Controlling for Time Invariant Heterogeneity

Yona Rubinstein

July 2016

Observables and Unobservables Confounding Factors

- The key problem we face in estimating *causal models* is that firms, individuals, governments and nations: (i) make policy *choices*; (ii) these *choices* are **correlated** with latent factors that affect the outcomes of interest - such as profits, wages health etc. - regardless of their choices.
- So far the key to causal inference was to control for observed confounding factors; that is controlling for "*observables*".
- For example we use measures of schooling to control for the potential selection into the financial sector on workers' productivity.
- Yet, what if important confounders are **unobserved**?
- In this note we start our journey to a variety of strategies to control a bias on "*unobservables*".
- In this lecture we consider econometric strategies that use data with a time dimension to **control** for *unobserved time* invariant omitted variables.

The Structural Model

- To illustrate let's consider the example of the finance wage premium.
- Specifically, let Y_{it} equal the (in logs) wages of worker i at time t and let F_{it} denote whether she works in finance $(F_{it} = 1)$ or outside of the financial sector $(F_{it} = 0)$.
- Let's further assume, for simplicity of illustration, that the *causal model* exhibits the following functional form:

$$Y_{it} = \beta_0 + \beta_F F_{it} + U_{it} \tag{1}$$

- The *causal impact* of finance on log hourly wage is **common** and equals to β_F.

The Structural Model (cont.)

- U_{it} stands for person-year influences on wages that the econometrician **does not observe** in **her data** and therefore are *omitted* from the model.
- This **aggregated** *omitted variable* (U_{it}) combines many factors that can be decomposed into two sources:
  - First: *time invariant components* - such as cognitive (C_i) and non-cognitive traits (NC_i) - that influence earnings;
  - Second: *time variant components* - **zero mean** *person-specific* "shocks" to wages (ε_{it}) that might **fluctuate** from *year* to *year*.
- Therefore, we can express U_{it} as:

$$U_{it} = \beta_C C_i + \beta_{NC} NC_i + \varepsilon_{it}, \tag{2}$$

- β_C and β_{NC} are the *causal impacts* of **cognitive** and **non-cognitive** *traits* on log hourly wage rates (Y_{it}) respectively.
- Important: note that we implicitly assume that the impact of cognitive and non-cognitive traits is the same for all type of jobs!!

Time Variant and Time Invariant Unobservables

- Note that C_i and NC_i are *unobserved* by the econometrician; yet these are **observed** (at least partially) by employers and **influence wages**.
- Assuming that the impact of cognitive and non-cognitive traits is the same for all type of jobs – that is γ_C and γ_{NC} do not vary across jobs – we can sum all unobserved time invariant factors into one aggregate person specific *time invariant* component θ_i – known as person *fixed effect*:

$$\theta_i = \beta_C C_i + \beta_{NC} NC_i. \tag{3}$$

- The parameter θ_i, person's i *fixed effect*, captures *ALL* **person-specific** *time invariant* factors that influence wages (or any other outcome) and are *unobserved* by the econometrician.

Individual Fixed Effects and Shocks

- We can now express U_{it} as a combination of (i) person fixed effect (θ_i) (ii) and person-specific fluctuations (ε_{it}) around her mean:

$$U_{it} = \theta_i + \varepsilon_{it}. \tag{4}$$

- Note that **by construction** the sum of ε_{it} for each work equals to zero.
- Yet, this **does not** mean that these shocks sum to zero when we add them for each person separately (i) during the years that they work in the financial sector and (ii) when they work in all other sectors.
- That is, while

$$\varepsilon_i = 0,$$

it might be that:

$$\begin{aligned} \varepsilon_{0i} &\neq 0 \\ \varepsilon_{1i} &\neq 0 \end{aligned}$$

The OLS Regression Coefficient: Confounding Treatment and Selection Effects

- A key concern is that individuals with better cognitive and non-cognitive traits, "better" fixed effect, are more likely to work in the finance industry than others.
- To illustrate consider identifying the *causal impact* of finance by estimating the following model using OLS:

$$Y_{it} = b_0 + b_F F_{it} + e_{it}. \tag{5}$$

- Note that in this case we use b_F, the OLS regression coefficient, to infer regarding the causal impact of finance on log hourly wages.
- Therefore, when we aim at estimating causal relationships using OLS we refer to the OLS regression coefficients as OLS estimates for the causal parameters
- We denote those by $\hat{\beta}^{OLS}$
- So in our case b_F is $\hat{\beta}_F^{OLS}$.

The OLS Regression Coefficient: Confounding Treatment and Selection Effects

- Assuming that we observe the population of workers then the estimated regression coefficient b_{F}. provides the gap in mean log hourly wages between those who work in the financial industry and all other workers.
- The gap in mean log hourly wages, the OLS estimate for the causal impact of working in the financial sector on log hourly wage $\left(\hat{\beta}_F^{OLS}\right)$, can be decomposed into three factors:

$$
\begin{array}{ccccccccc}
b_F & \equiv & \hat{\beta}_F^{OLS} & = & \beta_F & + & (\theta_1 - \theta_0) & + & (\varepsilon_1 - \varepsilon_0) \\
 & & & & \Downarrow & & \Downarrow & & \Downarrow \\
 & & & & \text{ATE} & + & \text{(selection on "fixed effects")} & + & \text{(selection on "shocks")}
\end{array} \tag{6}
$$

- I elaborate that in the next slides.

Selection on Unobservables

- The regression coefficient b_F equals to:

$$b_F = \frac{cov\left(Y_{it}, F_{it}\right)}{var(F_{it})} \tag{7}$$

- Let's express Y_{it} in terms of the causal model, that is as $Y_{it} = \beta_0 + \beta_F F_{it} + \theta_i + \varepsilon_{it}$.
- In that case:

$$b_F = \frac{cov\left(Y_{it}, F_{it}\right)}{var(F_{it})} = \frac{cov\left(\left(\beta_0 + \beta_F F_{it} + \theta_i + \varepsilon_{it}\right), F_{it}\right)}{var(F_{it})} \tag{8}$$

- The above also equals to:

$$b_F = \beta_F + \frac{cov\left(\theta_i, F_{it}\right)}{var(F_{it})} + \frac{cov\left(\varepsilon_{it}, F_{it}\right)}{var(F_{it})} \tag{9}$$

- Any intuition about the terms above?

Selection Bias

- Note that:

$$\frac{cov(\theta_i, F_{it})}{var(F_{it})} = (\theta_1^1 - \theta_0^0) = (\theta_1 - \theta_0) \tag{10}$$

$$\frac{cov(\theta_i, F_{it})}{var(F_{it})} = (\varepsilon_1^1 - \varepsilon_0^0) = (\varepsilon_1 - \varepsilon_0)$$

- The first term $(\theta_1^1 - \theta_0^0)$: the gap between the average *fixed effects* of workers in finance and the average fixed effects of all others.
- The second term $(\varepsilon_1^1 - \varepsilon_0^0)$: the gap between the average wage *person-specific shocks* of workers when working finance and the average *person-specific shocks* when working in all other industries.
- By substituting the above into (9) we obtain that:

$$\begin{array}{ccccccccc} b_F & \equiv & \hat{\beta}_F^{OLS} & = & \beta_F & + & (\theta_1 - \theta_0) & + & (\varepsilon_1 - \varepsilon_0) \\ & & & & \Downarrow & & \Downarrow & & \Downarrow \\ & & & & \text{ATE} & + & \text{(selection on "fixed effects")} & + & \text{(selection on "shocks")} \end{array} \tag{11}$$

Controlling for Selection on Fixed Effects

- Is it possible to estimate the causal impact of finance on wages (β_F) in observational data?
- Yes it is! (or yes we can!)
- The key *structural assumptions* which allow us to condition out selection bias are:
  - that the unobserved C_i and NC_i are indeed fixed.
  - person-specific wages shocks are zero mean independent from their sector of employment.
- If these assumptions hold we can identify the causal impact of working in the financial sector on log hourly wages (β_F) in observational panel data, that is using a repeated observations on individuals.
- Specifically, there are two ways:
  - **Fixed Effects Model**
  - **Differencing** from period (year) to the next.

The Fixed Effect Model

- Let's consider that we have *panel data*, i.e., repeated observations on individuals.
- Let's further assume that the *causal model* exhibits the following functional form:

$$Y_{it} = \beta_0 + \beta_F F_{it} + \delta t_{it} + U_{it}, \qquad (12)$$

- δ reflects the effect of time (t) on log hourly earnings. If time has no effect then $\delta = 0$.
- The error term U_{it} is a combination of person fixed effect (θ_i) and person-specific fluctuations (ε_{it}) around her mean:

$$U_{it} = \theta_i + \varepsilon_{it}. \qquad (13)$$

- The *causal impact* of finance on log hourly wage is **common** and equals to β_F.
- We are interested in estimating the causal impact of finance on log hourly wages.
- The *parameter of interest* is β_F.

The Fixed Effect Model (cont.)

- We can re-write the model in equation (12) in terms of deviations from each work mean.
- The mean log hourly of each work is given by:

$$Y_i = \beta_0 + \beta_F F_i + U_i. \tag{14}$$

- By subtracting this mean from we obtain for each individual her hourly wage in year t relative to her life time average:

$$Y_{it} - Y_i = (\beta_0 - \beta_0) + \delta\,(t_{it} - t_i) + \beta_F\,(F_{it} - F_i) + (\theta_i - \theta_i) + (\varepsilon_{it} - \varepsilon_i) \tag{15}$$

- Note that three terms cancel: (i) the constant, (ii) persons' fixed effects and (iii) the mean shock. Since the term $(t_{it} - t_i) = 1$ as we move from one year to the next:

$$Y_{it} - Y_i = \delta + \beta_F\,(F_{it} - F_i) + \varepsilon_{it}, \tag{16}$$

- Hence, the deviations from means eliminates the unobserved individual fixed effects!
- The above (16) is known as the *Fixed Effects Model*.

Estimating Fixed Effects Model using OLS

- The PRF that corresponds to the fixed effects model is equation (16) is:

$$Y_{it} - Y_i = b_0 + b_F \left(F_{it} - F_i \right) + e_{it}, \tag{17}$$

- The OLS estimate of the Fixed Effect Model, that is b_F^{FE}, equals to:

$$\begin{array}{ccccccc} b_F & = & \hat{\beta}_F^{FE} & = & \beta_F & + & (\varepsilon_1 - \varepsilon_0) \\ & & & & \Downarrow & & \Downarrow \\ & & & & \text{ATE} & + & \text{(selection on "shocks")} \end{array} \tag{18}$$

- Assuming no selection on wage shocks, that $(\varepsilon_1 - \varepsilon_0) = 0$, then the *Fixed Effects* model provides an unbiased estimate of the causal impact of working in the financial industry on log hourly wages:

$$\hat{\beta}_F^{FE} = \beta_F. \tag{19}$$

When Should We Use Fixed Effects Model?

- The Fixed Effects Model, hereafter FE, has pros and cons.
- It is a useful tool if the following holds:

1. *Selection on person fixed effects*: for instance, the more able college graduate are more likely to be employed in the financial sector.
2. *No selection on unobservable that might change over time*: for instance, workers are not switching to finance in their productive years and out of finance when they are less productive (effort?).
3. *Accurate measures of treatment status*: for instance, changes in industry indicators reflect a real switch from one industry to another rather than measurement errors.

- When the above holds, the FE model provides an unbiased estimate of the average treatment effect (on the treated!).
- Users should keep in mind that (2) and (3) are not trivial conditions.

First Differences

- Alternatively we could condition out person fixed effect by differencing two consecutive periods.
- By subtracting last year from this year outcomes from we obtain for each individual the change in her hourly wage in year t:

$$Y_{it} - Y_{it-1} = \delta + \beta_F \left(F_{it} - F_{it-1} \right) + \left(\varepsilon_{it} - \varepsilon_{it-1} \right), \tag{20}$$

- Which can also be writyen as:

$$\Delta Y_{it} = \delta + \beta_F \Delta F_{it} + \Delta \varepsilon_{it} \tag{21}$$

- By estimating equation (21) using OLS we obtain:

$$\begin{array}{cccccc} \hat{\beta}_F^{DIF} & = & \beta_F & + & \Delta \varepsilon_{i,switch} \\ & & \Downarrow & & \Downarrow \\ & & \text{ATE} & + & \text{(the average change in} \\ & & & & \text{"shocks" when switching)} \end{array} \tag{22}$$

Fixed Effects vs. First Differences

- At first glance both methods look identical. In fact when we have two observations per unit the two models are identical.
- Yet, the methods differ when we follow observations over more than two periods.
- In this case the FE differs from DIF on two **key issues**:

1. *Measurement errors* in the exact timing of *treatment* (F_{it}): when the exact timing of treatment is measured with errors, for instance, recorded a year later than it actually happened.
2. The *exogeneity* of treatment to outcome shocks (the ε_{it}): the extent that treatment is correlated with past and future shocks.

- The FE method performs better than DIF when the timing of treatment is measured with errors
- The FE method requires stronger assumptions regarding the correlation between the outcome shocks and the treatment than DIF.

Fixed Effects vs. First Differences (cont.)

- *Measurement errors* in the exact timing of *treatment* (F_{it}):
  - The DIF method is very sensitive to measurement errors in the timing of treatment since it identifies the effect from the change in outcomes when treatment presumably started;
  - The FE method compares the conditional mean of all periods after with the conditional mean of all periods before.
  - Therefore the latter is more robust to measurement errors in treatment timings.
- The *exogeneity* of treatment to outcome shocks (the ε_{it}):
  - the FE method requires that the treatment status in period t to be **uncorrelated** with **past**, **present**, and **future** *outcome shocks*. This is known as *strict exogeneity*.of treatment.
  - The DIF method requires treatment to be **uncorrelated** with current and lagged values of the outcome shocks. This is known as *exogeneity*.of treatment.
  - The latter is a sub set of the first requirement and therefore a weaker restriction than *strict exogeneity*.

Take Home Message

- We can use panel data to control for selection on time invariant latent factors.
- The repeated observations allow us to difference out the fixed effects.
- Since the *FE* method uses all periods before, during and after treatment it is less sensitive to the *DIF* method to measurement errors in the exact timing of treatment.
- Yet, for the same reason the *FE* method requires stronger assumption regarding the correlation between future, current and past outcome shocks and the treatment status.
- In the next set of slides we'll practice these methods in the context of the finance wage premium.