

## 1

### (a)

t-test, 1 restriction.

### (b)

t-test, 1 restriction.

### (c)

F-test, 2 restrictions.

### (d)

F-test, 4 restrictions.

### (e)

F-test, 1 restriction.

## 2

Expect restricted residual sum of squares to be bigger. Restriction in model means less flexibility to fit data, hence larger residuals and larger sum of squared residuals.

## 3

Restricted regression is obtained by substituting the null hypothesis into the original model.  $H_0 : \beta_3 + \beta_4 = 1$  and  $\beta_5 = 1$ . So,  $\beta_3 = 1 - \beta_4$  and  $\beta_5 = 1$ . Substitute into the original equation:

$$y_t = \beta_1 + \beta_2 x_{2t} + (1 - \beta_4)x_{3t} + \beta_4 x_{4t} + 1x_{5t} + u_t$$

$$y_t = \beta_1 + \beta_2 x_{2t} + x_{3t} - \beta_4 x_{3t} + \beta_4 x_{4t} + x_{5t} + u_t$$

$$y_t - x_{3t} - x_{5t} = \beta_1 + \beta_2 x_{2t} + \beta_4 (x_{4t} - x_{3t}) + u_t$$

Define  $y_t^* = y_t - x_{3t} - x_{5t}$  and  $x_t^* = x_{4t} - x_{3t}$ . Restricted regression:  $y_t^* = \beta_1 + \beta_2 x_{2t} + \beta_4 x_t^* + u_t$ .  
F-test statistic is given by:

$$F = \frac{(RSS_R - RSS_{UR})/r}{RSS_{UR}/(n - k)}$$

Here,  $RSS_R = 102.87$ ,  $RSS_{UR} = 91.41$ ,  $n = 96$ ,  $r = 2$ ,  $k = 5$ .

$$F = \frac{(102.87 - 91.41)/2}{91.41/(96 - 5)} = \frac{11.46/2}{91.41/91} = \frac{5.73}{1.0045} \approx 5.70$$

Degrees of freedom for F-distribution are  $df_1 = r = 2$  and  $df_2 = n - k = 91$ . Critical value  $F_{0.05, 2, 91} \approx 3.10$  (from F-distribution table or software). Since  $F = 5.70 > 3.10$ , we reject the null hypothesis at 5% significance level.

Conclusion: Reject  $H_0$ . There is statistically significant evidence against the null hypothesis  $H_0 : \beta_3 + \beta_4 = 1$  and  $\beta_5 = 1$ .

## 4

### (a)

t-ratios are calculated by dividing the coefficient estimate by its standard error. For  $S_i$ :  $t = 0.801/0.147 \approx 5.45$ . For  $MB_i$ :  $t = 0.321/0.136 \approx 2.36$ . For  $PE_i$ :  $t = 0.164/0.420 \approx 0.39$ . For  $BETA_i$ :  $t = -0.084/0.120 \approx -0.70$ .

For  $S_i$  and  $MB_i$ , t-ratios are relatively large (in absolute value), suggesting they are statistically significant at conventional levels (e.g., 5% or 10%). For  $PE_i$  and  $BETA_i$ , t-ratios are small, suggesting they are not statistically significant.

Conclusion: Size and market-to-book ratio seem to have significant effects on stock returns. Price/earnings ratio and beta do not appear to have significant effects.

### (b)

Based on the results,  $PE_i$  and  $BETA_i$  could be considered for deletion due to their low t-ratios and statistical insignificance.

### (c)

The coefficient on  $BETA_i$  is -0.084. If beta increases from 1 to 1.2, the change in beta is  $1.2 - 1 = 0.2$ . Expected effect on stock return = coefficient of  $BETA_i$  \* change in  $BETA_i$  =  $-0.084 * 0.2 = -0.0168$ . Expected stock return would decrease by approximately 0.0168 or 1.68 percentage points.

(d)

The sign on beta is negative, which is not as expected from simple CAPM theory. CAPM suggests a positive relationship between beta and expected return, meaning higher beta should lead to higher expected return. Possible reasons for unexpected negative sign: 1. Other variables are not controlled for properly, or the model form is incorrect. 2. Beta may be correlated with other regressors, affecting its estimated coefficient. 3. This particular sample might show an unusual relationship. 4. CAPM is a simplified model and may not perfectly describe real-world stock returns, especially when other factors are included in the regression.

## 5

Dep. Variable:	YearlyEarningsNOK	R-squared:	0.664			
Model:	OLS	Adj. R-squared:	0.653			
Method:	Least Squares	F-statistic:	63.10			
Date:	Sat, 08 Feb 2025	Prob (F-statistic):	1.27e-22			
Time:	14:17:43	Log-Likelihood:	-1242.3			
No. Observations:	100	AIC:	2493.			
Df Residuals:	96	BIC:	2503.			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P>  t	[0.025	0.975]
Intercept	-3.523e+04	4.1e+04	-0.859	0.393	-1.17e+05	4.62e+04
Education	3.22e+04	2358.007	13.657	0.000	2.75e+04	3.69e+04
Age	-426.6183	480.995	-0.887	0.377	-1381.386	528.149
Sex	2.284e+04	1.24e+04	1.836	0.069	-1853.300	4.75e+04
	Omnibus:	0.722	Durbin-Watson:	2.062		
	Prob(Omnibus):	0.697	Jarque-Bera (JB):	0.845		
	Skew:	0.175	Prob(JB):	0.656		
	Kurtosis:	2.716	Cond. No.	349.		

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

## 6

As stated before, for each parameter  $\beta_i$ , we test  $H_0 : \beta_i = 0$  vs  $H_1 : \beta_i \neq 0$ . We use the t-test for this. The regression output provides the t-statistics and p-values for each coefficient.

## 7

The only parameter statistically significant at the 5% level is Education.

The coefficient for Education is approximately 32200. This means that, holding Age and Sex constant, each additional year of education is associated with an increase in yearly earnings of approximately 32200 NOK, on average. This effect is statistically significant at the 5% level. The coefficient for Sex is approximately 22840, with a p-value of 0.069. If we consider a 10% significance level, Sex would be considered marginally significant. The positive coefficient suggests that, holding Education and Age constant, males (Sex=1) earn approximately 22840 NOK more per year than females (Sex=0), on average. However, note that at the stricter 5% level of significance, we would not conclude that Sex has a statistically significant effect. Age and Intercept are not statistically significant even at the 10% level.

## 8

The regression model (prediction model) for salary is:

$$\widehat{YearlyEarnings} = -35230 + 32200 \times Education - 426.6 \times Age + 22840 \times Sex$$
$$\widehat{YearlyEarnings} \approx 437958.2$$

## 9

Violation of uncorrelated error terms with explanatory variables can occur due to omitted variable bias. This happens when a variable that affects Yearly Earnings and is correlated with Education, Age, or Sex is left out of the regression.

Two variables that could address this issue are:

- Experience: More experienced individuals generally earn more, and experience is likely correlated with age.
- Occupation/Industry: Different occupations and industries have different pay scales. Occupation and industry might be correlated with education and sex (certain fields are dominated by one sex or require specific education levels). Omitting these can lead to correlated errors.

## 10

$R^2$  (0.664) indicates that 66.4% of the variation in Yearly Earnings is explained by Education, Age, and Sex. This is a moderate value, suggesting the model captures a fair amount of earnings variation.  $R^2$ 's main disadvantage is that it always increases with added variables, potentially leading to overfitting. To counter this, we use Adjusted  $R^2$  (0.653). Adjusted  $R^2$

penalizes adding unnecessary variables, providing a better measure for comparing models with different numbers of predictors. While  $R^2$  is useful, it doesn't guarantee a good model or causality, only goodness of fit within the sample.