Internally valid treatment effect studies cannot be used out-of-the-box to predict the effect of the treatment in a new context. Policymakers often, however, only have treatment effect data from other places and times and may wish to use this data to predict the effect of a policy in their own context.

Assume there exists a conditional treatment distribution, $P(\tau|\,\omega)$, that is stable across time and space. ω could consist of, for example, the parameters of the underlying data generating process. Although this ω might exist, very often we both do not know what variables it consists of and/or cannot even measure those variables if we did know them.

Thus, it is more helpful to ask questions about an alternative set of variables, $\mathbf{x}$, to which we do have access. We can also consider transforming $\mathbf{x}$ via a function $f(\cdot)$. By definition:

$$P(\tau|f(\mathbf{x})) = \int P(\tau|f(\mathbf{x}),\mathbf{z})P(\mathbf{z}\,|f(\mathbf{x}))\;d\,\mathbf{z}$$

For some unobserved variables $\mathbf{z}$. We can consider the follow possibilities:

1. If $f(\mathbf{x})$ directly recovered ω (for instance, when $\omega \subset \mathbf{x}$ and $f(\cdot)$ is a variable selection function), then $P(\tau|f(\mathbf{x}))$ would be invariant across time and space and could be used for prediction in any context after being observed in any other context. This characterizes the problem and technique described in Rojas-Carulla et al. 2018.

2. If $f(\mathbf{x}) \subset \omega$, then we can characterize $\mathbf{z} \coloneqq \omega \setminus f(\mathbf{x})$, and we know that $P(\tau|f(\mathbf{x}),\mathbf{z})$ will be stable across all time and space and our measured function, $P(\tau|f(\mathbf{x}))$, will be stable across the time and space over which $P(\mathbf{z}\,|f(\mathbf{x}))$ is stable.

3. If ω consists entirely of latent variables (as often must be the case in any social science problem where the prediction consists of an individuals choice whose "true" causal factors must be said to lie deep in their minds), then by definition $\mathbf{x} \cap \omega \neq 0$. Thus, $f(\cdot)$ should not be considered a variable selection problem, but rather a latent variable recovery problem. In this case, if $f(\mathbf{x}) \subset \omega$ we have the previous scenario.