A Protocol for Factor Identification

by

Kuntara Pukthuanthong
University of Missouri
&
Richard Roll
UCLA & Caltech

The Q-Group
Phoenix, October 10, 2016

Evidence about Multiple Factors

- From 2001 through 2010, the monthly total return on the S&P 500 had an annualized standard deviation of ***16.3%***. The average volatility for the S&P's constituents was ***36.1%***

- Monthly total return correlation between the S&P 500 and Barclay's Bond Aggregate Index ***-0.0426***. The return correlations between these two indexes and the Goldman Sachs Commodity index were ***0.266*** and ***0.0113***, respectively

Implications

- First empirical fact implies the existence of ***common*** underlying systematic influences, (or “risk drivers” or “factors”) that limit diversification within an asset class; otherwise diversified portfolios would have much smaller volatilities
- Second fact intimates the presence of ***multiple*** systematic factors; otherwise diversified portfolios would be more correlated across asset classes, countries, and sectors

General Agreement About

- Multiple factors drive asset returns
- Their identities remain subject to wide debate
- We suggest a protocol for their identification
- The goal here is normative: to come up with an acceptable procedure whenever a new factor candidate is nominated

Principles that Apply

- Factor movements should not be easily predictable
- A characteristic such as firm size, or anything else known in advance, cannot be a factor
- However, characteristics can be related to mean returns either
  - because they happen to align with factor loadings
  - or because they represent arbitrage opportunities

Principles that Apply (Cont.)

- Pervasive factors, those with risk premiums
  - Generate the covariances among returns on real assets, those held in the aggregate market portfolio, whose risk cannot be eliminated
  - Do not necessarily generate covariances of any assets in zero net supply, such as bonds and derivatives

Real Assets only need to be explained by factors

- 90% of previous research uses equities
  - But leveraged equities are hybrids
  - Use unlevered firms, real estate, commodities, human capital
- No need to explain derivatives or bonds
  - But bond or derivative portfolios could serve as proxies for factors
- Non-linear factor models are mainly about assets held in zero net supply

Factor models based on characteristics

- E.g., Size, book/market, momentum, beta, modified duration, beta estimation risk, even the letter beginning the firm's name
- These might happen to be associated with factor loadings
  - Long/short portfolio might proxy for underlying factor
  - It is unidentified

Example: Book/Market

- Fama and French (1992)
  - Portfolio long higher book/market stocks and short lower book/market stocks
  - Proxy some underlying factors; much subsequent debate about their identity
- Daniel and Titman (1997)
  - Book/market is a characteristic, not a factor
  - Stocks with higher book/market ratios, but not higher loadings on the Fama/French portfolio, had higher returns on average,…but this doesn't prove they are not related to a different underlying factor

Priced characteristics

- Alluring to investors (if they are not risks)
- Goyal (2011) remarks that "…it is especially easy to check for pricing of characteristics…" (Fama/ MacBeth method)
- Subrahmanyam (2010) identifies more than 50 cross-sectional characteristics that have been used to predict returns in the finance literature
- Harvey, et al., 2016, identify 316 factor candidates of all types

Factors and the (possibly conditional) covariance matrix

- If the world were very simple, there would be one sure-fire method to extract good factor proxies
  - Linear combinations of factors would reveal themselves in the principal components analysis (PCA) of the covariance matrix of observed returns on <u>real</u> assets
  - In a really perfect world, there would be only a few principle components with large eigenvalues (a lot fewer than assets)
  - If this were true, factor proxies could be extracted from rather small subsets of assets

A necessary condition

- A factor candidate must be related to the principal components of the covariance matrix of returns
- Moskowitz (2003) checks three candidates, size, book/market, and momentum
  - Size is
  - Book/market might be
  - Momentum is not

Difficulties in factor extraction from the covariance matrix

- It produces only estimate for linear combinations of the true underlying factors, not the factors themselves
- It's compromised by non-stationarity since there is no plausible reason why the number of factors or their relative importance should be constant through time
- It extracts true risk drivers, pervasive non-diversifiable factors (or linear combinations thereof) along with diversifiable factors, such as industry factors, that are not associated with risk premiums
- Our protocol suggests a cure for each difficulty

Perceptions, not actual macro values

- Underlying drivers ***cannot*** be the infrequently-published official numbers about macro-economic variables because market prices move around much too rapidly
- Perceptions could include
  - rational anticipations of change in macro conditions that are truly pervasive such as real output growth, real interest rates, inflation, energy, etc., and
  - behavior-driven pervasive shocks in confidence or risk perceptions such as panics, liquidity crises, etc.

Previous factor measurement methods

- Statistical, such as principle components or factor analysis, (e.g., Roll and Ross [1980], Connor and Korajczyk [1988])
- Macro-economic variables that seem likely to be pervasive, pre-whitened, industrial production, inflation, and so on, (e.g., Chen, Roll and Ross [1986])
- Characteristics that are empirically related to average returns (e.g., Fama/French [1992], Carhart [1997]
- A lesser known but simpler approach: a handful of rather heterogeneous indexes or tradable portfolios

Pluses and Minuses of factor measurement methods

- Statistical: sound theoretically but stationarity required
- Macro-economic variables: the best theoretically but low frequency
- Characteristics: accommodate non-stationarity but based on no theory whatsoever
- Heterogeneous indexes or tradable portfolios: easy and good for non-stationarity but heterogeneity is a question

ETFs

- ETFs are diversified portfolios or derivatives-based equivalents
- Returns must be driven mainly by underlying factors; i.e., by high-frequency changes in market perceptions of macro-economic conditions
- Their idiosyncratic volatility should be relatively small
- They are generally liquid, transparent, and cheap to trade
- Their variety across several asset classes suggests a healthy degree of heterogeneity; but beware of using several in the same asset class
- This is unexplored territory as of now

Our Protocol: Steps toward Necessary and Sufficient Conditions

- To be judged a factor, a factor candidate must survive a series of 6 hurdles that represent the necessary conditions
- To be a "priced" factor, a candidate must satisfy the six necessary conditions plus a sufficient condition
- We describe the 7 hurdles below and give numerical examples for each in the paper

Protocol step #1

- Most efficient to use real asset returns
- Easiest step #1
  - Collect returns on unleveraged equities for firms that do not engage in hedging or use of derivatives for other purposes
  - These will be the ultimate dependent variables
  - One could also use other real assets; e.g., real estate (unleveraged), human capital, etc.
- Danger: some industries might be excluded

Around 1100 US firms have little debt; Here's a sample of 30 that we'll use as an example

Permno	CUSIP	Ticker	Company Name
10026	46603210	JJSF	J & J Snack Foods Corp
10044	77467810	RMCF	Rocky Mountain Chocolate Fac Inc
10100	27135100	ALRN	American Learning Corp
10107	59491810	MSFT	Microsoft Corp
10138	74144T10	TROW	T Rowe Price Group Inc
10163	98385710	XRIT	X Rite Inc
10200	75991610	RGEN	Repligen Corp
10239	57755200	BWINB	Baldwin & Lyons Inc
10258	15117B10	CLDX	Celldex Therapeutics Inc
10259	82656510	SIGM	Sigma Designs Inc
10272	87910110	TKLC	Tekelec
10299	53567810	LLTC	Linear Technology Corp
10302	23280610	CY	Cypress Semiconductor Corp
10318	57665200	BCPC	Balchem Corp
10355	23391210	DJCO	Daily Journal Corp
10363	00163U10	AMAG	A M A G Pharmaceuticals Inc
10382	46224100	ASTE	Astec Industries Inc
10395	63890410	NAVG	Navigators Group Inc
10397	95075510	WERN	Werner Enterprises Inc
10463	76091110	REFR	Research Frontiers Inc
10530	58958410	VIVO	Meridian Bioscience Inc
10547	18482P10	CLFD	Clearfield Inc
10550	74265M20	PDEX	Pro Dex Inc Colo
10644	88337510	TGX	Theragenics Corp
10645	51179510	LAKE	Lakeland Industries Inc
10656	44461000	ACET	Aceto Corp
10781	65066100	BKSC	Bank South Carolina Corp
10812	42234710	HTLD	Heartland Express Inc
10838	23282830	CYTR	Cytrx Corp
10853	76017410	RENT	Rentrak Corp

Protocol step #2

- Collect a set of K factor candidates that consist of two subsets:
- "Spanning" portfolios such as J ETFs, where J is a half-dozen or so, that are as heterogeneous as possible
- K-J factor candidates of interest, e.g.,
  - Size, book/market, momentum
  - Any of the 50 or so documented in Subrahmanyam (2010), or 316 from Harvey, et al.

The ETFs below have ten-year histories
January 2002 through December 2011

Permno	Ticker	Company	ETF name
89187	EPP	iShares	MSCI Pacific ex-Japan
88405	IEV	iShares	S&P Europe 350 Index
83215	EWH	iShares	MSCI Hong Kong Index
83225	EWJ	iShares	MSCI Japan Index
88290	ILF	iShares	S&P Latin America 40 Index
88294	IYR	iShares	Dow Jones US Real Estate
84398	SPY	SPDR	S&P 500

Protocol step #3

- Using the selected set of real asset returns from step #1, calculate the <u>conditional</u> real return covariance matrix V_t (N x N)
  - One suggested method for calculating V_t is the innovative technique of Ledoit, Santa-Clara, and Wolf (2003)
  - It provides <u>conditional</u> covariance matrices that are positive semi-definite, usually time varying, and typically somewhat persistent

Protocol step #4

- From each estimated LSW covariance matrix, V_t, extract the real return eigenvectors corresponding to the L largest eigenvalues; we call these eigenvectors "risk portfolios"

Protocol step #5

- With the K factor candidates from step #2 and the L “risk portfolios” from step #4, estimate the conditional cross-covariance matrix, which we denote C_t
  - It will have K rows and L columns (i.e., K x L); the entry in the i^{th} row and j^{th} column being the covariance between factor candidate i and risk portfolio j
- It will also be convenient to retain the conditional covariance matrix of the factor candidates, $V_{f,t}$ (K x K) and the conditional covariance matrix of the real return risk portfolios, $V_{e,t}$ (L x L.) LSW delivers all three conditional matrices, C_t, $V_{f,t}$ and $V_{e,t}$.

The unconditional covariance matrix with dimension 20 X 20 is reported below for ten principle components extracted from firms with minimal debt, shown in Table 1, seven ETFs shown in Table 2, the Fama-French (1992) SMB and HML factors, and the Carhart (1997) momentum factor (Mom). The sample periods spans 120 months from January 2002 through December 2011. For clarity, the matrix is reported with standard deviations along the diagonal and correlations off diagonal. The principle components are orthogonal to each other by construction and are normalized to have unit variance. For the other ten assets, the factor candidates, standard deviations are shown in natural units of percent per month. .

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	EPP	IEV	EWH	EWJ	ILF	IYR	SPY	SMB	HML	Mom
PC1	1.00																			
PC2	0.00	1.00																		
PC3	0.00	0.00	1.00																	
PC4	0.00	0.00	0.00	1.00																
PC5	0.00	0.00	0.00	0.00	1.00															
PC6	0.00	0.00	0.00	0.00	0.00	1.00														
PC7	0.00	0.00	0.00	0.00	0.00	0.00	1.00													
PC8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00												
PC9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00											
PC10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00										
EPP	0.58	0.00	0.16	-0.14	0.01	0.02	0.22	0.12	0.08	0.06	6.67									
IEV	0.61	-0.08	0.18	-0.15	0.00	0.06	0.27	0.14	0.10	0.05	0.89	6.14								
EWH	0.53	-0.04	0.13	-0.18	-0.04	0.01	0.21	0.05	0.04	0.05	0.86	0.77	6.54							
EWJ	0.44	-0.02	0.23	-0.05	-0.05	0.04	0.12	0.04	0.08	0.07	0.67	0.67	0.55	5.18						
ILF	0.56	0.07	0.21	-0.22	0.03	-0.08	0.26	0.02	0.15	0.03	0.80	0.74	0.77	0.55	7.68					
IYR	0.48	-0.09	0.31	-0.13	-0.04	0.16	0.21	0.01	0.12	0.05	0.70	0.70	0.58	0.60	0.57	7.50				
SPY	0.71	0.00	0.21	-0.18	-0.05	0.04	0.25	0.16	0.12	0.03	0.85	0.90	0.74	0.63	0.75	0.74	4.62			
SMB	0.56	-0.06	0.24	-0.01	0.08	0.14	0.21	-0.11	-0.02	0.09	0.29	0.27	0.29	0.30	0.40	0.43	0.32	2.49		
HML	0.01	-0.06	0.22	0.16	0.08	0.13	0.22	-0.02	0.16	0.07	0.26	0.26	0.04	0.35	0.14	0.48	0.23	0.21	2.51	
Mom	-0.36	0.05	-0.15	0.06	0.06	0.04	-0.06	0.08	-0.02	0.02	-0.28	-0.39	-0.31	-0.27	-0.37	-0.41	-0.44	-0.08	-0.16	5.42

Estimated conditional covariance matrix for December 2006, a month roughly halfway through the overall 2002-2011 sample. Conditional covariance matrices are constructed each month using the flexible multivariate procedure of Ledoit, Santa Clara, and Wolf (2003). Each conditional covariance matrix has dimension 20 X 20 and portrays ten principle components extracted from firms with minimal debt, seven ETFs, the Fama-French (1992) SMB and HML factors, and the Carhart (1997) momentum factor. For clarity, the example monthly matrix below is displayed with standard deviations (%) along the diagonal and correlations off diagonal. The first ten assets are the principle components and are unconditionally orthogonal to each other by construction and scaled to have unit unconditional variance. Standard deviations for the other ten assets, the factor candidates, are in natural units of percent per month.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	EPP	IEV	EWH	EWJ	ILF	IYR	SPY	SMB	HML	Mom
PC1	0.89																			
PC2	-0.09	0.91																		
PC3	0.18	0.00	0.90																	
PC4	0.10	-0.03	-0.03	0.97																
PC5	0.01	-0.01	0.06	-0.08	0.92															
PC6	-0.02	-0.04	-0.01	-0.02	-0.09	0.96														
PC7	0.07	-0.06	0.05	-0.12	0.08	0.06	0.85													
PC8	0.12	-0.04	0.00	0.02	-0.13	0.08	0.10	0.92												
PC9	-0.02	-0.06	0.07	-0.05	-0.07	0.04	0.18	-0.02	0.91											
PC10	0.21	-0.05	0.03	-0.01	-0.05	0.00	0.03	0.04	0.06	1.00										
EPP	0.61	-0.01	0.27	-0.11	0.03	-0.02	0.25	0.06	0.13	0.16	5.42									
IEV	0.61	-0.09	0.31	-0.09	-0.10	-0.01	0.28	0.14	0.20	0.25	0.82	3.66								
EWH	0.61	-0.03	0.23	-0.18	0.42	0.02	0.19	0.06	0.09	0.12	0.70	0.51	6.02							
EWJ	0.52	0.06	0.29	-0.07	-0.07	-0.04	0.10	0.03	0.13	0.06	0.50	0.48	0.50	3.51						
ILF	0.60	0.00	0.35	-0.22	-0.04	-0.12	0.29	-0.04	0.18	0.16	0.68	0.68	0.54	0.44	6.39					
IYR	0.64	-0.14	0.31	-0.05	-0.07	0.08	0.19	0.06	0.14	0.04	0.65	0.61	0.55	0.50	0.54	6.00				
SPY	0.61	-0.09	0.29	-0.14	-0.13	0.00	0.29	0.05	0.18	0.21	0.74	0.85	0.47	0.44	0.66	0.69	2.69			
SMB	0.34	-0.09	0.37	0.00	0.09	0.18	0.31	-0.16	0.02	0.12	0.25	0.30	0.27	0.39	0.43	0.40	0.36	2.39		
HML	-0.04	-0.05	0.14	0.33	-0.08	0.14	0.13	0.02	0.01	-0.06	0.12	0.12	0.01	0.27	-0.01	0.17	-0.01	0.01	1.71	
Mom	-0.12	0.08	-0.05	0.12	0.09	0.21	0.14	-0.03	-0.02	0.17	0.13	-0.07	0.16	-0.03	-0.08	-0.11	-0.08	0.17	0.06	2.23

Protocol step #6

- Using the matrix C_t from step #5, compute canonical correlations between the factor candidates and the risk portfolios (eigenvectors)
- This involves first finding two sets of weights so that the correlation is maximum between the weighted factor candidates and weighted risk portfolios

Protocol Step #6 (Cont.)

- The vectors a and b are chosen to maximize the correlation, which is given by

$$\rho = \frac{a_t' C_t b_t}{\{(a_t' V_{f,t} a_t)(b_t' V_{e,t} b_t)\}^{1/2}}$$

Protocol Step #6 (Cont.)

- The maximum correlation occurs when

$$a_t = V_{f,t}^{-1/2} g_t$$

- And g_t is the maximum eigenvector in the following matrix:

$$V_{f,t}^{-1/2} C_t V_{e,t}^{-1} C_t' V_{e.t}^{-1/2}$$

- Vector b is proportional to g_t
- There are min(L,K) pairs of orthogonal canonical variables sorted from the highest correlation to the lowest

Step #6 (continued)

- Each correlation can be transformed into a variable that is asymptotically distributed as Chi-Square under the null hypothesis that the true correlation is zero

- This provides a method of testing whether the factor candidates as a group are related to the covariance matrix of real returns

- By examining the relative sizes of the weightings in a_t, one can obtain an insight into which factor candidates, if any, are more related to real return covariances

Intuition of canonical correlation

- The true underlying drivers of real returns are undoubtedly changes in perceptions about macroeconomic variables; very high frequency movements

- But factor candidates and the risk portfolios (eigenvectors) need not be identical to a particular macro perception, but they should each be linear combinations of the underlying macro perceptions

Intuition of canonical correlation

- Instead, each candidate or risk portfolio is some (possibly) different linear combination of all the pertinent macro variables
  - This is the well-known “rotation” problem in principal component or factor analysis

- A good factor candidate must be related to some linear combination of the risk portfolios, which depend only on the covariance matrix

"Rejected" factor candidates

- Some factor candidates may not pass the necessary condition hurdles
- These might actually be more interesting to investors than true risk factors
- If a rejected factor candidate is reliably associated with the mean returns on a diversified portfolio of assets, whether or not they are real assets, an arbitrage is possible

Profiting from "rejected" factors

- To construct a profitable low-risk portfolio, factor loadings are required on the accepted factors and on the rejected factor
- Then a long/short position is engineered from the rejected factors loadings (e.g., the characteristics) such that zero factor exposures are structured against every one of the accepted factors
- This essentially eliminates priced risky factor exposure and produces a positive cash flow with little risk if the resulting idiosyncratic volatility is minimal

Canonical Correlations between Factor Candidates, (seven ETFs plus the Fama-French (1992) SMB and HML factors, and the Carhart (1997) momentum factor), versus ten principle components computed from a random sample of thirty stocks with minimal debt. The first column reports the ten canonical correlations computed from the unconditional covariance matrix. The second column reports the time series means of canonical correlations from the conditional covariance matrices computed with the flexible multivariate procedure of Ledoit, Santa Clara, and Wolf (2003). The unconditional canonical correlations are sorted in descending order by their estimated squares. The right-most column gives Newey-West T-statistics (ten lags) for the mean correlations in the second colunn.

Unconditional	Mean Conditional	Newey-West T-Statistic
0.921	0.840	85.68
-0.517	-0.199	-2.74
0.442	0.271	5.10
0.350	0.073	2.05
-0.317	0.104	3.62
-0.232	0.060	2.32
0.130	0.041	2.17
-0.096	0.018	1.42
0.082	0.009	1.56
-0.016	-0.003	-0.89

Weightings are reported below for the two canonical variates that yield the largest canonical correlation. The first canonical variate (left panel) is a weighted average of factor candidates, (seven ETFs listed plus the Fama-French (1992) SMB and HML factors, and the Carhart (1997) momentum factor. The second canonical variate (right panel) is a weighted average of ten principle components computed from a random sample of thirty stocks with minimal debt.

	Unconditional Weighting	Mean Conditional Weighting	Newey-West T-Statistic		Unconditional Weighting	Mean Conditional Weighting	Newey-West T-Statistic
	Factor Candidates				Principle Components		
EPP	-0.0113	-0.0079	-1.592	PC1	0.8919	0.9073	36.76
IEV	0.0275	0.0005	0.065	PC2	-0.0292	0.0268	5.346
EWH	-0.0100	0.0914	16.12	PC3	0.2592	0.2109	12.69
EWJ	-0.0057	0.0236	5.121	PC4	-0.1663	-0.2518	-11.55
ILF	-0.0014	0.0232	6.495	PC5	-0.0080	0.1957	8.460
IYR	-0.0126	0.0195	4.980	PC6	0.0861	0.0209	1.147
SPY	0.1670	0.0382	4.868	PC7	0.2960	0.1470	10.43
SMB	0.2044	0.0781	7.015	PC8	0.0793	-0.0363	-4.824
HML	-0.0479	-0.0724	-5.273	PC9	0.0663	0.1249	12.49
Mom	-0.0157	-0.0484	-4.479	PC10	0.0553	-0.0102	-1.850

Sufficient Conditions for a True Risk Factor

- Factor candidate must first pass all six steps that represent the necessary conditions
- Then run Fama/MacBeth with the real individual asset returns as dependent variables and the factor loadings (betas) as explanatory variables; do not use portfolios
- Employ instrumental variables for the betas to mitigate measurement error

Instruments for Betas

- The instrument must be highly correlated with true (but unknown) beta but be uncorrelated with the beta estimation error
- Two suggestions
  - Betas estimated from observations after the cross-section
  - Betas from interspersed observations before the cross-section; this method is less subject to variation in true betas and risk premiums

Quasi-Fama/MacBeth with IV

- Split the sample into three sub-samples
  - Observations 1, 4, 7, 10,…estimate betas
  - Observations 2, 5, 8, 11,…estimate beta instruments
  - Observations 3, 6, 9, 12,…do FM cross-sectional instrumental variable regressions
- The third sub-sample delivers asymptotically unbiased estimate of risk premiums
- One could also split sample into unequal sub-samples
- Is there anything wrong with also doing permutations of the above scheme; i.e., 2, 5, 8..; 3, 6, 9..; 1, 4, 7…??

Instrument difficulties

- No small sample results, asymptotics only
- Notoriously poor power with small samples
- Hence, in the sufficient conditions tests, we use 1,100 non-debt US stocks and 2,517 daily time series observation (2002-2012)
- The factor loadings from 5 positively significant ETFs plus SMB, HML & Mom from Canonical Correlations are explanatory variables

Exactly Matched (N&T) Simulations to Check Validity of New IV Procedure

Simulation Parameters

Factor	True Risk Premium (%/day)	Standard Deviation (%/day)	Betas: Cross-Sectional Mean	Betas: Standard Deviation
1	.04%	.4	1.0	.2
2	.03%	.3	Zero	.2
3	.02%	.3	Zero	.2
4	Zero	.2	Zero	.2
5	Zero	.2	Zero	.2
6	Zero	.2	Zero	.2

Idiosyncratic Shock
lognormal cross-sectional distribution

Mean of logs: 0.8	Standard deviation: 0.2

Panel B.
Estimated Risk Premiums

Factor	Mean IV Coefficient	T-Statistic
1	.0496%	5.86
2	.0275%	4.21
3	.0189%	2.86
4	.0011%	.234
5	-.0007%	-.139
6	-.0002%	-.223

Panel C.
Correlations between Betas and Their Instruments, sub-samples one and two

1	.891
2	.816
3	.827
4	.674
5	.704
6	.685

Summary Statistics from Alternate Observation Regressions

	Mean	Sigma	T-Stat	Skewness	Kurtosis	Maximum	Minimum
Regressions using sub-sample #1, observations 1, 4, 7…							
EWH	0.016	0.142	3.752	0.426	3.033	1.017	-0.513
EWJ	0.009	0.149	2.057	-0.039	3.705	1.072	-0.770
ILF	0.086	0.183	15.278	1.163	2.022	0.937	-0.422
IYR	0.005	0.210	0.804	0.801	4.837	1.333	-0.827
SPY	0.698	0.450	50.473	-0.151	-0.037	2.122	-0.803
SMB	0.827	0.540	49.822	0.139	-0.277	2.571	-0.859
HML	0.080	0.483	5.381	0.099	1.486	2.864	-1.468
MOM	-0.104	0.259	-13.062	-0.239	1.014	0.798	-1.125
Intercept	0.048	0.121	13.060	0.984	3.308	0.785	-0.324
Adj. R^2	0.207	0.159	42.483	0.578	-0.061	0.798	-0.006
Regressions using sub-sample #2, observations 2, 5, 8…							
EWH	-0.011	0.151	-2.279	-0.760	5.610	0.704	-0.985
EWJ	0.029	0.159	5.879	0.472	2.146	0.814	-0.586
ILF	0.066	0.177	12.148	1.047	2.342	0.879	-0.542
IYR	-0.007	0.219	-1.070	0.915	5.193	1.347	-0.842
SPY	0.763	0.440	56.385	-0.070	0.170	2.161	-1.098
SMB	0.797	0.571	45.388	0.122	-0.323	2.805	-1.150
HML	0.055	0.493	3.644	0.388	3.021	3.735	-1.493
MOM	-0.085	0.248	-11.174	-0.410	0.622	0.566	-1.099
Intercept	0.056	0.113	16.118	0.762	3.560	0.774	-0.404
Adj. R^2	0.219	0.162	44.053	0.499	-0.188	0.824	-0.005

Beta and Beta Instrument Correlations Across Samples

Correlations between Betas and Their Instruments

Factor	Sub-samples		
	1 & 2	1&3	2&3
EWH	0.111	0.078	0.173
EWJ	0.143	0.092	0.114
ILF	0.607	0.531	0.482
IYR	0.628	0.617	0.602
SPY	0.702	0.698	0.701
SMB	0.765	0.761	0.781
HML	0.631	0.586	0.565
MOM	0.510	0.518	0.478

Risk Premiums and Average Factor Means

Panel B.
Estimated Risk Premiums

Factor	Mean IV Coefficient (%/day)	T-Statistic
ILF	0.0586	0.758
IYR	-0.0083	-0.148
SPY	0.0089	0.215
SMB	0.0160	0.864
HML	0.0053	0.241
MOM	0.1544	2.924

Panel C.
Factor Returns, 2002-2011

Factor	Mean (%/day)	T-Statistic	Excess Kurtosis
ILF	0.09337	2.055	13.702
IYR	0.05709	1.308	12.625
SPY	0.02087	0.757	11.519
SMB	0.01482	1.244	3.886
HML	0.1160	0.987	7.165
MOM	0.0059	0.279	9.213

Work in Progress

- Non-stationary estimates of betas and beta instruments
- Simulations to check test power of necessary conditions of protocol
- More work investigating the IV procedure

Thanks for your kind attention