Conditional Tests of Factor Augmented Asset Pricing Models with Human Capital and Housing: Some New Results

Olga Klinkowska, Autonomous University of Barcelona

1. Introduction

The static Capital Asset Pricing Model (CAPM) developed by Sharpe (1964) and Lintner (1965) and generalized by Black (1972) is the first important capital asset pricing model. In spite of many shortcomings and a lack of empirical support is still widely cited in the asset-pricing field and very often compared to other models. Moreover the CAPM and its theoretical justification account for a root of many models, which were proved to work pretty well in practical applications. This paper contributes to this trend in the empirical asset pricing literature.

According to the CAPM the risk of an asset is measured by the co-movements of the asset's return and the return on wealth portfolio—the portfolio of all the assets in the economy. Poll (1977) points that this portfolio is not observed making the CAPM model not testable. In practical applications one uses a proxy of the return on the market portfolio. But which proxy is "the best"?

Moreover in the real world investors make their decisions conditional on all the information available to them. Hansen and Richard (1987) note that these information might not be observable making tests of conditional pricing models impossible. The best what can be done is to use a subset of the unobservable information set of financial market participants. But which variables should belong to this subset?

In this paper I respond to the Roll's critique and Hansen and Richard's critique. I extend the return on wealth portfolio to include return on human capital and housing. These two assets were shown to be important components of the aggregate wealth in the economy. Moreover I propose to include all the information that we can observe as conditioning information in a conditional asset-pricing model. It still will be only the subset of real information set of the agents but the best we can create. I use the dynamic factor methodology according to which the information in a large number of economic variables can be summarized by only a few estimated common factors. I extract common factors and use some of them as conditioning variables. I estimate the asset pricing models using both the Fama MacBeth and the GMM estimation techniques. I evaluate their goodness of pricing cross-sectional average excess stock returns using formal χ^2 tests and informal parameters like R^2 (and R^2 adjusted), root mean square error and the Hansen Jagannathan distance. My results suggest that both better proxy for a total wealth portfolio and better representation of unobservable conditional information set of investors significantly improves the performance of the simple CAPM model. Moreover the conditional models proposed in this paper are also shown to be competitive to popular Fama French three factor model, which was proved to have an extraordinary empirical performance.

The idea of including human capital and real estate in the market portfolio is not new, however, using dynamic factor analysis to generate conditional information set in asset pricing context has not been deeply studied yet. Already Jagannathan and Wang (1966) observed that human capital, although not directly measurable, is an important part of the total wealth in the economy. In the proxy for the return on the market portfolio they include the return on human capital, which they measure by the growth rate in per capita labour income. In this paper I follow them and use the same proxy for the return on human capital. Heaton and Lucas (2000) consider the variant of the model proposed by Jagannathan and Wang (1966) in which the human capital is determined by two components: the value of future wage income and the value of future proprietary income. They construct the returns to these two components of human capital using the growth rates in aggregate wage income and aggregate nonfarm proprietary income.

Kullmann (2003) points that also real estate constitutes a significant portion of the aggregate wealth. She extends the proxy for the wealth portfolio and incorporates commercial and residential real estate. I do not differentiate between these two types of real estate and include only residential wealth. As a proxy for the return on residential properties she uses the change in the median price of existing homes as reported by NAR[1]. I use different proxy: the change in OFHEO House Price Index. Additionally my empirical study is conducted for quarterly frequency while Kullmann uses monthly data. To my knowledge Kullmann (2003) is the only paper that looks at the importance of housing capital in the simple CAPM framework. The results received in this paper complement those obtained by Kullmann.

I introduce the dynamics into the CAPM model by explicitly modelling the coefficients in the stochastic discount factor as dependent on the current period information. This technique has already been applied by Lettau and Ludvigson (2001) who used consumption wealth ratio as a conditioning variable.

Applications of estimated common factors are quite common in the macroeconomic literature but scarce in empirical finance literature. Ludvigson and Ng (2007b) use dynamic factor approach to model empirical conditional risk-return relation of excess stock market returns. In another paper (Lud-

[1] National Association of Realtors

vigson and Ng (2007a)) they use the same methodology to analyze bond risk premium. Bai (2007) derives from the common factors new state variables and shows that innovations in these state variables account for the cross-section of expected returns. To my knowledge these are the main applications of dynamic factors in the asset pricing literature. My paper contributes then to this field with the new use of common factors.

The paper proceeds as follows. Section 2 discusses the theoretical justification of the models. I present the asset-pricing model and expand it to include human capital and housing wealth. Moreover I describe how I introduce conditional information. Section 3 describes the data used in empirical study and their sources. I show how I construct the measures of the return on human capital and residential properties and I describe a large set of macroeconomic and financial variables from which I extract conditioning variables. In Section 4 I extract common factors by principal component methodology. I choose which estimated factors I use as conditional variables and describe them. Section 5 explains the estimation techniques used in practical applications. I discuss the Fama MacBeth and GMM approaches and present the evaluation criteria. Section 6 discusses obtained results and comments on them. Section 7 concludes.

2. Theoretical Model

In this section I present the theoretical model for which I conduct the empirical analysis. I describe the theory that explains the pricing mechanism and justify why the model gives some interesting results.

2.1. CAPM and its enrichment

The static Capital Asset Pricing Model (CAPM) developed by Sharpe (1964) and Lintner (1965) and generalized by Black (1972) is the first important capital asset pricing model. It has very sound theoretical basis relying on the mean-variance efficiency concept. The model states that the expected return on any risky asset i is proportional to the asset's systematic risk. This risk is measured by beta β_i^m, defined as a covariance of the asset's return with the return on the market portfolio normalized by the variance of the market portfolio:

$$E\left(R_{i,t+1}\right) = E\left(R_{0,t}\right) + \lambda_m \beta_i^m$$

where $\beta_i^m = \dfrac{Cov\left(R_{i,t+1}, R_{m,t+1}\right)}{Var\left(R_{m,t+1}\right)}$ and $E(R_{0,t})$ is the return on zero-beta portfolio.

The same expressed in the language of Euler equation is given by $E(m_{t+1}R_{i,t+1}) = 1$, where $m_{t+1} = a + b\,R_{\mathrm{m},t+1}$ and is a stochastic discount factor (SDF hereafter). It is important to underline here that this specification is unconditional. Cochrane (2005) points that when the parameters a and b in the SDF are constant then conditional pricing $E_t(m_{t+1}R_{i,t+1}) = 1$ is equivalent to unconditional pricing $E(m_{t+1}R_{i,t+1}) = 1$. It makes then no difference to look at conditional or unconditional pricing equation.

According to the theory the market portfolio is defined as a portfolio of all risky assets that exist in the market and it represents total wealth in the economy. The return on market portfolio is unknown and in empirical applications plausible proxies are used. The classical ones are the returns on broad-based stock market indices, which include all or most of the assets traded on stock exchange. Roll (1977) criticizes such approach and states that any proxy is poor and cannot provide an accurate representation of the entire market. True wealth portfolio would necessarily include every available asset like commodities, real estate, precious metals etc. It is impossible therefore to create the market portfolio and calculate the return on it. In spite of this the research world still uses proxies and tries to construct them in the best possible way.

Mayers (1972) points out that the human capital accounts for the substantial part of the total wealth in the economy. Jagannathan and Wang (1966) are the first to extend the proxy for the market return to include also a measure of the return on human capital. This improves the explanation power of the classical CAPM model. The R^2 increases from 0.014 to 0.305 when the return on human capital is included in the market return.[2]

Heaton and Lucas (2000) use the Survey of Consumer Finances to examine the cross-sectional variation in the composition of the household's wealth. Their analysis shows that the real estate is an extremely large component of individuals' financial wealth as well as total wealth. Kullmann (2003) takes this into account and includes the return on real estate into the market return. She also differentiates between residential and commercial real estates and justifies this by the fact that these are two different asset classes and they need to be treated separately. Adding the two types of real estate into the market portfolio and including their returns in the market return increases the R^2 of the model to 0.48 from 0.14 for the classical CAPM. When also the proxy for the return on human capital is included, then the R^2 of such a model rises to 0.49.[3]

The idea of looking closer at the wealth portfolio and specifying the particular and significant components of this wealth forms the basis of my modelling strategy.

Human capital is a non-marketable asset and identifying the return on it may seem a challenging task. It is difficult to define human capital. Economists, while referring to it, usually have in mind properties like: productive

---

[2] This is calculated for 100 size and pre-beta sorted portfolios, in the period July 1963–December 1990.

[3] This is calculated for 100 size and pre-beta sorted portfolios, in the period January 1972––December 1999.

skills acquired by employees, technical knowledge obtained through education process and experience, capabilities of the individuals who use them to improve their productivity and increase their income. The natural trend is then that highly qualified households receive greater income. One can easily notice the strong relation between human capital and labour. Many early economic theories referred to human capital simply as labour, one of the three factors of production. It seems then justified that the proxy for the return on human capital should be related to the labour income. Jagannathan and Wang (1966) point that the monthly per capita income in the U.S. from salaries and wages was about 63% of the total income in the period January 1959–December 1992. This suggests that the human capital contributes significantly to the total wealth.

Real estate is another important constituent of total wealth. Heaton and Lucas (2000) report that owner-occupied real estate constitutes 33,3% of total national wealth and accounts for the largest fraction of it. Since 1995 the homeownership rate in U.S. increased by around 8% making the housing even more important asset. Residential real estate differs substantially from other assets. It combines the flow of housing services (the homeowner gets to live in the house instead of renting it—it is the consumption good) and an investment good (return on the housing equity—it is the saving vehicle). Moreover housing wealth is less liquid than any other financial wealth and exposes the homeowner to the idiosyncratic household-specific risk. The Census Bureau of U.S. reports that around 70% of households own their residences. All this suggests that the housing represents a significant portion of total wealth.

The first basic asset-pricing model that I test is a model in which the wealth portfolio is extended to include the human capital wealth and the residential housing wealth as these are important components of the total wealth. Differently from Kullmann's approach I do not distinguish between the residential and commercial real estate and I include only broadly-defined residential housing.[4] I assume that the market portfolio is a linear function of its main components:

$$R_{m,t+1} = \varphi_0 + \varphi_{vw} R_{vw,t+1} + \varphi_{hc} R_{hc,t+1} + \varphi_{hs} R_{hs,t+1}$$

[4] As a return on commercial real estates Kullmann (2003) uses the return on NAREIT EREIT index, which represents the total return on a portfolio of all Equity REITs traded on stock exchange. Equity REITs are publicly traded companies that own and often operate income-producing real estate. However EREITs account only for around 3% of commercial real estate according to Case Glaeser and Parker (2000) and might not be an accurate measure of the return on commercial properties. Additionally the empirical results of Kullmann (2003) show that in most cases the estimated risk factors related to commercial properties are not significant.

Such a specification leads to the following unconditional pricing equation:

$$E\left(R_{i,t+1}\right) = E\left(R_{0,t}\right) + \lambda_{vw}\beta_i^{vw} + \lambda_{hc}\beta_i^{hc} + \lambda_{hs}\beta_i^{hs}$$

This specification can be viewed as a three-factor model where each factor represents the return on different part of total wealth portfolio. As a result β^{hc} represents the risk related to human capital wealth and β^{hs} represents the risk related to housing wealth. The pricing kernel equivalent to this expected beta representation is the following:

$$m_{t+1} = a + b_{vw}R_{vw,t+1} + b_{hc}R_{hc,t+1} + b_{hs}R_{hs,t+1}$$

One of the shortcomings of the CAPM and its proposed extension is the fact that this is a one period model and does not include any dynamics. When applied period by period it produces risk premium, which are constant over time. It fails to take into account the time-varying investment opportunities and does not account for the intertemporal hedging component of asset demand. All this is great limitation of the model. Many empirical studies[5] show that stock excess returns can be forecasted, suggesting that risk premium vary over time. Lettau and Ludvigson (2001) show that they are high in bad times when risk or risk aversion is high and low in good times when the risk or risk aversion is low. This reflects the fact that agents, when making investment decisions, take into account all the information which allow them to make good decisions. In the next subsection I explain the way of introducing dynamics into this model.

2.2. Conditional Factor Model

Cochrane (2005) proposes

a very simple and beautiful solution to the problem raised by conditioning information.

This solution is to explicitly model the coefficients in the SDF as dependent on the current period information in a linear way:

$$a_t = a_0 + a_z z_t$$
$$b_t = b_0 + b_z z_t$$

The SDF $m_{t+1} = a_t + b_t' f_{t+1}$, where $f_{t+1} = \left[R_{vw,t+1}, R_{hc,t+1}, R_{hs,t+1},\right]'$, has then time varying coefficients and such a specification is able to produce time-

[5] E.g. Shiller (1984), Campbell and Shiller (1988), Fama and French (1988), Fama and French (1989), Campbell (1991), Fama and Schwert (1977), Hodrick (1992), Lamont (1998)], Ferson and Harvey (1999).

-varying risk premium. Now it is the conditional pricing equation that matters since agents make investment decisions conditional on information available to them at this time. The variable z_t reflects this information. Such approach allows to obtain model with time-varying risk premium in which factors price assets conditionally but which has little theoretical structure.

For given pricing kernel the conditional beta representation is given by:

$$E_t\left(R_{i,t+1}\right) = R_{0,t} + \lambda_t'\beta_{it} \tag{1}$$

for $i = 1, \ldots, N$, where

$$\begin{aligned}\beta_{it} &= Var_t\left(f_{t+1}, f_{t+1}'\right)^{-1} Cov_t\left(R_{i,t+1}, f_{t+1}\right)\\ \lambda_t &= -R_{0,t} Var_t\left(f_{t+1}, f_{t+1}'\right) b_t\end{aligned} \tag{2}$$

It is obvious now that risk prices λ_t and measures of riskiness β_{it} are now time-dependent and the expected returns are conditional on available information. The linear conditional factor model can be expressed as a scaled multifactor model with constant coefficients:

$$\begin{aligned}m_{t+1} &= \left(a_0 + a_z z_t\right) + \left(b_{0,vw} + b_{z,vw} z_t\right) R_{vw,t+1} + \left(b_{0,hc} + b_{z,hc} z_t\right) R_{hc,t+1} +\\ &+ \left(b_{0,hs} + b_{z,hs} z_t\right) R_{hs,t+1}\end{aligned} \tag{3}$$

$$\begin{aligned}m_{t+1} &= a_0 + a_z z_t + b_{0,vw} R_{vw,t+1} + b_{z,vw}\left(z_t R_{vw,t+1}\right) + b_{0,hc} R_{hc,t+1} +\\ &+ b_{z,hc}\left(z_t R_{hc,t+1}\right) + b_{0,hs} R_{hs,t+1} + b_{z,hs}\left(z_t R_{hs,t+1}\right)\end{aligned}$$

$$m_{t+1} = a_0 + c'\widetilde{f}_{t+1}$$

where
$\widetilde{f}_{t+1} = \left[z_t \;\; R_{vw,t} \;\; z_t \;\; R_{vw,t} \;\; z_t \;\; R_{hc,t} \;\; z_t \;\; R_{hc,t} \;\; z_t \;\; R_{hs,t} \;\; z_t \;\; R_{hs,t}\right]'$ and
$c = \left[a_z \;\; b_{0,vw} \;\; b_{z,vw} \;\; b_{0,hc} \;\; b_{z,hc} \;\; b_{0,hs} \;\; b_{z,hs}\right]'$.
For the new pricing factors $\widetilde{f}_{t+1}$ account: the instrument z_t, the original factors f_{t+1} and the original factors scaled with the conditional variable z_t: $z_t f_{t+1}$. The advantage of such representation is that now the SDF has constant coefficients and this multifactor scaled model can be tested using unconditional moments since, as noted in the previous section, in this case the conditional pricing is equivalent to unconditional pricing. The unconditional multifactor beta representation for this SDF is the following:

$$E\left(R_{i,t+1}\right) = E\left(R_{0,t}\right) + \widetilde{\lambda}'\widetilde{\beta}_i \tag{4}$$

for $i = 1, \ldots, N$, where

$$\widetilde{\beta}_i = Var\left(\widetilde{f}_t, \widetilde{f}_t'\right)^{-1} Cov\left(R_{it}, \widetilde{f}_t\right)$$
$$\widetilde{\lambda} = -E\left(R_{0,t}\right) Var\left(\widetilde{f}_t, \widetilde{f}_t'\right) c$$

It is important to underline here that for the scaled model the lambdas $\widetilde{\lambda}$ do not have the casual interpretation as the risk prices. This is because the scaled model is derived from a conditional one, which has a conditional beta representation with time-varying risks β_{it} and their prices λ_t, given by equation (2). There is no simple relation between time-varying risk prices l_t and coefficients $\widetilde{\lambda}$. I apply methodology that allows me to estimate a scaled unconditional model of the form (4). It is possible to uncover the estimate of b_t but still without making any additional assumptions it is not possible to uncover time-varying risk prices λ_t for the factors f_{t+1}. Facing this difficulty I concentrate on unconditional moments and evaluate the model on the basis of its unconditional pricing abilities.

2.3. Conditioning variable

An important decision is the choice of the conditioning variable z_t. Agents make their investment decision conditional on all the information available for them in the moment of decision-making. The role of the conditioning variable is then to sum up and express all this information. It should give some suggestions on how the asset prices and returns are going to behave in the future. The instrument z_t should be a good predictor of assets' returns and so far only these possibilities have been explored in empirical research.[6] However, such approach creates some potential inconveniences. First of all a large body of the literature demonstrates that there are more than one variables that predict stock returns. Valuation ratios (like dividend price ratio or earnings price ratio), interest rates, time or default spreads, cay ratio (consumption wealth ratio) were proved to possess good forecasting abilities. A simple question is then which variable to choose. Choosing one of them means that the information included in the others will not be used. Moreover there may appear some problems of temporal instability in the forecasting relations. There are periods in which the forecasting abilities of some predictors may be weakened or may disappear at all. A classical example is the behaviour of dividend price ratios in the mid–90-ties when they were heavily going down but it was not reflected in lower stock returns as expected. Additional critique was presented by Hansen and Richard (1987) who pointed that conditioning information of agents are not observable and, by analogy to Roll's critique, the conditional assets-pricing model is not testable. In empirical applications what is used, is always a proxy and as a consequence this proxied information set is always a subset of the agents' information set.

[6] E.g. Ferson and Harvey (1999), Santos and Veronesi (2006), Lettau and Ludvigson (2001).

I propose to make this proxy as good as possible by including information coming from a large set of macroeconomic and financial variables instead of just one variable. This information is reflected in dynamic factors. The research on dynamic factor models prove that the information in a large number of economic time series can be summarized by only a few estimated factors. This gives a possibility to explore a much richer set of instruments, which is more likely to span the unobservable information set of financial market participants. It is then more probable to reduce the effect of the omitted-information estimation bias. Moreover summarizing information from a large set of variables allows to eliminate the arbitrariness in the choice of the conditioning variable and may provide a robustness against the problem of temporal instability. Improved representation of the agents' information set is then my response to the Hansen and Richard's critique.

2.4. Tested models

For the sake of completeness and in order to facilitate the assessment of the models I examine the performance of the following models.

- CAPM model:

$$E\left(R^{e}_{i,t+1}\right) = \gamma + \lambda_{vwf}\beta^{vwf}_{i}$$

$$m_{t+1} = 1 + b_{vwf}\left(R^{vw}_{t+1} - R^{f}_{t+1}\right)$$

- vw-hc-hs model:

$$E\left(R^{e}_{i,t+1}\right) = \gamma + \lambda_{vw}\beta^{vw}_{i} + \lambda_{hc}\beta^{hc}_{i} + \lambda_{hs}\beta^{hs}_{i}$$

$$m_{t+1} = 1 + b_{vw}R^{vw}_{t+1} + b_{hc}R^{hc}_{t+1} + b_{hs}R^{hs}_{t+1}$$

- conditional vw-hc-hs models:

$$E\left(R^{e}_{i,t+1}\right) = \gamma + \lambda_{z}\beta^{z}_{i} + \lambda_{vw}\chi\beta^{vw}_{i} + \lambda_{zvw}\beta^{zvw}_{i} + \lambda_{hc}\beta^{hc}_{i} + \lambda_{zhc}\beta^{zhc}_{i} +$$
$$+\lambda_{hs}\beta^{hs}_{i} + \lambda_{zhs}\beta^{zhs}_{i}$$

$$m_{t+1} = 1 + z_{t} + b_{vw}R^{vw}_{t+1} + b_{zvw}\left(z_{t}R^{zvw}_{t+1}\right) + b_{hc}R^{hc}_{t+1} + b_{zhc}\left(z_{t}R^{zhc}_{t+1}\right) +$$
$$+b_{hs}R^{hs}_{t+1} + b_{zhs}\left(z_{t}R^{hs}_{t+1}\right)$$

I evaluate as well the performance of the Fama French three-factor model (FF) that has been empirically successful for comparison:

$$E\left(R^{e}_{i,t+1}\right) = \gamma + \lambda_{vwf}\beta^{vwf}_{i} + \lambda_{smb}\beta^{smb}_{i} + \lambda_{hml}\beta^{hml}_{i}$$

$$m_{t+1} + 1 + b_{vwf}\left(R^{vw}_{t+1} - R^{f}_{t+1}\right) + b_{smb}R^{smb}_{t+1} + b_{hml}R^{hml}_{t+1}$$

3. Data

In this study I use data sampled at a quarterly frequency and the sample period is 1975Q2–2006Q4. This gives me 127 time series observations.

It has become standard in the empirical asset pricing literature to see how well the model prices the 25 Fama French portfolios[7]. This is understandable because of the importance of size and value anomalies and I follow this stream in the literature. However Lewellen et al. (2007) point that these portfolios have a strong factor structure, which means that the group of some factors can explain nearly all of the time-series variation in portfolio returns. This makes it likely that the betas on at least any two proposed factors will line up with the portfolios' expected returns. In order to break this structure they propose to include portfolios that do not correlate so strongly with *SMB* and *HML* factors and evaluate the model in terms of how well it prices all the 25 Fama French portfolios and added portfolios at the same time. Following this advice I use as test assets the following: 3 size portfolios, 3 book-to-market value portfolios and 10 industry portfolios (total of 16 portfolios). All the data on the returns of the portfolios are taken from the Kenneth French's web page.[8] I use the excess returns of the test portfolios over the three-month T-Bill rate taken from the CRSP[9] from the WRDS[10] database.

As a measure of the return on market portfolio I use the return on a value-weighted index, which includes all the assets traded in NYSE, AMEX and NASDAQ. The data on the index are taken from the CRSP from the WRDS database.

To proxy the return on human capital I follow Jagannathan and Wang (1996) and I define the return on human capital as a growth rate in per capita labour income L_t:

$$R_t^{hc} = \frac{L_t + L_{t-1}}{L_{t-1} + L_{t-2}} - 1 \tag{5}$$

The data on labour income come from the NIPA Tables published by U.S. Bureau of Economic Analysis (BEA). The main concern is how the labour income is defined. Jagannathan and Wang (1996) treat the labour income as a difference between total personal income and personal dividend income. Moreover they divide this difference by the total U.S. population so they get per capita labour income. Apart from that they use monthly observations and apply a one-lag timing convention justifying this that it is consistent with the fact the monthly labour income data are typically published with a one-month delay. However with quarterly data this is not the case. I do not follow

---

7 (abbr.: FF25 portfolios) These are value-weighted returns for the intersection of 5 size and 5 book-to-market B/M equity portfolios traded on NYSE, AMEX and NASDAQ.

8 http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

9 Centre for Research in Security Prices.

10 Wharton Research Data Services.

then Jagannathan and Wang's timing convention and instead assume that at the end of the period the income is already known.

To proxy the return on housing I use the net change in the OFHEO[11] House Price Index (HPI):

$$R_t^{hs} = \frac{HPI_t}{HPI_{t-1}} - 1 \tag{6}$$

The OFHEO HPI for U.S. is a broad measure of the movement of single-family house prices in the U.S. It serves as an indicator of house price trends in U.S. and reflects the cost of structure and land, simultaneously controlling for the quality of the house. The HPI is a weighted, repeat-sales index—it measures average price changes in repeat sales or refinancing on the same physical properties. The calculation methodology uses the repeat valuation framework, which is widely considered as the most accurate way to measure valuation changes in housing markets over time. When a specific property is for example resold the new sale price is matched to the property's first sale price. These two price points for a specific property are called a "sale pair". The difference in the sale pair is measured and recorded. The HPI is based then on repeat transactions, which helps to control for the differences in the quality of the houses comprising the sample used for statistical estimation. For this reason the HPI is described as a "constant quality" house price index. The index is calculated using mortgage transaction data provided by Fannie Mae and Freddie Mac. The HPI is calculated quarterly and published with around 2-month lag. It is available since 1975–Q1.

I estimate common factors from a large panel of 126 macroeconomic and financial time series.[12] The data are obtained from the Global Insights Basic Economics Database and the Conference Board's Indicators Database. They were chosen to represent wide categories of economic and financial time series: real output and income, employment and hours, real retail, manufacturing and sales data, international trade, consumer spending, housing starts, inventories and inventory sales ratios, orders and unfilled orders, compensation and labour costs, capacity utilization measures, price indexes, interest rates and interest rate spreads, stock market indicators and foreign exchange measures. The raw data are transformed in order to assure stationarity and before extraction of the factors they are also standardized. The complete list of all the series used in this study as well as the details of the transformations are given in the Appendix.

---

[11] Office for Federal Housing Enterprise Oversight.

[12] This is a subset of the 132 variables used by Stock and Watson (2002a), Stock and Watson (2002b), Stock and Watson (2004), Ludvigson. and Ng (2007a).

4. Extraction of the common factors

I use the methodology of dynamic factor analysis for large datasets in order to obtain few factors that can effectively represent information included in a large panel of variables. In this approach it is assumed that each variable i in a large dataset of N variables has the factor structure of the form:

$$x_{it} = \lambda_t' f_t + e_{it}$$

where f_t is a $r \times 1$ vector of $r \times 1$ latent common factors and λ_i is a corresponding $r \times 1$ vector of latent factor loadings and e_{it} is an idiosyncratic error term. It is important to underline here that the number of latent common factors is substantially lower than the number of variables in a dataset that is $r << N$ so that there exist few factors (r) that can express the information contained in many variables (N). Moreover the cross-sectional dimension N is large, and might be larger than the time-series dimension T, that is $N, T >> 0$. In this paper $N = 126$ and $T = 127$. As the common factors are not observable I use the methodology of principal components analysis to estimate these factors and replace f_t by $\hat{f}_t$, which is a vector of estimated first r principal components. Estimated in such a way common factors are orthogonal. Jushan and Ng (2002) propose the following criteria that allow consistently estimate the number of factors r:

$$\hat{r} = \arg \min_{0 \le k \le kmax} PCP(k) \quad \text{with} \quad PCP(k) = S(k) + k\bar{\sigma}^2 g(N,T)$$

$$\hat{r} = \arg \min_{0 \le k \le kmax} IC(k) \quad \text{with} \quad IC(k) = \ln\big(S(k) + kg(N,T)\big)$$

where $S(k) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \hat{\lambda}_i^k \hat{f}_t^k)^2$ is the average sum of squared residuals when k factors are estimated $\bar{\sigma}^2 = S(kmax)$, with $kmax$ prespecified and $g(N, T)$ is a penalty function.[13] According to these criteria the first 8 principal components can consistently span the same space as the real factors f_t. Figure 1 shows that these 8 estimated common factors explain around 60% of the variance of the dataset composed of 126 macro and financial variables.

The next step is to choose only these common factors that have predictive power for excess stock returns. Additional advantage when checking for predictive abilities of the estimated factors is the fact that they are not persistent. Table 1 below shows that the autocorrelation coefficients are not high,

[13] Jushan and Ng (2002) present the conditions that the penalty function needs to satisfy and give some examples of penalty functions. One of them is $g(N,T) = \frac{N+T}{NT}\ln\left(\frac{NT}{N+T}\right)$. For more examples please look into Jushan and Ng (2002).

the highest one is 0.62 for the first estimated factor. For comparison, valuation ratios like dividend price of earnings price ratios exhibit the persistency of magnitude 0.90. The problem with strong autocorrelation is that it can cause a finite-sample bias in the coefficients of predictive regressions estimated by OLS. However, when potential forecasters are weakly autocorrelated the OLS estimates are trustworthy. It is possible then to run forecasting regressions by OLS and evaluate the predictive power of the candidate forecasters by looking at the their individual significance, R^2 and R^2 adjusted and some information criteria[14].

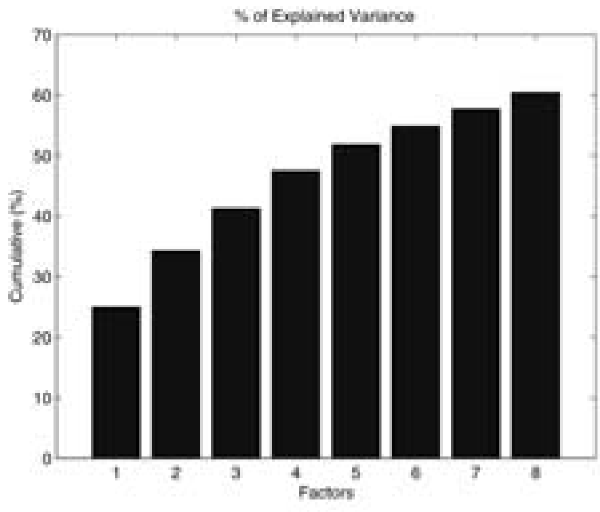

Figure 1.

The cumulative percentage of the variance in the large panel of macro and financial data explained by the factors

Table 1.

The first order correlation coefficients for common factors

Factors	AR1($\hat{f}_i$)
$\hat{f}_1$	0.62
$\hat{f}_2$	0.44
$\hat{f}_3$	–0.15
$\hat{f}_4$	0.06
$\hat{f}_5$	0.52
$\hat{f}_6$	–0.27

[14] It is however not possible to judge the sign of the coefficients as this can be done for purely defined predictors because the potential candidates are now the factors that are influenced by all the variables in the large dataset.

Factors	AR1($\hat{f}_i$)
$\hat{f}_7$	0.28
$\hat{f}_8$	0.41

In the large literature on the predictive power for excess stock returns it is pretty common to take simple historical average excess return as a benchmark and then investigate the consequence of adding a candidate predictor. I follow this practice and run regressions of the form:

$$r^e_{t+1} = \gamma_0 + \gamma_i \hat{f}_{i,t} + \varepsilon_{t+1} \qquad (7)$$

where $i = 1, \ldots, 8$ so each estimated common factor is taken as a regression one at a time and r^e_{t+1} are quarterly returns on $S\&P500$ Index in excess of three-month T-bill rate. I assess the predictive power of each common factor in terms of its significance, R^2 and R^2 adjusted. As the factors are orthogonal, the same estimates of coefficients can be obtained by regressing excess return on all the factors at one time.[15] However then it is not possible to asses each factor in terms of what fraction of variance in excess returns it explains. The results of the individual regressions are presented in the Table 2.

Table 2.

Results of the individual predictive regressions given by (7) estimated by OLS

	$\hat{\gamma}_i$	t-stat	p-value	R^2	R^2adj
$\hat{f}_1$	–0.29	–3.03	0.001	0.047	0.042
$\hat{f}_2$	–0.29	–1.81	0.036	0.017	0.012
$\hat{f}_3$	–0.19	–1.04	0.149	0.006	0.000
$\hat{f}_4$	0.16	0.83	0.205	0.004	–0.002
$\hat{f}_5$	0.24	1.04	0.149	0.006	0.000
$\hat{f}_6$	0.22	0.79	0.214	0.003	–0.002
$\hat{f}_7$	–0.16	–0.54	0.293	0.002	–0.004
$\hat{f}_8$	–0.13	–0.44	0.329	0.001	–0.004

We can observe that only the first two common factors $\hat{f}_1$ and $\hat{f}_1$ are statistically important. Moreover R^2 and R^2 adjusted for these factors are significantly higher than for others. I also check if the information summarized by these two factors is not duplicated by other variables, which in the predictive literature already have the status of good forecasters. In order to do this I run the following predictive regressions:

[15] In case of multiple regression, not reported here, estimated coefficients are approximately the same as those reported in the Table 2.

$$r^e_{t+1} = \theta_{1,0} + \theta_{1,3} x_t + \varepsilon_{1,t+1}$$

$$r^e_{t+1} = \theta_{2,0} + \theta_{2,1} \hat{f}_{1t} + \theta_{2,3} x_t + \varepsilon_{2,t+1}$$

$$r^e_{t+1} = \theta_{3,0} + \theta_{3,2} \hat{f}_{2t} + \theta_{3,3} x_t + \varepsilon_{3,t+1}$$

$$r^e_{t+1} = \theta_{4,0} + \theta_{4,1} \hat{f}_{1t} + \theta_{4,2} \hat{f}_{2t} + \theta_{4,3} x_t + \varepsilon_{4,t+1}$$

where $x_t \in$ {*divided price ratio* (*dp*), *earnings price ratio* (*ep*), *default spread* (*ds*), *term spread* (*ts*), *cay*}. I compare this model with the benchmark, which is its restricted version with x_t as the only regressor. The results of these regressions are presented in the Table 3.

Table 3.

The results of predictive regressions with additional predictors as regressors

	const	$\hat{f}_{1t}$	$\hat{f}_{2t}$	dp	ep	ds	ts	coy	R^2	R^2odj
(1)	3.89***	–0.29***	–0.29**						0.065	0.059
(2)	16.17***			3.65***					0.039	0.035
(3)	15.37***	–0.26***		3.25***					0.078	0.073
(4)	19.08***		–0.41***	4.51***					0.074	0.069
(5)	18.32***	–0.25***	–0.40***	4.09***					0.111	0.106
(6)	13.62***				3.49***				0.035	0.029
(7)	12.43***	–0.26***			3.06**				0.074	0.069
(8)	18.21***		–0.51***		5.13***				0.082	0.077
(9)	16.88***	–0.25***	4.66***		4.66***				0.117	0.112
(10)	0.83					3.11***			0.031	0.025
(11)	2.12*	–0.23**				1.81			0.056	0.051
(12)	1.08		–0.24*			2.86**			0.043	0.038
(13)	2.44**	–0.24***	–0.26**			1.47			0.070	0.065
(14)	3.73***						0.09***		0.047	0.042
(15)	3.78***	–0.21**					0.07**		0.067	0.062
(16)	3.75***		–0.13				0.09***		0.050	0.049
(17)	3.81***	–0.23***	–0.19				0.05*		0.074	0.069
(18)	3.85***							1.37***	0.065	0.059
(19)	3.85***	–0.23***						1.18***	0.093	0.088
(20)	3.85***		–0.25**					1.33***	0.078	0.073
(21)	3.85***	–0.23***	–0.26**					1.14***	0.107	0.102

We can observe that in almost all the specifications the estimated common factors are statistically important. Moreover the R^2 and R^2adj statistics significantly increase when one of the factors or both are included as additional regressors. This proves that indeed they do not duplicate the information already included in each of the prespecified regressor x_t. Relying on the presented analysis I choose the first and the second estimated common factors $\hat{f}_{1t}$ and $\hat{f}_{2t}$ as those that will serve as instrument z_t in conditional linear asset pricing model specified in equation (3).

It would be desirable to give an economic interpretation to the predictive factors. However, this might not be easy since there can be many estimates of factors as they are only identifiable up to an $r \times x$ matrix. It is then inaccurate to look at the values of estimated factors. Moreover principal component methodology produces factors, which are orthogonal, so they carry different information and do not overlap. The most important however, is that, by construction, common factors are influenced in different degrees by all the variables, which constitute the large panel of the data. This means that they do not explicitly reflect only one macro or financial variable. Nevertheless I briefly describe the two common factors in terms of how they relate to the variables included in the large dataset. Following Ludvigson and Ng (2007a), I investigate this relation by looking at the fraction of variance in each variable of the large dataset explained by each of the two estimated factors $\hat{f}_{1t}$ and $\hat{f}_{2t}$ individually. In other words I regress each of the 126 variables ont $\hat{f}_{1t}$ and $\hat{f}_{2t}$, one at a time, and report R^2 statistics. For each factor then I depict marginal R^2, which are the calculated R^2 statistics, in the form of bar charts. The 126 macro and financial variables are grouped by economic categories and labeled using numbered ordering given in the Appendix.

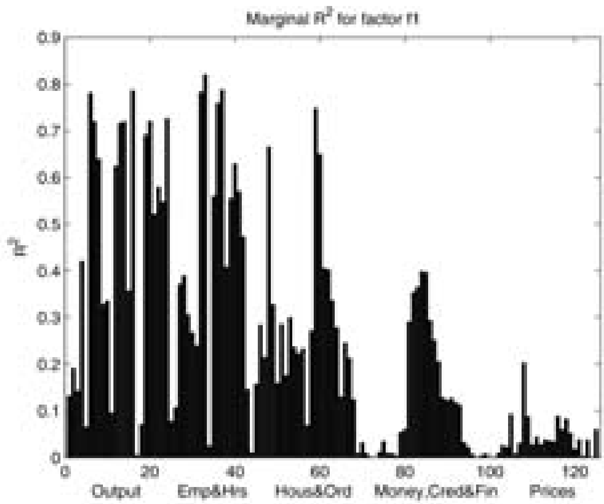

Figure 2.
Marginal R^2 for factor $\hat{f}_{1t}$

Figure 2 shows the marginal R^2 for the first common factor. We can observe that this factor is heavily related to macroeconomic variables, which measure industrial production, employment, housing and new manufacturing orders, but there is little correlation with financial variables and prices. On the other hand from the Figure 3 we can conclude that the second common factor loads mainly on financial variables like interest rates and interest rate spreads as well as on prices. But there is little correlation with macroeconomic variables. It is interesting to confront these conclusions with the results of predictive regressions. When both the second common factor $\hat{f}_{2t}$ and spread rate (either default spread or time spread) are included in one regression, we can see that either $\hat{f}_{2t}$ or spread rate become insignificant and R^2 does not increase so much as in case of other prespecified predictors. The reason of this is probably the fact tha $\hat{f}_{2t}$ reflects mainly information included in interest rates and interest rate spreads among which there are default and time spreads. It should not be the surprising that $\hat{f}_{2t}$ and rate spreads duplicate some information.

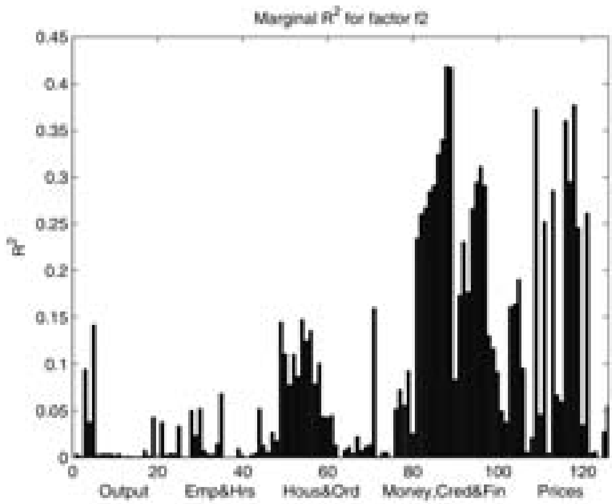

Figure 3.
Marginal R^2 for factor $\hat{f}_{2t}$

5. Estimation Methodology

I estimate the different asset pricing model specifications using two procedures: the Fama-MacBeth procedure proposed by Fama and MacBeth (1973) and the GMM procedure.

5.1. Fama MacBeth procedure

The Fama-MacBeth (FM hereafter) procedure is a very simple and intuitive approach widely used in many empirical applications of asset pricing models in spite of some shortcomings. Lettau and Ludvigson (2001) point that it can be especially useful in cases when there is a moderate number of time series observations (T) but still one wants to check how the model prices a relatively large number of assets (N). The procedure has two stages; in the first stage the time series regressions are run and the betas are estimated:

$$R^e_{it} = a_0 + \beta'_i f_t + \varepsilon_{it} \tag{8}$$

for $i = 1, ..., N$ and $t = 1, ..., T$.

The first stage of FM produces the beta estimates $\hat{\beta} = \text{cov}(R_t, f'_t)var(f_t f'_t)^{-1}$. Given betas in stage two one runs the cross-sectional regressions of portfolio returns on the betas at each time period in the sample:

$$R^e_t = \gamma_t + \hat{\beta}\lambda_t + \alpha_t \tag{9}$$

Cochrane (2005) points that one can run this regression with or without a constant as the theory says that the constant or zero-beta excess return should be zero. Moreover, according to econometric theory, including the constant does not influence the consistency of the estimated parameters lambdas λ_t. In the empirical literature on the asset pricing models the cross-sectional regression is usually run with constant and then authors either comment on the constant estimate or leave it without any explanation. In this paper I include the constant in the estimation. I provide a detailed discussion and the influence of including the constant along with the results.

For the simplicity of notation let $X = [1 \ \ \hat{\beta}]$ and $\theta_t = [\gamma_t \ \ \lambda'_t]'$.

The second stage FM procedure results in a time series of coefficient estimates $\{\hat{\theta}\}^t_{t=1}$ and pricing errors $\{\hat{\alpha} = R^e_t - \hat{\gamma}_t - \hat{\beta}\hat{\lambda}\}^t_{t=1}$. The final estimates are the time series averages of the estimated cross-sectional coefficients and pricing errors[16]:

$$\hat{\theta} = E_T(\hat{\theta}_t)$$

$$\hat{\alpha} = E_T(\hat{\alpha}_\tau)$$

Intuitively and as Fama and MacBeth suggest, one can use the standard deviations of the cross-sectional time series estimates to generate the sampling errors for final estimates:

$$cov(\hat{\theta}) = \frac{1}{\mathrm{T}} E_T[(\hat{\theta}_t - \hat{\theta})(\hat{\theta}_t - \hat{\theta})']$$

$$cov(\hat{\alpha}) = \frac{1}{\mathrm{T}} E_T[(\hat{\alpha}_t - \hat{\alpha})(\hat{\alpha}_t - \hat{\alpha})']$$

[16] $E_T(\cdot) = \frac{1}{T}\sum_{t=1}^{T}(\cdot)$.

It is important to emphasize here the error-in-variable problem that the FM approach suffers from. The betas used in the cross-sectional regressions as explanatory variables are in fact not fixed—they are estimates of the true unknown betas and, although consistent, they are estimated with some error. Shanken (1992) shows that this cannot be disregarded as it matters even asymptotically and proposes corrected (Shanken correction hereafter) asymptotic standard errors. The correction factor due to Shanken is:

$$Sh = (1 + \hat{\lambda}' \hat{\Sigma}_f^{-1} \hat{\lambda})$$

where $\hat{\Sigma} = E_T[(f_t - E_T(f_t))(f_t - E_T(f_t)')]$ is the sample variance-covariance matrix of the factors. The corrected errors of the estimated lambdas λ and pricing errors α are the following:

$$cov_{Sh}(\hat{\theta}) = \frac{1}{T}[(X'X)^{-1} X' \hat{\Sigma} X (X'X)^{-1} * Sh + \widetilde{\Sigma}_f]$$
$$cov_{Sh}(\hat{\alpha}) = \frac{1}{T}(I_N - X(X'X)^{-1} \hat{\Sigma}(I_N - X(X'X)^{-1} X')* Sh$$

where $\hat{\Sigma}_f$ is a matrix with a leading column and row of zeros and $\hat{\Sigma}_f$ in the lower right corner, $\widetilde{\Sigma} = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \hat{\Sigma}_f \end{bmatrix}$, and $\hat{\Sigma}_f$ is the sample variance-covariance matrix of the residuals, $\hat{\Sigma} = E_T(\varepsilon_t \varepsilon_\tau')$, where $\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} & \varepsilon_{2t} & \dots & \varepsilon_{Nt} \end{bmatrix}'$.

Jagannathan and Wang (1998) show that the FM procedure does not necessarily overstate the precision of the standard errors in the presence of conditional heteroskedasticity so I do not totally reject Fama-MacBeth (uncorrected) standard errors but I report both of them.

To evaluate the goodness of the models estimated with FM procedure I use both formal and informal criteria. The formal criterion is testing the null hypothesis weather all the pricing errors $\hat{\alpha}$ are jointly zero (the Wald test). Estimates of the pricing errors and their variance-covariance matrix allow me to calculate the statistics $\hat{\alpha}' \, cov(\hat{\alpha})^{-1} \hat{\alpha}$, which asymptotically follows the X^2_{N-L} distribution, where N is the number of tested portfolios and L is the number of factors. However I use this test with caution because the null of zero pricing errors may not be rejected, not because of small pricing errors but because of their high sampling error. What is more, for this reason, I use the test only to test one particular model and not to compare different models. Lettau and Ludvigson (2001) cite several investigations (Burnside and Eichenbaum (1996), Hansen, Heaten and Yaron (1996)) that find that the tests, which rely on the estimate of the variance-covariance matrix of pricing errors, have poor small-sample properties. This can be especially unfavourable when the number of tested portfolios is high relative to the time series sample size. So complementary to this formal criterion, I also use two additional however informal criteria: the root mean square pricing error $RMSE = \sqrt{\frac{1}{N}\hat{\alpha}'\hat{\alpha}}$ and the R^2 and adjusted R^2 of the cross-sectional regression.

The *RMSE* simply reflects how "big" is the average pricing error but, conversely to the test statistics $\hat{\alpha}'cov(\hat{\alpha})^{-1}\hat{\alpha}$, does not take into account its variance-covariance matrix. It is then a natural complement to the formal *chi square* test and it can be very useful while comparing the different asset pricing models. As for the R^2 (simple and adjusted) these are informative summary statistics, which reflect how well the model fits the data.

$$R^2 = 1 - \frac{(\overline{R}_i^e - X\hat{\theta})'(\overline{R}_i^e - X\hat{\theta})}{(\overline{R}_i^e - \ddot{R}^e)'(\overline{R}_i^e - \ddot{R}^e)}$$

$$adjR^2 = 1 - (1 - R^2)\frac{N-1}{N-L-1}$$

where $\overline{R}_i^e = E_T(R_{it}^e)$, $\ddot{R}^e = E_N(R_i^e)$.

Lewellen et al. (2007) warn not to rely on the R^2 criterion very deeply and show that high R^2 may provide a weak support for the model. Burnside (2007) points that R^2 may be not bounded between 0 and 1 unless a constant is included in the second stage of the FM procedure (which is the case in this paper) and the predicted returns include the constant. However, the probability limit of the R^2 statistic is 1, no matter if a constant is included or not. In spite of these shortcomings the R^2 is still a basis of the assessment of different models and it is widely used in empirical literature.

5.2. GMM procedure

The basic asset pricing equation for excess returns is:

$$E\left(m_t R_t^e\right) = 0 \tag{10}$$

where $R_t^e = \left[R_{1t}^e \quad R_{2t}^e \quad \dots \quad R_{Nt}^e\right]'$. m_t is a stochastic discount factor (SDF) which is linear in factors and given by[17]:

$$m_t \equiv m_t(b) = 1 + f_f'b$$

[17] As noted by Cochrane (2005) when tested assets are excess returns then the pricing equation (10) does not identify the mean of the SDF as $E\left(m_t R_t^e\right) = E\left(km_t R_t^e\right) = 0$ for any constant k. One has to then normalize one of the unknown parameters (usually a) in $m_t = a + f_t'b$. The normalization does not affect the test statistics based on pricing errors so I follow Cochrane's suggestion and set $a = 1$.

One can treat the asset pricing equation as the theoretical moment conditions $g(b) = E(u_t)$ defining:

$$u_t \equiv u_t(b) = m_t R_t^e \tag{11}$$

It is straightforward then to map the concept of asset pricing equation into standard GMM framework. GMM estimation is based on minimizing a quadratic form of the sample moment conditions of the model. The moment conditions are the pricing errors of the model and GMM naturally minimizes a linear combination of sample pricing errors. The vector of unknown parameters b of the SDF m_t is determined by solving the GMM minimization criterion:

$$\hat{b} = \arg\min_{\{b\}} g_T(b)' W g_T(b)$$

where sample moment conditions are defined as $g_T(b) = E_T(u_t) = E_T\left(m_t R_t^e\right)$ and W is some positive definite weighting matrix. The GMM theory gives also the variance-covariance matrix of the estimated parameters:

$$var(\hat{b}) = \frac{1}{T}\left(d'\mathrm{W}d\right)^{-1} d'WSWd'\left(d'\mathrm{W}d\right)^{-1} \tag{12}$$

where d is the derivative of the moment conditions with respect to the parameters: $d = \frac{\partial g_T(b)}{\partial b'}$ and S is the spectral density matrix given by $S = \sum_{j=-\infty}^{\infty} E\left(u_t u'_{t-j}\right)$ where u_t are given in equation (11), estimated as $\hat{S} = E_T(\hat{u}_t\, \hat{u}'_t)$.

The choice of the weighting matrix is still a matter of considerable debate in the literature. This choice is crucial because the weighting matrix specifies how important are particular moments or linear combination of moments in the minimization. Among different weighting matrices there are three of a particular interest: the identity matrix $W = I$, the inverse of the variance-covariance matrix of moment conditions $W = S^{-1}$ (efficient weighting matrix) and the inverse of the second-moment matrix of the returns $W = E\left(R_t R_t'\right)^{-1}$ (Hansen Jagannathan or HJ weighting matrix). In empirical literature all the three variants are used, however they may work better or worse and one has to use them with caution. It has to be stressed here that the choice of the weighting matrix does not affect the consistency of the estimated parameters but it has an impact on their efficiency. In many cases it turns out to be wise to trade some degree of the efficiency for the robustness to model misspecifications.

"If you don't know which weighting matrix to use, take the identity matrix" says the informal advice. Often this suggestion is a very good choice.[18] The identity matrix gives to each of the moment conditions the same value and all the moments are equally important in the minimization. This makes the GMM estimation with the identity matrix useful while comparing different models

[18] In the GMM literature it is common to refer to the first stage GMM when $W = I$.

as in all model the same pricing errors are assigned the same weights. The first stage estimates may give up some of their asymptotic efficiency but they are still consistent and can be more robust to statistical and economical problems. Moreover Altonji and Segal (1996) show that the first-stage GMM estimates with the identity matrix are also much more robust to small-sample problems than the GMM estimates in which the weighting matrix was estimated.

In the conventional two-stage GMM[19] Hansen (1982) advocates using the inverse of the estimated variance-covariance matrix of moment conditions as a weighting matrix. He calls it an efficient weighting matrix because the obtained estimates are asymptotically efficient. The choice of such weighting matrix is based then on statistical considerations. Efficient GMM focuses on the well-measured moments. The S^{-1} matrix gives more importance to these pricing errors, which have lower variance (low $var\left(m_t R_{it}^e\right)$), as they are estimated with higher precision. In practice it means that GMM will focus to price best such linear combinations of the portfolios that have small return variance. If among the portfolios there are a nearly-risk free portfolios then GMM will concentrate on them and their moment condition will be assigned significant weights. This property of efficient GMM makes it useless in model comparisons. In different models the efficient matrices are different and as a result they may value the pricing errors differently. In one model more importance can be assigned to price the small-size portfolios while in another, the value portfolios. Obviously one cannot compare the goodness of the models as they concentrate on different assets. One of the serious shortcomings of using efficient weighting matrix is also its poor small-sample properties. The estimate of S from the first stage is of low quality. This turns out to be even more harmful when there is a small number of time series observations relative to the cross sectional sample size. Cochrane (2005) remarks that when the number of moment conditions (cross sectional sample size) is more than around 1/10 of the number of time series observations then the S estimates tend to become unstable and nearly singular. This may drastically decrease the quality of the second stage GMM estimates and influence the Wald test as it may "improve" the χ^2 statistics not by lowering the pricing errors but by "blowing up" their variance-covariance matrix S^{-1} .

Hansen and Jagannathan (1997) propose to use an alternative weighting matrix: the inverse of the second moment matrix of the returns $W = E(R_t R_t')^{-1}$. They show that the square root of the minimum value of the GMM objective

[19] GMM is very often done in a two stage manner. The identity matrix is used in the first stage as a weighting matrix $W = I$ to minimize the objective function $g_T(\theta)'Wg_T(\theta)$. The choice of $W = I$ is recommended because it treats all the moment conditions equally. If one has some particular information on the importance of moment conditions one can reflect this in the structure of the weighting matrix. However often it is not the case and "the best" is to start with identity matrix. The first stage results in the estimator $\hat{\theta}$ (consistent but not efficient) which is the used to estimate the spectral density matrix $S = T\,\mathrm{var}\left(g(\theta)\right)$. The inverse of the spectral density matrix S^{-1} is then used as a weighting matrix (called efficient weighting matrix) in the second stage GMM. It results in the estimator $\hat{\hat{\theta}}$, which is consistent and efficient.

function with the proposed weighting matrix has some interesting economic interpretation. Suppose that $b = \arg\min_{\{b\}} g_T(b)'E(R_T R_T')^{-1} g_T(b)$. Then

$$HJd(\hat{b}) = \sqrt{g_T(\hat{b})'E(R_t R_t')^{-1} g_T(\hat{b})} \tag{13}$$

is the minimum distance between the stochastic discount factor estimated from the model as $\hat{m}_t \equiv m_t(\hat{b}) = 1 + \hat{b}' f_t$ and the space of true discount factors $\mathcal{M}$. So when $W = E(R_t R_t')^{-1}$ then the GMM objective function is the square of the distance and the GMM estimates are found by making this distance as small as possible. The HJ distance has also another interpretation: it is a maximum absolute pricing error per unit norm or a maximum mispriced Sharpe ratio for any test portfolio. Jagannathan and Wang (1996) derive the asymptotic test for the HJ distance statistic to check the hypothesis that the distance is zero. They show that the statistic $T * HJd^2$ is asymptotically distributed as a weighted sum of $N - L$ identically and independently distributed random variables each with $\chi^2(1)$ distribution, where N is the number of tested portfolios (moment conditions) and L is the number of factors (parameters to estimate). This allows to obtain the p-values of the HJ statistic through simulations.[20] The HJ distance statistic and its p-value account for additional criteria of the model assessment, next to $RMSE$. What's more HJ weighting matrix remains constant for different model specifications so the distance measure can be directly comparable across models. However, similar to the efficient GMM weighting matrix, the second moment matrix of the returns may also suffer from poor small-sample properties. It is also even more nearly singular than the spectral density matrix. However its influence on the χ^2 statistics is not so obvious now because this has to be calculated using the formula from equation (15).

Both the efficient and HJ weighting matrices have to be estimated before one can use them in the GMM optimization. Cochrane (2005) points that when an estimated weighting matrix is used then test portfolios are not in fact the original portfolios but their linear combinations. The initial portfolios are usually formed on some economically interesting characteristics such as size, book to market value or industry. However, their linear combinations, especially those that involve strong long and short positions, may lose some of these properties. In the end the GMM estimation with estimated weighting matrix will concentrate on pricing some potentially strange and artificial portfolios combined from the original ones that may not be of any economic interest. A remedy for this is to use the identity matrix as this does not influence the structure of the tested portfolios and leaves the characteristics untouched. However the efficient and HJ weighting matrices are superior to the identity matrix in another respect. The GMM estimates with

[20] The detailed prescription for how to run simulations and get the p-values is described in Jagannathan and Wang (1996) in Appendix C.

$W = E(R_t R_t')^{-1}$ or $W = S^{-1}$ are invariant to the initial choice of portfolios. But when the identity matrix is used to weight the moment conditions then this is no longer true and the results depend on the initial portfolio selection. This property addresses the critique raised by Kandel and Starmbaugh (1995) and Roll and Ross (1994).

In the view of the above discussion I conduct the GMM estimation with the identity matrix and the HJ matrix as weighting matrices. Using identity-weighting matrix gives robust estimates and does not distort the properties of the original portfolios. On the other hand using HJ weighting matrix allows to assess the model with additional criterion in the form of HJ distance and is insensitive to the choice of portfolios. Both facilitate the model comparisons.

Once the GMM parameters $\hat{b}$ are estimated one can start to evaluate the model. The main GMM based test is the χ^2 test of the over identifying restrictions. It allows to test if all the pricing errors (moment conditions) are jointly significant. The GMM J statistic is the following:

$$Jst = g_T'\left[var(g_T)\right]^{+} g_T \sim \chi^2_{N-L} \tag{14}$$

where $[]^{+}$ denotes the pseudo-inverse since the variance-covariance matrix of the moment conditions is singular with rank $N-L$. This can be estimated using the "longer"[21] formula:

$$var(g_T) = \frac{1}{T}\left(I - d'\left(d'Wd\right)^{-1} d'W\right) S \left(I - d'W\left(d'Wd\right)^{-1} d\right)' \tag{15}$$

I use the J statistic only to evaluate a particular model but I do not compare the statistics between the models. For comparisons, I use the root mean square pricing error $RMSE = \sqrt{\frac{1}{N} g_T(\hat{b})' g_T(\hat{b})}$. Moreover when using HJ weighting matrix I calculate for each model the HJd statistics and their p-values and I also compare them across models.

Some comments have to be made here concerning the equivalence between the beta pricing models and the linear models for the discount factor. In case of excess returns and normalized SDF this equivalence is the following:

$$m_t = 1 + f_t' b \Leftrightarrow E(R_t^e) = \beta\lambda \tag{16}$$

$$\beta = cov(R_t^e f_t') var(f_t f_t')^{-1} \tag{17}$$

[21] The "longer" formula is valid for any positive definite weighting matrix. When the efficient weighting matrix is used then this formula simplifies to $var(g_T) = \frac{1}{T}(S - d(d'S^{-1}d)^{-1}d')$ and the J statistic is $Jst = Tg_T'S^{-1}g_T$.

$$\lambda = -\frac{var(f_t f_t')b}{E(m_t)} \tag{18}$$

It is invariant to the normalization of the SDF. Another subtle difference is between the pricing errors coming from the FM regression and GMM estimation. In FM procedure what is minimized is the square of the pricing errors α and in GMM estimation—the weighted square of the pricing errors[22] g_T. The FM pricing errors and the GMM pricing errors are not the same. The relation between them is:

$$g_T = \alpha E(m_t) \tag{19}$$

The FM and GMM procedure give exactly the same results only when the identity matrix is used in the GMM estimation and the normalization is such that $E(m_t) = 1$. Otherwise the two procedures may give different results. In the empirical asset pricing literature it is common to use GMM and then report also lambdas λ and their standard errors calculated by delta method. However one has to bear in mind that in GMM estimation it is the g_T that is minimized, not α. So while the model estimated by GMM may exhibit an excellent fit in terms of GMM pricing errors α it may also give a very poor fit in terms of FM pricing errors g_T. And vice versa: model estimated by FM regression may work well in terms of FM pricing errors α and it may be of poor quality in terms of GMM pricing errors g_T. This also motivates the use of both FM and GMM estimation techniques.

6. Results and Comments

I present and evaluate the results of the estimations in two respects: if including human capital and housing improves empirical performance of the model and if introducing dynamics in the form of extracted common factor as an instrument matters. Additionally I compare obtained outcomes with those of Fama French three-factor model. Fama and French specify three factors that affect assets' excess returns: excess marker return $R_{t+1}^{vw} - R_{t+1}^{f}$, size premium R_{t+1}^{smb} and growth premium R_{t+1}^{hml}. These factors proxy for unobserved macroeconomic risks, however it is still not clear which risks. Nevertheless, the model was proved to be very successful in practice and it is common in empirical finance literature to take it as a benchmark.

[22] I call pricing errors both α and g_T when it is clear from the context which one is of the interest. When it is not the case I use the name FM pricing errors and GMM pricing errors to differentiate between them.

6.1. Estimation Method: Fama MacBeth regressions

Tables 4 and 5 show the results of the Fama MacBeth regressions. We can easily observe that just including human capital and real estate improves unconditional pricing of simple CAPM model. R^2 and R^2adj increase significantly for both groups of test assets. Augmented model explains 52% of variation in expected excess returns of 25 Fama French portfolios while simple CAPM only 9%. For 16 mixed portfolios the improvement is even stronger: augmented model explains 65% of variation in their expected excess returns while simple CAPM only 0.2%. In Tables in the Appendix we can also notice that for both groups of test assets the loadings λ_{HS} on real estate component of total wealth are positive and significant. This means that the housing part of wealth is important in pricing assets and carries significant and positive risk premium. The coefficient on human capital λ_{HC} is negative but not significantly in opposite to the same coefficient on value-weighted portfolio, which is negative but significantly. This effect is quite common in empirical literature on asset pricing when 25 Fama French portfolios are used as test assets. In both simple CAPM and Fama French three factor model the risk premium on market portfolio are as well significantly negative.

Table 4.

Fama MacBeth regressions 25 FF portfolios: comparison

	CAPM	vw-hc-hs	$\hat{f}_1$	$\hat{f}_2$	FF
RMSE	0.73	0.53	0.31	0.38	0.35
R^2	0.09	0.52	0.83	0.75	0.79
R^2adj	0.05	0.45	0.76	0.65	0.76
χ^2stat	75.82	58.87	47.23	43.64	62.27
p-value	0.000	0.000	0.000	0.000	0.000
Sh χ^2stat	74.54	28.33	10.82	12.43	50.57
p-value	0.000	0.131	0.860	0.770	0.000
Sh	1.02	*2.08*	4.36	3.51	1.23

It is interesting also to look at χ^2 statistic. For 25 Fama French portfolios its value is high 58.87, indicating that augmented model produces statistically high pricing errors (significantly different from zero). However, when we account for the fact that betas used in the second stage were estimated and calculate corrected χ^2 statistic it is then lower enough to be insignificant indicating insignificant pricing errors. Nevertheless I have to be cautious here because there are two forces that lower χ^2 statistic: lower pricing errors or their higher variances. For augmented model the pricing errors are indeed lower that for simple CAPM but their corrected variance-covariance matrix $var(\alpha)$ is blown up by the Shanken factor 2.08 making χ^2 statistic even lower (and in consequence we do not reject that pricing errors are different from zero). And the Shanken factor for simple CAPM is 1.02. Why looking only at corrected χ^2 statistic can be misleading, we can see clearly by comparison of the model augmented with human capital and housing with Fama French three factor model. The FF model has lower *RMSE* that the augmented model but according to corrected χ^2 statistic still produces significant pricing errors. This is because its $var(\alpha)$ is not increased by the Shanken factor, which for Fama French model is only 1.23. For augmented model the Shanken correction significantly decreases the precision of estimated variance-covariance matrix of pricing errors and in consequence the χ^2 statistic is low. On the other hand when 16 mixed portfolios are tested, the augmented model is not rejected nor by uncorrected neither by corrected χ^2 test while simple CAPM is rejected at 5% (although not at 1%) by both tests. Bai (2007) makes some interesting comments on the Shanken correction factor. She points that the correction is large for models that include scaled macroeconomic factors, which are more persistent than the mimicking portfolio factors. This observation is also quite common in asset pricing literature.

Table 5.

Fama MacBeth regressions for 16 mixed portfolios: comparison

	CAPM	vw-hc-hs	$\hat{f}_1$	$\hat{f}_2$	FF
RMSE	0.37	0.22	0.19	0.17	0.25
R^2	0.002	0.65	0.71	0.77	0.56
R^2adj	–0.07	0.56	0.46	0.57	0.45
χ^2stat	24.23	12.90	12.73	12.81	16.80
p-value	0.043	0.38	0.12	0.12	0.16
Sh χ^2stat	24.23	9.02	7.61	5.17	15.80
p-value	0.043	0.70	0.47	0.74	0.20
Sh	1.00	1.43	1.67	2.48	1.06

Another conclusion that we can draw from Tables 4 and 5 is that including common factors $\hat{f}_1$ and $\hat{f}_2$ as instruments in conditional pricing model also improves the empirical performance of the model. R^2 statistics increase significantly when instruments are included in the pricing equation. The conditional model with $\hat{f}_1$ as a conditioning variable explains 83% of variation in expected excess returns of 25 Fama French portfolios while augmented CAPM 52% and Fama French three factor model 79% (with $\hat{f}_2$ as a conditioning variable this fraction is lower 75% but still higher than that of augmented CAPM). For 16 mixed portfolios the improvement is not so strong: the model with $\hat{f}_1$ as a conditioning variable explains 71% of variation in their expected excess returns while augmented CAPM 65% and Fama French three factor model 56% (with $\hat{f}_2$ as a conditioning variable this fraction is slightly higher 77%). In terms of R^2adj conditional models with $\hat{f}_1$ and $\hat{f}_2$ are better than augmented CAPM (R^2adj are 76% and 65% for conditional models and 45% for unconditional augmented CAPM) but not necessarily better that Fama French model (for which R^2adj is 76%) when 25 Fama French models are tested. For 16 mixed portfolios the conclusions are different: both conditional models are better that Fama French three factor model (46% and 57% vs. 45%) but only $\hat{f}_2$ as a conditioning variable improves unconditional pricing (R^2adj for augmented CAPM is 56%).

For 25 Fama French portfolios uncorrected χ^2 test rejects both conditional models as well as Fama French three factor model as producing too large pricing errors but corrected χ^2 test approves conditional models with factors $\hat{f}_1$ and $\hat{f}_2$ as instruments. Again the reasons are probably not lower pricing errors (*RMSE* for conditional models are only slightly lower than for other models) but high Shanken factors (4.36 for $\hat{f}_1$ and 3.51 for $\hat{f}_2$), which blow up the variance covariance matrix of pricing errors. For 16 mixed portfolios nor corrected neither uncorrected χ^2 test rejects conditional pricing models with $\hat{f}_1$ and $\hat{f}_2$ as carrying conditional information.

Additional comment should be also made on the constant in simple CAPM, CAPM augmented with human capital and housing and Fama French three factor model. For these three models the constant is positive, significant and plays a crucial role in the model's fit. As it was noted earlier the theory says that the constant or zero-beta excess return should be zero. This phenomenon is also not so rare in empirical applications and it has already been observed many times. The constant indicates by how much the model misprices the risk free rate. For example when the FF25 portfolios are tested the constant in the Fama French 3 factor model is 4,91%, which means that the model misprices the risk free rate by 4,91% quarterly and this is a lot. Burnside (2007) suggests that a measurement error in the estimated betas and resulting downward bias in the estimated lambdas as well as a liquidity premium in T-bills can account for the explanation of a positive pricing error for the risk free rate. Nevertheless it is not common in the literature to reject the model because of significant constant. Nevertheless the asset pricing literature does not concentrate on the explanation of the high significance of the constant and this may be an interesting topic to cover in the future research.

In the Appendix I also visualize observed excess returns in comparison with excess returns predicted by different models. Figures 4–8 depict realized vs. predicted expected returns of 25 Fama French portfolios and Figure 9–13 do the same for 16 mixed portfolios. In general all the figures confirm previous conclusions.

6.2. Estimation Method: GMM, $\mathbf{W} = \boldsymbol{I}$ and $\mathbf{W} = \left(\boldsymbol{ERR}\right)^{-1}$

Tables 6 and 7 show the results of the GMM estimation for two different weighting matrices: the identity matrix I_N and the inverse of second moments of excess returns matrix $E(R^e R^{e\prime})$. It should be pointed here that the general conclusions remain unaffected by the type of weighting matrix used in the estimation.

When comparing simple CAPM model with this augmented with human capital and housing we can observe that the augmented model always produces significantly lower *RMSE* and *HJd*. Moreover almost in all the cases apart from one (when test portfolios are 25 Fama French portfolios and $E(R^e R^{e\prime})$ is used as a weighting matrix) the χ^2 test indicates that the pricing errors are not statistically different from zero and the augmented model cannot be rejected. When $W = E(R^e R^{e\prime})$ is used, additional test of HJ distance show that HJ distance is indeed statistically small (p-value of the HJ statistics is 0.34 for 25 Fama French portfolios and 0.71 for 16 mixed portfolios) and again proving that the augmented model cannot be rejected. All these findings confirm the general conclusion that extending the return on total wealth portfolio so that it includes human capital and housing as I do in this paper indeed influences the empirical performance of the simple CAPM model. In fact the augmented model is also pretty competitive to Fama French three factor model. In almost all the cases it has lower *RMSE* and *HJd* (only in case when test portfolios are 25 Fama French portfolios and $E(R^e R^{e\prime})$ is used as a weighting matrix the *RMSE* of Fama French three factor model 0.38 and is lower that *RMSE* of augmented CAPM 0.50). Moreover for 25 Fama French portfolios the Fama French model is rejected by both χ^2 test and *HJ* test as producing to high pricing errors and too high HJ distance.

Table 6.

GMM estimation for 25 Fama French portfolios: comparison

	Model	CAPM	vw-hc-hs	$\hat{f}_1$	$\hat{f}_1$	FF
W = 1	RMSE	1.01	0.15	0.12	0.13	0.38
	χ^2stat	85.67	30.22	21.58	15.18	65.04
	p-value	0.00	0.11	0.25	0.65	0.00
	HJd	5.02	0.75	0.61	0.65	1.85
$W = E(RR')^{-1}$	RMSE	1.14	0.50	0.46	0.47	0.39
	χ^2stat	88.94	41.96	30.35	26.84	64.87
	p-value	0.00	0.006	0.03	0.08	0.00
	HJd	0.66	0.20	0.19	0.20	0.55
	HJstat	55.29	5.18	4.56	4.83	39.19
	p-value	0.00	0.34	0.39	0.28	0.003

An interesting observation is that when GMM estimation technique is used, the superiority of the conditional models over unconditional is not so clear as it is in case of Fama MacBeth regressions. The conditional models with estimated common factors $\hat{f}_1$ and $\hat{f}_2$ as instruments always produce lower *RMSE* and *HJd* with comparison to augmented CAPM model. But the differences in the values of these measures are not high (for 16 mixed portfolios as test assets and $W = E(R^e R^{e\prime})$ the *RMSE* are actually the same 0.09). The conditional models are never rejected in terms of χ^2 test and *HJd*: they are proved to generate statistically small pricing errors and HJ distances for both groups of test assets and both weighting matrices. For 25 Fama French portfolios the conditional models are also superior to Fama French model in terms of χ^2 test and *HJd* as the Fama French model is rejected by both tests as producing to high pricing errors and HJ distance. For 16 mixed portfolios none of these models is rejected by the two formal criteria but the conditional models are characterized by lower *RMSE* and *HJd* than Fama French three factor model.

Table 7.

GMM estimation for 16 mixed portfolios: comparison

	Model	CAPM	vw-hc-hs	$\hat{f}_1$	$\hat{f}_2$	FF
W = 1	RMSE	0.58	0.09	0.08	0.07	0.38
	χ^2stat	27.27	10.74	4.61	5.50	17.57
	p-value	0.04	0.63	0.86	0.79	0.17
	HJd	2.34	0.38	0.32	0.28	1.52
$W = E(RR')^{-1}$	RMSE	0.63	0.09	0.09	0.09	0.42
	χ^2stat	26.57	10.57	4.65	9.35	18.44
	p-value	0.03	0.65	0.86	0.41	0.14
	HJd	0.42	0.12	0.09	0.11	0.31
	HJstat	22.76	1.72	1.04	1.64	12.51
	p-value	0.06	0.71	0.76	0.44	0.27

6.3. Comments

Presented in this section results confirm two main hypotheses of this paper: one that extending market portfolio by human capital and housing indeed matters in asset pricing model and the second that introducing dynamics in the form of extracted common factor as an instrument also improves empirical performance of asset pricing models. Both types of improvements are proved to be of importance in practical applications. Moreover the proposed conditional models can be viewed as competitive to the Fama French three factor model as they were shown to work better.

7. Conclusions

There are two main difficulties in examining the empirical support for the classical static CAPM model. One of them is the fact that the total wealth portfolio is not observable and the second that the real world is dynamic and not static. In practical applications it is quite often assumed that the return on the total wealth portfolio can be well proxied by the return on broadly--defined stock market index. Moreover while dealing with the cross-section of assets returns, it is also convenient to regard risk premium as time-invariant.

I this paper I try to deal with both weaknesses. I extend the proxy for the return on market portfolio so that, next to the return on stock market index it also includes the return on human capital and the return on housing as these two types of wealth constitute the significant portion of the total wealth. I follow Jagannathan and Wang (1996) and proxy the return on human capital by the growth rate in per capita labour income. As a representative of the return on housing I use the change in the OFHEO House Price Index, which, to the best of my knowledge, has not been yet used in empirical literature on asset pricing. Moreover by introducing dynamics into the asset pricing model I make risk premium time-variant, which more reasonably reflects the reality. An important novelty is the fact that I represent the conditioning information by the common factors which I estimate using dynamic factor methodology. In this way I explore a much richer set of instruments, which is more likely to span the unobservable information set of financial market participants.

Obtained results prove that real estate are important in empirical asset pricing. They confirm as well that more accurate proxy for the return on total wealth portfolio indeed matters. The CAPM augmented with human capital and housing works significantly better than simple CAPM. Moreover when dynamics is taken into account this augmented model gives even better results—it can explain around 80% of the variation in the cross-section of excess returns on 25 Fama French portfolios. For different test assets the conditional model is also superior to the Fama French model. I show then that it is important to use a good representative of the conditional information set of the agents as it can significantly improve the empirical performance of the asset pricing models.

References

Altonji, J. G. and L. M. Segal (1996). "Small-sample bias in gmm estimation of covariance structures". *Journal of Business and Economic Statistics* 14 (July), 353–366.

Bai, J. (2007). "Predictive macro indices as asset pricing risk factors". *Working paper. Graduate School of Business*, University of Chicago.

Black, F. (1972, July). "Capital market equilibrium with restricted borrowing. *Journal of Business* 4 5, 444–55.

Burnside, A. C. (2007). "Empirical asset pricing and statistical power in the presence of weak risk factors". *NBER Working Paper* (13357).

Burnside, A. C. and M. Eichenbaum (1996). "Small-sample properties of gmmbased wald tests". *Journal of Business and Economic Statistics* 14 (July), 294–308.

Campbell, J. Y. (1991, March). "A variance decomposition for stock returns". *Economic Journal* 101, 157–79.

Campbell, J. Y. and R. J. Shiller (1988). "The dividend-price ratio and expectations of future dividends and discount factors". *Review of Financial Studies* 1, 195–228.

Case, K. E., E. L. Glaeser, and J. A. Parker (2000). "Real estate and the macroeconomy". *Brookings Papers on Economic Activity* 2000 (2), 119–162.

Cochrane, J. (2005). *Asset Pricing*. Princeton University Press: Princeton NY.

Fama, E. and K. French (1988, October). "Dividend yields and expected stock returns". *Journal of Financial Economics* 22,3–25.

Fama, E. and K. French (1989). "Business conditions and expected returns on stock and bonds". *Journal of Financial Economics* 25,23–49.

Fama, E. and J. MacBeth (1973). "Risk, return, and equilibrium: Empirical tests". *Journal of Political Economics* 81 (June–July), 607–636.

Fama, E. and G. W. Schwert (1977). "Human capital and capital market equilibrium. *Journal of Financial Economics* 4,95–125.

Ferson, W. E. and C. R. Harvey (1999, August). "Conditioning variables and the cross section of stock returns". *Journal of Finance* 54, 1325–60.

Hansen, L. P. (1982). "Large sample properties of generalized method of moments estimators". *Econometrica* 50 (July), 1029–1054.

Hansen, L. P., J. Heaton, and A. Yaron (1996). "Finite-sample properties of some alternative gmm estimators". *Journal of Business and Economic Statistics* 14 (July), 262–280.

Hansen, L. P. and R. Jagannathan (1997). "Assessing specification errors in stochastic discount factor models". *Journal of Finance* 52 (June), 557–590.

Hansen, L. P. and S. F. Richard (1987). "The role of conditioning information in deducting testable restrictions implied by dynamic asset pricing models". *Econometrica* 55, 587–614.

Heaton, J. and D. Lucas (2000). "Portfolio choice and asset prices: The importance of entrepreneurial risk". *Journal of Finance* 55, 1163–1198.

Hodrick, R. J. (1992). "Dividend yields and expected stock returns: Alternative procedures for inference and measurement". *Review Financial Studies* 5 (3), 357–86.

Jagannathan, R. and Z. Wang (1996). "The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51 (March), 3–53.

Jagannathan, R. and Z. Wang (1998). "An asymptotic theory for estimating beta-pricing models using cross-sectional regression". *Journal of Finance* 53 (August), 1285–1309.

Jushan, B. and S. Ng (2002). "Determining the number of factors in approximate factors models". *Econometrica* 10 (1), 191–221.

Kandel, S. and R. F. Stambaugh (1995). "Portfolio inefficiency and the cross-section of expected returns". *Journal of Finance* 50 (March), 157–184.

Kullmann, C. (2003). "Real estate and its role in asset pricing". *Unpublished Working Paper* University of British Columbia, Vancouver.

Lamont, O. (1998, October). "Earnings and expected returns". *Journal of Finance* 53, 1563–87.

Lettau, M. and S. Ludvigson (2001). "Resurrecting conditional (C)CAPM: A cross-sectional test when risk premium are time-varying". *Journal of Political Economics* 109, 1238–1287.

Lewellen, J., S. Nagel, and J. Shanken (2007). "A skeptical appraisal of asset-pricing tests". *NBER Working Paper*.

Lintner, J. (1965, December). "Security prices, risk, and maximal gains from diversification". Journal of Finance 20, 587–615.
Ludvigson, S. and S. Ng (2007a). "Macro factors in bond risk premium. *Working Paper*, New York University.
Ludvigson, S. C. and S. Ng (2007b). "The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics* 83, 171–222.
Mayers, D. (1972). "Nonmarketable assets and capital market equilibrium under uncertainty". *Studies in the Theory of Capital Markets*. Ed. Jensen, Praeger, 223–248.
Roll, R. (1977). "A critique of the asset pricing theory's tests: Part l. *Journal of Financial Economics* 4, 129–176.
Roll, R. and S.A. Robs (1994). "On the cross-sectional relation between expected returns and betas". *Journal of Finance* 49 (March), 101–121.
Santos, T. and P. Veronesi (2006). "Labor income and predictable stock returns". *The Review of Financial Studies* 19 (1), 1–44.
Shanken, J. (1992). "On the estimation of beta pricing models". *Review of Financial Studies* 5, 1–34.
Sharpe, W. F. (1964, September). "Capital asset prices: A theory of market equilibrium under conditions of risk". *Journal of Finance* 19, 425–42.
Shiller, R. J. (1984). "Stock prices and social dynamics". *Brookings Papers Economic Activity* (2), 457–98.
Stock, J. H. and M. W. Watson (2002a). "Forecasting using principal components from a large number of predictors". *Journal of the American Statistical Association* 97, 1167–1179.
Stock, J. H. and M. W. Watson (2002b). "Macroeconomic forecasting using diffusion indexes". *Journal of Business and Economic Statistics* 20 (2), 147–162.
Stock, J. H. and M. W. Watson (2004). "Forecasting with many predictors". *Unpublished Paper*, Princeton University.

Appendix A. Tables

*—significant at 10% significance level
**—significant at 5% significance level
***—significant at 1% significance level

A.1. Fama MacBeth regressions; Test Assets: 25 Fama French portfolios

Table A.1.1.

CAPM model, in %

CAPM	const	rm-rf
λ-lambda	3.96	–1.07
t-stat	3.69***	–0.86
Sht-stat	3.66***	–0.86

Table A.1.2.

vw-hc-hs model, in %

CAPM	coast	vw	hc	hs
λ-lambda	5.12	–3.10	–0.07	0.89
t-stat	4.59***	–2.43**	–0.29	2.59***
Sht-stat	3.18***	–1.84**	–0.21	1.83**

Table A.1.3.

Conditional CAPM model with cond. variable $cv = \hat{f}_1$

CAPM	coast	$\hat{f}_1$	vw	$\hat{f}_1 *$vw	hc	$\hat{f}_1 *$hc	hs	$\hat{f}_1 *$hs
λ-lambda	3.21	0.27	–1.38	–2.94	–0.30	–0.69	–0.51	0.20
t-stat	2.87***	0.80	–1.05	–1.15	–1.43**	–1.02	–1.86**	0.30
Sht-stat	1.38*	0.39	–0.57	0.83	–0.72	–0.50	–0.93	0.15

Table A.1.4.

Conditional CAPM model with cond. variable $cv = \hat{f}_2$

CAPM	const	$\hat{f}_2$	vw	$\hat{f}_2 *$vw	hc	$\hat{f}_2 *$hc	hs	$\hat{f}_2 *$hs
λ-lambda	4.86	–0.31	–2.91	5.81	–0.15	–1.17	0.61	–0.92
t-stat	4.64***	–1.15	–2.33**	2.18**	–0.69	–1.95**	1.82**	–2.10**
Sht-stat	2.47**	–0.64	–1.43*	1.21	–0.39	–1.08	0.99	–1.17

Table A.1.5.

Fama French three-factor model, in %

CAPM	coast	rm-rf	smb	hml
λ-lambda	4.91	–2.84	0.74	1.40
t-stat	4.27***	–2.12**	1.59*	2.54***
Sht-stat	3.85***	–1.97**	1.58*	2.53***

A.2. Fama MacBeth regressions; Test Assets: 3 size portfolios, 3 book-to-market value portfolios and 10 industry portfolios

Table A.2.1: CAPM model, in %

CAPM	coast	rm-rf
λ-lambda	2.20	0.07
t-stat	2.64***	0.07
Sht-stat	2.64***	0.07

Table A.2.2.

rm-hc-hs model, in %

CAPM	const	vw	hc	hs
λ-lambda	1.59	0.76	0.38	0.54
t-stat	1.86**	0.64	1.37*	2.23**
Sht-stat	1.55*	0.57	1.15	1.90**

Table A.2.3.

Conditional CAPM model with cond. variable $cv = \hat{f}_1$

CAPM	coast	$\hat{f}_1$	vw	$\hat{f}_1 * $vw	hc	$\hat{f}_2 * $hc	hs	$\hat{f}_2 * $hs
λ-lambda	2.15	–0.07	0.09	2.53	0.50	–0.30	0.48	0.02
t-stat	1.58*	–0.18	0.06	0.46	1.48*	–0.32	1.93**	–0.01
Sht-stat	1.22	–0.14	0.05	0.36	1.16	–0.25	1.53*	–0.01

Table A.2.4.

Conditional CAPM model with cond. variable $cv = \hat{f}_2$

CAPM	coast	$\hat{f}_2$	vw	$\hat{f}_2 * $vw	hc	$\hat{f}_2 * $hc	hs	$\hat{f}_2 * $hs
λ-lambda	1.41	–0.36	0.82	1.36	0.32	–0.28	0.79	–0.42
t-stat	1.79*	–0.66	0.74	0.47	1.10	–0.34	2.43**	–0.47
Sht-stat	1.14	–0.42	0.55	0.30	0.71	–0.22	1.58*	–0.30

Table A.2.5.

Fama French three-factor model, in %

CAPM	coast	rm-rf	smb	hml
λ-lambda	3.18	–1.03	0.72	0.36
t-stat	2.52**	–0.70	1.45*	0.60
Sht-stat	2.44**	–0.69	1.44*	0.60

Appendix B. Figures

Each figure consists of two graphs, which show real vs. predicted expected excess returns. The graphs on the left show 25 Fama French portfolios sorted jointly on size and book-to-market. Assets are denoted as ij for $i, j = 1, \ldots, 5$, where $i = 1$ for the smallest companies and $i = 5$ for the biggest companies and $j = 1$ for growth firms (low book-to-market values) and $j = 5$ for value firms (high book-to-market values). The graphs on the right show 3 size portfolios ($s1$, $s2$, $s3$), 3 book-to-market portfolios ($b1$, $b2$, $b3$) and 10 industry portfolios. The abbreviations for the industry portfolios are the following: *Nd*—Nondurables, *Dr*—Durables, *Mn*—Manufacturings, *En*—Energy, *HT*—HighTech, *Tl*—Telecomunication, *Sh*—Shops, *Hl*—Health, *Ut*—Utilities, *Ot*—Others.

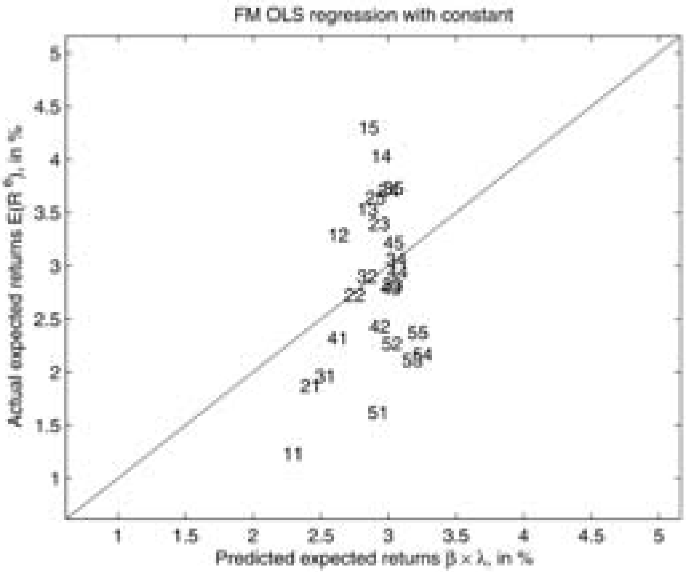

Figure 4.

Simple CAPM, 25 Fama French portfolios

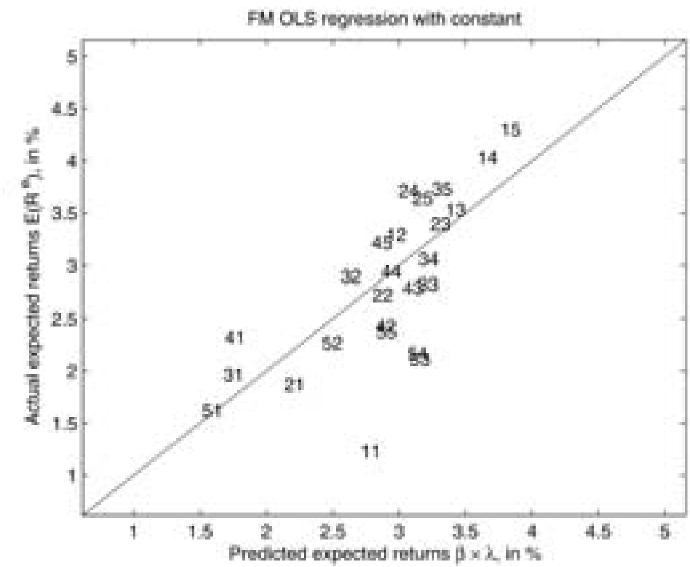

Figure 5.

Unconditional CAPM augmented with human capital and housing, 25 Fama French portfolios

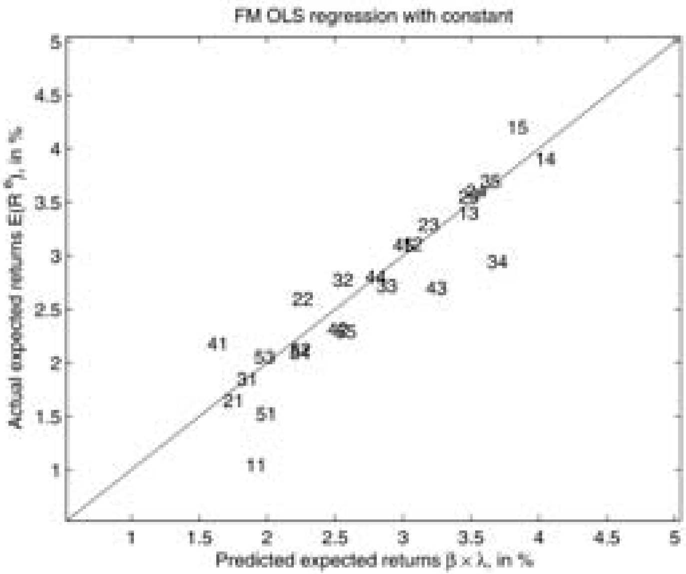

Figure 6.

Conditional CAPM with human capital and housing, instrument $z_t = \hat{f}_{1t}$, 25 Fama French portfolios.

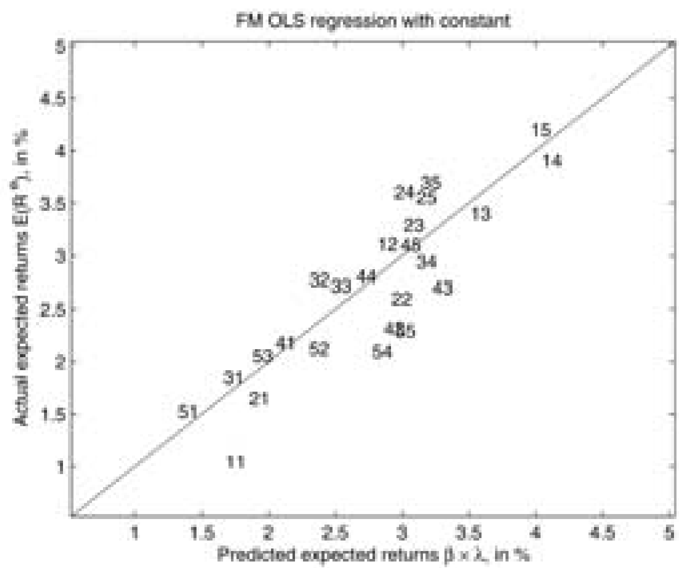

Figure 7.

Conditional CAPM with human capital and housing, instrument $z_t = \hat{f}_{2t}$, 25 Fama French portfolios

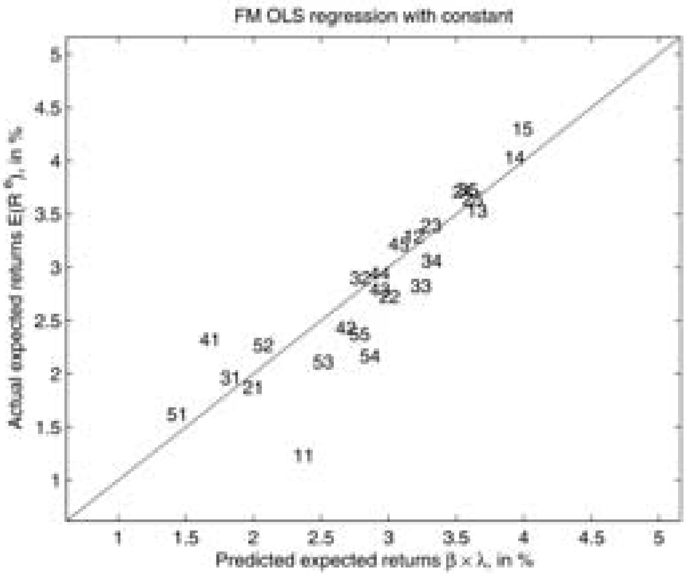

Figure 8.

Fama French three factor model, 25 Fama French portfolios.

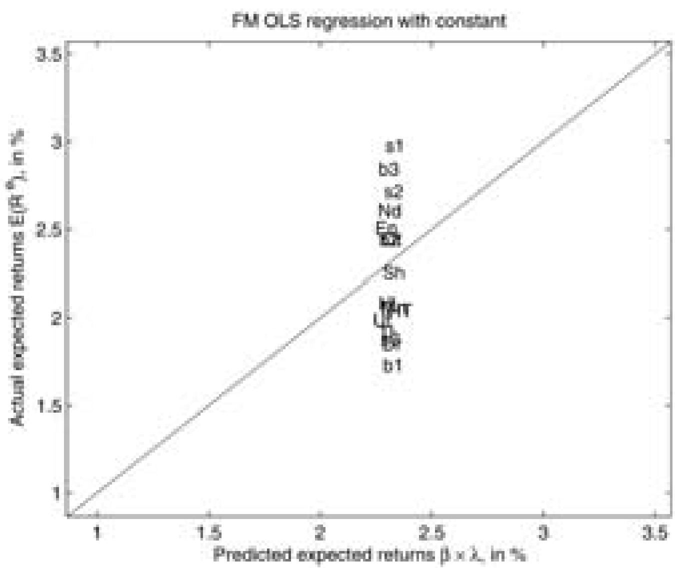

Figure 9.

Simple CAPM, 16 mixed portfolios

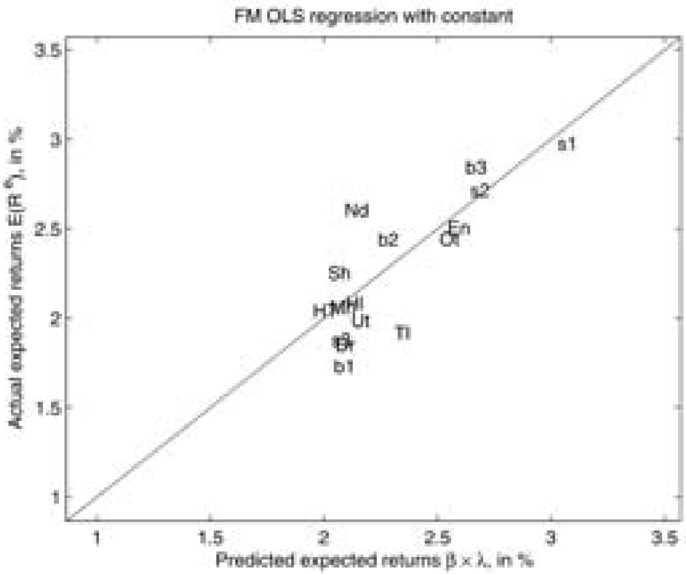

Figure 10.

Unconditional CAPM with human capital and housing, 16 mixed portfolios

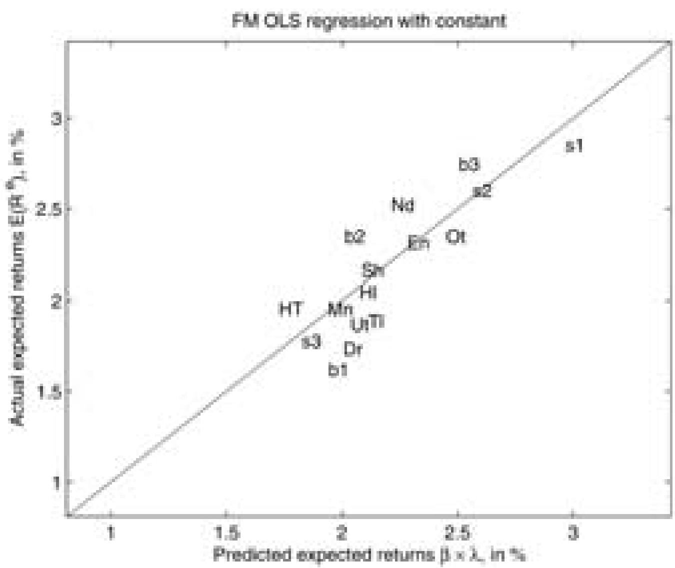

Figure 11.

Conditional CAPM with human capital and housing, instrument $z_t = \hat{f}_{1t}$, 16 mixed portfolios

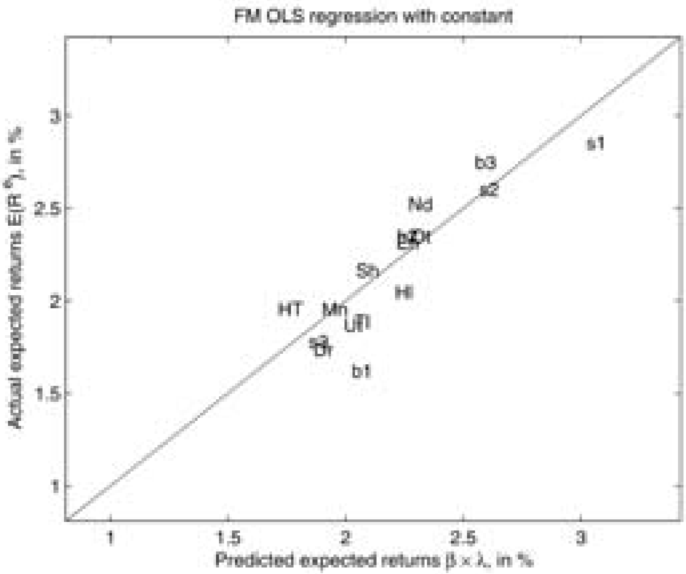

Figure 12.
Conditional CAPM with human capital and housing, instrument $z_t = \hat{f}_{2t}$, 16 mixed portfolios

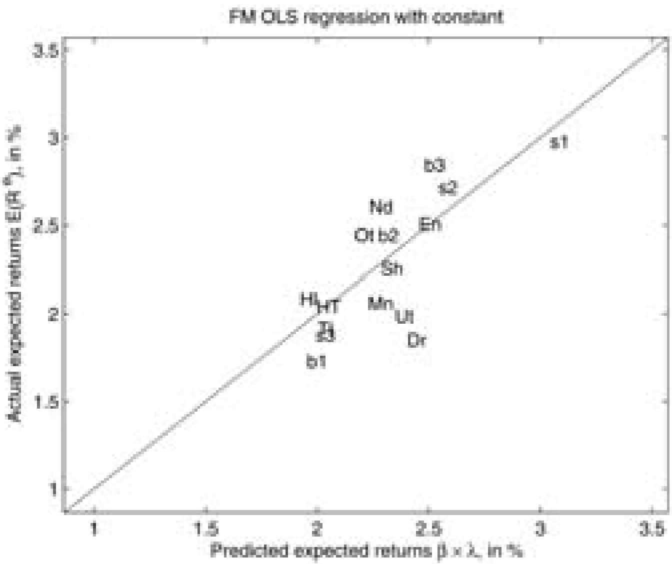

Figure 13.
Fama French three factor model, 16 mixed portfolios

Abstract

Conditional Tests of Factor Augmented Asset Pricing Models with Human Capital and Housing: Some New Results

In this paper I develop the asset pricing model in which the wealth portfolio is enriched with human capital and housing capital. These two types of capital account for a significant portion of the total wealth. Additionally I introduce dynamics into the model and represent conditioning information by common factors estimated with dynamic factor methodology. In this way I can use more accurate representative of the unobservable information set of the investors. Obtained results prove that indeed better proxy for market return matters. Moreover conditional models show promising empirical performance and often price the cross-section of excess equity returns better than the Fama French three factor model.