	−0.800	−0.067
$\hat{f}_4$	−0.303	0.118	0.253	−0.128	0.185	−0.417	−0.194	0.092
$\hat{f}_5$	0.306	0.179	−0.365	0.026	0.046	−0.474	−0.009	0.183
$\hat{f}_6$	0.103	−0.140	0.321	0.179	−0.398	0.008	0.050	0.177
$\hat{f}_7$	0.064	−0.023	0.125	0.004	0.743	0.088	−0.078	0.100
$\hat{f}_8$	−0.241	0.073	−0.023	0.111	−0.057	0.119	−0.052	0.689

TABLE 12.3

Long Run Correlation between $\hat{f}_t$ and $\hat{g}_t$

	$\hat{g}_1$ **Output**	$\hat{g}_2$ **Labor**	$\hat{g}_3$ **Housing**	$\hat{g}_4$ **Mfg.**	$\hat{g}_5$ **Money**	$\hat{g}_6$ **Finance**	$\hat{g}_7$ **Prices**	$\hat{g}_8$ **Stocks**	R^2
$\hat{f}_1$	0.447	0.536	0.215	0.066	−0.008	0.140	−0.002	−0.038	0.953
$\hat{f}_2$	0.548	−0.466	0.296	0.299	0.031	−0.536	−0.135	0.266	0.689
$\hat{f}_3$	0.100	0.026	−0.152	−0.036	−0.007	0.211	−0.390	−0.026	0.935
$\hat{f}_4$	−0.925	0.699	0.491	−0.242	0.004	−0.444	−0.077	−0.064	0.723
$\hat{f}_5$	0.682	0.417	−0.624	−0.135	−0.000	−0.488	0.018	0.146	0.790
$\hat{f}_6$	0.070	−0.357	0.467	−0.098	−0.294	0.144	0.061	0.100	0.490
$\hat{f}_7$	0.226	−0.252	0.136	−0.095	0.540	0.325	−0.080	0.180	0.692
$\hat{f}_8$	−0.986	0.447	−0.224	0.167	0.025	0.313	−0.049	0.905	0.797

Reported are estimates of $A_{r.0}$, obtained from the regression: $\hat{f}_{rt} = A_{r.0}\hat{g}_t + \sum_{i=1}^{p-1} A_{r.i}\Delta g_{t-i} + e_t$ with $p = 4$.

and $\hat{g}_t$, we first standardize $\hat{f}_t$ and $\hat{g}_t$ to have unit variance. We then consider the regression

$$\hat{f}_{rt} = a + A_{r.0}\hat{g}_t + \sum_{i=1}^{p-1} A_{r.i}\Delta\hat{g}_{t-i} + e_{it},$$

where for $r = 1, \ldots, 8$ and $i = 0, \ldots, p-1$, $A_{r.i}$ is a 8×1 vector of coefficients summarizing the dynamic relation between $\hat{f}_{rt}$ and lags of $\hat{g}_t$. The coefficient vector $A_{r.0}$ summarizes the long-run relation between $\hat{g}_t$ and $\hat{f}_t$. Table 12.3 reports results for $p = 4$, along with the R^2 of the regression. Except for $\hat{f}_6$, the current value and lags of $\hat{g}_t$ explain the principal components quite well. While it is clear that $\hat{f}_1$ is a real activity factor, the remaining $\hat{f}$s tend to load on variables from different categories. Tables 12.2 and 12.3 reveal that $\hat{g}_t$ and $\hat{f}_t$ reduce the dimensionality of information in the panel of data in different ways. Evidently, the $\hat{f}_t$s are weighted averages of the $\hat{g}_t$s and their lags. This can be important in understanding the results to follow.

12.4 Predictive Regressions

Let $\hat{H}_t \subset \hat{h}_t$, where $\hat{h}_t$ is either $\hat{f}_t$ or $\hat{g}_t$. Our predictive regression can generically be written as

$$rx_{t+1}^{(n)} = \alpha'\hat{H}_t + \beta' CP_t + \epsilon_{t+1}. \tag{12.8}$$

Equation 12.8 allows us to assess whether $\widehat{H}_t$ has predictive power for excess bond returns, conditional on the information in CP_t. In order to assess whether macro factors $\widehat{H}_t$ have unconditional predictive power for future returns, we also consider the restricted regression

$$rx_{t+1}^{(n)} = \alpha'\widehat{H}_t + \epsilon_{t+1}. \tag{12.9}$$

Since $\hat{F}_t$ and $\hat{G}_t$ are both linear combinations of $x_t = (x_{1t}, \ldots x_{Nt})'$, say $F_t = q_F' x_t$ and $G_t = q_G' x_t$, we can also write Equation 12.8 as

$$rx_{t+1}^{(n)} = \alpha^{*\prime} x_t + \beta' CP_t + \epsilon_{t+1}$$

where $\alpha^{*\prime} = \alpha_F' q_F'$ or $\alpha_G' q_G'$. The conventional regression Equation 12.1 puts a weight of zero on all but a handful of x_{it}. When $\hat{H}_t = \hat{F}_t$, q_F is related to the k eigenvectors of $xx'/(NT)$ that will not, in general, be numerically equal to zero. When $\hat{H}_t = \hat{G}_t$, q_G and thus α^* will have many zeros since each column of $\hat{G}_t$ is estimated using a subset of x_t. Viewed in this light, a factor augmented regression with PCA down-weights unimportant regressors. A FAR estimated using blocks of data sets put some but not all coefficients on x_t equal to zero. A conventional regression is most restrictive as it constrains almost the entire α^* vector to zero.

As discussed earlier, factors that are pervasive in the panel of data x_{it} need not have predictive power for $rx_{t+1}^{(n)}$, which is our variable of interest. In Ludvigson and Ng (2007), $\hat{H}_t = \hat{F}_t$ was determined using a method similar to that used in Stock and Watson (2002b). We form different subsets of $\hat{f}_t$, and/or functions of $\hat{f}_t$ (such as $\hat{f}_{1t}^2$). For each candidate set of factors, $\widehat{F}_t$, we regress $rx_{t+1}^{(n)}$ on $\widehat{F}_t$ and CP_t and evaluate the corresponding in-sample BIC and $\bar{R}^2$. The in-sample BIC for a model with k regressors is defined as

$$\text{BIC}_{in}(k) = \hat{\sigma}_k^2 + k\frac{\log T}{T},$$

where $\hat{\sigma}_k^2$ is the variance of the regression estimated over the entire sample. To limit the number of specifications we search over, we first evaluate r univariate regressions of returns on each of the r factors. Then, for only those factors found to be significant in the r univariate regressions, we evaluate whether the squared and the cubed terms help reduce the BIC criterion further. We do not consider other polynomial terms, or polynomial terms of factors not important in the regressions on linear terms.

In this chapter, we again use the BIC to find the preferred set of factors, but we perform a systematic and therefore much larger search. Instead of relying on results from preliminary univariate regressions to guide us to the final model, we directly search over a large number of models with different numbers of regressors. We want to allow excess bond returns to be possibly nonlinear in the eight factors and hence include the squared terms as candidate regressors. If we additionally include all the cubic terms, and given that we have eight factors and CP to consider, we would have over thirteen million (2^{27}) potential models. As a compromise, we limit our candidate regressor set to eighteen variables: $(\hat{f}_{1t}, \ldots, f_{8t}; \hat{f}_{1t}^2, \ldots, f_{8t}^2; \hat{f}_{1t}^3, CP_t)$. We also restrict the maximum number of predictors to eight. This leads to an evaluation of 106,762 models.[5]

[5] This is obtained by considering $C_{18,j}$ for $j = 1, \ldots, 8$, where $C_{n,k}$ denotes choosing k out of n potential predictors.

The purpose of this extensive search is to assess the potential impact on the forecasting analysis of fishing over large numbers of possible predictor factors. As we show, the factors chosen by the larger, more systematic, search are the same as those chosen by the limited search procedure used in Ludvigson and Ng (2007). This suggests that data mining does not in practice unduly influence the findings in this application, since we find that the same few key factors always emerge as important predictor variables regardless of how extensive the search is.

It is well known that variables found to have predictive power in-sample do not necessarily have predictability out of sample. As discussed in Hansen (2008), in-sample overfitting generally leads to a poor out-of-sample fit. One is less likely to produce spurious results based on an out-of-sample criterion because a complex (large) model is less likely to be chosen in an out-of-sample comparison with simple models when both models nests the true model. Thus, when a complex model is found to outperform a simple model out of sample, it is stronger evidence in favor of the complex model. To this end, we also find the best among 106,762 models as the minimizer of the out-of-sample BIC. Specifically, we split the sample at $t = T/2$. Each model is estimated using the first $T/2$ observations. For $t = T/2 + 1, \ldots, T$, the values of predictors in the second half of the sample are multiplied into the parameters estimated using the first half of the sample to obtain the fit, denoted $\hat{r}x_{t+12}$. Let $\tilde{e}_t = rx_{t+12} - \hat{r}x_{t+12}$ and $\tilde{\sigma}_k^2 = \frac{1}{T/2}\sum_{t=T/2+1}^{T} \tilde{e}_t^2$ be the out-of-sample error variance corresponding to model j. The out-of-sample BIC is defined as

$$\text{BIC}_{\text{out}}(j) = \log \tilde{\sigma}_j^2 + \frac{\dim_j \log(T/2)}{T/2},$$

where dim_j is the size of model j. By using an out-of-sample BIC selection criterion, we guard against the possibility of spurious overfitting. Regressors with good predictive power only over a subsample will not likely be chosen. As the predictor set may differ depending on whether the CP factor is included (i.e., whether we consider Equations 12.8 and 12.9), the two variable selection procedures are repeated with CP excluded from the potential predictor set. Using the predictors selected by the in- and the out-of-sample BIC, we reestimate the predictive regression over the entire sample. In the next section, we show that the predictors found by this elaborate search are the same handful of predictors found in Ludvigson and Ng (2007) and that these handful of macroeconomic factors have robust significant predictive power for excess bond returns beyond the CP factor.

We also consider as predictor a linear combination of $\hat{h}_t$ along the lines of Cochrane and Piazzesi (2005). This variable, denoted $\hat{H}8_t$ is defined as $\hat{\gamma}'\hat{h}_t^+$ where $\hat{\gamma}$ is obtained from the following regression:

$$\frac{1}{4}\sum_{n=2}^{5} rx_{t+1}^{n} = \gamma_0 + \gamma'\hat{h}_t^+, \tag{12.10}$$

with $\hat{h}_t^+ = (\hat{h}_{1t}, \ldots, \hat{h}_{8t}, \hat{h}_{1t}^3)$. The estimates are as follows:

	$h_t = \hat{f}_t$		$h_t = \hat{g}_t$	
	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$
h_1	−1.681	−4.983	0.053	0.343
h_2	0.863	3.009	−1.343	−2.593
h_3	−0.018	−0.203	−0.699	−1.891
h_4	−0.626	−2.167	0.628	1.351
h_5	−0.264	−1.463	−0.001	−0.012
h_6	−0.720	−2.437	−0.149	−0.691
h_7	−0.426	−2.140	−0.018	−0.210
h_8	0.665	3.890	−0.418	−2.122
h_1^3	0.115	3.767	0.049	1.733
cons	0.900	2.131	0.764	1.518
$\bar{R}^2$	0.261		0.104	

Notice that we could also have replaced $\hat{h}_t$ in the above regression with $\hat{H}_t$, where $\hat{H}_t$ comprises predictors selected by either the in- or the out-of-sample BIC. However, $\hat{H}8_t$ is a factor-based predictor that is arguable less vulnerable to the effects of data mining because it is simply a linear combination of all the estimated factors.

Tables 12.4 to 12.7 report results for maturities of 2, 3, 4, and 5 years. The first four columns of each table are based on the static factors (i.e., $\hat{H}_t = \hat{F}_t$), while columns 5 to 8 are based on the dynamic factors (i.e., $\hat{H}_t = \hat{G}_t$). Of these, columns 1, 2, 5, and 6 include the CP variable, while columns 3, 4, 7, and 8 do not include the CP. Columns 9 and 10 report results using $\hat{F}8$ with and without CP and columns 11 and 12 do the same with $\hat{G}8$ in place. Our benchmark is a regression that has the CP variable as the sole predictor. This is reported in last column, i.e., column 13.

12.4.1 Two-Year Returns

As can be seen from Table 12.4, the CP alone explains 0.309 of the variance in the 2-year excess bound returns. The variable $\hat{F}_8$ alone explains 0.279 (column 10), while $\hat{G}_8$ alone explains only 0.153 of the variation (column 12). Adding $\hat{F}8$ to the regression with the CP factor (column 9) increases $\bar{R}^2$ to 0.419, and adding $\hat{G}8$ (column 11) to CP yields an $\bar{R}^2$ of 0.401. The macroeconomic factors thus have nontrivial predictive power above and beyond the CP factor.

We next turn to regressions when both the factors and CP are included. In Ludvigson and Ng (2007), the static factors $\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}, \hat{f}_{8t}$, and CP are found to have the best predictive power for excess returns. The in-sample BIC still finds the same predictors to be important, but adds $\hat{f}_{6t}$ and $\hat{f}_{5t}^2$ to the predictor list. It is, however, noteworthy that some variables selected by the BIC have individual t statistics that are not significant. The resulting model has an $\bar{R}^2$ of 0.460 (column 1). The out-of-sample BIC selects smaller models and finds $\hat{f}_1, \hat{f}_8, \hat{f}_5^2, \hat{f}_1^3$, and the CP to be important regressors (column 2).

TABLE 12.4

Regressions $rx^{(2)}_{t+1} = a + \alpha' \hat{H}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{H}=\hat{F}$				$\hat{H}=\hat{G}$				$\hat{H}=\hat{F}$		$\hat{H}=\hat{G}$		
	In	Out	In	Out	In	Out	In	Out					
$\hat{H}$	1	2	3	4	5	6	7	8	9	10	11	12	13
$\hat{H}_1$	−0.761	−0.793	−0.935	−0.931	-	-	0.147	0.170	-	-	-	-	-
tstat	−5.387	−4.848	−5.748	−5.449	-	-	2.947	2.623	-	-	-	-	-
$\hat{H}_2$	-	-	0.325	0.326	−0.494	−0.627	−0.699	−0.646	-	-	-	-	-
tstat	-	-	2.663	2.520	−3.151	−3.623	−2.905	−3.062	-	-	-	-	-
$\hat{H}_3$	-	-	-	-	−0.492	-	−0.532	−0.487	-	-	-	-	-
tstat	-	-	-	-	−4.813	-	−2.889	−3.012	-	-	-	-	-
$\hat{H}_4$	−0.291	-	−0.399	−0.399	-	-	0.186	-	-	-	-	-	-
tstat	−2.716	-	−3.103	−2.974	-	-	1.039	-	-	-	-	-	-
$\hat{H}_6$	−0.151	-	−0.281	−0.280	0.137	-	−0.163	-	-	-	-	-	-
tstat	−1.322	-	−1.949	−1.795	1.679	-	−1.594	-	-	-	-	-	-
$\hat{H}_7$	−0.128	-	−0.143	−0.144	-	-	-	-	-	-	-	-	-
tstat	−1.577	-	−1.517	−1.365	-	-	-	-	-	-	-	-	-
$\hat{H}_8$	0.240	0.241	0.302	-	−0.136	-	−0.164	-	-	-	-	-	-
tstat	2.981	3.297	3.575	-	−1.562	-	−1.997	-	-	-	-	-	-
$\hat{H}^2_2$	-	-	-	-	-	−0.100	-	-	-	-	-	-	-
tstat	-	-	-	-	-	−2.147	-	-	-	-	-	-	-
$\hat{H}^2_4$	-	-	-	-	−0.074	-	−0.121	−0.118	-	-	-	-	-
tstat	-	-	-	-	−3.165	-	−3.167	−3.076	-	-	-	-	-
$\hat{H}^2_5$	−0.080	−0.110	-	-	-	-	-	-	-	-	-	-	-
tstat	−2.468	−2.925	-	-	-	-	-	-	-	-	-	-	-
$\hat{H}^2_6$	-	-	-	-	−0.086	−0.083	−0.084	−0.080	-	-	-	-	-

(*continued*)

TABLE 12.4 (Continued)

Regressions $rx^{(2)}_{t+1} = a + \alpha' \hat{H}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{H} = \hat{F}$				$\hat{H} = \hat{G}$				$\hat{H} = \hat{F}$		$\hat{H} = \hat{G}$		
	In	Out	In	Out	In	Out	In	Out					
$\hat{H}$	1	2	3	4	5	6	7	8	9	10	11	12	13
tstat	-	-	-	-	−6.245	−6.804	−3.642	−3.176	-	-	-	-	-
$\hat{H}^3_1$	0.044	0.047	0.057	0.056	0.019	-	-	-	-	-	-	-	-
tstat	2.912	2.887	3.081	3.338	2.254	-	-	-	-	-	-	-	-
CP	0.385	0.411	-	-	0.452	0.433	-	-	0.336	-	0.413	-	0.455
tstat	5.647	6.981	-	-	7.488	7.738	-	-	4.437	-	6.434	-	8.836
$\hat{H}8$	-	-	-	-	-	-	-	-	0.332	0.482	0.427	0.544	-
tstat	-	-	-	-	-	-	-	-	4.336	7.212	3.880	3.493	-
$\bar{R}^2$	0.460	0.430	0.283	0.258	0.477	0.407	0.200	0.192	0.419	0.279	0.401	0.153	0.309

Note: The table reports estimates from OLS regressions of excess bond returns on the lagged variables named in column 1. The dependent variable rx^n_{t+1} is the excess log return on the n year Treasury bond. $\widehat{H}_t$ denotes a set of regressors formed from consisting of functions of $\hat{f}_t$ or $\hat{g}_t$ where $\hat{f}_t$ is a set of eight factors estimated by the method of principal components, and $\hat{g}_t$ is a vector of eight dynamic factors estimated by Bayesian factors. The panel of data used in estimation consists of 131 individual series over the period 1964:1 to 2007:12. $\hat{H}8_t$ is the single factor constructed as a linear combination of the eight estimated factors and $\hat{f}^3_1$. CP_t is the Cochrane and Piazzesi (2005) factor that is a linear combination of five forward spreads. Newey and West (1987) corrected t-statistics have lag order 18 months and are reported in parentheses. A constant is always included in the regression even though its estimate is not reported in the table.

TABLE 12.5

Regressions $rx^{(3)}_{t+1} = a + \alpha' \hat{H}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{H} = \hat{F}$				$\hat{H} = \hat{G}$								
	In	**Out**	**In**	**Out**	**In**	**Out**	**In**	**Out**	$\hat{H} = \hat{F}$		$\hat{H} = \hat{G}$		
$\hat{H}$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
$\hat{H}_1$	−1.232	−1.280	−1.624	−1.592	-	-	-	-	-	-	-	-	-
tstat	−5.079	−4.581	−5.553	−5.479	-	-	-	-	-	-	-	-	-
$\hat{H}_2$	−0.028	-	0.694	0.703	−0.782	−1.094	−1.259	−1.056	-	-	-	-	-
tstat	−0.147	-	2.851	2.982	−2.805	−3.773	−2.983	−3.092	-	-	-	-	-
$\hat{H}_3$	-	-	-	-	−0.807	-	−0.843	−0.734	-	-	-	-	-
tstat	-	-	-	-	−4.297	-	−2.667	−2.548	-	-	-	-	-
$\hat{H}_4$	−0.423	-	−0.588	−0.592	-	-	0.421	-	-	-	-	-	-
tstat	−2.193	-	−2.518	−2.496	-	-	1.225	-	-	-	-	-	-
$\hat{H}_6$	−0.433	-	−0.598	−0.590	-	-	−0.356	-	-	-	-	-	-
tstat	−1.890	-	−2.294	−2.269	-	-	−2.006	-	-	-	-	-	-
$\hat{H}_7$	−0.338	-	−0.360	−0.342	-	-	-	-	-	-	-	-	-
tstat	−2.138	-	−2.109	−1.989	-	-	-	-	-	-	-	-	-
$\hat{H}_8$	0.389	0.428	0.550	0.553	−0.308	-	−0.329	-	-	-	-	-	-
tstat	2.593	3.190	3.718	3.738	−2.018	-	−2.143	-	-	-	-	-	-
$\hat{H}_1^2$	-	-	0.156	-	-	-	-	-	-	-	-	-	-
tstat	-	-	0.854	-	-	-	-	-	-	-	-	-	-
$\hat{H}_2^2$	-	-	-	-	-	−0.208	-	-	-	-	-	-	-
tstat	-	-	-	-	-	−2.668	-	-	-	-	-	-	-
$\hat{H}_3^2$	0.111	-	-	-	-	-	-	-	-	-	-	-	-
tstat	1.999	-	-	-	-	-	-	-	-	-	-	-	-

(*continued*)

TABLE 12.5 (Continued)

Regressions $rx_{t+1}^{(3)} = a + \alpha' \hat{H}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{H} = \hat{F}$				$\hat{H} = \hat{G}$								
	In	**Out**	**In**	**Out**	**In**	**Out**	**In**	**Out**	$\hat{H} = \hat{F}$		$\hat{H} = \hat{G}$		
$\hat{H}$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
$\hat{H}_4^2$	-	-	-	-	−0.190	-	−0.250	−0.275	-	-	-	-	-
tstat	-	-	-	-	−3.925	-	−3.005	−3.622	-	-	-	-	-
$\hat{H}_5^2$	-	−0.161	-	-	-	-	-	-	-	-	-	-	-
tstat	-	−2.179	-	-	-	-	-	-	-	-	-	-	-
$\hat{H}_6^2$	-	-	-	-	−0.152	−0.147	−0.140	−0.127	-	-	-	-	-
tstat	-	-	-	-	−7.130	−6.883	−3.307	−2.551	-	-	-	-	-
$\hat{H}_7^2$	-	-	-	-	0.089	-	-	-	-	-	-	-	-
tstat	-	-	-	-	2.687	-	-	-	-	-	-	-	-
$\hat{H}_1^3$	0.095	0.086	0.141	0.106	0.032	-	0.031	-	-	-	-	-	-
tstat	3.235	3.204	2.922	3.445	2.233	-	1.942	-	-	-	-	-	-
CP	0.760	0.784	-	-	0.847	0.821	-	-	0.644	-	0.786	-	0.856
tstat	5.329	6.885	-	-	7.516	7.770	-	-	4.661	-	6.381	-	8.301
$\hat{H}8$	-	-	-	-	-	-	-	-	0.588	0.877	0.710	0.931	-
tstat	-	-	-	-	-	-	-	-	4.494	7.133	3.624	3.256	-
$\bar{R}^2$	0.455	0.424	0.268	0.267	0.475	0.418	0.182	0.167	0.432	0.277	0.404	0.135	0.328

TABLE 12.6

Regressions $rx_{t+1}^{(4)} = a + \alpha' \hat{H}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{H} = \hat{F}$				$\hat{H} = \hat{G}$				$\hat{H} = \hat{F}$		$\hat{H} = \hat{G}$		
	In	Out	In	Out	In	Out	In	Out					
$\hat{H}$	1	2	3	4	5	6	7	8	9	10	11	12	13
$\hat{H}_1$	−1.521	−1.521	−2.011	−2.050	-	-	-	-	-	-	-	-	-
tstat	−5.138	−4.149	−5.013	−5.290	-	-	-	-	-	-	-	-	-
$\hat{H}_2$	-	-	1.069	1.069	−0.952	−1.342	−1.619	−1.601	-	-	-	-	-
tstat	-	-	3.028	3.095	−2.680	−3.754	−2.812	−2.848	-	-	-	-	-
$\hat{H}_3$	-	-	-	-	−1.036	-	−1.080	−1.078	-	-	-	-	-
tstat	-	-	-	-	−4.127	-	−2.486	−2.401	-	-	-	-	-
$\hat{H}_4$	−0.436	-	−0.689	−0.681	-	-	0.590	0.452	-	-	-	-	-
tstat	−1.595	-	−1.957	−1.978	-	-	1.221	0.927	-	-	-	-	-
$\hat{H}_5$	-	-	−0.321	-	-	-	-	-	-	-	-	-	-
tstat	-	-	−1.475	-	-	-	-	-	-	-	-	-	-
$\hat{H}_6$	−0.668	-	−0.889	−0.889	-	-	−0.605	-	-	-	-	-	-
tstat	−2.160	-	−2.522	−2.449	-	-	−2.333	-	-	-	-	-	-
$\hat{H}_7$	−0.534	–	−0.535	−0.541	-	-	-	-	-	-	-	-	-
tstat	−2.401	-	−2.222	−2.209	-	-	-	-	-	-	-	-	-
$\hat{H}_8$	0.578	0.636	0.820	0.822	−0.474	-	−0.521	-	-	-	-	-	-
tstat	2.820	3.365	3.935	3.914	−2.344	-	−2.277	-	-	-	-	-	-
$\hat{H}_1^2$	-	−0.146	-	-	-	-	-	-	-	-	-	-	-
tstat	-	−0.770	-	-	-	-	-	-	-	-	-	-	-
$\hat{H}_2^2$	-	-	-	-	-	−0.284	-	-	-	-	-	-	-
tstat	-	-	-	-	-	−2.934	-	-	-	-	-	-	-

(*continued*)

TABLE 12.6 (Continued)

Regressions $rx_{t+1}^{(4)} = a + \alpha' \hat{H}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{H} = \hat{F}$				$\hat{H} = \hat{G}$								
	In	Out	In	Out	In	Out	In	Out	$\hat{H} = \hat{F}$		$\hat{H} = \hat{G}$		
$\hat{H}$	1	2	3	4	5	6	7	8	9	10	11	12	13
$\hat{H}_3^2$	0.177	-	-	-	-	-	-	-	-	-	-	-	-
tstat	2.527	-	-	-	-	-	-	-	-	-	-	-	-
$\hat{H}_4^2$	-	-	-	-	−0.262	-	−0.354	−0.367	-	-	-	-	-
tstat	-	-	-	-	−3.692	-	−2.976	−3.552	-	-	-	-	-
$\hat{H}_5^2$	-	−0.228	-	-	-	-	-	-	-	-	-	-	-
tstat	-	−2.309	-	-	-	-	-	-	-	-	-	-	-
$\hat{H}_6^2$	-	-	-	-	−0.231	−0.227	−0.219	−0.189	-	-	-	-	-
tstat	-	-	-	-	−6.923	−9.811	−4.375	−3.248	-	-	-	-	-
$\hat{H}_7^2$	-	-	-	-	0.148	0.104	-	-	-	-	-	-	-
tstat	-	-	-	-	3.258	2.233	-	-	-	-	-	-	-
$\hat{H}_1^3$	0.131	0.081	0.142	0.148	0.037	-	0.036	-	-	-	-	-	-
tstat	3.436	1.483	3.938	3.602	1.964	-	1.599	-	-	-	-	-	-
CP	1.115	1.158	-	-	1.238	1.219	-	-	0.955	-	1.150	-	1.235
tstat	6.077	7.028	-	-	7.821	8.197	-	-	4.765	-	6.417	-	8.224
$\hat{H}8$	-	-	-	-	-	-	-	-	0.777	1.204	0.864	1.188	-
tstat	-	-	-	-	-	-	-	-	4.474	7.247	3.388	3.061	-
$\bar{R}^2$	0.473	0.441	0.263	0.260	0.496	0.445	0.171	0.155	0.452	0.273	0.416	0.114	0.357

TABLE 12.7

Regressions $rx^{(5)}_{t+1} = a + \alpha' \hat{H}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{H} = \hat{F}$				$\hat{H} = \hat{G}$								
	In	**Out**	**In**	**Out**	**In**	**Out**	**In**	**Out**	$\hat{H} = \hat{F}$		$\hat{H} = \hat{G}$		
$\hat{H}$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
$\hat{H}_1$	−1.653	−1.373	−2.214	−2.277	0.308	-	0.326	-	-	-	-	-	-
tstat	−4.723	−3.686	−4.503	−4.819	1.701	-	2.049	-	-	-	-	-	-
$\hat{H}_2$	-	-	1.355	1.355	−1.145	−1.573	−1.928	−1.609	-	-	-	-	-
tstat	-	-	3.111	3.195	−2.653	−3.691	−2.760	−2.994	-	-	-	-	-
$\hat{H}_3$	-	-	-	-	−1.161	-	−1.199	−1.003	-	-	-	-	-
tstat	-	-	-	-	−3.615	-	−2.224	−2.021	-	-	-	-	-
$\hat{H}_4$	−0.516	-	−0.818	−0.805	-	-	0.654	-	-	-	-	-	-
tstat	−1.478	-	−1.861	−1.881	-	-	1.128	-	-	-	-	-	-
$\hat{H}_5$	-	-	−0.523	-	-	-	-	-	-	-	-	-	-
tstat	-	-	−1.969	-	-	-	-	-	-	-	-	-	-
$\hat{H}_6$	−0.856	-	−1.120	−1.120	-	-	−0.678	-	-	-	-	-	-
tstat	−2.150	-	−2.566	−2.462	-	-	−2.049	-	-	-	-	-	-
$\hat{H}_7$	−0.686	-	−0.685	−0.694	-	-	-	-	-	-	-	-	-
tstat	−2.479	-	−2.321	−2.299	-	-	-	-	-	-	-	-	-
$\hat{H}_8$	0.702	0.725	0.985	0.988	−0.563	-	−0.608	-	-	-	-	-	-
tstat	2.756	3.292	3.956	3.907	−2.217	-	−2.156	-	-	-	-	-	-
$\hat{H}_1^2$	-	−0.563	-	-	-	-	-	-	-	-	-	-	-

(continued)

TABLE 12.7 (Continued)

Regressions $rx^{(5)}_{t+1} = a + \alpha' \hat{H}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{H} = \hat{F}$				$\hat{H} = \hat{G}$				$\hat{H} = \hat{F}$		$\hat{H} = \hat{G}$		
	In	Out	In	Out	In	Out	In	Out					
$\hat{H}$	1	2	3	4	5	6	7	8	9	10	11	12	13
tstat	-	−3.037	-	-	-	-	-	-	-	-	-	-	-
$\hat{H}^2_2$	-	-	-	-	-	−0.339	-	-	-	-	-	-	-
tstat	-	-	-	-	-	−2.955	-	-	-	-	-	-	-
$\hat{H}^2_3$	0.204	-	-	-	-	-	-	-	-	-	-	-	-
tstat	2.327	-	-	-	-	-	-	-	-	-	-	-	-
$\hat{H}^2_4$	-	-	-	-	−0.357	-	−0.465	−0.466	-	-	-	-	-
tstat	-	-	-	-	−4.429	-	−3.497	−3.684	-	-	-	-	-
$\hat{H}^2_6$	-	-	-	-	−0.269	−0.279	−0.253	−0.234	-	-	-	-	-
tstat	-	-	-	-	−6.235	−9.685	−4.407	−3.596	-	-	-	-	-
$\hat{H}^2_7$	-	-	-	-	0.179	-	-	-	-	-	-	-	-
tstat	-	-	-	-	3.221	-	-	-	-	-	-	-	-
$\hat{H}^3_1$	0.150	-	0.160	0.170	-	-	-	-	-	-	-	-	-
tstat	3.310	-	3.893	3.440	-	-	-	-	-	-	-	-	-
CP	1.316	1.394	-	-	1.457	1.413	-	-	1.115	-	1.359	-	1.453
tstat	5.603	6.985	-	-	7.237	7.409	-	-	4.370	-	5.969	-	7.576
$\hat{H}8$	-	-	-	-	-	-	-	-	0.938	1.437	0.955	1.338	-
tstat	-	-	-	-	-	-	-	-	4.542	7.281	3.078	2.854	-
$\bar{R}^2$	0.435	0.392	0.251	0.245	0.453	0.408	0.152	0.135	0.422	0.259	0.377	0.097	0.330

Among the dynamic factors, $\hat{g}_2$ (labor market), $\hat{g}_8$ (stock market), $\hat{g}_6^2$ (bonds and foreign exchange) along with CP are selected by both BIC procedures as predictors (columns 5 and 6). Interestingly, the output factor $\hat{g}_1$ is not significant when the CP is included. The out-of-sample BIC has an $\bar{R}^2$ of 0.407, showing that there is a substantial amount of variation in the 2-year excess bond returns that can be predicted by macroeconomic factors. The in-sample BIC additionally selects $\hat{g}_{3t}$, $\hat{g}_{6t}$ and some higher-order terms with an $\bar{R}^2$ of 0.477. Thus, predictive regressions using $\hat{f}_t$ and $\hat{g}_t$ both find a factor relating to real activity ($\hat{f}_{1t}$ or $\hat{g}_{1t}$) and one relating to the stock market ($\hat{f}_{8t}$ or $\hat{g}_{18}$) to have significant predictive power for 2-year excess bond returns.

Results when the regressions do not include the CP variable are in columns 3, 4, 7, and 8. Evidently, $\hat{f}_2$ is now important according to both the in- and out-of-sample BIC, showing that the main effect of CP is to render $\hat{f}_2$ redundant. Furthermore, the out-of-sample BIC now selects a model that is only marginally more parsimonious than that selected by the in-sample BIC. The regressions with $\hat{F}$ alone have an $\bar{R}^2$ of 0.283 and 0.258, respectively, slightly less than what is obtained with CP as the only regressor.

Regressions based on the dynamic factors are qualitatively similar. The factors $\hat{g}_1$, $\hat{g}_3$, and $\hat{g}_4$, found not to be important when CP is included are now selected as relevant predictors when CP is dropped. Without CP, the dynamic factors selected by the in-sample BIC explain 0.2 of the 1-year-ahead variation in excess bond returns, while the more parsimonious model selected by the out-of-sample BIC has an $\bar{R}^2$ of 0.192. These numbers are lower than what we obtain in columns (3) and (4) using $\hat{F}_t$ as predictors.

It is important to stress that we consider the two sets of factor estimates not to perform a horse race of whether the PCA or the Bayesian estimator is better. The purpose instead is to show that macroeconomic factors have predictive power for excess bond returns irrespective of the way we estimate the factors. Although the precise degree of predictability depends on how the factors are estimated, a clear picture emerges. At least 20% of the variation in excess bound returns can be predicted by macroeconomic factors even in the presence of the CP factor.

12.4.2 Longer Maturity Returns and Overview

Tables 12.5 to 12.7 report results for returns with maturity of 3, 4, and 5 years. Most of the static factors found to be useful in predicting $rx_{t+1}^{(2)}$ by the in-sample BIC remain useful in predicting the longer maturity returns. These predictors include $\hat{f}_{1t}$, $\hat{f}_{4t}$, $\hat{f}_{6t}$, $\hat{f}_{7t}$, $\hat{f}_{8t}$, $\hat{f}_{1t}^3$, and CP. Of these, $\hat{f}_{1t}$, $\hat{f}_{8t}$, and CP are also selected by the out-of-sample BIC procedure. The nonlinear term $\hat{f}_{1t}^3$ is an important predictor in equations for all maturity returns except the 5 years. The factors add at least 10 basis points to the $\bar{R}^2$ with CP as the sole predictor.

The dynamic factors found important in explaining 2-year excess return are generally also relevant in regressions for longer maturity excess returns. The in-sample BIC finds $\hat{g}_{2t}$, $\hat{g}_{3t}$, $\hat{g}_{8t}$, $\hat{g}_{4t}^2$, $\hat{g}_{6t}^2$ along with the CP to be important

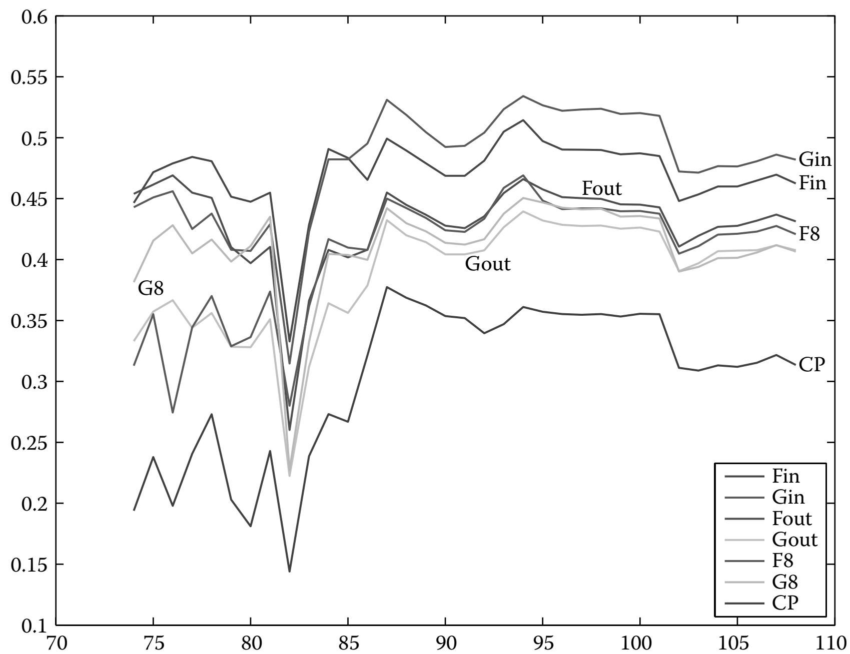


FIGURE 12.9
Adjusted R-squares, with CP. Fin and Gin are the $\bar{R}^2$ from rolling estimation of Equation 12.8, with predictors selected by the in-sample BIC. Fout and Gout use predictors selected by the out-of-sample BIC. F8 and G8 use a linear combination of eight factors as predictors, where the weights are based on Equation 12.10.

in regressions of all maturities. The output factor is again not significant in regressions with 3- and 4-year maturities. It is marginally significant in the 5-year maturity, but has the wrong sign. While $\hat{g}_8$ was relevant in the 2-year regression, it is not an important predictor in the regressions for longer maturity returns. The out-of-sample BIC finds dynamic factors from the labor market ($\hat{g}_{2t}$), the bond and foreign exchange markets ($\hat{g}_{6t}$). Together, these factors have incremental predictive power for excess bond returns over CP, improving the $\bar{R}^2$ by slightly less than 10 basis points.

The relevance of macroeconomic variables in explaining excess bond returns is reinforced by the results in columns 10 and 12, which show that a simple linear combination of the eight factors still adds substantial predictive power beyond the CP factor. This result is robust across all four maturities considered, noting that the coefficient estimate on $\hat{H}8$ increases with the holding period without changing the statistical significance of the coefficient.

To see if the predictability varies over the sample, we also consider rolling regressions. Starting with the first regression that spans the sample 1964:1 to 1974:12, we add 12 monthly observations each time and record the $\bar{R}^2$. Figure 12.9 shows the $\bar{R}^2$ for regressions with CP included. Apart from a notable drop around the 1983 recession, $\bar{R}^2$ is fairly constant. Figure 12.10

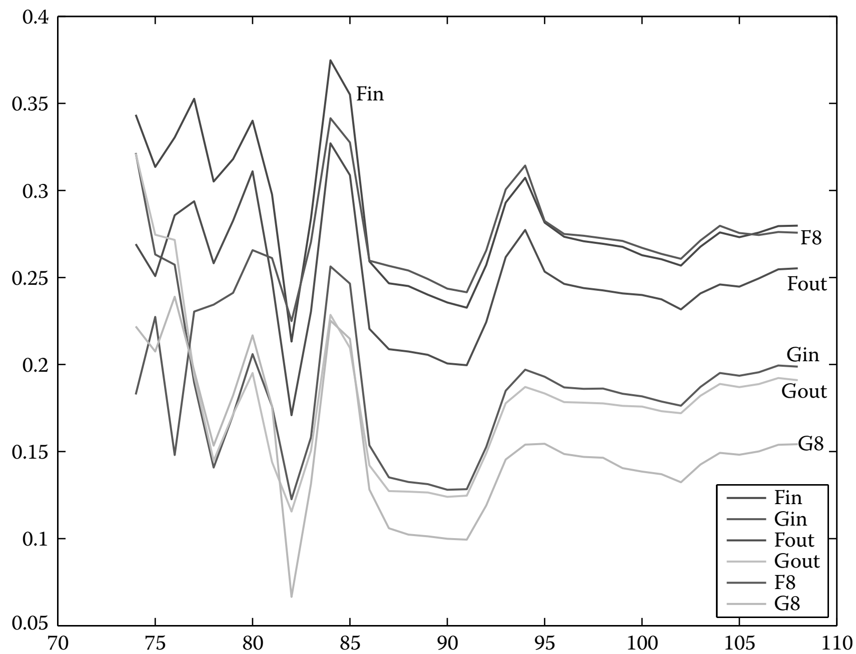


FIGURE 12.10
Adjusted R-squares, without CP. Fin and Gin are the $\bar{R}^2$ from rolling estimation of (8), with predictors selected by the in-sample BIC. Fout and Gout use predictors selected by the out-of-sample BIC. F8 and G8 use a linear combination of eight factors as predictors, where the weights are based on (10).

depicts the $\bar{R}^2$ for regressions without CP. Notice that the $\bar{R}^2$ that corresponds to $\hat{F}8_t$ tends to be 15 basis points higher than $\hat{G}8_t$. As noted earlier, each of the eight $\hat{f}_t$ is itself a combination of the current and lags of the eight $\hat{g}_t$. This underscores the point that imposing a structure on the data to facilitate interpretation of the factors comes at the cost of not letting the data find the best predictive combination possible.

The results reveal that the estimated factors consistently have stronger predictive power for one- and multi-year ahead excess bond returns. The most parsimonious specification has just two variables — $\hat{H}8$ and CP_t — explaining over 40% of the variation in rx^n_{t+1} of every maturity. A closer look reveals that the real activity factor $\hat{f}_{1t}$ is the strongest factor predictor, both numerically and statistically. As $\hat{g}_{1t}$ tends not to be selected as predictor, this suggests that the part of $\hat{f}_{1t}$ that has predictive power for excess bond returns is derived from real activity other than output. However, the dynamic factors $\hat{g}_{2t}$ (labor market) and $\hat{g}_{3t}$ (housing) have strong predictive power. Indeed, $\hat{f}_{1t}$ is highly correlated with $\hat{g}_{2t}$ and the coefficients for these predictors tend to be negative. This means that excess bond returns of every maturity are countercyclical, especially with the labor market. This result is in accord with the models of Campbell (1999) and Wachter (2006), which posit that forecasts of excess returns should be countercyclical because risk aversion is low in good times and high in recessions. We will subsequently show that yield risk premia, which are based on forecasts of excess returns, are also countercyclical.

12.5 Inference Issues

The results thus far assume that N and T are large and that $\sqrt{T}/N$ tends to zero. In this section, we first consider the implication for factor augmented regressions when $\sqrt{T}/N$ may not be small as is assumed. We then examine the finite sample inference issues.

12.5.1 Asymptotic Bias

If excess bond returns truly depend on macroeconomic factors, then consistent estimates of the factors should be better predictors than the observed variables because these are contaminated measures of real activity.[6] An appealing feature of PCA is that if $\sqrt{T}/N \to 0$ as $N, T \to \infty$, then $\hat{F}_t$ can be treated in the predictive regression as though it were F_t. To see why this is the case, consider again the infeasible predictive regression, dropping the observed predictors W_t for simplicity. We have

$$\begin{aligned} rx^n_{t+1} &= \alpha_F^{+\prime} F_t + \epsilon_{t+1} \\ &= \alpha'_F \hat{F}_t + \alpha'_F (HF_t - \hat{F}_t) + \epsilon_{t+1}, \end{aligned}$$

where $\alpha_F = \alpha H^{-1}$, and H is a $r \times r$ matrix defined in Bai and Ng (2006a). Let $S_{\hat{F}\hat{F}} = T^{-1} \sum_{t=1}^{T} \hat{F}_t \hat{F}'_t$. Then

$$\sqrt{T}(\hat{\alpha}_F - \alpha_F) = \hat{S}^{-1}_{\hat{F}\hat{F}} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{F}_t \epsilon_{t+1} \right) + S^{-1}_{\hat{F}\hat{F}} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{F}_t (HF_t - \hat{F}_t) \right) \alpha_F . \tag{12.11}$$

But $T^{-1}\hat{F}'(FH' - \hat{F}) = O_p(\min[N, T]^{-1})$, a result that follows from Bai (2003). Thus if $\sqrt{T}/N \to 0$, the second term is negligible. It follows that

$$\sqrt{T}(\hat{\alpha}_F - \alpha_F) \xrightarrow{d} N(0, \text{Avar}(\hat{\alpha}_F)),$$

where

$$\text{Avar}(\hat{\alpha}_F) = \text{plim } S^{-1}_{\hat{F}\hat{F}} \widehat{\text{Avar}}(g_t) S^{-1}_{\hat{F}\hat{F}},$$

$\widehat{\text{Avar}}(g_t)$ is an estimate of the asymptotic variance of $g_{t+1} = \hat{\epsilon}_{t+1} \hat{F}_t$.

[6] Moench (2008) finds that factors estimated from a large panel of macroeconomic data explain the short rate better than output and inflation.

Consider now the case when $\sqrt{T}$ is comparable to N. Although the first term on the right-hand side of Equation 12.11 is mean zero, the second term is a $O_p(1)$ random variable that may not be mean zero. This generates a bias in the asymptotic distribution for $\hat{\alpha}_F$.

PROPOSITION 12.1 *Suppose assumptions A–E of Bai and Ng (2006a) hold and let $\hat{F}_t \subset \hat{f}_t$, where $\hat{f}_t$ are the principal component estimates of f_t, $x_{it} = \lambda'_{it} f_t + e_{it}$. Let $\hat{\alpha}_F$ be obtained from least squares estimation of the FAR $y_{t+h} = \alpha'_F \hat{F}_t + e_{t+h}$. An estimate of the bias in $\hat{\alpha}_F$ is*

$$\hat{B}_1 \approx -S^{-1}_{\hat{F}\hat{F}} \left(\frac{1}{NT} \sum_{t=1}^{T} \widehat{\text{Avar}_t}(\hat{F}_t) \right) \hat{\alpha}_F,$$

where $\text{Avar}_t(F_t) = V^{-1}\Gamma_t V^{-1}$, *$V$ is a $r \times r$ diagonal matrix of the eigenvalues of $(N \cdot T)^{-1} x x'$, and $\Gamma_t = \lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} E(\lambda_i \lambda'_j e_{it} e_{jt})$. Let $\hat{\alpha}^B_F = \hat{\alpha}_F - \hat{B}_1$ be the biased corrected estimate. Then*

$$\sqrt{T}(\hat{\alpha}^B_F - \alpha_F) \xrightarrow{d} N\Big(0, \text{Avar}(\hat{\alpha}_F)\Big).$$

The asymptotic variance for the bias corrected estimator is the same as $\hat{\alpha}_F$. Proposition 1 makes use of the fact that

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \hat{F}_t (HF_t - \hat{F}_t)' &= \frac{1}{T}\sum_{t=1}^{T} (\hat{F}_t - HF_t)(HF_t - \hat{F}_t)' + HF_t(HF_t - \hat{F}_t) \\
&= -E\left[\frac{1}{T}\sum_{t=1}^{T} (\hat{F}_t - HF_t)(HF_t - \hat{F}_t)'\right] + o_p(1) \\
&= -\frac{1}{NT}\sum_{t=1}^{T} \text{Avar}(\hat{F}_t) + o_p(1).
\end{aligned}$$

The estimation of $\text{Avar}_t(\hat{F}_t)$ was discussed in Bai and Ng (2006a). If $E(e^2_{it}) = \sigma^2$ for all i and t, $\text{Avar}_t(F_t)$ is the same for all t. Although Γ_t will depend on t if e_{it} is heteroskedastic, a consistent estimate of Γ_t can be obtained for each t when the errors are not cross-sectionally correlated, i.e., $E(e_{it} e_{jt}) = 0$. Alternatively, if $E(e_{it} e_{jt}) = \sigma_{ij} \neq 0$ for some or all t, panel data permit an estimate of $\text{Avar}(\hat{F}_t)$ that does not depend on t even when the e_{it} are cross-sectionally correlated. This estimator of Γ_t, referred to as CS-HAC in Bai and Ng (2006a), will be used later.

As this result on bias is new, we consider a small Monte Carlo experiment to gauge the magnitude of the bias as N and T changes. We consider a model with $r = 1$ and 2 factors. We assume $\lambda_i \sim N(0, 1)$ and $F_t \sim N(0, 1)$. These are only simulated once. Samples of $x_{it} = \lambda_i F_t + e_{it}$ and $y_t = \alpha' F_t + \epsilon_t$ are obtained by simulating $e_{it} \sim \sigma N(0, 1)$ and $\epsilon_t \sim N(0, 1)$ for $i = 1, \ldots N$, $t = 1, \ldots T$. We let $\alpha = 1$ when $r = 1$ and $\alpha = (1, 2)$ when $r = 2$. We consider three values of σ.

The smaller σ is, the more informative are the data for the factors. The results are as follows:

Estimated Bias for $\hat{\alpha}_1$

DGP: $y_t = F_t'\alpha + \epsilon_t, \quad x_{it} = \lambda_i F_t + e_{it}$

	σ = 1				σ = 4			
***r* = 1**	***T* = 50**	**100**	**200**	**500**	**50**	**100**	**200**	**500**
$N = 50$	−0.025	−0.020	−0.022	−0.019	−0.171	−0.156	−0.210	−0.242
100	−0.009	−0.009	−0.009	−0.012	−0.107	−0.107	−0.115	−0.138
200	−0.004	−0.004	−0.005	−0.004	−0.058	−0.058	−0.068	−0.071
500	−0.002	−0.002	−0.002	−0.002	−0.024	−0.030	−0.031	−0.034
***r* = 2**								
50	0.014	−0.035	0.026	0.017	0.002	−0.244	−0.077	−0.124
100	−0.020	0.003	−0.018	−0.020	0.116	−0.170	−0.056	−0.158
200	−0.010	0.001	0.007	−0.009	−0.104	−0.036	0.077	−0.092
500	−0.005	0.002	−0.004	0.001	−0.047	−0.043	0.028	0.031

As the true value of α is one, the entries can also be interpreted as percent bias. For large N and T, the bias is quite small and ignoring the sampling error in $\hat{F}_t$ should be inconsequential. Bias is smaller when $T/N = c$ than when $N/T = c$ for the same $c > 1$, confirming that the factors are more precisely estimated when there are more cross-section units to wash out the idiosyncratic noise. However, when σ is large and the data are uninformative about the factors, the bias can be well over 10% and as large as 20%. In such cases, the bias is also increasing in the number of estimated factors.

12.5.2 Bias When the Predictors Are Functions of $\hat{f}_t$

Our predictive regression has two additional complications. First, some of our predictors are powers of the estimated factors. Second, $\hat{F}8_t$ is a linear combination of a subset of $\hat{f}_t$ and $\hat{f}_{1t}^3$, which is a nonlinear function of $\hat{f}_{1t}$. To see how to handle the first problem, consider the case of the scalar predictor, $\hat{m}_t = m(\hat{f}_{1t})$ and let $m_t = m(Hf_{1t})$ where m takes its argument to the power b. The factor augmented regression becomes

$$y_t = \alpha_F' \hat{m}_t + \alpha_F'(m_t - \hat{m}_t) + \epsilon_t,$$

where $\alpha_F = \alpha H^{-b}$. The required bias correction is now of the form

$$B_2 = S_{\hat{m}\hat{m}'}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\hat{m}\Big(m_t - \hat{m}_t\Big)'\alpha_F\right).$$

But since m is continuous in $\hat{f}_{1t}$,

$$m(\hat{f}_{1t}) = m(f_t) + m_{\hat{f}_1,t}(\hat{f}_{1t} - Hf_{1t}),$$

P1: NARESH CHANDRA
November 3, 2010 16:42 C7035 C7035˙C012

where $m_{\hat{f},t} = \frac{\partial \hat{m}_t}{\partial \hat{f}_{1t}}|_{\hat{f}_{1t}=Hf_{1t}}$. We have

$$\hat{m}_t - m_t = b(Hf_{1t})^{b-1}(\hat{f}_{1t} - Hf_{1t}) = O_p(\min[N, T]^{-1}).$$

Given the foregoing result, it is then straightforward to show that

$$T^{-1}\sum_{t=1}^{T}\hat{m}_t(m_t - \hat{m}_t)' = \left[T^{-1}\sum_{t=1}^{T} m_{\hat{f}_1,t}\text{Avar}(\hat{f}_{1t})m'_{\hat{f}_1,t}\right] + o_p(1).$$

Extending the argument to the case when m_t is a vector leads to the bias correction

$$\hat{B}_2 = -S^{-1}_{\hat{m}\hat{m}'}\left(T^{-1}\sum_{t=1}^{T} m_{\hat{F},t}\text{Avar}(\hat{F}_{1t})m'_{\hat{F},t}\right)\alpha_F.$$

Finally, consider the predictive regression

$$y_t = \alpha'_F \hat{M}_t + \epsilon_t,$$

where $\hat{M}_t = \hat{\gamma}_0 + \hat{\gamma}'\hat{m}_t$. The bias can be estimated by

$$\hat{B}_3 = \hat{\gamma}'\hat{B}_2\hat{\gamma}.$$

In our application, $\hat{\gamma}$ is obtained from estimation of Equation 12.10.

While in theory, these bias corrections are required only when $\sqrt{T}/N$ does not tend to zero, in finite samples, the bias correction might be desirable even when $\sqrt{T}/N$ is small. We calculate the biased corrected estimates for two specifications of the predictive regressions. The first is when the predictors are selected by the in-sample BIC (column 1 of Tables 12.4 to 12.7). As this tends to lead to a larger model, the bias is likely more important. The second is when $\hat{F}8_t$ is used as predictor (column 9 of Tables 12.4 to 12.7), which is the most parsimonious of our specifications. Note that the observed predictor CP is not associated with first-step estimation error. As such, this predictor does not contribute to bias.

Reported in Table 12.8 are results using the CS-HAC, which allows the idiosyncratic errors to be cross-sectionally correlated. Results when the errors are heteroskedastic but cross-sectionally uncorrelated are similar. The results indicate that the bias is quite small. For the present application, the effect of the bias correction is to increase the absolute magnitude of the coefficient estimates in the predictive regressions. The t-statistics (not reported) are correspondingly larger. The finding that the macroeconomic factors have predictive power for excess bond returns is not sensitive to the assumption underlying the asymptotically validity of the FAR estimates.

12.5.3 Bootstrap Inference

According to asymptotic theory, heteroskedasticity and autocorrelation consistent standard errors that are asymptotically $N(0, 1)$ can be used to obtain

TABLE 12.8

Biased Corrected Estimates: $rx_{t+1}^{(n)} = a + \alpha' \hat{F}_t + \beta' CP_t + \epsilon_{t+1}$

$\hat{F}$	$n = 2$		$n = 3$		$n = 4$		$n = 5$	
$\hat{H}_1$	−0.761	-	−1.232	-	−1.521	-	−1.653	-
$\tilde{\alpha}$	−0.785	-	−1.277	-	−1.576	-	−1.724	-
bias	0.024	-	0.045	-	0.054	-	0.072	-
$\hat{H}_2$	-	-	−0.028	-	-	-	-	-
$\tilde{\alpha}$	-	-	−0.059	-	-	-	-	-
bias	-	-	0.032	-	-	-	-	-
$\hat{H}_4$	−0.291	-	−0.423	-	−0.436	-	−0.516	-
$\tilde{\alpha}$	−0.307	-	−0.454	-	−0.472	-	−0.564	-
bias	0.016	-	0.031	-	0.036	-	0.048	-
$\hat{H}_6$	−0.151	-	−0.433	-	−0.668	-	−0.856	-
$\tilde{\alpha}$	−0.168	-	−0.468	-	−0.710	-	−0.912	-
bias	0.018	-	0.035	-	0.042	-	0.055	-
$\hat{H}_7$	−0.128	-	−0.338	-	−0.534	-	−0.686	-
$\tilde{\alpha}$	−0.145	-	−0.372	-	−0.573	-	−0.737	-
bias	0.017	-	0.034	-	0.039	-	0.051	-
$\hat{H}_8$	0.240	-	0.389	-	0.578	-	0.702	-
$\tilde{\alpha}$	0.225	-	0.355	-	0.542	-	0.654	-
bias	0.016	-	0.033	-	0.036	-	0.048	-
$\hat{H}_3^2$	-	-	0.111	-	0.177	-	0.204	-
$\tilde{\alpha}$	-	-	0.114	-	0.181	-	0.209	-
bias	-	-	−0.004	-	−0.004	-	−0.006	-
$\hat{H}_5^2$	−0.080	-	-	-	-	-	-	-
$\tilde{\alpha}$	−0.078	-	-	-	-	-	-	-
bias	−0.003	-	-	-	-	-	-	-
$\hat{H}_1^3$	0.044	-	0.095	-	0.131	-	0.150	-
$\tilde{\alpha}$	0.045	-	0.096	-	0.133	-	0.153	-
bias	−0.001	-	−0.002	-	−0.002	-	−0.003	-
CP	0.385	0.336	0.760	0.644	1.115	0.955	1.316	1.115
$\tilde{\alpha}$	0.381	0.343	0.760	0.660	1.108	0.980	1.306	1.147
bias	0.004	−0.007	–	−0.016	0.007	−0.026	0.010	−0.032
$\hat{H}8$	-	0.332	-	0.588	-	0.777	-	0.938
$\tilde{\alpha}$	-	0.342	-	0.607	-	0.802	-	0.972
bias	-	−0.010	-	−0.019	-	−0.025	-	−0.035

Note: The bias unadjusted estimates are reported in columns 1 and 9 of Tables 12.4 to 12.7, respectively.

TABLE 12.9

Bootstrap Estimates When $\hat{H}_t = \hat{F}_t$: Regression $rx_{t+1}^{(n)} = \alpha'\hat{F}_t + \beta' CP_t + \epsilon_{t+1}$

			Bootstrap		Bootstrap under the Null	
	$\hat{\alpha}$	**bias**	**95% CI**	**99% CI**	**95% CI**	**99% CI**
				***n* = 2**		
$\hat{H}_1$	−0.761	0.012	(−1.143 −0.343)	(−1.071 −0.399)	(−0.021 −0.015)	(−0.021 −0.016)
$\hat{H}_4$	−0.291	−0.006	(−0.554 −0.031)	(−0.508 −0.073)	(−0.003 0.003)	(−0.002 0.003)
$\hat{H}_6$	−0.151	−0.002	(−0.467 0.166)	(−0.408 0.100)	(−0.015 0.016)	(−0.015 0.016)
$\hat{H}_7$	−0.128	−0.004	(−0.285 0.027)	(−0.258 −0.010)	(−0.008 0.011)	(−0.007 0.009)
$\hat{H}_8$	0.240	0.004	(0.054 0.425)	(0.088 0.404)	(−0.011 0.010)	(−0.010 0.008)
$\hat{H}_5^2$	−0.080	0.003	(−0.187 0.040)	(−0.170 0.015)	(−0.010 −0.003)	(−0.009 −0.003)
$\hat{H}_1^3$	0.044	−0.001	(0.010 0.076)	(0.016 0.071)	(−0.000 0.000)	(−0.000 0.000)
CP	0.385	−0.003	(0.262 0.516)	(0.276 0.490)	(0.003 0.009)	(0.003 0.008)
$\bar{R}^2$	0.460		(0.237 0.523)	(0.261 0.500)	(0.019 0.045)	(0.021 0.042)
				***n* = 3**		
$\hat{H}_1$	−1.232	0.027	(−1.914 −0.506)	(−1.797 −0.655)	(−0.021 −0.015)	(−0.021 −0.016)
$\hat{H}_2$	−0.028	−0.017	(−0.574 0.505)	(−0.486 0.426)	(−0.001 0.005)	(−0.000 0.005)
$\hat{H}_4$	−0.423	−0.004	(−0.881 0.030)	(−0.811 −0.050)	(−0.003 0.003)	(−0.003 0.003)
$\hat{H}_6$	−0.433	0.012	(−0.969 0.093)	(−0.870 0.024)	(−0.014 0.015)	(−0.013 0.014)
$\hat{H}_7$	−0.338	−0.002	(−0.585 −0.094)	(−0.549 −0.140)	(−0.009 0.010)	(−0.007 0.009)
$\hat{H}_8$	0.389	−0.002	(0.082 0.669)	(0.140 0.632)	(−0.009 0.008)	(−0.008 0.007)
$\hat{H}_3^2$	0.111	−0.003	(−0.046 0.250)	(−0.006 0.221)	(0.000 0.002)	(0.000 0.002)
$\hat{H}_1^3$	0.095	−0.002	(0.034 0.145)	(0.046 0.136)	(0.000 0.001)	(0.000 0.001)
CP	0.760	−0.001	(0.546 0.980)	(0.582 0.935)	(0.003 0.009)	(0.003 0.008)
$\bar{R}^2$	0.455		(0.280 0.559)	(0.303 0.533)	(0.013 0.035)	(0.014 0.032)

(continued)

TABLE 12.9 (Continued)

Bootstrap Estimates When $\hat{H}_t = \hat{F}_t$: Regression $rx^{(n)}_{t+1} = \alpha'\hat{F}_t + \beta' CP_t + \epsilon_{t+1}$

	$\hat{\alpha}$	bias	Bootstrap 95% CI	Bootstrap 99% CI	Bootstrap under the Null 95% CI	Bootstrap under the Null 99% CI
			***n* = 4**			
$\hat{H}_1$	−1.521	0.047	(−2.488 −0.480)	(−2.323 −0.617)	(−0.021 −0.015)	(−0.021 −0.016)
$\hat{H}_4$	−0.436	0.001	(−1.048 0.178)	(−0.958 0.090)	(−0.004 0.003)	(−0.003 0.003)
$\hat{H}_6$	−0.668	−0.002	(−1.410 0.131)	(−1.297 0.002)	(−0.014 0.015)	(−0.013 0.014)
$\hat{H}_7$	−0.534	0.004	(−0.942 −0.178)	(−0.849 −0.230)	(−0.009 0.010)	(−0.007 0.008)
$\hat{H}_8$	0.578	0.004	(0.119 1.022)	(0.206 0.957)	(−0.010 0.009)	(−0.009 0.007)
$\hat{H}_3^2$	0.177	−0.001	(−0.031 0.375)	(0.002 0.339)	(0.000 0.002)	(0.000 0.002)
$\hat{H}_1^3$	0.131	−0.003	(0.055 0.206)	(0.068 0.189)	(0.000 0.001)	(0.000 0.001)
CP	1.115	−0.006	(0.820 1.401)	(0.861 1.348)	(0.003 0.009)	(0.003 0.009)
$\bar{R}^2$	0.473		(0.277 0.567)	(0.303 0.545)	(0.014 0.036)	(0.016 0.034)
			***n* = 5**			
$\hat{H}_1$	−1.653	0.026	(−2.832 −0.429)	(−2.648 −0.606)	(−0.021 −0.015)	(−0.021 −0.016)
$\hat{H}_4$	−0.516	−0.004	(−1.306 0.321)	(−1.190 0.169)	(−0.003 0.003)	(−0.003 0.003)
$\hat{H}_6$	−0.856	0.011	(−1.870 0.190)	(−1.666 0.012)	(−0.014 0.014)	(−0.013 0.014)
$\hat{H}_7$	−0.686	0.012	(−1.182 −0.119)	(−1.071 −0.244)	(−0.007 0.010)	(−0.006 0.009)
$\hat{H}_8$	0.702	−0.004	(0.139 1.286)	(0.224 1.160)	(−0.009 0.008)	(−0.009 0.007)
$\hat{H}_3^2$	0.204	0.000	(−0.059 0.491)	(−0.017 0.419)	(0.000 0.002)	(0.001 0.002)
$\hat{H}_1^3$	0.150	−0.001	(0.051 0.242)	(0.069 0.232)	(0.000 0.001)	(0.000 0.001)
CP	1.316	−0.009	(0.896 1.723)	(0.945 1.663)	(0.003 0.009)	(0.003 0.008)
$\bar{R}^2$	0.435		(0.225 0.518)	(0.251 0.488)	(0.015 0.036)	(0.016 0.033)

robust t-statistics for the in-sample regressions. Moreover, provided $\sqrt{T}/N$ goes to zero as the sample increases, the $\widehat{F}_t$ can be treated as observed regressors, and the usual t-statistics are valid (Bai and Ng 2006a). To guard against inadequacy of the asymptotic approximation in finite samples, we consider bootstrap inference in this section.

To proceed with a bootstrap analysis, we need to generate bootstrap samples of $rx_{t+1}^{(n)}$, and thus the exogenous predictors Z_t (here just CP_t), as well as of the estimated factors $\widehat{F}_t$. Bootstrap samples of $rx_{t+1}^{(n)}$ are obtained in two ways: first by imposing the null hypothesis of no predictability, and second, under the alternative that excess returns are forecastable by the factors and conditioning variables studied above. The use of monthly bond price data to construct continuously compounded annual returns induces an MA(12) error structure in the annual log returns. Thus, under the null hypothesis that the expectations hypothesis is true, annual compound returns are forecastable up to an MA(12) error structure, but are not forecastable by other predictor variables or additional moving average terms.

Bootstrap sampling that captures the serial dependence of the data is straightforward when, as in this case, there is a parametric model for the dependence under the null hypothesis. In this event, the bootstrap may be accomplished by drawing random samples from the empirical distribution of the residuals of a $\sqrt{T}$ consistent, asymptotically normal estimator of the parametric model, in our application a twelfth-order moving average process. We use this approach to form bootstrap samples of excess returns under the null. Under the alternative, excess returns still have the MA(12) error structure induced by the use of overlapping data, but estimated factors $\widehat{F}_t$ are presumed to contain additional predictive power for excess returns above and beyond that implied by the moving average error structure.

To create bootstrapped samples of the factors, we re-sample the $T \times N$ panel of data, x_{it}. For each i, we assume that the idiosyncratic errors e_{it} and the errors u_t in the factor process are AR(1) processes. Least squares estimation of $\widehat{e}_{it} = \rho_i \widehat{e}_{it-1} + v_{it}$ yields the estimates $\widehat{\rho}_i$ and $\widehat{v}_{it}$, $t = 2, \ldots, T$, recalling that $\widehat{e}_{it} = x_{it} - \widehat{\lambda}_i' \hat{f}_t$. These errors are then re-centered. To generate a new panel of data, for each i, $\widehat{v}_{it}$ is re-sampled (while preserving the cross-section correlation structure) to yield bootstrap samples of $\widehat{e}_{it}$. In turn, bootstrap values of x_{it} are constructed by adding the bootstrap estimates of the idiosyncratic errors, $\widehat{e}_{it}$, to $\widehat{\lambda}_i' \widehat{F}_t$. Applying the method of principal components to the bootstrapped data yields a new set of estimated factors. Together with bootstrap samples of CP_t created under the assumption that it is an AR(1), we have a complete set of bootstrap regressors in the predictive regression.

Each regression using the bootstrapped data gives new estimates of the regression coefficients. This is repeated B times. Bootstrap confidence intervals for the parameter estimates and $\bar{R}^2$ statistics are calculated from $B = 10,000$ replications. We compute 90th and 95th percentiles of $\hat{\beta}_F$ and $\hat{\alpha}_F$, as well as the bootstrap estimate of the bias. This also allows us to compare the adequacy of our calculations for asymptotic bias considered in the previous subsection. The exercise is repeated for 2-, 3-, 4-, and 5-year excess bond returns.

To conserve space, the results in Table 12.9 are reported only for the largest model (corresponding to column 1 of Tables 12.4 to 12.7). The results based on bootstrap inference are consistent with asymptotic inference. In particular, the magnitude of predictability found in the historical data is too large to be accounted for by sampling error of the size we currently have. The coefficients on the predictors and factors are statistically different from zero at the 95% level and are well outside the 95% confidence interval under the null of no predictability. The bootstrap estimate of the bias on coefficients associated with the estimated factors are small, and the $\bar{R}^2$ are similar in magnitude to what was reported in Tables 12.4 to 12.7.

12.5.4 Posterior Inference

In Tables 12.4 to 12.7, we have used the posterior mean of G_t in the predictive regression computed from 1000 draws (taken from a total of 25,000 draws) from the posterior distribution of G_t. The $\hat{\alpha}$ do not reflect sampling uncertainty about G_t. To have a complete account of sampling variability, we estimate the predictive regressions for each of the 1000 draws of G_t. This gives us the posterior distribution for α as well as the corresponding t-statistic.

Reported in Table 12.10 are the posterior mean of α_G along with the 5% and 95% percentage points of the t-statistic. The point estimates reported in Tables 12.4 to 12.7 are very close to the posterior means. Sampling variability from having to estimate the dynamic factors has little effect on the estimates of the factor augmented regressions.

So far we find that macroeconomic factors have nontrivial predictive power for bond excess returns and that the sampling error induced by $\hat{F}_t$ or $\hat{G}_t$ in the predictive regressions are numerically small. Multiple factors contribute to the predictability of excess returns, so it is not possible to infer the cyclicality of return risk premia by observing the signs of the individual coefficients on factors in forecasting regressions of excess returns. But Tables 12.4 to 12.7 provide a summary measure of how the factors are related to future excess returns by showing that excess bond returns are high when the linear combinations of all factors, $\hat{F}8_t$ and $\hat{G}8_t$, are high. Figures 12.11 and 12.12 show that $\hat{F}8_t$ and $\hat{G}8_t$ are in turn high when real activity (as measured by industrial production growth) is low. The results therefore imply that excess returns are forecast to be high when economic activity is slow or contracting. That is, return risk premia are countercyclical. This is confirmed by the top panels of Figures 12.13 and 12.14, which plot return risk premia along with industrial production growth. The bottom panels of these figures show that the factors contribute significantly to the countercyclicality of risk-premia. Indeed, when factors are excluded (but CP_t is included), risk-premia are a-cyclical. Of economic interest is whether yield risk-premia are also countercyclical. We now turn to such an analysis.

TABLE 12.10

Posterior Mean: $rx_{t+1}^{(n)} = a + \alpha' \hat{G}_t + \beta' CP_t + \epsilon_{t+1}$

$\hat{F}$	**n = 2**		**n = 3**		**n = 4**		**n = 5**	
$\hat{H}_1$	-	-	-	-	-	-	0.288	-
$t_{.05}$	-	-	-	-	-	-	1.275	-
$t_{.95}$	-	-	-	-	-	-	1.912	-
$\hat{H}_2$	−0.506	-	−0.801	-	−0.976	-	−1.159	-
$t_{.05}$	−3.676	-	−3.239	-	−3.140	-	−3.099	-
$t_{.95}$	−2.942	-	−2.622	-	−2.477	-	−2.397	-
$\hat{H}_3$	−0.456	-	−0.746	-	−0.959	-	−1.074	-
$t_{.05}$	−5.335	-	−4.749	-	−4.616	-	−3.302	-
$t_{.95}$	−4.050	-	−3.637	-	−3.482	-	−3.374	-
$\hat{H}_6$	0.139	-	-	-	-	-	-	-
$t_{.05}$	1.819	-	-	-	-	-	-	-
$t_{.95}$	1.712	-	-	-	-	-	-	-
$\hat{H}_8$	−0.139	-	−0.309	-	−0.473	-	−0.561	-
$t_{.05}$	−1.872	-	−2.366	-	−2.622	-	−2.523	-
$t_{.95}$	−1.332	-	−1.732	-	−1.994	-	−1.863	-
$\hat{H}_4^2$	−0.070	-	−0.183	-	−0.253	-	−0.348	-
$t_{.05}$	−2.395	-	−2.982	-	−2.920	-	−3.713	-
$t_{.95}$	−2.787	-	−3.319	-	−3.089	-	−3.681	-
$\hat{H}_6^2$	−0.086	-	−0.154	-	−0.235	-	−0.274	-
$t_{.05}$	−5.427	-	−6.109	-	−6.109	-	−5.559	-
$t_{.95}$	−6.629	-	−7.223	-	−6.838	-	−6.138	-
$\hat{H}_7^2$	-	-	0.087	-	0.146	-	0.178	-
$t_{.05}$	-	-	2.408	-	2.866	-	2.852	-
$t_{.95}$	-	-	2.404	-	3.006	-	2.914	-
$\hat{H}_1^3$	0.019	-	0.032	-	0.037	-	-	-
$t_{.05}$	2.092	-	2.090	-	1.836	-	-	-
$t_{.95}$	2.346	-	2.357	-	2.095	-	-	-
CP	0.452	0.416	0.845	0.790	1.236	1.155	1.456	1.365
$t_{.05}$	7.200	6.334	7.285	6.300	7.568	6.348	7.012	5.900
$t_{.95}$	7.566	6.919	7.641	6.770	7.926	6.760	7.331	6.262
$\hat{H}8$	-	0.428	-	0.712	-	0.867	-	0.959
$t_{.05}$	-	3.330	-	3.096	-	2.888	-	2.610
$t_{.95}$	-	4.316	-	4.033	-	3.803	-	3.489
$\bar{R}^2_{0.95}$	0.471	0.399	0.469	0.403	0.489	0.415	0.448	0.377
$\bar{R}^2_{0.05}$	0.469	0.397	0.467	0.401	0.488	0.413	0.446	0.375

Note: Reported are the mean estimates when a predictive regression is run for each draw of G_t. Estimates when the regressors are the posterior mean of the G_t are reported in columns 5 and 10 of Tables 12.4 to 12.7, respectively.

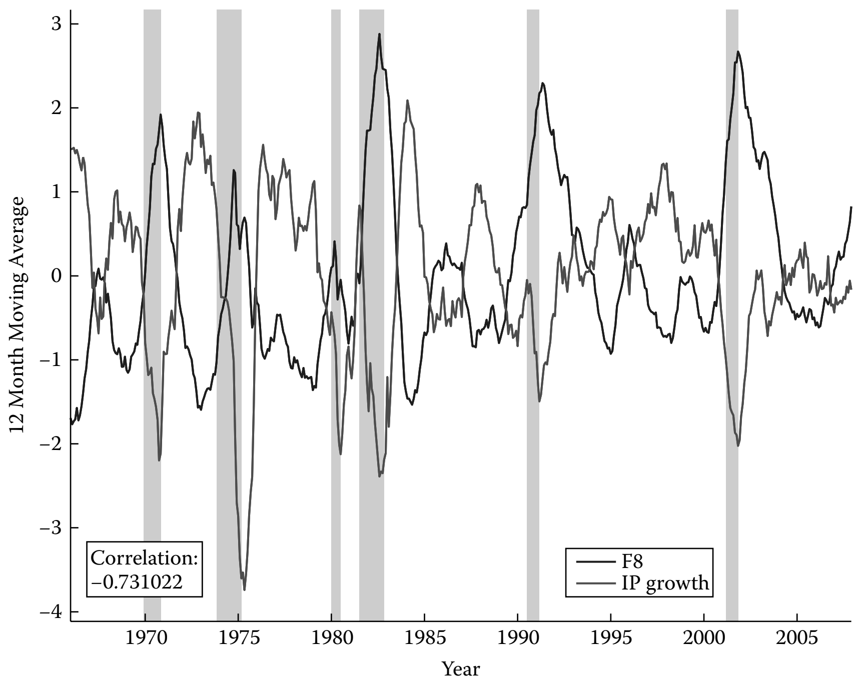


FIGURE 12.11
F8 and IP Growth

12.6 Countercyclical Yield Risk Premia

The *yield risk premium* or *term premium* should not be confused with the term spread, which is simply the difference in yields between the n-period bond and the one-period bond. Instead, the yield risk premium is a component of the the n-period yield:

$$y_t^{(n)} = \underbrace{\frac{1}{n} E_t\left(y_t^{(1)} + y_{t+1}^{(1)} + \cdots + y_{t+n-1}^{(1)}\right)}_{\text{expectations component}} + \underbrace{\varkappa_t^{(n)}}_{\text{yield risk premium}}. \tag{12.12}$$

Under the expectations hypothesis, the yield risk premium, $\varkappa_t^{(n)}$, is assumed constant.

It is straightforward to show that the yield risk premium is identically equal to the average of expected future return risk premia of declining maturity:

$$\varkappa_t^{(n)} = \frac{1}{n}\left[E_t\left(rx_{t+1}^{(n)}\right) + E_t\left(rx_{t+2}^{(n-1)}\right) + \cdots + E_t\left(rx_{t+n-1}^{(2)}\right)\right]. \tag{12.13}$$

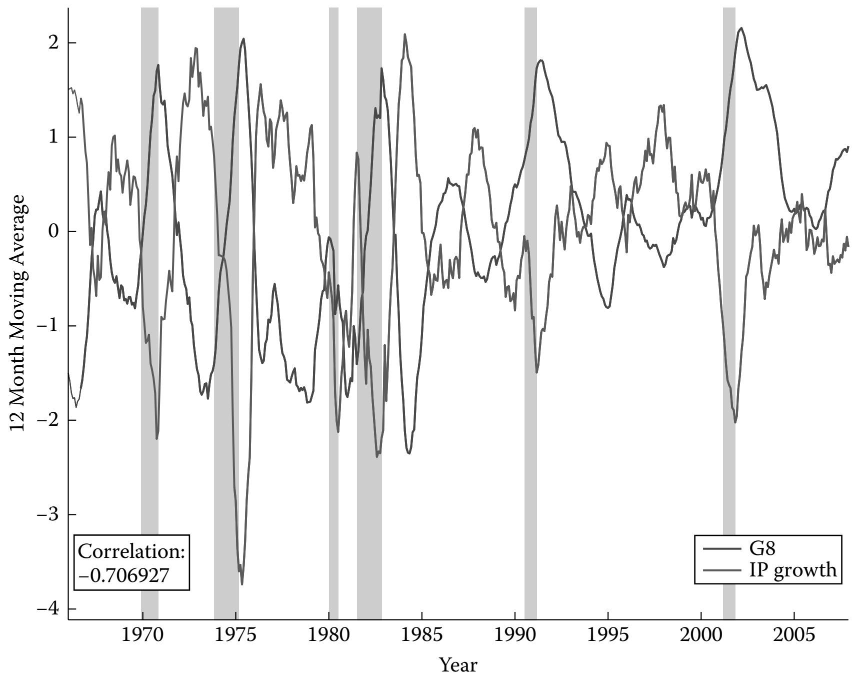


FIGURE 12.12
G8 and IP Growth.

To form an estimate of the risk premium component in yields, $\varkappa_t^{(n)}$, we need estimates of the multistep ahead forecasts that appear on the right-hand side of Equation 12.13. Denote estimated variables with "hats." Then

$$\widehat{\varkappa}_t^{(n)} = \frac{1}{n}\left[\widehat{E}_t\left(rx_{t+1}^{(n)}\right) + \widehat{E}_t\left(rx_{t+2}^{(n-1)}\right) + \cdots + \widehat{E}_t\left(rx_{t+n-1}^{(2)}\right)\right], \tag{12.14}$$

where $\widehat{E}_t(\cdot)$ denotes an estimate of the conditional expectation $E_t(\cdot)$ formed by a linear projection. As estimates of the conditional expectations are simply linear forecasts of excess returns, multiple steps ahead our earlier results for the FAR have direct implications for risk premia in yields.

To generate multistep ahead forecasts we estimate a monthly pth-order vector autoregression (VAR). The idea behind the VAR is that multistep ahead forecasts may be obtained by iterating one-step ahead linear projections from the VAR. The VAR vector contains observations on excess returns, the Cochrane–Piazzesi factor, CP_t and $\hat{H}_t$, where $\hat{H}_t$ are the estimated factors (or a linear combination of them). Let

$$Z_t^U \equiv \left[rx_t^{(5)}, rx_t^{(4)}, ..., rx_t^{(2)}, CP_t, \hat{H}8_t\right]'$$

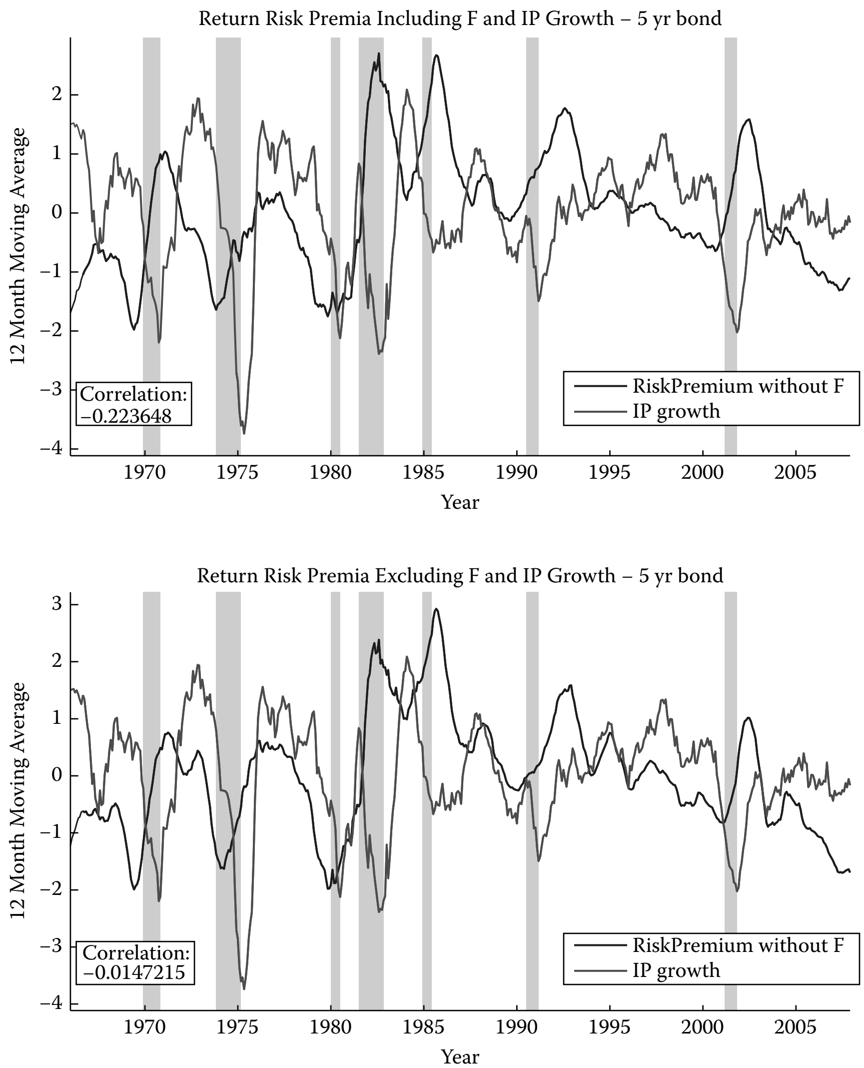


FIGURE 12.13
Return Risk Premia.

where $\hat{H}8$ is either $\hat{F}8$ or $\hat{G}8$. For comparison, we will also form bond forecasts with a restricted VAR that excludes the estimated factors, but still includes CP_t as a predictor variable:

$$Z_t^R \equiv \left[rx_t^{(5)}, rx_t^{(4)}, \ldots, rx_t^{(2)}, CP_t\right]'.$$

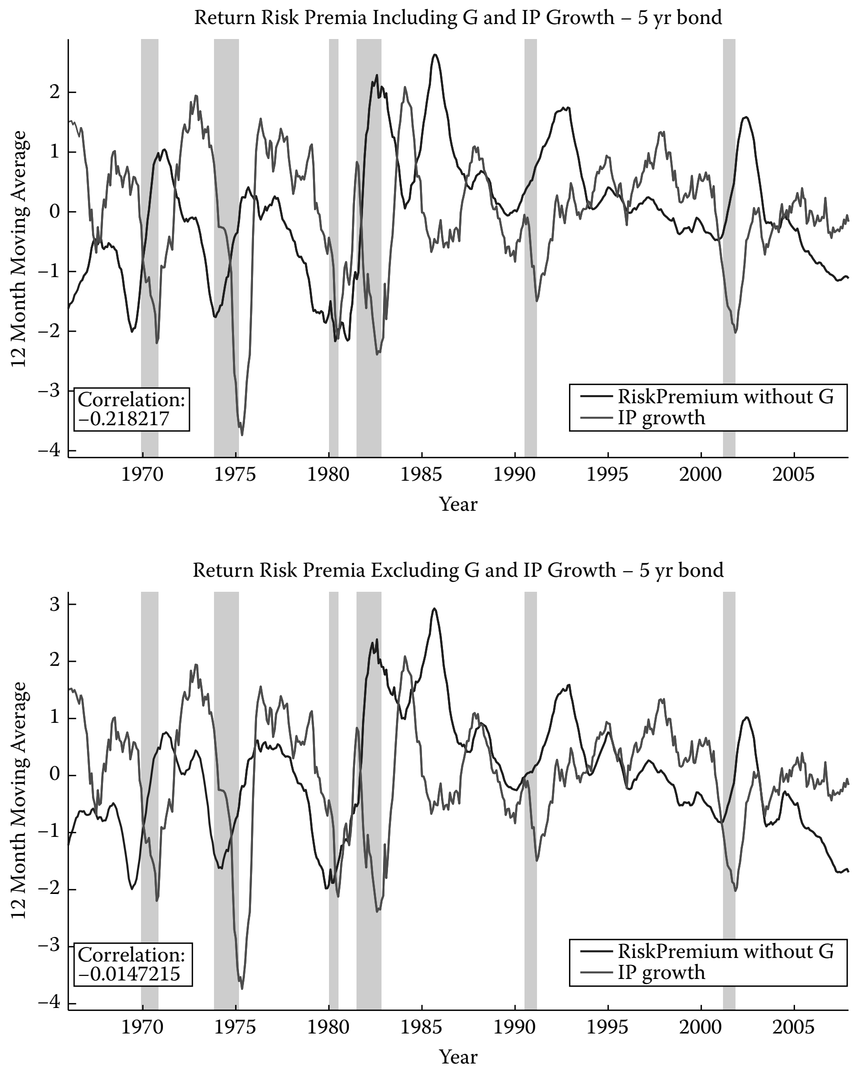


FIGURE 12.14
Return Risk Premia.

We use a monthly VAR with $p = 12$ lags, where, for notational convenience, we write the VAR in terms of mean deviations[7]:

$$Z_{t+1/12} - \boldsymbol{\mu} = \boldsymbol{\Phi}_1 (Z_t - \boldsymbol{\mu}) + \boldsymbol{\Phi}_2 (Z_{t-1/12} - \boldsymbol{\mu}) + \cdots + \boldsymbol{\Phi}_p (Z_{t-11/12} - \boldsymbol{\mu}) + \varepsilon_{t+1/12}. \quad (12.15)$$

[7] This is only for notational convenience. The estimation will include the means.

Let k denote the number of variables in Z_t. Then Equation 12.15 can be expressed as a $VAR(1)$:

$$\boldsymbol{\xi}_{t+1/12} = \mathbf{A}\boldsymbol{\xi}_t + \mathbf{v}_{t+1/12}, \tag{12.16}$$

where,

$$\underset{(kp\times 1)}{\boldsymbol{\xi}_{t+1/12}} \equiv \begin{bmatrix} Z_t - \boldsymbol{\mu} \\ Z_{t-1/12} - \boldsymbol{\mu} \\ \cdot \\ \cdot \\ \cdot \\ Z_{t-11/12} - \boldsymbol{\mu} \end{bmatrix} \qquad \underset{(kp\times 1)}{\mathbf{v}_t} \equiv \begin{bmatrix} \varepsilon_{t+1/12} \\ \mathbf{0} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{0} \end{bmatrix}$$

$$\underset{(kp\times kp)}{\mathbf{A}} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \boldsymbol{\Phi}_3 & \cdot & \cdot & \boldsymbol{\Phi}_{p-1} & \boldsymbol{\Phi}_p \\ \mathbf{I}_n & \mathbf{0} & \mathbf{0} & \cdot & \cdot & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} & \cdot & \cdot & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdot & \cdot & \mathbf{I}_n & \mathbf{0} \end{bmatrix}.$$

Multistep ahead forecasts are straightforward to compute using the first-order VAR:

$$E_t\boldsymbol{\xi}_{t+j/12} = \mathbf{A}^j\boldsymbol{\xi}_t.$$

When $j = 12$, the monthly VAR produces forecasts of 1-year ahead variables, $E_t\boldsymbol{\xi}_{t+1} = \mathbf{A}^{12}\boldsymbol{\xi}_t$; when $j = 24$, it computes 2-year ahead forecasts, and so on. Define a vector ej that picks out the jth element of $\boldsymbol{\xi}_t$, i.e., $e1'\boldsymbol{\xi}_t \equiv rx_t^{(5)}$. In the notation above, we have $e1_{(kp\times 1)} = [1, 0, 0, \ldots, 0]'$, $e2_{(kp\times 1)} = [0, 1, 0, \ldots, 0]'$, analogously for $e3$ and $e4$. Thus, given estimates of the VAR parameters $\mathbf{A}$, we may form estimates of the conditional expectations on the right-hand side of Equation 12.14 using the VAR forecasts of return risk premia. For example, the estimate of the expectation of the 5-year bond, 1 year ahead, is given by $\widehat{E}_t(rx_{t+1}^{(5)}) = e1'\mathbf{A}^{12}\boldsymbol{\xi}_t$; the estimate of the expectation of the 4-year bond, 2 years ahead, is given by $\widehat{E}_t(rx_{t+2}^{(4)}) = e2'\mathbf{A}^{24}\boldsymbol{\xi}_t$, and so on.

Letting $\hat{H}_t = \hat{F}5_t$ where $\hat{F}5_t$ is a linear combination of $\hat{f}_{1t}$, $\hat{f}_{1t}^3$, $\hat{f}_{3t}$, $\hat{f}_{4t}$, and $\hat{f}_{8t}$. we showed in Ludvigson and Ng (2007) that both yield and return risk premia are more countercyclical and reach greater values in recessions than in the absence of $\hat{H}_t$. Here, we verify that this result holds up for different choices of $\hat{H}_t$. To this end, we let $\hat{H}_t$ be the static and dynamic factors selected by the out-of-sample BIC. These two predictor sets embody information in fewer factors than the ones implied by the in-sample BIC, $\hat{H}8$, or $F5_t$ used in Ludvigson and Ng (2007). The point is to show that a few macroeconomic factors are enough to generate an important difference in the properties of risk premia. Specifically, without $\hat{F}_t$ in Z_t^U, the correlation between the estimated return risk premium and IP growth is −0.014. With $\hat{F}_t$ in Z_t^U, the correlation

P1: NARESH CHANDRA
November 3, 2010 16:42 C7035 C7035˙C012

is −0.223. These correlations are −0.045 and −0.376 for yield risk premia. With $\hat{G}_t$ in Z_t^U, the correlation of IP growth with return and yield risk premium are −0.218 and −0.286, respectively. Return and yield risk premia are thus more countercyclical when the factors are used to forecast excess returns.

Figure 12.15 shows the 12-month moving average of risk-premium component of the 5-year bond yield. As we can see, yield risk premia were particularly high in the 1982–1983 recession, as well as shortly after the 2001 recession. Figure 12.16 shows the yield risk premia estimated with and without using $\hat{F}_t$ to forecast excess returns, while Figure 12.17 shows a similar picture with and without $\hat{G}_t$. The difference between the risk premia estimated with and without the factors is largest around recessions. For example, the yield risk premium on the 5-year bond estimated using the information contained in $\hat{F}_t$ or $\hat{G}_t$ was over 2% in the 2001 recession, but it was slightly below 1% without $\hat{G}_t$. The return risk premia (not reported) show a similar pattern.

When the economy is contracting, the countercyclical nature of the risk factors contributes to a steepening of the yield curve even as future short-term rates fall. Conversely, when the economy is expanding, the factors contribute to a flattening of the yield curve even as expectations of future short-term rates rise. This implies that information in the factors is ignored. Too much variation in the long-term yields is attributed to the expectations component in recessions. Information in the macro factors are thus important in accurate decomposition of risk premia, especially in recessions.

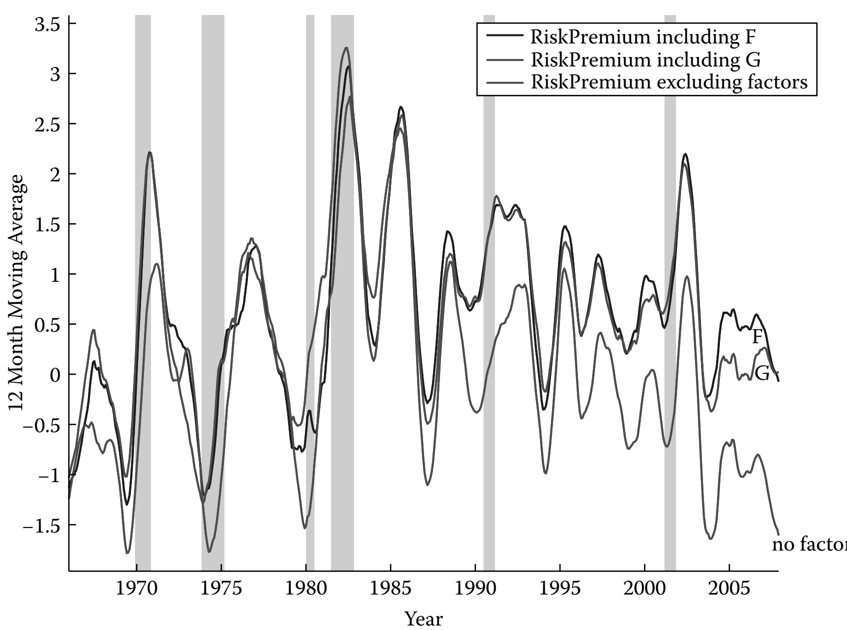


FIGURE 12.15
Yield Risk Premium with and without factors −5 yr bond.

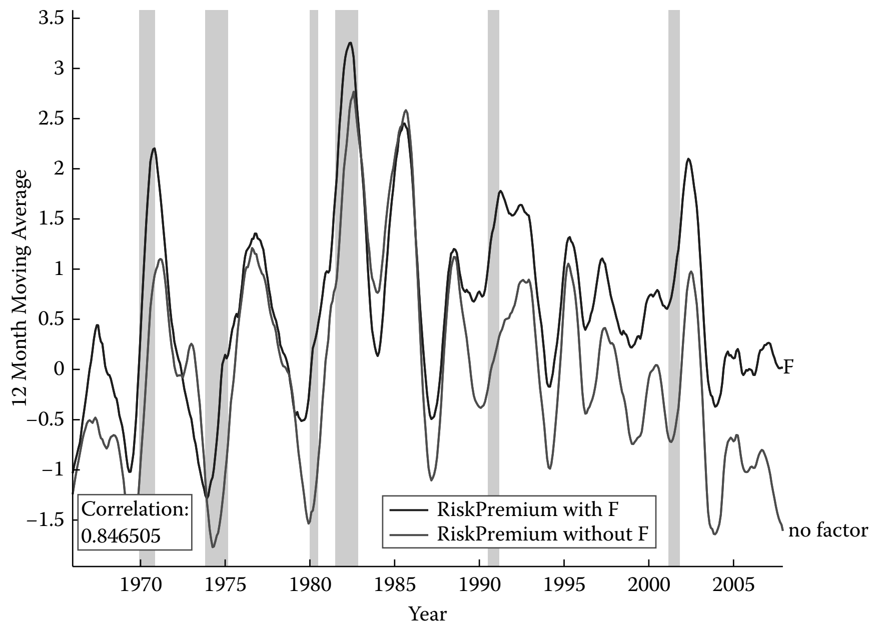


FIGURE 12.16
Yield Risk Premia Including and Excluding F −5 yr bond.

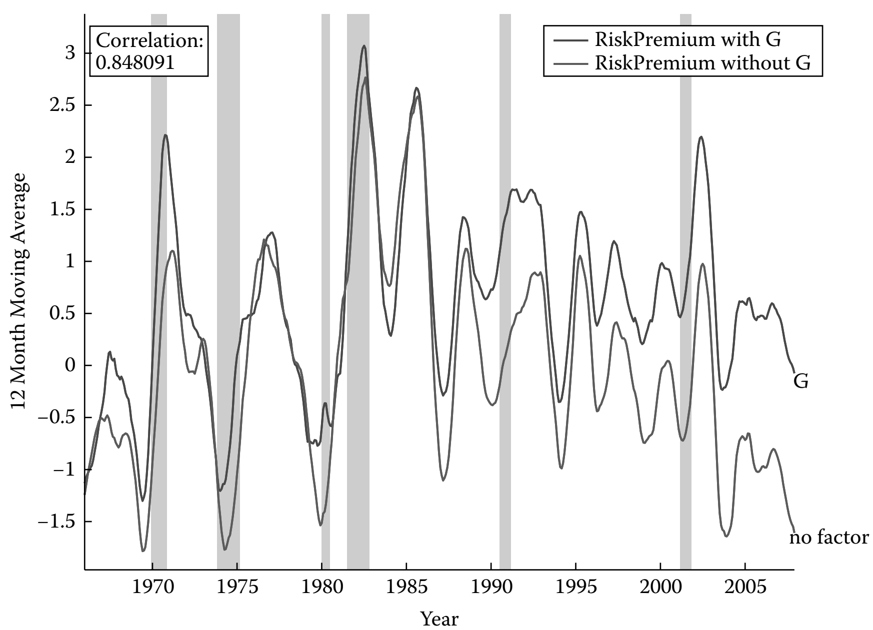


FIGURE 12.17
Yield Risk Premia Including and Excluding G −5 yr bond.

12.7 Conclusion

There is a good deal of evidence that excess bond returns are predictable by financial variables. Yet, macroeconomic theory postulates that it is real variables relating to macroeconomic activity that should forecast bond returns. This chapter presents robust evidence in support of the theory. Macroeconomic factors, especially the real activity factor, has strong predictive power for excess bond returns even in the presence of financial predictors. Our analysis consists of estimating two sets of factors and a comprehensive specification search. We also account for sampling uncertainty that might arise from estimation of the factors. While the estimated risk premia without using the macro factors to forecast excess returns are acyclical, both bond returns and yield risk premia are countercyclical when the factors are used. The evidence indicate that investors seek compensation for macroeconomic risks associated with recessions.

12.8 Acknowledgment

We thank Jushan Bai for helpful suggestions and Matt Smith for excellent research assistance. We also thank the Conference Board for providing us with some of the data. Financial support from the National Science Foundation (Grant No. 0617858 to Ludvigson and SES-0549978 to Ng) is gratefully acknowledged. Ludvigson also acknowledges financial support from the Alfred P. Sloan Foundation and the CV Starr Center at NYU. Any errors or omissions are the responsibility of the authors.

Data Appendix

This appendix lists the short name of each series, its mnemonic (the series label used in the source database), the transformation applied to the series, and a brief data description. All series are from the Global Insights Basic Economics Database, unless the source is listed (in parentheses) as TCB (The Conference Board's Indicators Database) or AC (author's calculation based on Global Insights or TCB data). In the transformation column, ln denotes logarithm, Δ ln and Δ^2 ln denote the first and second difference of the logarithm, lv denotes the level of the series, and Δ lv denotes the first difference of the series. The data are available from 1959:01 to 1997:12.

P1: NARESH CHANDRA
November 3, 2010 16:42 C7035 C7035˙C012

Group 1: Output and Income

No.	Gp	Short Name	Mnemonic	Tran	Descripton
1	1	PI	ypr	Δln	Personal Income (AR, Bil. Chain 2000 $) (TCB)
6	1	IP: total	ips10	Δln	Industrial Production Index–Total Index
7	1	IP: products	ips11	Δln	Industrial Production Index–Products, Total
8	1	IP: final prod	ips299	Δln	Industrial Production Index–Final Products
9	1	IP: cons gds	ips12	Δln	Industrial Production Index–Consumer Goods
10	1	IP: cons dble	ips13	Δln	Industrial Production Index–Durable Consumer Goods
11	1	IP: cons nondble	ips18	Δln	Industrial Production Index–Nondurable Consumer Goods
12	1	IP: bus eqpt	ips25	Δln	Industrial Production Index–Business Equipment
13	1	IP: matls	ips32	Δln	Industrial Production Index–Materials
14	1	IP: dble matls	ips34	Δln	Industrial Production Index–Durable Goods Materials
15	1	IP: nondble matls	ips38	Δln	Industrial Production Index–Nondurable Goods Materials
16	1	IP: mfg	ips43	Δln	Industrial Production Index–Manufacturing (Sic)
17	1	IP: res util	ips307	Δln	Industrial Production Index–Residential Utilities
18	1	IP: fuels	ips306	Δln	Industrial Production Index–Fuels
19	1	NAPM prodn	pmp	lv	Napm Production Index (Percent)
20	1	Cap util	utl11	Δlv	Capacity Utilization (SIC-Mfg) (TCB)

Group 2: Labor Market

No.	Gp	Short Name	Mnemonic	Tran	Descripton
21	2	Help wanted indx	lhel	Δlv	Index Of Help-Wanted Advertising In Newspapers (1967=100;Sa)
22	2	Help wanted/emp	lhelx	Δlv	Employment: Ratio; Help-Wanted Ads:No. Unemployed Clf
23	2	Emp CPS total	lhem	Δln	Civilian Labor Force: Employed, Total (Thous.,Sa)
24	2	Emp CPS nonag	lhnag	Δln	Civilian Labor Force: Employed, Nonagric.Industries (Thous.,Sa)
25	2	U: all	lhur	Δlv	Unemployment Rate: All Workers, 16 Years &
26	2	U: mean duration	lhu680	Δlv	Unemploy.By Duration: Average(Mean)Duration In Weeks (Sa)
27	2	U < 5 wks	lhu5	Δln	Unemploy.By Duration: Persons Unempl.Less Than 5 Wks (Thous.,Sa)

No.	Gp	Short Name	Mnemonic	Tran	Descripton
28	2	U 5–14 wks	lhu14	Δln	Unemploy.By Duration: Persons Unempl.5 To 14 Wks (Thous.,Sa)
29	2	U 15 + wks	lhu15	Δln	Unemploy.By Duration: Persons Unempl.15 Wks + (Thous.,Sa)
30	2	U 15–26 wks	lhu26	Δln	Unemploy.By Duration: Persons Unempl.15 To 26 Wks (Thous.,Sa)
31	2	U 27+ wks	lhu27	Δln	Unemploy.By Duration: Persons Unempl.27 Wks + (Thous,Sa)
32	2	UI claims	claimuii	Δln	Average Weekly Initial Claims, Unemploy. Insurance (Thous.) (TCB)
33	2	Emp: total	ces002	Δln	Employees On Nonfarm Payrolls: Total Private
34	2	Emp: gds prod	ces003	Δln	Employees On Nonfarm Payrolls–Goods-Producing
35	2	Emp: mining	ces006	Δln	Employees On Nonfarm Payrolls–Mining
36	2	Emp: const	ces011	Δln	Employees On Nonfarm Payrolls–Construction
37	2	Emp: mfg	ces015	Δln	Employees On Nonfarm Payrolls–Manufacturing
38	2	Emp: dble gds	ces017	Δln	Employees On Nonfarm Payrolls–Durable Goods
39	2	Emp: nondbles	ces033	Δln	Employees On Nonfarm Payrolls–Nondurable Goods
40	2	Emp: services	ces046	Δln	Employees On Nonfarm Payrolls–Service-Providing
41	2	Emp: TTU	ces048	Δln	Employees On Nonfarm Payrolls–Trade, Transportation, And Utilities
42	2	Emp: wholesale	ces049	Δln	Employees On Nonfarm Payrolls–Wholesale Trade.
43	2	Emp: retail	ces053	Δln	Employees On Nonfarm Payrolls–Retail Trade
44	2	Emp: FIRE	ces088	Δln	Employees On Nonfarm Payrolls–Financial Activities
45	2	Emp: Govt	ces140	Δln	Employees On Nonfarm Payrolls–Government
(46)	2	Emp-hrs nonag	a0m048	Δln	Employee Hours In Nonag. Establishments (AR, Bil. Hours) (TCB)
47	2	Avg hrs	ces151	lv	Avg Weekly Hrs of Prod or Nonsup Workers On Private Nonfarm Payrolls–Goods-Producing
48	2	Overtime: mfg	ces155	Δlv	Avg Weekly Hrs of Prod or Nonsup Workers On Private Nonfarm Payrolls–Mfg Overtime Hours
49	2	Avg hrs: mfg	aom001	lv	Average Weekly Hours, Mfg. (Hours) (TCB)
50	2	NAPM empl	pmemp	lv	Napm Employment Index (Percent)

No.	Gp	Short Name	Mnemonic	Tran	Descripton
129	2	AHE: goods	ces275	$\Delta^2 ln$	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls–Goods-Producing
130	2	AHE: const	ces277	$\Delta^2 ln$	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls–Construction
131	2	AHE: mfg	ces278	$\Delta^2 ln$	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls–Manufacturing

Group 3: Housing

No.	Gp	Short Name	Mnemonic	Tran	Descripton
51	3	Starts: nonfarm	hsfr	ln	Housing Starts:Nonfarm(1947–58);Total Farm & Nonfarm(1959–)(Thous.,Saar)
52	3	Starts: NE	hsne	ln	Housing Starts:Northeast (Thous.U.)S.A.
53	3	Starts: MW	hsmw	ln	Housing Starts:Midwest(Thous.U.)S.A.
54	3	Starts: South	hssou	ln	Housing Starts:South (Thous.U.)S.A.
55	3	Starts: West	hswst	ln	Housing Starts:West (Thous.U.)S.A.
56	3	BP: total	hsbr	ln	Housing Authorized: Total New Priv Housing Units (Thous.,Saar)
57	3	BP: NE	hsbne*	ln	Houses Authorized By Build. Permits:Northeast(Thou.U.)S.A
58	3	BP: MW	hsbmw*	ln	Houses Authorized By Build. Permits:Midwest(Thou.U.)S.A.
59	3	BP: South	hsbsou*	ln	Houses Authorized By Build. Permits:South(Thou.U.)S.A.
60	3	BP: West	hsbwst*	ln	Houses Authorized By Build. Permits:West(Thou.U.)S.A.

Group 4: Consumption, Orders and Inventories

61	4	PMI	pmi	lv	Purchasing Managers' Index (Sa)
62	4	NAPM new ordrs	pmno	lv	Napm New Orders Index (Percent)
63	4	NAPM vendor del	pmdel	lv	Napm Vendor Deliveries Index (Percent)
64	4	NAPM Invent	pmnv	lv	Napm Inventories Index (Percent)
65	4	Orders: cons gds	a1m008	Δln	Mfrs' New Orders, Consumer Goods And Materials (Mil. $) (TCB)
66	4	Orders: dble gds	a0m007	Δln	Mfrs' New Orders, Durable Goods Industries (Bil. Chain 2000 $) (TCB)
67	4	Orders: cap gds	a0m027	Δln	Mfrs' New Orders, Nondefense Capital Goods (Mil. Chain 1982 $) (TCB)
68	4	Unf orders: dble	a1m092	Δln	Mfrs' Unfilled Orders, Durable Goods Indus. (Bil. Chain 2000 $) (TCB)

69	4	M&T invent	a0m070	Δln	Manufacturing And Trade Inventories (Bil. Chain 2000 $) (TCB)
70	4	M&T invent/sales	a0m077	Δlv	Ratio, Mfg. And Trade Inventories To Sales (Based On Chain 2000 $) (TCB)
3	4	Consumption	cons-r	Δln	Real Personal Consumption Expenditures (AC) (Bill $) pi031/gmdc
4	4	M&T sales	mtq	Δln	Manufacturing And Trade Sales (Mil. Chain 1996 $) (TCB)
5	4	Retail sales	a0m059	Δln	Sales Of Retail Stores (Mil. Chain 2000 $) (TCB)
132	4	Consumer expect	hhsntn	Δlv	U. Of Mich. Index Of Consumer Expectations(Bcd-83)

Group 5: Money and Credit

No.	Gp	Short Name	Mnemonic	Tran	Descripton
71	5	M1	fm1	$\Delta^2 ln$	Money Stock: M1(Curr,Trav.Cks,Dem Dep,Other Ck'able Dep)(Bil$,Sa)
72	5	M2	fm2	$\Delta^2 ln$	Money Stock:M2(M1+O'nite Rps,Euro$,G/P&B/D & Mmmfs&Sav& Sm Time Dep(Bil$,Sa)
73	5	Currency	fmscu	$\Delta^2 ln$	Money Stock: Currency held by the public (Bil$,Sa)
74	5	M2 (real)	fm2-r	Δln	Money Supply: Real M2, fm2/gmdc (AC)
75	5	MB	fmfba	$\Delta^2 ln$	Monetary Base, Adj For Reserve Requirement Changes(Mil$,Sa)
76	5	Reserves tot	fmrra	$\Delta^2 ln$	Depository Inst Reserves:Total, Adj For Reserve Req Chgs(Mil$,Sa)
77	5	Reserves nonbor	fmrnba	$\Delta^2 ln$	Depository Inst Reserves:Nonborrowed,Adj Res Req Chgs(Mil$,Sa)
78	5	C&I loans	fclnbw	$\Delta^2 ln$	Commercial & Industrial Loans Outstanding + NonFin Comm. Paper (Mil$, SA) (Bci)
79	5	C&I loans	fclbmc	lv	Wkly Rp Lg Com'l Banks:Net Change Com'l & Indus Loans(Bil$,Saar)
80	5	Cons credit	ccinrv	$\Delta^2 ln$	Consumer Credit Outstanding–Nonrevolving(G19)
81	5	Inst cred/PI	ccipy	Δlv	Ratio, Consumer Installment Credit To Personal Income (Pct.) (TCB)

Group 6: Bond and Exchange Rates

86	6	Fed Funds	fyff	Δlv	Interest Rate: Federal Funds (Effective) (% Per Annum,Nsa)
87	6	Comm paper	cp90	Δlv	Commercial Paper Rate
88	6	3 mo T-bill	fygm3	Δlv	Interest Rate: U.S.Treasury Bills, Sec Mkt,3-Mo.(% Per Ann,Nsa)
89	6	6 mo T-bill	fygm6	Δlv	Interest Rate: U.S.Treasury Bills, Sec Mkt,6-Mo.(% Per Ann,Nsa)
90	6	1 yr T-bond	fygt1	Δlv	Interest Rate: U.S.Treasury Const Maturities,1-Yr.(% Per Ann,Nsa)
91	6	5 yr T-bond	fygt5	Δlv	Interest Rate: U.S.Treasury Const Maturities,5-Yr.(% Per Ann,Nsa)
92	6	10 yr T-bond	fygt10	Δlv	Interest Rate: U.S.Treasury Const Maturities,10-Yr.(% Per Ann,Nsa)
93	6	Aaa bond	fyaaac	Δlv	Bond Yield: Moody's Aaa Corporate (% Per Annum)
94	6	Baa bond	fybaac	Δlv	Bond Yield: Moody's Baa Corporate (% Per Annum)
95	6	CP-FF spread	scp90F	lv	cp90-fyff (AC)
96	6	3 mo-FF spread	sfygm3	lv	fygm3-fyff (AC)
97	6	6 mo-FF spread	sfygm6	lv	fygm6-fyff (AC)
98	6	1 yr-FF spread	sfygt1	lv	fygt1-fyff (AC)
99	6	5 yr-FF spread	sfygt5	lv	fygt5-fyff (AC)
100	6	10 yr-FF spread	sfygt10	lv	fygt10-fyff (AC)
101	6	Aaa-FF spread	sfyaaac	lv	fyaaac-fyff (AC)
102	6	Baa-FF spread	sfybaac	lv	fybaac-fyff (AC)
103	6	Ex rate: avg	exrus	Δln	United States;Effective Exchange Rate(Merm)(Index No.)
104	6	Ex rate: Switz	exrsw	Δln	Foreign Exchange Rate: Switzerland (Swiss Franc Per U.S.$)
105	6	Ex rate: Japan	exrjan	Δln	Foreign Exchange Rate: Japan (Yen Per U.S.$)
106	6	Ex rate: UK	exruk	Δln	Foreign Exchange Rate: United Kingdom (Cents Per Pound)
107	6	EX rate: Canada	exrcan	Δln	Foreign Exchange Rate: Canada (Canadian $ Per U.S.$)

Group 7: Prices

108	7	PPI: fin gds	pwfsa	$\Delta^2 ln$	Producer Price Index: Finished Goods (82=100,Sa)
109	7	PPI: cons gds	pwfcsa	$\Delta^2 ln$	Producer Price Index: Finished Consumer Goods (82=100,Sa)
110	7	PPI: int materials	pwimsa	$\Delta^2 ln$	Producer Price Index:Intermed Mat.Supplies & Components(82=100,Sa)
111	7	PPI: crude matls	pwcmsa	$\Delta^2 ln$	Producer Price Index: Crude Materials (82=100,Sa)
112	7	Spot market price	psccom	$\Delta^2 ln$	Spot market price index: bls & crb: all commodities(1967=100)

113	7	PPI: nonferrous materials	pw102	$\Delta^2 ln$	Producer Price Index: Nonferrous Materials (1982=100, Nsa)
114	7	NAPM com price	pmcp	lv	Napm Commodity Prices Index (Percent)
115	7	CPI-U: all	punew	$\Delta^2 ln$	Cpi-U: All Items (82–84=100,Sa)
116	7	CPI-U: apparel	pu83	$\Delta^2 ln$	Cpi-U: Apparel & Upkeep (82–84=100,Sa)
117	7	CPI-U: transp	pu84	$\Delta^2 ln$	Cpi-U: Transportation (82–84=100,Sa)
118	7	CPI-U: medical	pu85	$\Delta^2 ln$	Cpi-U: Medical Care (82–84=100,Sa)
119	7	CPI-U: comm.	puc	$\Delta^2 ln$	Cpi-U: Commodities (82–84=100,Sa)
120	7	CPI-U: dbles	pucd	$\Delta^2 ln$	Cpi-U: Durables (82–84=100,Sa)
121	7	CPI-U: services	pus	$\Delta^2 ln$	Cpi-U: Services (82–84=100,Sa)
122	7	CPI-U: ex food	puxf	$\Delta^2 ln$	Cpi-U: All Items Less Food (82–84=100,Sa)
123	7	CPI-U: ex shelter	puxhs	$\Delta^2 ln$	Cpi-U: All Items Less Shelter (82–84=100,Sa)
124	7	CPI-U: ex med	puxm	$\Delta^2 ln$	Cpi-U: All Items Less Midical Care (82–84=100,Sa)
125	7	PCE defl	gmdc	$\Delta^2 ln$	Pce, Impl Pr Defl:Pce (2000=100) (AC) (BEA)
126	7	PCE defl: dlbes	gmdcd	$\Delta^2 ln$	Pce, Impl Pr Defl:Pce; Durables (2000=100) (AC) (BEA)
127	7	PCE defl: nondble	gmdcn	$\Delta^2 ln$	Pce, Impl Pr Defl:Pce; Nondurables (2000=100) (AC) (BEA)
128	7	PCE defl: service	gmdcs	$\Delta^2 ln$	Pce, Impl Pr Defl:Pce; Services (2000=100) (AC) (BEA)

Group 8: Stock Market

No.	Gp	Short Name	Mnemonic	Tran	Descripton
82	8	S&P 500	fspcom	Δln	S&P's Common Stock Price Index: Composite (1941–43=10)
83	8	S&P: indust	fspin	Δln	S&P's Common Stock Price Index: & Industrials (1941–43=10)
84	8	S&P div yield	fsdxp	Δlv	S&P's Composite Common Stock: Dividend Yield (% Per Annum)
85	8	S&P PE ratio	fspxe	Δln	S&P's Composite Common Stock: &Price-Earnings Ratio (%,Nsa)

References

Aguilar, G., and M. West. 2000. Bayesian Dynamic Factor Models and Portfolio Allocation. *Journal of Business and Economic Statistics* 18:338–357.

Ang, A., and M. Piazzesi. 2003. A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics* 50:745–787.

Bai, J. 2003. Inferential Theory for Factor Models of Large Dimensions. *Econometrica* 71(1):135–172.

Bai, J., and S. Ng. 2002. Determining the Number of Factors in Approximate Factor Models. *Econometrica* 70(1):191–221.

——— 2006a. Confidence Intervals for Diffusion Index Forecasts and Inference with Factor-Augmented Regressions. *Econometrica* 74(4):1133–1150.

——— 2006b. Forecasting Economic Time Series Using Targeted Predictors. *Journal of Econometrics*, forthcoming.

——— 2008. Large Dimensional Factor Analysis. *Foundations and Trends in Econometrics* 3(2):89–163.

Boivin, J., and S. Ng. 2005. Undertanding and Comparing Factor Based Forecasts. *International Journal of Central Banking* 1(3):117–152.

Brandt, M. W., and K. Q. Wang. 2003. Time-Varying Risk Aversion and Unexpected Inflation. *Journal of Monetary Economics* 50:1457–1498.

Brillinger, D. 1981. *Time Series: Data Analysis and Theory*. San Francisco: Wiley.

Campbell, J. Y. J. H. C. 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy* 107:205–251.

Carter, C. K., and R. Kohn. 1994. On Gibbs Sampling for State Space Models. *Biometrika* 81(3):541–533.

Cochrane, J. H., and M. Piazzesi. 2005. Bond Risk Premia. *The American Economic Review* 95(1):138–160.

Connor, G., and R. Korajzcyk. 1986. Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis. *Journal of Financial Economics* 15:373–394.

DeMol, C., D. Giannone, and L. Reichlin. 2006. Forecasting Using a Large Number of Predictors: Is Bayesian Regression a Valid Alternative to Principal Components. *ECB Working Paper 700*.

Duffie, G. 2008. Information in (and not in) the Term Structure. Mimeo, Johns Hopkins University, Baltimore, MD.

Fama, E. F., and R. H. Bliss. 1987. The Information in Long-Maturity Forward Rates. *American Economic Review* 77(4):680–692.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin. 2005. The Generalized Dynamic Factor Model, One Sided Estimation and Forecasting. *Journal of the American Statistical Association* 100:830–840.

Fruhwirth-Schnatter, S. 1994. Data Augmentation and Dynamic Linear Models. *Journal of Time Series Analysis* 15:183–202.

Geweke, J., and G. Zhou. 1996. Measuring the Pricing Error of the Arbitrage Pricing Theory. *Review of Financial Studies* 9(2):557–87.

Hansen, P. 2008. In-Sample and Out-of-Sample Fit: Their Joint Distribution and Its Implications for Model Selection. Stanford University, Stanford, CA.

Kim, C., and C. Nelson. 2000. *State Space Models with Regime Switching*. Cambridge, MA: MIT Press.

Kose, A., C. Otrok, and C. Whiteman. 2003. International Business Cycles: World Region and Country Specific Factors. *American Economic Review* 93(4):1216–1239.

Kozicki, S., and P. Tinsley. 2005. Term Structure Transmission of Monetary Policy. Working Paper 05–06, Fed. Reserve Bank of Kansas City, Kansas City, MO.

Lopes, H., and M. West. 2004. Bayesian Model Assessment in Factor Analysis. *Statistical Sinica* 14:41–87.

Ludvigson, S., and S. Ng. 2007. Macro Factors in Bond Risk Premia. *Review of Fiancial Studies*, forthcoming.

Moench, E. 2008. Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach. *Journal of Econometrics* 46:26–43.

Pagan, A. 1984. Econometric Issues in the Analysis of Regressions with Generated Regressors. *International Economic Review* 25:221–247.

Piazzesi, M., and E. Swanson. 2004. Futures Prices as Risk-Adjusted Forecasts of Monetary Policy. *NBER Working Paper No. 10547*.

Stock, J. H., and M. Watson. 1989. New Indexes of Coincident and Leading Economic Indications. In *NBER Macroeconomics Annual 1989*. O. J. Blanchard and S. Fischer (ed.). Cambridge, MA: MIT Press.

——— 2002a. Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association* 97:1167–1179.

——— 2002b. Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics* 20(2):147–162.

——— 2005. Implications of Dynamic Factor Models for VAR Analysis. NBER WP 11467.

Wachter, J. 2006. A Consumption Based Model of the Term Structure of Interest Rates. *Journal of Financial Economics* 79:365–399.