Jerome has always been a good target – for corr anyway. My main model from the v3 data used Jerome as the primary target (but not the only one) and didn't use nomi at all. And interestingly, that model was always good/consistent at corr and bad/mediocre at TC (probably averaged zero TC, or slightly negative). Retired now so I don't know how it would have done on recent data.

Still, it isn't that weird that things don't stay the same, or they don't behave the way they did in training – the relationships encoded (inherently) are non-stationary. (I'm guessing, not knowing how they are made.) If neutralizing risk involves assumptions about this being related to that in a certain way…well, that may not stay the same. I'm wondering if some targets are more complex than others which also makes them more brittle (so vanilla methods may be more robust). Is a target that is "less risky" (supposedly) also counting on more relationships between features/returns that need to hold in order for that to be true?