

ECONOMICS 7344 – MACROECONOMIC THEORY II, Spring 2008

Measurement Error in Regression and Friedman’s version of the Permanent Income Model.

Keynes argued that saving tends to be an increasing function of income (which may lead to insufficient aggregate demand as aggregate income increase) based on consumption of families, where it was observed that in the regression

$$c_{it} = \alpha + \gamma y_{it} + u_{it} ,$$

where c_{it} and y_{it} is the consumption and income of family i at period t , the estimated slope coefficient $\hat{\gamma}$ is significantly lower than 1.

It was, however, also observed that when the consumption function was estimated on aggregate data

$$c_t = a + by_t + w_t ,$$

a coefficient (\hat{b}) near 1 was estimated.

How could this be? Milton Friedman gave a clever answer building on economic and econometric reasoning (an answer that still, in a more developed form, is the benchmark model of consumer analysis). Milton Friedman was later given the Nobel Prize “for his achievements in the fields of consumption analysis, monetary history and theory, and for his demonstration of the complexity of stabilization policy.”

Friedman argued that consumers rationally do not adjust consumption to their current income but rather adjust consumption to the average income over a period. The logic is simply that when utility functions are concave an (approximately) constant level of consumption lead to higher life-time utility than having consumption adjust up and down with temporary income fluctuations. To be explicit, assume that the income of a consumer is

$$y_{it} = y_{it}^p + e_{it} ,$$

where e_{it} is iid noise terms (in statistical language) interpreted (in economic language) as temporary income shocks (an inheritance, a lottery win, a spell of unemployment, etc.)

and y_{it}^p is the “permanent income” of consumer i . Friedman gave an *ad hoc* definition of permanent income—something like the typical income of a consumer when temporary components are removed. Only later was a rigorous definition developed by Hansen and Sargent (following the formulation of Hall) —the logic of the argument below will, however, hold for both the *ad hoc* and the rigorous definition of permanent income. Now assume that

$$c_{it} = \alpha + \beta y_{it}^p + v_{it} .$$

Since the e_{it} terms are temporary while the “permanent income” term y_{it}^p only varies slowly with time the average over time $\frac{1}{T} \sum y_{it} = \frac{1}{T} \sum y_{it}^p + \frac{1}{T} \sum e_{it}$, and (by the law of large numbers) $\frac{1}{T} \sum e_{it}$ becomes negligible when T is large. Friedman argued that taking the average over just 3 years would be a reasonable approximation (the next couple of weeks we will do all this more rigorously).

This explains why the slope of the consumption function is higher when one uses aggregate data at least if we assume, as seems reasonable, that the permanent income components y_{it}^p are *not* independent across individuals. (The part of peoples salaries that is not due to lottery wins, sabbaticals, sickness, etc. on average follows the business cycle.) In this case we have that aggregate consumption

$$\sum_{i=1}^N c_{it} = a + b \sum_{i=1}^N y_{it} + \sum_{i=1}^N u_{it}$$

where N is the (large) number of individuals in the economy. This implies (by the law of large numbers) that aggregate consumption per capita $c_t = \frac{1}{N} \sum_{i=1}^N c_{it}$ satisfies

$$c_t = a + b \frac{1}{N} \sum_{i=1}^N (y_{it}^p + e_{it}) + \frac{1}{N} \sum_{i=1}^N u_{it} \approx a + b \frac{1}{N} \sum_{i=1}^N y_{it}^p + w_t$$

where w_t is an aggregate error term and where the temporary income terms disappears due to the averaging over a large number of people. In conclusion, we will not expect downward bias due to measurement error in the regression

$$c_t = a + by_t + w_t ,$$

where $y_t = \frac{1}{N} \sum_{i=1}^N y_{it}^p$ is aggregate income, simply because in the calculation of aggregate income by averaging across individuals has killed off the temporary income terms.

Note: There are other reasons (having to do with “unit root econometrics”) why it is not good practise to regress the *level* of consumption on the *level* of income—so it is advisable to instead regress the consumption growth rate on the income growth rate—but that is a topic for a time series course.

In econometric terms y_{it} can be considered equal to permanent income plus a measurement error. It is easy to show that the OLS estimator is “biased towards 0” when the regressor is measured with error so we will do it explicitly:

Bias of the OLS estimator when the regressor is measured with error. Consider a regression model of form

$$y_i = \alpha + \beta x_i + u_i .$$

Under the standard OLS assumptions (x_i fixed, $E u_i = 0$, $E u_i u_j = 0$ when $i \neq j$ and constant variance of the u_i s) the efficient OLS-estimator of β (based on N observations) is

$$\hat{\beta} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

Now since

$$y_i - \bar{y} = \alpha + \beta x_i + u_i - (\alpha + \beta \bar{x} + \bar{u}) = \beta(x_i - \bar{x}) + (u_i - \bar{u}) ,$$

we have

$$\hat{\beta} = \frac{\sum (x_i - \bar{x})(\beta(x_i - \bar{x}) + (u_i - \bar{u}))}{\sum (x_i - \bar{x})^2}$$

or

$$\hat{\beta} - \beta = \frac{\sum (x_i - \bar{x})(u_i - \bar{u})}{\sum (x_i - \bar{x})^2} = \frac{\frac{1}{N} \sum (x_i - \bar{x})(u_i - \bar{u})}{\frac{1}{N} \sum (x_i - \bar{x})^2} .$$

For $N \rightarrow \infty$ we have $\frac{1}{N} \sum (x_i - \bar{x})(u_i - \bar{u}) \rightarrow 0$ and $\frac{1}{N} \sum (x_i - \bar{x})^2 \rightarrow \text{var}(x)$, so the right hand side converges to zero, i.e., the OLS estimator is *consistent* ($\hat{\beta} \rightarrow \beta$).

If x_i is measured with error this consistency result doesn't hold. Assume

$$x_i^* = x_i + e_i ,$$

where e_i is a “classical measurement error” where $E e_i = 0$, $E e_i e_j = 0; i \neq j$ and $E e_i u_j = 0; \forall i, j$. Now, if you regress y on x^* using the OLS formula, $\hat{\beta}$ will be biased towards zero; i.e. $E|\hat{\beta}| < E|\beta|$.

This is easy to demonstrate: We have

$$\hat{\beta} = \frac{\sum (x_i^* - \bar{x}^*)(\beta(x_i - \bar{x}) + (u_i - \bar{u}))}{\sum (x_i^* - \bar{x}^*)^2}$$

$$= \frac{\beta \frac{1}{N} \sum (x_i - \bar{x})(x_i - \bar{x}) + \beta \frac{1}{N} \sum (e_i - \bar{e})(x_i - \bar{x}) + \frac{1}{N} \sum (x_i^* - \bar{x}^*)(u_i - \bar{u})}{\frac{1}{N} \sum (x_i - \bar{x} + e_i - \bar{e})^2}$$

where the second and third terms in the numerator converges to 0 by the law of large numbers. We then have

$$\hat{\beta} \approx \beta \frac{\frac{1}{N} \sum (x_i - \bar{x})^2}{\frac{1}{N} \sum (x_i - \bar{x})^2 + \frac{1}{N} \sum (e_i - \bar{e})^2 + \frac{1}{N} \sum ((x_i - \bar{x})(e_i - \bar{e}))} \rightarrow \beta \frac{\text{var}(x)}{\text{var}(x) + \text{var}(e)}$$

This demonstrates that $\hat{\beta}$ converges to the true β times a term smaller than 1.