OLS/MLR Assessment I: Goodness-of-fit

SST

SSE

SSR

MSE

RMSE

R^2

adjR^2

- Review of SLR Assessment (*predicteds* v. *actuals*)
- R-sq shortcoming in MLR models: *Just showing up!*
- A Quick Comparison of SLR and MLR Assessment – Not much that's new!
- MLR Goodness-of-Fit: Adjusted R-squared
- … and adding and subtracting RHS variables
- Comparing MLR Models I: Goodness-of-Fit metrics in action

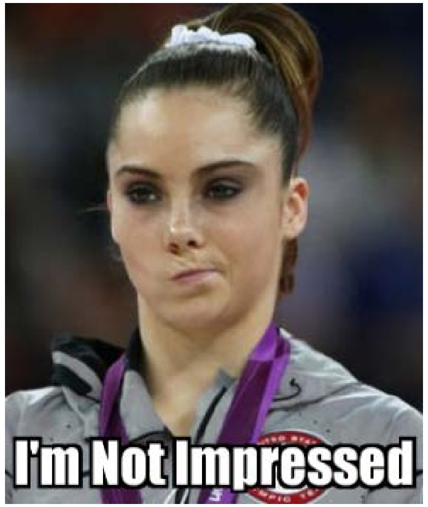
OLS/SLR Assessment Review: Predicteds v. Actuals

- MSE (Mean Squared Error): $MSE = \dfrac{SSR}{n-2}$ … an average squared residual, sort of…
- RMSE (Root Mean Squared Error): $RMSE = \sqrt{MSE}$ … sort of an average residual, but more like a square root of an average squared residual, sort of…
- R^2 (Coefficient of Determination): $R^2 = 1 - \dfrac{SSR}{SST} = \dfrac{SSE}{SST} = \dfrac{S_{\hat{y}\hat{y}}}{S_{yy}} = \rho_{xy}^2 = \rho_{\hat{y}y}^2$ … proportion of the variance of the *actuals explained* by the *predicteds*, as well as the correlation (squared) between the *predicteds* and the *actuals*.
- Usefulness: The MSE and RMSE metrics are not in standardized units, making it difficult to interpret the magnitudes. But R^2, which ranges from zero to one, is standardized to some extent, making it perhaps more useful in assessing the performance of the model:

R^2: $0 \le R^2 \le 1$… closer to one is better…. closer to zero, not so much

R-sq Shortcoming: *Gives variables credit for just showing up!*

- R^2 gives credit to variables for *just showing up*… irrespective of their explanatory power!
- ***What drives the result***: Remember that OLS coefficients are always found by minimizing SSRs. And so when RHS variables are added to a MLR model, SSRs will typically decrease, or at worst, stay the same. But they can never increase (and R^2 can never decrease) since you can't do a poorer job of minimizing SSRs when you have one more RHS variable to work with.
  - If the new variable has an OLS/MLR coefficient of zero, then the new variable has added nothing (no explanatory content) to the model, and SSRs and R^2 are unchanged.
  - Alternatively, if the new coefficient is non-zero (and uniquely defined) when minimizing SSRs, then SSRs will necessarily have decreased, and R^2 increases.
- ***What usually happens*:** When new explanatory variables are added to a model their coefficients will typically be non-zero and R^2 will typically increase. So no one should be impressed if R^2 increases when new RHS variables are added to the MLR analysis… that's entirely to be expected.



SLR v. MLR Assessment: *Not much that's new!*

	SLR	**MLR**
Sum Squares	SST = SSE + SSR	SST = SSE + SSR
R^2 (Coefficient of Determination) (w/ intercept term)	$R^2 = 1 - \frac{SSR}{SST} = \frac{SSE}{SST} = \rho_{xy}^2 = \rho_{\hat{y}y}^2$	$R^2 = 1 - \frac{SSR}{SST} = \frac{SSE}{SST} = \rho_{\hat{y}y}^2$
	$R^2 = \frac{SampleVar(predicted)}{SampleVar(actual)}$	$R^2 = \frac{SampleVar(predicted)}{SampleVar(actual)}$
Degrees of freedom (*dofs*)	$dofs = n - 2$	$dofs = n - k - 1$
MSE	$MSE = \frac{SSR}{dofs} = \frac{SSR}{n-2}$	$MSE = \frac{SSR}{dofs} = \frac{SSR}{n-k-1}$
RMSE	$RMSE = \sqrt{MSE}$	$RMSE = \sqrt{MSE}$
Adjusted R^2		$\bar{R}^2 = 1 - \frac{SSR}{SST}\frac{n-1}{n-k-1}$ $= 1 - \frac{SSR/(n-k-1)}{SST/(n-1)} = 1 - \frac{MSE}{S_{yy}}$

adj R-sq and MSE when changing RHS variables

$\bar{R}^2$

- $$adjR^2 = \bar{R}^2 = 1 - \left(\frac{SSR}{SST}\right)\left(\frac{n-1}{dofs}\right) = 1 - \frac{(n-1)}{SST}\left[\frac{SSR}{dofs}\right]$$
- As you add explanatory variables to the model, only the terms in the square brackets (*SSR* and *dofs*) are changing, both are typically declining. $\bar{R}^2$ will increase or decrease depending on the relative rates of change of *SSRs* and *dofs*:
  - If the decline in SSRs is faster than the decline in *dofs*, then $\left[\frac{SSR}{dofs}\right]$ will decline and $\bar{R}^2$ will increase with the additional explanatory variables.
  - But if the decline in SSRs is slower than the decline in *dofs*, then $\left[\frac{SSR}{dofs}\right]$ will increase , and $\bar{R}^2$ will decrease.
  - *Adjusted* R^2 will increase if SSRs are dropping faster than *dofs*

adj R-sq and MSE when changing RHS variables, cont'd

- *adj R^2* is always bounded above by R^2, and by 1:
  - $\bar{R}^2 < R^2 \leq 1$ for $k > 0$, since $\dfrac{(n-1)}{(n-k-1)} = \dfrac{(n-1)}{dofs} > 1$ and so $\left[\dfrac{SSR}{SST}\right]\left[\dfrac{n-1}{dofs}\right] > \left[\dfrac{SSR}{SST}\right]$.
- Since $\bar{R}^2 = 1 - \dfrac{SSR/(n-k-1)}{SST/(n-1)} = 1 - \dfrac{MSE}{S_{yy}}$, *adjusted R^2* and *MSE* will always move in opposite directions when S_{yy} is fixed.
- Accordingly, the two goodness-of-fit metrics (*adjusted R^2* and *MSE/RMSE)* are effectively redundant in the sense that knowing the movements patterns of one tells you the movements of the other.

R-sq, adj R-sq and MSE when changing RHS variables, cont'd

```
. esttab, r2 ar2 scalar (rmse) compress
```

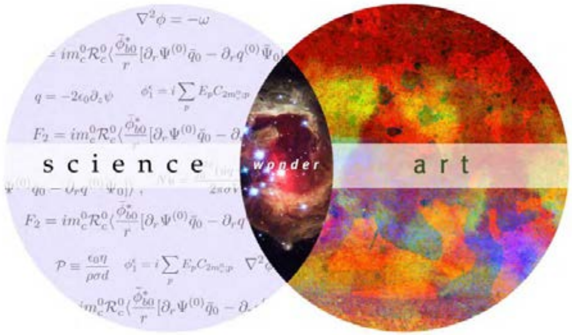
	(1) Brozek	(2) Brozek	(3) Brozek	(4) Brozek
hgt	-0.650***	-0.118	-0.131	-0.138
	(-6.29)	(-1.43)	(-1.51)	(-1.55)
wgt	0.187***	-0.120***	-0.108**	-0.100*
	(14.48)	(-5.41)	(-3.18)	(-2.52)
abd		0.880***	0.883***	0.898***
		(15.19)	(15.13)	(12.62)
hip			-0.0564	-0.0723
			(-0.49)	(-0.58)
chest				-0.0348
				(-0.38)
_cons	31.16***	-32.66***	-28.64**	-25.86*
	(4.51)	(-5.01)	(-2.71)	(-2.01)
N	252	252	252	252
R-sq	**0.4614**	**0.7210**	**0.7213**	**0.7215**
adj. R-sq	**0.4571**	**0.7177**	**0.7168**	**0.7158**
rmse	**5.7109**	**4.1184**	**4.1248**	**4.1320**

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

As Advertised!

Choosing the Best Model: *Art & Science*

- Comparing the performance of MLR models ***is as much art as science*** … and in truth, we typically look at a number of different aspects/properties of the model. But certainly *adj R-sq* and *RMSE* are in the conversation.
- Choosing between models depends in part on the goals of the analysis:
  - **Forecasting models** (*less is more*; focus on out-of-sample forecasting, and don't over-fit the data)
  - **Behavioral models** (*parsimony preferred*; the challenging art form)
  - **Favorite coefficient models** (*more is more*; focus on the favorite coefficient… and don't worry about the other aspects of the model… other than making sure that you really have included every possible relevant explanatory variable, and accordingly that you have minimized the possibility of omitted variable impact/bias)



OLS/MLR Assessment I - GOFs: *TakeAways*

- The SLR Goodness of Fit (GOF) metrics (R-sq, MSE and RMSE) extend to MLR models with the only change being that in computing MLR MSEs you now divide SSRs by n-k-1, the degrees of freedom (dofs) in the MLR model.
- And the SLR interpretations carry over to MLR models: R-sq is the proportion of the variation of the LHS variable explained by the model, MSE is almost an average squared residual, and RMSE is almost an average residual.
- However: R-sq is not so useful in evaluating MLR models, as R-sq will almost always increase as you add RHS variables to a model. Or put differently: Added RHS variables get R-sq credit just *for showing up*.
- A new GOF metric, adjusted R-sq, offers a response… and only increases when the new RHS variables have some significant explanatory power (in the form of significantly reduced SSRs)
- Adj R-sq increases with the added RHS variables, if the decrease in SSRs is larger than the decrease in *dofs*.
- MSE and adj R-sq will move in opposite directions so long as n and SST are unchanged. They are both useful in choosing between models… as well, the type/goal of the model (*forecasting*, *behavioral* or *favorite coefficient*) will also guide that choice.

onwards… to OLS/*MLR Analytics II*