Why MLR?

There are two main reasons for using multilevel regression models (sometimes referred to as multilevel models, random coefficient models, or mixed models). The first is purely statistical. Ordinary least square regression assumes that observations (errors really) are independent of one another. So, whenever data points are clustered together (e.g., family members, geographic regions, coworkers, patients), the independence assumption will be violated. Its violation leads to underestimation of standard errors and p-values (i.e., too many Type I errors).

Barcikowski (1981) showed that even a very small intraclass correlation coefficient, a measure of clustering, can lead to alpha rates above the nominal value.

Multilevel models were developed to take clustering into account and they estimate regression models without inflation of Type I errors.
A second rationale of multilevel models is theoretical. Researchers have always been interested in phenomena that occur at different theoretical levels. Psychologists may be interested students in the classroom social environment. Sociologists, for example, are interested in neighborhoods as well as individual residents. Economists have theories about work environment as well as individual worker productivity. Health researchers are not only concerned about individual patient care but also the health care system’s impact on health.

Multilevel modeling allows researchers to investigate the interconnections between these micro and macro levels.

**Micro to micro**

```
X ---------> Y
```

**Macro to macro**

```
Z ---------> Y
```

**Macro and micro to micro**

```
Z
  ---------> Y

X ---------> Y
```

**Micro to micro depends on macro**

```
Z
  ---------> Y

X ---------> Y
```

---

Figure concept borrowed from Snijders & Bosker (1999; 2012).