Consider a sample of N individuals, with J judges. Each judge is assigned n_j individuals, with $\sum_j n_j = N$.

Define the first-stage outcome (dismissal) as X_i for person i. Let Judge_i denote the judge that is assigned to person i.

Then, let $1(\text{Judge}_i = j)$ be an indicator for person i's judge equal to j. Hence, mechanically, $\sum_{j=1}^{J} 1(\text{Judge}_i = j) = 1$.

Consider that the judges are randomly assinged. We want to use this in a first-stage regression of the following:

$$X_i = \pi_0 + \sum_{j=1}^{J-1} \pi_j 1(\text{Judge}_i = j) \tag{0.1}$$

Note that mechanically, π_0 will be EXACTLY equal $N^{-1} \sum_i X_i 1(\text{Judge}_i = J)$, or the average (non-leave-one-out) mean leniency. Similarly, pi_j is just $N^{-1} \sum_i X_i 1(\text{Judge}_i = j) - N^{-1} \sum_i X_i 1(\text{Judge}_i = J)$, or the relative leniency.

This is $J - 1$ instruments!

Our predicted values for X_i is just $\bar{X}(j)$, the simple average for a given judge j who was assigned to i.

Now consider that we are overidentified. If we implelemented jackknife IV (Imbens Angrist Krueger), our predicted values would be our predicted values for X_i still uses the SAME approach, but now "leave-outs" our own observation.

Finally, in Dobbie Goldsmith-PInkham and Yang + other leniency appraoches, they diference out their "own-location", because the first stage result controls for fixed effects. E.g., consider α_l, a location fixed effect that is necessary for identification. Consider the following first stage:

$$X_i = \pi_0 + \sum_{j=1}^{J-1} \pi_j 1(\text{Judge}_i = j) + \alpha_l \tag{0.2}$$

Now, consider the residual regression – projecting onto α_l will subtract the own location averages (across all judges within that location).