SMS STRATEGIC HUMAN CAPITAL INTEREST GROUP SEMINAR ON EMPIRICAL METHODS: EDITED CHAT FROM THE ZOOM

QUESTIONS TO EVAN STARR (Panel data methods)

Question 1. Denisa Mindruta: You started by saying that we rarely plot our data---at least in the published papers. Two questions 1) do you recommend this in any published paper appendix or rather for examining the data before we do any analysis? 2) and should it be part of how we evaluate papers as reviewers/ readers?
Answer. Good questions. On 1) I'm very comfortable with abductive work where we look at the data and figure out what's going on and let that drive subsequent analyses (if one has big data they could split the data, develop H's in one, and test in the other). What I'm not ok with is pretending we developed the theory first when we really didn't. On 2) It always gives me more confidence as an editor to see the binned scatterplots as opposed to plots of fitted values. It's just more complete and harder to game.

Question 2. Sekou Bermiss: What is a command using which stats software?
Answer. FYI: Here's the link to the "binsreg" command in Stata
https://nppackages.github.io/binsreg/

Question 3. Thomaz Teodorovicz: How should we interpret the new difference-in-difference (DiD) literature when using only never-treated as control groups (e.g. limit sample to windows around the treatment in the treatment group and match each treated unit to a never-treated control unit). Think about the famous million-dollar plant paper by Greenstone et al
Answer. The "never treated" group helps immensely because it is a "clean control". The problems mostly arise from comparing later treated groups to earlier treated groups.

Question 4. Thomaz Teodorovicz: Related to above, what if we focus on estimating dynamic effects (panel-data event-study/dynamic DiD), rather than a single Beta parameter?

Answer. Yes, you're still in trouble here. The reason is because in some cases it will look like the control is growing faster than the treatment, causing you to get the wrong sign on the event study.

Question 5. Denisa Mindruta: If we do this purging--instead of modeling somehow the staggered adoption--would this be forcing us to limit our theory development? (I could be wrong about what purging means)

Answer. I'm not sure what you mean by modeling. The Goodman-Bacon decomposition theorem says that when you run OLS with two-way fixed effects, you get a weighted average of all the 2x2 diff-in-diff coefficients. When you have heterogeneous effects, or effects that grow over time, some of those 2x2 diff-in-diff estimates are going to be severely biased. So the "purging" is about throwing away the 2x2 diff-in-diff estimates that cause the bias. It has nothing to do with "theory" per se and does not limit theory in any way.

Question 6. Rhett Brymer: All your examples yield underestimates of the true effect. Is this always the case? Or can non-purging also create issues of overestimation of the effects?

Answer. Yes, it can cause underestimation. It depends on who has the largest effects, whether they grow over time, and when each unit is adopted.

Question 7. Thomaz Teodorovicz: What about bin scatter vs. lowess plots?

Answer. Binned scatterplots can incorporate smoothing within and across bins, or be fully non-parametric….so in my view binned scatterplots > lowess (since lowess by default smooths).

QUESTIONS TO RORY ECKARDT (multi-level modeling)

Question 8. David Kryscynski: Forgive me if I am misunderstanding, but it seems that you are modeling the relationship between KSAOs and performance with different assumptions about what that relationship looks like, but it seems that the "emerged" HCR is not directly measured. Do you have insights into how to actually measure the HCR more directly?
Answer. These modeling approaches are focused on capturing the part of the HCR that has emerged to a higher level of analysis. [see discussion in the video for additional detail on this answer].

Question 9. Rhett Brymer: You may be getting to this, but I'm wondering if there are theoretical reasons to choose any one of these four different empirical approaches.
Answer. The differences between the modeling approaches likely are more methodological and stem from the nature of data one has access to in a particular study context.

Question 10. Clinton Chadwick: Rory, it seems like emergence is defined by positive complementarities (2+2>4), but is there room in your thinking for interactions amongst system elements that reduce an outcome of interest (i.e., dis-synergies)?
Answer. Yes, and I think this is an interesting and important area for future work. This is noted in Eckardt et al. (2021 JOB) with regard to the potential impact of negative unit affect. Brymer and Hitt (2019) also discuss the notion of dis-complementarities.

Question 12. Federica De Stefano: +1 on Clint: do these methodologies also allow us to look at a case in which the combination of individual KSAOs not only leads to "no emergence" but even to depletion of value?

David Kryscynski: I also second Clint's question on "dis-synergies"

Answer. The dis-synergies should show up as reductions to the intercepts and/or slopes of the KSAO-performance association and should be captured using many of these techniques as well.

Question 12: Denisa Mindruta: You don't have to address this now (it may not be a quick clarification question) but is there a discussion in the literature about "ecological fallacy" as we make inferences based on a multi-level analysis?
Answer. Rousseau (1985) discusses ecological fallacies for multilevel approaches in general and Ployhart and Moliterno (2011) also discuss the two common types of concerns (cross-level fallacies and contextual fallacies) in this area and relate it to HCR. Dansereau and Yammarino also discuss this in their work on the within and between analysis (WABA) multilevel technique.

Question 13. Sekou Bermiss: Most of these models are based on a very precise measure of individual KSAOs. I am curious how often people are finding this data (i.e., proprietary firm data)?
Answer. It can be hard to find the detailed individual and group level data needed to look at emergent phenomena. This is a point Steve Kozlowski has noted previously and is one of the reasons why we likely see so little empirical research on emergence. Nonetheless, I suspect this will get better over time with increased availability of big data, wearable sensors, and digital traces. Steve Kozlowski has a number of good articles that talk about the challenges of data collection with emergence and some suggested solutions.

Question 14. Denisa Mindruta: Can someone quickly clarify why do these analytical models use the terms "emergence" and "enabling"
Answer. This is in reference to the emergence enabling mechanisms discussed by Ployhart and Moliterno (2011).

Question 15. Evan Starr: Should we think of these as largely correlational exercises? How do we think about counterfactuals and identifying assumptions in these complicated models?
Answer. There is certainly a need for more work on potential endogeneity issues in multilevel methods. Dan McNeish and John Antonakis have some papers on this, but much more work is needed.

Question 16. Federica De Stefano: Do you expect audiences coming from different traditions (i.e. HR and Strategy) will have different opinions on the most appropriate model approach?

Answer. It would not surprise me if macro-oriented researchers prefer the Cobb-Douglas modeling approach as it emphasizes outcomes that uniquely reside at a higher level of analysis.

Question 17. Clinton Chadwick: Because they're based on individual HC, it seems like these techniques are more suited to picking up the "amplification" phenomenon than transformation of HC, especially collective HCRs.

Answer. These techniques emphasize the amplification aspect as it is most consistent with the general notion of emergence and also has the potential to be most relevant from a strategic standpoint. These points are discussed in more detail in Eckardt and Jiang (2019) and Eckardt, Crocker, and Tsai (2021 IJHRM). The Eckardt and Jiang (2019) chapter also discusses a multilevel technique to capture the transformation aspect of HCRE.

Question 18. Abhi Moulick: What software are you using for ABM?
Answer. We are using R.

Question 19. David Kryscynski: Individuals come together to form a group level resource. The group level resource has some potential to create value. But, there is a difference between the group potential to create value and the extent to which it actually creates value (shoutout to Clint's 2017 AMR). These multilevel methods seem to be using actual observed performance to back into insights about the complementarities, but might not be capturing the underlying potential in the group HCR.

Clinton Chadwick: DK, I was thinking the same thing. What these methods can do is infer HCRs from the existence of performance synergies (and I agree that these methods are improvements over what we've been doing). But one would assume that there are HCRs present in some circumstances, but they don't lead to superadditive outcomes. i.e., I don't know how to directly measure HCRs empirically, so we are inferring them from the presence of effects on outcome variables of interest.

David Kryscynski: I agree Clint! It may be that inferring from observable performance is the best we can do for now, but, as Rory indicated, theory is way under-specified regarding how

emergence happens. So, maybe this is a nice opportunity for both theoretical and empirical work moving forward.

Answer. In most empirical settings, I am not sure we can identify a priori what the conceptual maximum potential value creation is for a given human capital resource. I think it's also an interesting idea to think if such maximum potential exists – if we go back to Penrose, the services/performance an organization derives from its resources is likely only limited by the creativity of managers. Nonetheless, this is an advantage of ABM approaches as we can identify a theoretical maximum performance that can be achieved by a given group and then look at the degree to which the different mechanisms involved with increasing deployment of individual human capital (e.g., motivation, use of limited cognitive/attention capacity) can work toward achieving this potential maximum.

Question 20: I liked your explanation of the different ways to study HCRE, but it would *really helpful* to those us who are new to HCRE, if you could take an empirical example and discuss the positives and negatives to using the different methods.

Answer. I agree that there is a need for future work to do additional validation of the HCRE measurement techniques and discuss the positives and negatives of the current options.

Rory put together a reference list for his talk. You can find it here:
https://cdn.strategicmanagement.net/uploads/60c2/HCREMethodsRefList-10June2021.pdf

Here is also a link to a website with a list of additional readings on multilevel methods.
https://sites.google.com/a/binghamton.edu/multilevel-modeling/

QUESTIONS TO DENISA (two-sided matching)

Question 21. Evan Starr : Under the logic in Fox (2010), nobody ever makes a mistake. Is that a problem when labor is an experience good?

Answer. Yes, there is an assumption that observed hiring decisions are the outcome of an equilibrium matching process. This assumption can hold more strongly in some settings (e.g. CEO hiring, where firms and individuals compete for better matches based on well-defined preferences) than other settings (e.g. hiring of low-skill employees under labor shortages, etc). However, Fox's approach does not apply for estimating dissolution of matches and neither to so-called search-and-matching models for labor market.

Question 22. Evan Starr: How do matching models like this account for firm size? When two people can join the same firm? If the models don't account for this, but it is a reality in the data, then what does it mean for how we interpret the output?

Answer. The estimation procedure fully incorporates these instances: Fox's maximum estimator applies to one to one matching, to one to many matching (e.g. when one firm hires multiple employees) and to many to many matching (e.g. a VC firm makes multiple investments at the same time, while startups receive investment from multiple investors, including a VC syndicate).

The more interesting question (at least from a theory standpoint), is how to think/model one to many matching: For example, do we want to estimate complementarities between workers hired at the same time? This could be interesting when looking at top management teams. Or do we consider only complementarities between the firm and each hire (independently of the other hiring decisions), or between the VC and each startup (independently of the other startups)?

Question 23. Clinton Chadwick: I wonder how much these approaches can be enriched by OB/HR theory and empiricism on matching? I wish I knew that literature better, but there is a ton of work on at least superficially similar themes, such as Person-Org fit.

Answer. Great point, I think there are many opportunities here to combine the formal matching theory with the Person-Org fit literature.

Question 24. Rhett Brymer: The information you have about the set of potential partners on the other side of the market seems really important. Could you comment on the informational assumptions in these methods and the degree to which they can be manipulated?

Answer. The method implicitly assumes that observed hires are the outcome of a large-scale matching process, and not the outcome of idiosyncratic preferences or idiosyncratic informational advantage. These idiosyncratic aspects are usually relegated to the error term. However, if someone has a good theory on why some subsets of hires are subject to different market constraints (and good data) I think these situations can be modeled. I would need some concrete examples here to provide more advice, but my hunch is that one could always impose some constraints when writing the inequalities that underlie Fox's estimation approach.

Further, you are right that there are informational and market friction assumptions. A related question though is whether the existing methods could fare better. In the 2016 SMJ paper (Mindruta, Moeen, Agarwal, 2016) we show that actually logit/probit are faring very poorly relative to the maximum score estimator even under market friction conditions. This is because as I said, it is very difficult to deal with settings where there is market sorting: it's difficult/impossible to recover mutual preferences if there is multidimensional sorting.

Question 25. Ulya Tsolmon: Is the 2-sided matching method able to determine the 2nd order, 3rd order, etc. cascading matching preferences? E.g., if my first-choice match is not available, then I might go with my 2nd or 3rd preference.

Answer. I assume you would be more interested to understand this issue in situations in which, as a researcher, you don't have access to the ex-ante preferences of the players. One way to deal with this question entails calculating the "match value" of the current match relative to other possible, but unrealized matches that involve the same agent. This can be easily done once the parameters of the matching function are estimated using Fox's methodology. Then, for each agent in the market, you could observe how its "match value" with the current partner compares with all other counterfactual match values. Further, assuming that the agent would always prefer matching with the partner with whom it creates the highest value, you could then see how far off from the "ideal" matching partner the actual partner is. You could also understand why the focal agent could not "get" the most preferred partner---in other words, you will know on which

dimensions the focal agent was “deficient” or weak and could not meet the expectations of its most preferred partner.