Recitation Notes 1

Konrad Menzel

September 18, 2006

1 Digression: Deriving the Slutsky Equation using Roy's Identity

As in the lecture, denote compensated and uncompensated demand for leisure as $l(p, w, \bar{y})$, and $l^c(p, w, \bar{u})$, respectively. Analogously to the other derivation, we start from

$$l(p, w, \bar{y}) = l^c(p, w, v(p, w, \bar{y}))$$

It turns out that it is more convenient to formulate this problem in terms of "excess" demand for leisure $l(p, w, \bar{y}) - T$ (i.e. beyond the initial time endowment of T hours), so that we don't have to worry about changes in the value of the full income $wT + \bar{y}$. In other words, we evaluate the consumer's trading options relative to an initial bundle $(l_0, x_0) = (T, \bar{y})$ when we allow her to trade leisure for the consumption good at a rate $\frac{w}{p}$, i.e. the real wage. Excess demand for leisure is equal to the negative of hours worked and must therefore be non-positive.
Taking derivatives with respect to $\bar{y}$, we get

$$\frac{\partial}{\partial \bar{y}} l(p, w, \bar{y}) = \frac{\partial}{\partial \bar{u}} l^c(p, w, v(p, w, \bar{y})) \frac{\partial}{\partial \bar{y}} v(p, w, \bar{y}) \qquad (*)$$

Differentiation with respect to w gives us

$$\begin{aligned}
\frac{\partial}{\partial w} l(p, w, \bar{y}) &= \frac{\partial}{\partial w} \Big[l(p, w, \bar{y}) - T \Big] = \frac{d}{dw} l^c(p, w, v(p, w, \bar{y})) \\
&= \frac{\partial}{\partial w} l^c(p, w, v(p, w, \bar{y})) + \frac{\partial}{\partial \bar{u}} l^c(p, w, v(p, w, \bar{y})) \frac{\partial}{\partial w} v(p, w, \bar{y}) \\
&\overset{\text{Roy's ID}}{=} \frac{\partial}{\partial w} l^c(p, w, v(p, w, \bar{y})) - \frac{\partial}{\partial \bar{u}} l^c(p, w, v(p, w, \bar{y})) \frac{\partial}{\partial \bar{y}} v(p, w, \bar{y}) \Big[l(p, w, \bar{y}) - T \Big] \\
&\overset{(*)}{=} \frac{\partial}{\partial w} l^c(p, w, v(p, w, \bar{y})) - \frac{\partial}{\partial \bar{y}} l(p, w, \bar{y}) \Big[l(p, w, \bar{y}) - T \Big]
\end{aligned}$$

Bringing the second summand to the left-hand side gives the Slutsky equation as derived in the lecture. We can also put this in terms of labor supply, where $h(p, w, \bar{y}) := T - l(p, w, \bar{y})$ is the number of hours worked, so that

$$\frac{\partial}{\partial w} h^c(p, w, v(p, w, \bar{y})) = \frac{\partial}{\partial w} h(p, w, \bar{y}) - \frac{\partial}{\partial \bar{y}} h(p, w, \bar{y}) h(p, w, \bar{y})$$

2 Measurement Error

In Labor Economics, we work a lot with survey data, which is often very messy, mainly for the following reasons:

1. the survey doesn't really measure the exact concept used in our economic framework (e.g. human capital)
2. subjects do not know the exact answer
3. response behavior is subject to a number of psychological biases (in particular if subjects don't understand the question well)

While there's not always an easy solution to this problem, the classical errors in variables framework is a good way of understanding about how bad data will affect the results of empirical analysis and suggests some ways to address the problem.

2.1 Notation

Want to estimate model / population regression

$$Y_i^* = \alpha + X_i^* \beta + \varepsilon_i$$

but the observed data is generated by the measurement equations

$$\begin{aligned} Y_i &= Y_i^* + \eta_i \\ X_i &= X_i^* + \nu_i \end{aligned}$$

Recall that for a bivariate regression of Y_i on X_i and a constant, the probability limit of the regression coefficient is

$$\text{plim}_N \hat{\beta}_N = \frac{\text{Cov}(X_i, Y_i)}{\text{Var}(X_i)} = \frac{\text{Cov}(X_i, \alpha + X_i^* \beta + \varepsilon_i + \eta_i}{\text{Var}(X_i)} = \frac{\text{Cov}(X_i, X_i^*)}{\text{Var}(X_i)} \beta + \frac{\text{Cov}(X_i, \varepsilon_i + \eta_i)}{\text{Var}(X_i)} \quad (1)$$

This holds mechanically without any further assumption on where the data actually comes from. The factor in front of the true coefficient,

$$\lambda := \frac{\text{Cov}(X_i, X_i^*)}{\text{Var}(X_i)} = \frac{\text{Var}(X_i) + \text{Cov}(X_i^*, \nu_i)}{\text{Var}(X_i^*) + \text{Cov}(X_i^*, \nu_i) + \text{Var}(\nu_i)}$$

is often referred to as the *reliability ratio*. If X_i^* and ν_i are uncorrelated, this coefficient gives the proportion of the variation in the observed X_i which actually consists of a "signal" about X_i^* (as opposed to the "noise" ν_i). With this basic formula, it is possible to assess the consequences of a wide range of possible forms of measurement error.
Reliability ratios λ of key variables (see Angrist and Krueger, 1999):

- cross-sectional earnings: ca. 0.7, annual changes in panel data: falls to 0.6
- educational attainment: ca. 0.9
- annual hours worked: about 0.6

Figure 1: Employer versus employee-reported log wages with OLS regression line (Source: Angrist and Krueger (1999), Handbook of Labor Economics ch. 23, p.1347)

The graph from Angrist and Krueger uses information on wages that was collected independently from employers and employees and matched for each employee later on. Since we have two independent measurements for the key variable, we are in a position to assess the overall quality of the data. If there wasn't measurement error in either variable, all points in the scatter plot should lie on the 45-degree line. You should also convince yourself that under the assumption of independent measurement errors the slope of the regression line (i.e. the regression coefficient) estimates the reliability ratio of employee-reported wages (which is on the x-axis). Running that regression after switching the role of the two measurements (i.e. switching the axis) will typically *not* give the inverse of the slope of the regression line, but we'd get back the reliability ratio of employer-reported wages, which will also be between zero and one. In this example, one can immediately see that the slope of the regression line is substantially less than 1.

2.2 Classical Measurement Error

With "classical" measurement error, we mean that η_i and ν_i are uncorrelated with each other and, most importantly, the true values X_i^* and Y_i^* (and therefore also uncorrelated with ε). From the expression for the probability limit of the OLS coefficient in (1), we can see that since under the classical errors in variables assumptions, ε_i and η_i are uncorrelated with X_i, the asymptotic bias of the OLS estimate reduces to $(1 - \lambda)\beta$.
You should actually try all the following variants of the measurement error problem by taking your favorite regression, and add some computer-generated noise to some variables or others in order to get a better intuition for the mechanics of the problem. In Stata you can add normally distributed random noise with standard deviation, say, 0.5 by typing

```
gen x = x_star + 0.5*invnorm(uniform())
```

2.2.1 Error in Dependent Variables Y_i

We can see right away from the formula for λ that if measurement error on Y_i is classical, and all other variables are correctly measured,

$$\lambda := \frac{\text{Cov}(X_i, X_i^*)}{\text{Var}(X_i)} = \frac{\text{Var}(X_i) + \text{Cov}(X_i^*, \nu_i)}{\text{Var}(X_i^*) + \text{Cov}(X_i^*, \nu_i) + \text{Var}(\nu_i)} = \frac{\text{Var}(X_i^*)}{\text{Var}(X_i^*)} = 1$$

so that there is no bias since $(1-\lambda)\beta = 0$. However, poor measurement of Y_i is equivalent to an increase in the variance of ε_i, and therefore increases the standard errors of the OLS coefficients.

2.2.2 Error in Explanatory Variables X_i

This is the most important case for practical purposes in empirical work. From the derivation above, we know that

$$\text{plim}_N \hat{\beta}_N = \lambda\beta$$

with

$$\lambda = \frac{\text{Cov}(X_i, X_i^*)}{\text{Var}(X_i)} = 1 - \frac{\text{Var}(\text{Var}(\nu_i))}{\text{Var}(X_i^*) + \text{Var}(\nu_i)} < 1$$

Therefore, with classical error in the regressor, the coefficient on the mismeasured variable is biased towards zero. This bias is commonly referred to as *attenuation bias*.
The basic intuition for this bias is relatively straightforward: assume we replace X_i^* with a computer-generated random number before we run the regression. We'd then expect the coefficient on that "junk" regressor to be zero. If we add some "signal" to that variable, we'd expect the coefficient to move in the direction of the true value of the parameter.
Since the biased estimate is equal to $\lambda\beta$, note that if we knew the signal-to-noise ratio of the mismeasured variable, we could construct λ, and correct our estimates. It turns out that most econometric solutions to the error-in-variables problem are based on this idea.

2.2.3 Problems with Additional Controls W_i

Now consider a small variation of the original model in which we also include another variable W_i in order to control for some part of the variation in Y_i, but aren't particularly interested in the coefficient on that control:

$$Y_i^* = \alpha + X_i^*\beta + W_i^*\gamma + \varepsilon_i$$

where W_i^* is also potentially mismeasured:

$$W_i = W_i^* + \zeta_i$$

We'd now like to reduce this problem to a bivariate regression in order to be able to fit it into our framework, and as you might recall from problem H on the review exercise, we can do that with a partitioned regression: regress Y_i on the residuals from a regression of X_i on W_i,

$$\tilde{X}_i := X_i - W_i\hat{\pi}_{XW} \equiv X_i - W_i\frac{\text{Cov}(X_i, W_i)}{\text{Var}(W_i)} = X_i - W_i\frac{\text{Cov}(X_i^*, W_i^*)}{\text{Var}(W_i)}$$

Note that for the coefficient estimate, it doesn't matter whether we orthogonalize Y_i as well (though it does matter when we compute standard errors). The probability limit of the OLS coefficient is now

$$\text{plim}_N \hat{\beta}_N = \frac{\text{Cov}(Y_i, \tilde{X}_i)}{\text{Var}(\tilde{X}_i)} = \frac{\text{Cov}(X_i^*\beta + W_i^*\gamma, X_i - W_i\hat{\pi}_{XW})}{\text{Var}(X_i - W_i\hat{\pi}_{XW})}$$

Now we can simplify the numerator to

$$\begin{aligned}
\text{Cov}(X_i^*\beta + W_i^*\gamma, X_i - W_i\hat{\pi}_{XW}) &= \Big[\text{Cov}(X_i^*, X_i) - \text{Cov}(X_i^*, W_i)\hat{\pi}_{XW}\Big]\beta + \Big[\text{Cov}(W_i^*, X_i) - \text{Cov}(W_i^*, W_i)\hat{\pi}_{XW}\Big]\gamma \\
&= \left[\text{Cov}(X_i^*, X_i) - \text{Cov}(X_i^*, W_i)\frac{\text{Cov}(X_i^*, W_i^*)}{\text{Var}(W_i)}\right]\beta \\
&\quad + \left[\text{Cov}(W_i^*, X_i^*) - \text{Cov}(W_i^*, W_i)\frac{\text{Cov}(X_i^*, W_i^*)}{\text{Var}(W_i)}\right]\gamma \\
&= \left[1 - R_{XW}^2\frac{\text{Var}(W_i^*)}{\text{Var}(W_i)}\right]\text{Cov}(X_i^*, X_i)\beta + \left[1 - \frac{\text{Var}(W_i^*)}{\text{Var}(W_i)}\right]\text{Cov}(X_i^*, W_i^*)\gamma \\
&= \left[1 - R_{XW}^2\frac{\text{Var}(W_i^*)}{\text{Var}(W_i)}\right]\text{Cov}(X_i^*, X_i^*)\beta + \frac{\text{Var}(\zeta_i)}{\text{Var}(W_i^*) + \text{Var}(\zeta_i)}\text{Cov}(X_i^*, W_i^*)\gamma
\end{aligned}$$

where R_{XW}^2 is the R-squared of a regression of W^* on X^*, and the denominator becomes

$$\begin{aligned}
\text{Var}(X_i - W_i\hat{\pi}_{XW}) &= \text{Var}(X_i) - 2\text{Cov}(X_i, W_i)\hat{\pi}_{XW} + \text{Var}(W_i)\hat{\pi}_{XW}^2 \\
&= \text{Var}(X_i) - \frac{\text{Cov}(X_i, W_i)^2}{\text{Var}(W_i)} = \left[1 - R_{XW}^2\frac{\text{Var}(W_i^*)}{\text{Var}(W_i)}\right]\text{Var}(X_i^*) + \text{Var}(\nu_i)
\end{aligned}$$

Now, in the case that there is no measurement error in W_i, we therefore get

$$\text{plim}_N\hat{\beta}_N = \frac{[1 - R_{XW}^2]\text{Var}(X_i^*)\beta}{[1 - R_{XW}^2]\text{Var}(X_i^*) + \text{Var}(\nu_i)} = \left[1 - \frac{\text{Var}(\nu_i)}{[1 - R_{XW}^2]\text{Var}(X_i^*) + \text{Var}(\nu_i)}\right]\beta$$

This means that in the presence of additional controls W_i, attenuation bias gets worse the stronger the true X_i^* is correlated with W_i^*, since the controls start picking up a larger share of the variation in Y_i that should actually have been attributed to X_i^*. In addition, we can say that adding controls will exacerbate attenuation bias unless the controls are uncorrelated with X_i^* - in which case we'd have no reason to control for them.
Now conversely, if there's no measurement error in X_i, but there is classical error in the controls,

$$\begin{aligned}
\text{plim}_N\hat{\beta}_N &= \frac{\left[1 - R_{XW}^2\frac{\text{Var}(W_i^*)}{\text{Var}(W_i)}\right]\text{Var}(X_i^*)\beta + \frac{\text{Var}(\zeta_i)}{\text{Var}(W_i^*)+\text{Var}(\zeta_i)}\text{Cov}(X_i^*, W_i^*)\gamma}{\left[1 - R_{XW}^2\frac{\text{Var}(W_i^*)}{\text{Var}(W_i)}\right]\text{Var}(X_i^*)} \\
&= \beta + \frac{\frac{\text{Var}(\zeta_i)}{\text{Var}(W_i^*)+\text{Var}(\zeta_i)}}{\left[1 - R_{XW}^2\frac{\text{Var}(W_i^*)}{\text{Var}(W_i)}\right]}\frac{\text{Cov}(X_i^*, W_i^*)\gamma}{\text{Var}(X_i^*)} \\
&= \beta + \frac{\text{Var}(\zeta_i)}{[1 - R_{XW}^2]\text{Var}(W_i^*) + \text{Var}(\zeta_i)}B_W
\end{aligned}$$

where B_W is the omitted variables bias that would obtain if we didn't control for W_i^* at all. Therefore, while measurement error in other controls doesn't cause attenuation bias on the coefficient of interest, the control absorbs only part of the omitted variables bias.

2.3 Group Means as Proxy Variables

Sometimes we are completely missing any microdata on a variable we would like to include in our analysis, but instead there's some aggregate data on that variable we'd like to use. More specifically, say we have

a microdata set without reliable information on wages X_i^*. Suppose we observe instead some covariates W_i (for instance dummies for type of work, industry, and specific metropolitan area), and we are able to obtain average wages for each relevant group from a different source. Then we might suspect after the previous discussion that using these group averages as proxies should be a source of attenuation bias (but the error $(X_i^* - \mathbb{E}[X_i|W_i])$ would have conditional mean zero, and therefore continue to be "classical"). It turns out that this isn't so (at least in the limit) since

$$\lambda = \frac{\text{Cov}\left(\mathbb{E}[X_i|W_i], X_i^*\right)}{\text{Var}\left(\mathbb{E}[X_i|W_i]\right)} = \frac{\text{Cov}\left(\mathbb{E}[X_i|W_i], \mathbb{E}[X_i^*|W_i]\right)}{\text{Var}\left(\mathbb{E}[X_i|W_i]\right)} = \frac{\text{Cov}\left(\mathbb{E}[X_i^*|W_i], \mathbb{E}[X_i^*|W_i]\right)}{\text{Var}\left(\mathbb{E}[X_i^*|W_i]\right)} = 1$$

so that there is no bias from measurement error in X_i. The intuitive reason for this is that using group means actually converts our estimation problem into a regression of group averages, for which the individual "measurement error" averages out after conditioning on W_i. We can also interpret this regression as an instrumental variables estimator which uses the group characteristics W_i as instruments, and therefore deals with the measurement error problem in X_i.

2.4 Classical Measurement Error in Panel Data

With panel data, our basic model becomes

$$Y_{it} = \alpha_i + X_{it}^*\beta + \varepsilon_{it}$$

where $X_{it} = X_{it}^* + \nu_{it}$, and $X_{it}^* = \tau X_{it-1}^* + \mu_{it}$ and $\nu_{it} = \varrho\nu_{i,t-1} + \xi_{it}$ follow AR(1) process with $\tau, \varrho \in (-1, 1)$, and we want to allow for α_i to be correlated with X_i^*.
One way of estimating this model is by running a pooled OLS regression of first differences $\Delta Y_{it} := Y_{i,t} - Y_{i,t-1}$ for $t = 2, .., T$ on $\Delta X_{it} := X_{it} - X_{i,t-1}$. The reliability ratio for the differenced data is

$$\begin{aligned}
\lambda &= \frac{\text{Cov}(\Delta X_{it}, \Delta X_{it}^*)}{\text{Var}(\Delta X_{it})} = \frac{\text{Var}(\Delta X_{it}^*)}{\text{Var}(\Delta X_{it}^*) + \text{Var}(\Delta\nu_{it})} \\
&= 1 - \frac{\text{Var}(\Delta\nu_{it}^*)}{\text{Var}(\Delta X_{it}^*) + \text{Var}(\Delta\nu_{it})} \\
&= 1 - \frac{(1-\varrho)\text{Var}(\nu_{it})}{(1-\tau)\text{Var}(X_{it}^*) + (1-\varrho)\text{Var}(\nu_{it})} \\
&= 1 - \frac{\text{Var}(\nu_{it})}{\frac{1-\tau}{1-\varrho}\text{Var}(X_{it}^*) + \text{Var}(\nu_{it})}
\end{aligned}$$

since

$$\begin{aligned}
\text{Var}(\Delta\nu_{it}) &= \text{Var}([\varrho - 1]\nu_{it-1} + \xi_{it}) = (1-\varrho)^2\text{Var}(\nu_{it}) + \text{Var}(\xi_{it}) \\
&= (1-\varrho)^2\text{Var}(\nu_{it}) + (1-\varrho^2)\text{Var}(\nu_{it}) = 2(1-\varrho)\text{Var}(\nu_{it})
\end{aligned}$$

where we use that

$$\text{Var}(\nu_{it}) = \text{Var}(\nu_{i0}) = \text{Var}\left(\sum_{s=-\infty}^{0} \varrho^s \xi_{is}\right) = \sum_{s=-\infty}^{0} \varrho^{2s}\text{Var}(\xi_{it}) = \frac{1}{1-\varrho^2}\text{Var}(\xi_{it})$$

This indicates that the estimator based on first-differences has more attenuation bias if and only if $\tau > \varrho$, i.e. if the "signal" is more correlated than the measurement error.
Alternatively, we can estimate the coefficients with a fixed effects regression. Since this is equivalent to

putting in individual-specific dummies, the formula for OLS with other controls derived above suggests that again, the bias must be worse than in pooled OLS. However, it also turns out that, if the panel is longer than 2 periods, if measurement error are less correlated than the true values of X_{it}^*, first differences shows a larger attenuation bias than fixed effects.
Actually, the difference between the biases in these two panel estimators can be used to recover the signal-noise ratio from which in turn we can reconstruct the correct λ, so that with panels of length greater than 2 periods, we can construct a consistent estimate in the presence of measurement error.[1]

2.5 Cures for Classical Measurement Error

2.5.1 Get Better Data

Sometimes the same data set allows for several different ways of constructing a variable of economic interest (e.g. instead of relying on responses on the number of weeks worked in a year, we could also subtract the number of weeks not worked from the total number of weeks). With several independent measures, we can also estimate reliability ratios, or decide which variables seem to be more appropriate for our purposes based on a priori considerations. Sometimes there are also ways of obtaining better quality data from other sources (this works particularly well for grouped data which only approximates some conditional expectation). By the reasoning in the section on grouped means as proxies, one could also think about replacing the poorly measured variable by a group average from a different data set.

2.5.2 Get the Reliability Ratio for the Messy Data

There may be information on the reliability ratio for a certain type of data from other sources/datasets or validation data. We can then simply apply the "trick" mentioned above of adjusting the estimate by dividing the attenuation bias out of our estimate: $\hat{\beta}_{adj} := \frac{\hat{\beta}_{LS}}{\lambda}$.

2.5.3 Instrumental Variables

In some cases, we observe a second - probably equally noisy - independent measurement of X_i^*,

$$Z_i = X_i^* + \xi_i$$

We can then use that measurement as an instrumental variable in a 2 stage least square procedure, which gives us a consistent estimate for the coefficient on X_i:

$$\text{plim}_N \hat{\beta}_{IV} = \frac{\text{Cov}(Z_i, Y_i)}{\text{Cov}(Z_i, X_i)} = \frac{\text{Cov}(X_i^*, Y_i)}{\text{Cov}(X_i^*, X_i)} = \frac{\text{Var}(X_i^*)\beta}{\text{Var}(X_i^*)} = \beta$$

Another way of thinking about what instrumental variables actually does in this case is going back to the relation between attenuation bias and the reliability ratio:

$$\text{plim}_N \hat{\beta}_{IV} = \frac{\text{Cov}(Z_i, Y_i)}{\text{Cov}(Z_i, X_i)} = \frac{\text{Cov}(Z_i, Y_i)/\text{Var}(Z_i)}{\text{Cov}(Z_i, X_i)/\text{Var}(Z_i)} = \text{plim}_N \frac{\hat{\beta}_{LS}}{\hat{\lambda}} = \frac{\lambda\beta}{\lambda} = \beta$$

Therefore, in the bivariate case the IV procedure is equivalent to first estimating the reliability ratio of the regressor by regressing the alternative measure on it ("first stage" regression), and then adjust the biased estimate from the "second stage" regression of Y_i on X_i by dividing through the estimated reliability ratio (in IV terminology "backing out the structural form").

[1] see Grilliches and Hausman (1984): "Errors in Variables in Panel Data"

Good choices for instrumental variables aren't limited to other measurements of the noisy regressor in a narrow sense, but in principle we can use any variable which is to some degree correlated with the signal - $\text{Cov}(Z_i, X_i^*) \neq 0$ - but not the noise or measurement error (you should convince yourself that the reasoning above still goes through). In the oldest version of this type of estimator, the instrument consisted actually of a single dummy variable according to which the sample was split in two groups, one of which had a higher average value for the mismeasured variable than the other. The "grouped means as proxies" estimator described above is another example for an IV estimator, using group characteristics as instrumental variables for the regressor of interest.

2.5.4 Panel Data

As a short note, there is a paper by Grilliches and Hausman (1984):"Errors in Variables in Panel Data", which you will see / will have seen in 14.382 which exploits the difference between the first-differences and the fixed-effects panel estimators to correct the bias of either estimator of β.

2.6 "Non-Classical" Measurement Error

2.6.1 Error in Limited Dependent Variables

This section is only supposed to give an important caveat: if the mismeasured variable doesn't have full support on the real line (e.g. wages have to be positive - though log wages don't have to), the classical error in variables assumptions aren't plausible for data close to the boundaries. For instance, if we look at a binary variable - e.g. whether a person has a college degree or not - the measurement error (the term used for discrete variables is *misclassification*) can't possibly be uncorrelated with the true value: if the true value of the dummy is "1", measurement error can only consist in reporting "0" instead, and vice versa, so that the measurement error is negatively correlated with the truth.
There are many more technical econometric papers that propose solutions to this problem.

2.6.2 Division Bias (Borjas, 1980)

Sometimes, an explanatory variable used in the analysis is a function of our data on the dependent variable. Suppose we want to estimate (uncompensated) labor supply elasticity with respect to hourly wages (which we assume to be exogenous for now)

$$\log h_i^* = \alpha + \beta \log w_i + \varepsilon_i$$

Unfortunately, we only have information on total income y_i and hours worked h_i for all individuals in our dataset. We could therefore run an alternative regression

$$\log h_i = \alpha + \beta[\log y_i - \log h_i] + \varepsilon_i$$

If all variables are measured accurately,

$$\log y_i - \log h_i = \log(w_i h_i^*) - \log h_i = \log w_i + \log h_i^* - \log h_i^* = \log w_i$$

so that the fact that our dependent variable also appears on the right-hand side of the equation doesn't represent a problem per se. However, if we allow for (classical) error in variables in h_i,

$$\log h_i = \log h_i^* + \eta_i$$

we actually end up estimating the equation

$$\log h_i = \alpha + \beta[\log y_i - \log h_i^* - \eta_i] + \varepsilon_i + \eta_i = \alpha + \beta(\log w_i - \eta_i) + \varepsilon_i + \eta_i$$

so that the OLS coefficient plims to

$$\begin{aligned}\text{plim}_N \hat{\beta}_N &= \frac{\text{Cov}(\log w_i - \eta_i, \log h_i)}{\text{Var}(\log w_i - \eta_i)} = \frac{\text{Cov}(\log w_i - \eta_i, \log w_i \beta + \varepsilon_i + \eta_i)}{\text{Var}(\log w_i) + \text{Var}(\eta_i)} \\ &= \frac{\text{Var}(\log w_i)\beta - \text{Var}(\eta_i)}{\text{Var}(\log w_i) + \text{Var}(\eta_i)} = \lambda\beta - \frac{\text{Var}(\eta_i)}{\text{Var}(\log w_i) + \text{Var}(\eta_i)}\end{aligned}$$

so that, in addition to classical measurement error, we have a negative bias from the fact that the measurement error affects both the dependent and the main explanatory variable. The problem here is that, other than in the classical errors in variables model, the measurement error in the regressor (log wages) is mechanically related to the error in the dependent variable. I.e. if we measure fewer than actual hours worked, we overestimate hourly wages and will therefore get an estimate of the labor supply elasticity which is biased downward.

References

[1] Angrist, J., and A. Krueger **(1999):** "Empirical Strategies in Labor Economics, *Handbook of Labor Economics 3*, ch.**23**

[2] Borjas, G. **(1980):** The Relationship between Wages and Weekly Hours of Work: the Role of Division Bias, *Human Resources* **15(3)**, 409-23

[3] Bound, J, Ch. Brown, and N. Mathiowetz **(2001):** Measurement Error in Survey Data, *Handbook of Econometrics* ch.**59**

[4] Grilliches, Z., and J. Hausman **(1984):** Errors in Variables in Panel Data, *NBER Technical Paper* **t37**