Empirical Asset Pricing - Midterm Exam

Maximilian Huber

March 26, 2018

1 Estimating Asset Pricing Models Using GMM

		1st Stage		Final Stage				
Specification		$\hat{\gamma}$	$\hat{\psi}$	$\hat{\gamma}$	$\hat{\psi}$	(std. err $\hat{\gamma}$)	(std. err $\hat{\psi}$)	J_T stat	% p-val
1			0.2001		0.2138		0.0426	3.7335	29.17
2			0.2415		1.5160		0.3381	15.636	0.35
3			0.1721		0.1585		0.0245	18.521	0.24
4		7.8984	0.1605	20.951	0.1184	11.905	0.0171	18.673	0.09
5		7.5051	0.1603	6.0217	0.169	1.8784	0.0222	29.212	2.25
6	2 step	2.0597	0.4812	1.8301	0.542	0.3437	0.0883	41.306	0.22
6	∞ step	2.0597	0.4812	0.2990	3.6839	0.1553	2.524	34.855	1.45
7			1.3825		1.1025		0.0129	29.536	5.80
8		30.808	0.7127	96.628	0.3841	27.465	0.0759	55.210	0.00
9		100.63	0.41	52.737	0.1306	9.1371	0.0228	52.539	1.25

My estimation routine evolves around β, ψ, and γ. Usually I choose a HAC variance estimation with 4 lags, i.e. +/- one year.

General Remarks

- From task 5 on, the HAC variance matrix of the moments has a condition number of over 1e8. This means we may loose up to 8 digits of precision!

1.1 ψ and Industry Stocks

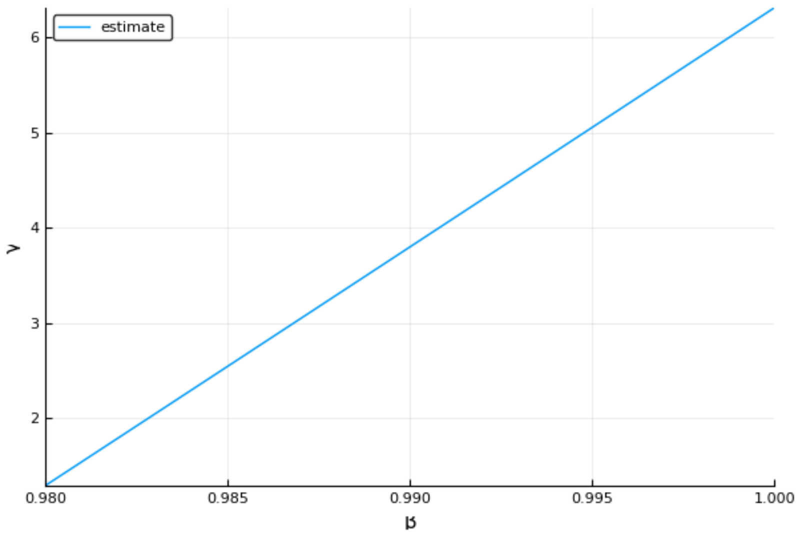
With β fixed at 1 and CRRA utility the estimation yields a γ around 5. This number is plausible, but using only industry portfolios might not be a good choice of test assets.

1.2 ψ, Industry Stocks and T-Bill

Including a risk-free asset depresses the risk aversion coefficient.

1.3 ψ and Value Portfolios

The relative risk aversion increases, the model is rejected by the J-test.
There appears to be a linear relation in the β and the estimate for γ:



The relative risk aversion γ subsumes parts of the impatience of the representative agent!

1.4 γ, ψ and Value Portfolios

The second column of the gradient of the sample moment function is almost zero, hence the second argument, γ, is not well estimated, due to lack of slope. A different value of γ does not change the sample moments much.

$$\nabla g_n(\theta) = \begin{bmatrix} 0.9977 & 0.0012 \\ 0.9645 & 0.001 \\ 0.9684 & 0.001 \\ 0.9368 & 0.0009 \\ 0.9187 & 0.0008 \\ 0.9428 & 0.0009 \end{bmatrix}$$

1.5 γ, ψ and Value Portfolios with lagged Consumption and Wealth Return

I had to restrict the GMM iteration to be a two-step procecure only. Iterating to convergence yielded an extremly high ψ.
Risk aversion and IES seem plausible!

1.6 γ, ψ and Value Portfolios and T-Bill with lagged Consumption, Wealth Return

I had to provide a more nuanced matrix for the first stage and do only one GMM iteration. Iterating until convergence yields implausibly high IES.
The risk aversion drops and the IES increases compared to #5. I did not expect that. I thought adding the risk-free asset would lay more focus on the risk-free rate puzzle (the inverse of the equity premium puzzle) and the IES would drop!

1.7 β, ψ and Value Portfolios and T-Bill with lagged Consumption and Wealth Return

The estimated β is 0.9806 in the first stage and 0.9781 after the final stage with a standard error of 0.002.

1.8 β, γ, ψ and Value Portfolios and T-Bill with more lagged Consumption

I had to switch to a Nelder-Mead algorithm and use a different matrix in the first stage. Then β is 0.84 in the first stage and 0.8455 after the final stage.

1.9 β, γ, ψ and Value Portfolios and T-Bill with more lagged Consumption and Wealth Return

β is 0.8711 in the first stage and 0.9844 after the final stage.

2 Forecasting Excess Stock Returns

2.1 Regressions

I chose the Newey-West lag to be the number of periods that are being aggregated minus one, e.g. the horizon one variance matrix is estimated without a lag. The reason is that I want to capture any autocovariance that is generated by aggregating multiple returns.
The pattern of the $\bar{R}^2$ is increasing with horizon.

2.1.a *cay* only

$r_{t+h,h}$ on cay_t			
Horizon	coeff	t-stat	$\bar{R}^2$
1	0.6397	14.375	0.0143
2	1.3309	12.896	0.0361
3	2.0517	12.374	0.0622
4	2.7777	11.470	0.0903
5	3.4264	10.951	0.1131
6	4.2060	10.739	0.1455
7	5.0345	10.588	0.1823
8	5.7702	11.479	0.2194

cay appears to be a predictor of return in all horizons.

- Remark on endogeneity: This regression suffers somewhat from an endogeneity problem! The y part of *cay* is arguably the left-hand-side of the regression, namely market return. Therefore assuming that consumption does not react much the a return shock, $\epsilon \uparrow$ implies $cay \downarrow$. Hence the coefficient is downwardly biased, hence my estimates are convervative and the argument of *cay* being a predictor goes through!

2.1.b Dividend-Price Ratio

$r_{t+h,h}$ on dp_t			
Horizon	coeff	t-stat	$\bar{R}^2$
1	0.0265	15.819	0.0115
2	0.0556	13.441	0.0309
3	0.0824	11.979	0.0490
4	0.1067	10.225	0.0650
5	0.1303	9.051	0.0806
6	0.1533	8.537	0.0951
7	0.1749	7.893	0.1086
8	0.1914	7.434	0.1196

The Dividend-Price ratio appears to be a predictor of return in all horizons.

- Remark on endogeneity: This regression suffers from the Mankiw-Shapiro/Stambaugh bias. Since dp_t is highly persistence and because a positive shock to $r_{t+h,t}$ also decreases dp_{t+h+1} this leads to an upward bias of the coefficient on dp_t. This notebook illustrates the bias.

2.1.c *cay*, Dividend-Price Ratio, *RREL*, *DEF* and *TRM*

	$r_{t+h,h}$ on										
	cay_t		dp_t		$RREL$		DEF		TRM		
Horizon	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	$\bar{R}^2$
1	0.6394	12.972	0.0273	10.937	-1.1081	-10.243	0.0050	2.2126	-0.0982	-7.5770	0.0392
2	1.3288	12.04	0.0538	8.6763	-1.6959	-4.1952	0.0132	2.3243	-0.0505	-1.7192	0.0718
3	2.0331	12.199	0.0777	8.1331	-2.3036	-3.0202	0.0173	2.1464	-0.0138	-0.26	0.1132
4	2.7343	12.732	0.0995	7.5634	-2.6611	-2.7353	0.0165	1.6384	-0.0174	-0.3076	0.1505
5	3.3182	12.567	0.1222	6.9383	-3.6028	-3.9473	0.0077	0.6843	-0.0033	-0.0653	0.1842
6	4.0528	11.8998	0.1371	6.4304	-2.8338	-4.1198	0.0131	1.0469	0.0196	0.3763	0.2114
7	4.7988	11.3502	0.1470	5.4155	-2.4969	-4.2001	0.0153	1.0662	0.1413	2.8511	0.2457
8	5.4648	12.267	0.1559	4.9077	-2.6224	-5.3412	0.0101	0.634	0.1593	2.8086	0.28

While *cay* appears to predict on all horizons equally strongly, *dp*'s strenght decays, but it is still siginificant. DEF as questionable predicitve power and TRM predicts in the very short term and faces a reversal in the longer term. This is interesting, an long-term interest rate hike leads to short term losses, but long term gains. $RREL$ appreas to predict well. Short-term interest rate hikes lead to losses.

2.2 VAR

- Remark: In the following derivation I do not follow the slides from the lecture. This approachs seems very straight forward to me.

The return h periods ahead, according to the VAR system is:

$$r_{t+h} = e^2 A^h x_t + e^2 A^{h-1}\epsilon_{t+1} + ... + e^2 \epsilon_{t+h}$$

where $e^2 = \big(0, 1, 0\big)$.
Summing over the terms $r_{t+1}, \dots, r_{t+h}$ gives rise to the multi-period return in the next h periods starting from t:

$$r_{t+h,t} = e^2\big((A^h + \dots + A^1)x_t\big) + e^2\big(A^{h-1}\epsilon_{t+1} + \dots + A^0(\epsilon_{t+h} + \dots + \epsilon_{t+1})\big)$$

The R-squared statistic is defines as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

where $SS_{tot} = \sum_{t=1}^{T-h}(r_{t+h,t} - \bar{r}_h)^2 = \sum_{t=1}^{T-h}\big(e^2(A^h + \dots + A^1)x_t + e^2\big(A^{h-1}\epsilon_{t+1} + \dots + A^0(\epsilon_{t+h} + \dots + \epsilon_{t+1})\big) - e^2(A^h + \dots + A^1)(x_t - \bar{x}_t)\big)^2 = \sum_{t=1}^{T-h}\big(e^2\big(A^{h-1}\epsilon_{t+1} + \dots + A^0(\epsilon_{t+h} + \dots + \epsilon_{t+1})\big) + e^2(A^h + \dots + A^1)\bar{x}_t\big)^2$ and $SS_{res}(A) = \sum_{t=1}^{T-h}\big(e^2\big(A^{h-1}\epsilon_{t+1} + \dots + A^0(\epsilon_{t+h} + \dots + \epsilon_{t+1})\big)\big)^2$.
I coded up R^2 as a function of the vectorized A matrix. Then I used automatic differentiation to calculate its derivative and applied the Delta Method.
The imputed R-square statistic is:

	R^2	
Horizon	coeff	(std. err)
2	0.0524	0.0014
4	0.1090	0.0055
8	0.2364	0.0189
12	0.3500	0.0320
16	0.4330	0.0426

3 Summary Statistics

The Campbell Shiller decomposition starts out from the defininition of the return on a stock: $R_{t+1}=\frac{P_{t+1}+D_{t+1}
Taking logs:

$$r_{t+1} = p_{t+1} - p_t + \log\Big(1 + \exp(\underbrace{d_{t+1} - p_{t+1}}_{dp_{t+1}})\Big)$$

The Taylor approximation of the last term around $\bar{dp}$ is:

$$\log\Big(1+\exp(dp_{t+1})\Big) \approx \underbrace{\log\Big(1+\exp(\bar{dp})\Big)}_{\kappa} + \underbrace{\frac{\exp(\bar{dp})}{1+\exp(\bar{dp})}}_{1-\rho}\Big(dp_{t+1}-\bar{dp}\Big) + o\Big(dp_{t+1}-\bar{dp}\Big)^2$$

where $\tilde{dp}$ is between dp_{t+1} and $\bar{dp}$ (this is the Lagrange form of the remainder).
Hence,

$$pd_t = \rho pd_{t+1} + \kappa - (1-\rho)\bar{dp} + \Delta d_{t+1} - r_{t+1}$$

subsituting forward and using the no-bubbles assumption yields:

$$pd_t = \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty}\rho^j\Big(\Delta d_{t+j+1} - r_{t+j+1}\Big)$$

or taking conditional expectations:

$$pd_t = \frac{\kappa}{1-\rho} + \mathbb{E}_t\sum_{j=0}^{\infty}\rho^j\Big(\Delta d_{t+j+1} - r_{t+j+1}\Big)$$

This equation allows me to do the variance decomposition. Using the unpredictability of dividend growth by the price-dividend ration (and the law of iterated expectations):

$$\mathbb{V}\Big[pd_t\Big] = \underbrace{\mathbb{C}\Big[\sum_{s=1}^{\infty}\rho^{s-1}\mathbb{E}_t\left[\Delta d_{t+s}\right], pd_t\Big]}_{=0} + \mathbb{C}\Big[-\sum_{s=1}^{\infty}\rho^{s-1}\mathbb{E}_t\left[r_{t+s}\right], pd_t\Big]$$

If the price-dividend ration has positive variance (which it has in the data), then there is predictability of future returns by the price-dividend ratio.
Now to answer the question, while pd_t does forecast a geometric sum of returns, it does not give specific information on the time series of the returns. For example, $\mathbb{E}_t\left[r_{t+s}\right]$ could be constant in s, or decreasing, or hump-shaped. Therefore, the price-dividend ratio is a summary statistic, but is by no means sufficient.