The Podcast *Quantitude*

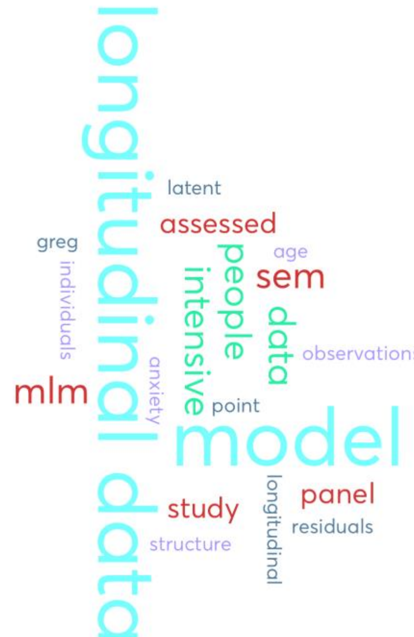
Greg Hancock & Patrick Curran

Season 4, Episode 9

Intensive Longitudinal Data: Be Careful What You Wish For

Published Tuesday November 15, 2022 • 52:16

SUMMARY KEYWORDS
model, longitudinal data, intensive, data, people, sem, mlm, panel, assessed, study, individuals, anxiety, observations, structure, age, residuals, longitudinal, latent, , point



Greg 00:04
Hi everybody. My name is Greg Hancock and along with my longitudinal friend, my intensive longitudinal friend, Patrick Curran, we make up quantity food. We're a podcast dedicated to all things quantitative ranging from the irrelevant to the completely irrelevant. In today's episode, we talked about that sweet spot between panel designs and time series designs, intensive longitudinal data, the logistical and analytical challenges, but more importantly, the tremendous potential along the way. We also mentioned campfire face plants, fear of horses, your spare eyelid, polishing your Nobel Prize, ripples in a pond. lick him, sir. August 14. The basketball effect, Southwest boarding number be 14 high school math teacher apologies. Open mic night, and Umbrella shopping. We hope you enjoy today's episode.

Patrick 00:56
So I think you share this with me but you and I are both incredibly fortunate that our teens, for some inexplicable reason seem to like hanging out with us. Yeah, there are phases, but I'd say on average, yes. The other night, we were having dinner out on the back deck.

Greg 01:15
Wait, is this the night where you had your accident?

Patrick 01:19
I used air quotes. I really appreciate you not raising this in a public forum where I would have to be deeply embarrassed. You can cut it but I'm putting it right back in. Yeah, we've established prior episodes. Yes, my quote, accident. So very briefly, from my back in the backyard. I texted Greg that I just fell into my camp fire. This is not figurative. I did not figuratively fall into the campfire.

Greg 01:55
I'm laughing and which makes me a horrible person.

Patrick 01:58
Folks, obviously you can't see Greg but he's kind of loosened it. Now. I've not seen him this happy in quite some time. I didn't get burned at all. I got scraped up pretty good. And I kid you not unwinding on my back working up at the stars. And my very first thought is Greg's gonna find this really funny.

Greg 02:17
I was so touched what I pictured based on your description was in Raiders of the Lost Ark. The guy's face melts toward the end of the movie. I pictured you like that texting me.

Patrick 02:34
I was like, Oh, this hurts so bad. Greg is gonna find this really funny.

Greg 02:39
And I did.

Patrick 02:40
And you did. Anyway, we're sitting around the table. And the kids have a shared friend group. They were in class and their friends started texting one another phrases that my kids would use in regular language. And they're these old timey ones that somehow they get from me. And the point of the story at the dinner table was they were thanking me for making them talk to like somebody who was born in 1850. I go immediately to defensive mode. I'm like, That is totally untrue. I don't do that at all. And they started rattling off. Yeah, your dollar shirt in a day late. Oh, Katie bar the door. What am I going to do? I got to see a man about a horse. Are we going to come tomorrow? Well, God Willing and the creek don't rise. It went on and on. And my favorite that I didn't even know I used they would come out and say yeah, I haven't studied for the test tomorrow. And I'd say well, at least you'll die standing up.

Greg 03:41
Were you raised in the Depression? Where did you get these?

Patrick 03:44
We have established that I would have made an ideal cowboy except for one minor thing. And that is I am terrified of horses. Really. I think any horse ever born either belongs in dog food or Elmer's glue. Horses terrify me. Other than that, I would have made an awesome cowboy. But no, I can't even explain where it came from. And some I can pin on my mom. So one that they raised was well be careful what you pray for. Because God might just give it to you. That one I can pin on mom. Okay. The Bob's your uncle falls under the one a little chin music.

Greg 04:31
Well, the last one that you said I like a lot actually. The gist of it is be careful what you wish for. And I have to say to be honest, that's a phrase that I utter a whole lot when someone has just successfully

defended their dissertation proposal. I will generally say some version of well, we've got good news and bad news. The good news is that you passed and the bad news is that you have to do it. I think they appreciate that phrase more and more as they get deeper into their dissertation.

Patrick 04:56
We all experienced that as a number of years ago by We're and I were really fortunate, and we got a multi million dollar grant from NIH. And we went out to celebrate, we each got a beer and sat across from each other. And we're just shell shocked. Because the correlate of be careful for what you wish for is, what were we thinking? It wasn't like trying to be funny. I mean, we were sitting there, and we were like, holy crap, remind me why we wanted this grant. Because now you gotta go do it. Yeah. So what got me thinking about all of this, besides the fact I fell into a campfire, and again, I want to thank you as a friend for holding that in the confidence that I shared it with.

Greg 05:44
Sure. And I think your left eyelid is gonna grow back.

Patrick 05:47
So that's why God gave me two. We had a family vacation. It was my family, my brother's family. My mom was there. My kid was four years old, running through the house stubbed their toe and runs to my mom crying. And my mom brushes the hair out of their eyes, and says, Well, sweetie, that's why God made shoes. And from the kitchen, my brother yells, that is where we came from. That's why Patrick and I are the way we are right there. So yep, that's why God made two eyelids. There you have a spare. So thank you for that. Yeah, you're welcome. Bauer and I had a chance to do a series of talks recently for the American Psychological Association. It was on intensive longitudinal data. And it was a lot of fun. JP Laurenceau also gave a presentation as part of the series. If folks are interested, we have on our centerstat.org webpage, there's a tab called resources. And we have all the downloads, we have links to the videos and whatnot. But we gave three 75-minute lectures on intensive longitudinal data. One of the early slide headings I have is be careful for what you ask. ILD which is the intensive longitudinal data ILD is one of those things where people say you have to gather intensive data, you have to do diary data, you have to do ecological momentary assessment. So you have to you have to you have to. And so you do this boatload of work, and then you get it and go back to the person and you say, Okay, well, what do I do with it? Now? It's like, oh, yeah, that's really complicated. I'm not sure. And it's one of those. Okay, you fed it? Now you own it? Yeah. What do you do when you put in blood, sweat and tears, and get this intensive data of day to day, hour to hour, minute to minute, and then by gosh, you got to do something with it.

Greg 07:49
So let's do a little bit of clarifying to start when people say they're going to do a longitudinal study, that actually means different things to different people, can we at least try to differentiate among panel data or time series data or intensive longitudinal? All of these seem to have some blurred lines here? So can we start trying to untangle that,

Patrick 08:10
I feel like we can think about three different clusters of design. One is what you referred to as the panel design. And this is old school man, this goes back back back back. What this is, is what a lot of us think about when we think about longitudinal data. That is you have a large sample of individuals, and they are all assessed at the same time, at maybe a one year increment or a two year increment, you get a sample of kids that range between 10 and 15 years of age at time one, and you assess them once a year for three years, we can think about it in the genesis of the term panels of data, we get the first wave, or the first panel, the second, the third, that's kind of bread and butter for a lot of areas of research.

Greg 09:00
And I wonder if it's bread and butter, because it's easy to get access to those data, I don't really mean easy, but it might be easier for people in education, it might be relatively easy to go gather data from kids every fall or every spring, where it might be the case that a school naturally gathers data at those particular times. So one of the things I kind of worry about with panel designs is why people choose the spacings that they do and whether or not it actually makes sense. Because I think sometimes people just do it out of convenience.

Patrick 09:31
Why do you rob banks? And it's that's where they keep the money. Why do you assessing you really in schools, that's where I can access the kids and they have the internal documentation that I can draw whatever that might be. I think that's part of it. I think part of it is those designs are highly amenable to particular kinds of research questions. So if you are interested in factors that are related with end of grade tests throughout elementary and middle school, Well, one end of grade test occurs at the end of grade three, and four and five. And so there's a nice natural kind of element to that, if you're looking at developmental trajectories of substance use from onset into regular use, and adolescents may be taking assessments every six months or every 12 months may be amenable to that. But part of it is a lot of us become unstuck in time and forget that we didn't always have an iPhone 10 in our pocket, not hyperbolically. Yeah. And iPhone has more computational power than the guidance system of the space shuttle, right. And all of us have some version of that within reach. And so we didn't have the technology to do these kind of daily assessments or things like that. I mean, people did this, don't get us wrong. Niall Bolger and Jean-Philippe Laurenceau have a wonderful textbook on this. And they talk about some of the history of these. It's not like we discovered intensive longitudinal data this past year. I mean, people have been doing this, but technology is another rate limiting step. You can't randomly ping people throughout the day in 1983.

Greg 11:13
This is starting to get at one of the questions that I asked people who say they want to do longitudinal designs, that they have maybe the hardest time answering and that is about the timing of this change process that they're trying to study, you know, so if, if I have someone who says we're going to gather data from people every fall, I always ask them why how does that align with the change process? How do you know that you're in the sweet spot that there has been enough time for the change that you want to study to occur but not too much time? It's a question that people have a really hard time answering. Now we have this kind of technology that you're describing, that allows us maybe to be a bit

more flexible at finding the sweet spot. But just because we gather a whole lot more data doesn't necessarily make that easier

Patrick 11:58
what swinging the pendulum all the way to the other end. So a panel design, I'm going to make up numbers, you have a sample of 300 kids that USS once a year for three years. Now we're going to swing it all the way to the other side to what might be called traditional time series data, where panel data has a very large sample of individuals assessed on a small number of repeated assessments. Time Series is exactly the opposite. What a classic time series is, may be one unit of observation that is followed 300 times well, you might ask yourself, why on earth would you ever do that? Well, imagine you were studying the closing price of the S&P500 index every day for a year. However many business days there are in that 300 and some let's just say 300. What you're primarily interested in is looking at forecasting, if we can try to understand these day to day to day fluctuations, and Ben Bernanke is rattling his sword about jumbo increase in the interest rate, what is that going to impact the S&P tomorrow or the day after the day after, and then we've all concluded that everyone listening made a massive vocational error by not going in to econ and marketing. Because what you do is then move your money around based on other information because you're forecasting some future state and then we give him the Nobel Prize for Economics. He got a freakin Nobel Prize. He's awesome. What comes in here is thinking about idiographic versus in nomothetic. Orientation to research. We talked about this before. I know right? idiographic and nomothetic.

13:48
You got your word a day calendar again.

Patrick 13:51
Of course, I never remember the definition. So I have my mnemonics. ideograph is idiot. It's one. You have one idiot. That's idiographic. And nomothetic is no more. I don't want any more sample, right? No most that what nomothetic is, is studying general laws. Idiographic is studying specific facts. All right? Well, if we have a time series that's focused on a single observation, we are highly limited in generalizing to other individuals that are similar to that one under study, we literally have a sample size of one. Now we have 300 observations, but we have very little basis for generalizing back to a population but swinging the pendulum back to the panel, and we have three or 400 subjects, so we have a much stronger ability to make generalizations back to the population. But we only have three repeated measures are for repeated measures. We may have done a horrible job in assessing how these constructs unfold over time. Going back to my kids list is, you got to pick the hill you're going to die on. So those are the bookends now in comes, why can't we all just get along? And let's take a little bit of one and a little bit of the other, why the enemies just get along, and we get intensive longitudinal data and what that is we have what we might think is a representative sample of individuals upon which we can draw generalized conclusions about the population. What is that number, I have no idea, maybe 50 people, maybe 80 people, maybe 100 people, not as much as a panel, but vastly bigger than a time series. But instead of having 300 repeated measures, or three repeated measures, maybe we have 15, or 20, or 40, maybe we have 40, repeated measures on ad individuals, that starts moving us into that interstitial region of intensive longitudinal data.

Greg 16:05
When I think about this with regard to generalizability, on the time series side, I usually think of it as that is the population that they want to study, right? If it was something to do with the gross national product of a particular country, then I assume you want to make inferences about that particular country. And so if there's any generalizability to be going on, it might be about what you could expect at other time points in the future. So this, we're trying to get generalizability across individuals as best we can. And we're getting a lot of benefit by having many, many measures, because we're gonna have a lot more sensitivity to study the process that's of interest. So I like this very much. But again, be careful what you wish for, as you're

Patrick 16:45
listening, think about what's near and dear to you in terms of theoretical questions, whatever that might be in whatever area you're studying, and think about causal influences, right, we're going to take a shot or whiskey and slam it on that saloon bar. And we're going to say my goal is to infer cause even from observational data, alright, so all of us want to do that. Don't be a wiener and say we're looking at associations, we're looking at predictability, just own it, take the shot or whiskey and say that I am inferring cause about what I'm studying, think about what Greg talked about in terms of that time, like so when I came up through the system, there was a lot of interest in the relation between negative effect and substance use. Sometimes it's called self medication model that the reason people drink is to self medicate with alcohol, have negative feelings and trying to make those go away. When I was on a research project, we looked at negative effect when a person was 15 years old, predicting alcohol use when they were 16 years old, knows that what the theory is predicting, because when you peel back a couple of layers on the onion, what the theory is really saying is, if you experience a negative effect during the day, are you more likely to drink that night? Right? And if you have had a drinking episode that night, are you more likely to experience negative affectivity? The following day, we don't like destroys between theoretical models and statistical models, because that is an abject threat to internal validity. Right? Internal validity is the extent to which you're making an accurate causal attribution about what your study. And if you look at do you feel depressed and anxious at age 15? And how much did you drink at age 16. That is not testing the theory as the theory lives.

Greg 18:46
I don't know if people actively think about it, or they just say, well, there's time when I think about that particular model that has negative effect. And then some version of substance use in my head, I picture this panel model with all kinds of little things back and forth. It's like this accordion that if you pulled it, you would see whether it's hourly or daily, all of this potential back and forth with all these little boxes that you never see. But what happens when someone does that study that goes from year one to year two, you've sort of collapsed all of those boxes. And it's just really not clear to me what you get out the other side, if you had some kind of stasis in a process would happen from one day to the next, then maybe you're going to get something that's informative when you look at negative effect, and you're one and substance use it year two, but when there are more interesting processes that are going on more locally with regard to time, I just don't even know what the heck you get out when you look at things at these broad broad time points. So the potential here is just tremendous. And it seems to be much more on point to the change process that you actually care about.

Patrick 19:54
That's how I really like thinking about it is how can we gather and evaluate data and know wave that is optimally consistent with what our theories predict there might be one year lags. Yeah, imagine that Ben Bernanke is home polishing his Nobel medal. So if we institute an increase in the prime interest rate today, what is the inflation rate going to be in six months? Right. Now, the very same question could be asked if the Fed implements an increase in the prime interest rate today, what is the stock market going to be tomorrow? Yeah, but it goes back to it all depends on the question, there's actually statistical term that corresponds to it. And then a colloquial one that I think about its decay, does an effect decay. And you can think about it in different ways, is dropping a rock into a quiet pool, and those waves go out. But if you wait long enough, the force of the waves spin themselves out, and it returns to a quiet pool. If you think those waves are related to other waves, but you wait until they are all gone. And then on the other side of the pool, you drop another rock in and say, Oh, see, there's no relation between my first rock and the second rock. And it's like, dude, there was a fundamental one, you just waited so long, that the proximal effects decayed away, it is not uninteresting to look out, on average, do people who experience higher levels of negative effect tend to on average, drink more alcohol, right, that's a between person effect, we don't want to give that up. What we also want to know is, if you're more anxious than you usually are, do you tend to drink more than you usually do? The classic example is, this afternoon, you have an anxiety reading of eight on the scale, we're using an I have an anxiety reading of 11, my anxiety is higher than yours is. But if my mean anxiety is 14, yeah, I'm actually less anxious than I usually am, you're having a good day. So what the intensive longitudinal data allows us to do is to simultaneously look at both of these issues. So we can look at these within person processes. But we also at least have the ability to look at the overall level of negative affectivity, the overall level of alcohol use, and see how those relate at all. But in cues of the country song. You're going to get this rectangular data matrix that is a country mile long, and you got to do something with it.

Greg 22:41
So this reminds me of what you did a number of years ago with your latent curve models with structured residuals, the idea of the structured residuals, helping us to try to separate out those things that are within processes from things that are between processes. And to me, when you die, I will say nice things about that work. In fact, I have it on my calendar

Patrick 23:03
on your calendar, you have one I'm gonna die, I would be kind of interested in that. You might

Greg 23:08
ask your family to keep August 14 Clear.

Patrick 23:13
I like that you bring up the SR model, if for no other reason than we can't use that with intensive longitudinal data. So let's think briefly about okay, you putting your blood sweat and tears and you've got this country mile long data file, maybe we have 80 people that range from 30 to 60, repeated observations, you got a boatload of data now that we got to deal with first 90% of sciences, data management, that is a tough go, right? I have greater fears of making a mistake in Data Management

than I do about fitting any model, because that's the greatest area where you can screw things up. Because in Data Management, the model is like the easy part.

Greg 23:57
Oh my gosh, yeah. And with this kind of data, in particular, it's just asking you to screw up.

Patrick 24:02
So think about any latent curve model, where we have repeated measures that are indicators on the growth factors. Every repeated measure that you have, you have to be able to have a manifest variable for that, that has to have a mean, and that has to have a variance to go into the SEM and the way that the latent curve model works. We have age 10, age 11, age 12, age 13, age 14, age 15, Bob's your uncle, we got everything. Now think about an assessment at 1028 at 115. And at 430. And for Greg, you have 830 to 1135 106. You can and very often have a situation in which no two people even share the same assessment period. Not that you don't have enough to compute a mean and a variance on it, but only one person in your entire data file was assessed at 10:16am. So the SEM and the latent curve model tap out, you're done in these kinds of data,

Greg 25:08
you can't even lie to yourself until you those were close, close enough to be in the same interval, as you start to think that these smaller and smaller intervals are where the action actually is,

Patrick 25:18
you just hit the nail on the head because I could take day assessments and night assessments and day assessments and night assessments. But then you start shooting yourself in the foot because you have a reason to collected at 1018 and 1142, and 155, whatever it might be. And now you're saying, Oh, it's just we're going to take the mean of everything that happened when the sun was up and the mean of everything that happened when the sun was down, and you're undermining the very nature of why you gathered the data in that way.

Greg 25:50
Yeah, but those issues that we talked about dichotomization here. But in this case, you're just losing information, valuable information about time

Patrick 25:57
well, and the tail wags the dog now, don't it right, because what you're doing is you're modifying your data to fit the modeling framework. Yeah. And that is bad, bad bad, we ain't doing that what we need to do is build a modeling framework that is consistent with the not only structure of the data, but with the motivating theoretical questions at hand. So before we start talking technical aspects, what is it about intensive longitudinal data that introduces issues that are not common in panel data, I think it's helpful to think about that there are unique features that we have to deal with, if you're assessing anxiety at 10:22am. And again, at 11:43am. Is the anxiety 1122 in some way more related to the anxiety at 1143. Then at 1022, and 315, right? When you start getting these highly densely distributed observations in time, it introduces a thing called serial correlation.

Greg 27:03
So could you clarify for us why it's not a thing in panel data, and it is now all of a sudden, the thing when we have the intensive data

Patrick 27:10
ripples in a pond, it's decay, not poppers corpse decay. So let's say that you're looking at anxiety and substance use at age 15. And age 16. Anxiety at age 15. And anxiety at age 16, are going to have some correlation. Obviously, we can't fit a growth model if those don't covary. But we are going to model those covariances at the level of the latent growth process. So if we have age 10 to 15, and we're fitting a linear growth model, we have a random intercept, we have a random slope, we have a covariance between intercept and slope. All right, let's teleport back to the mind's eye of path tracing rules. Remember, we were playing with guinea pigs at some point?

Greg 27:55
That's right. So we'll write who had the Fetish

Patrick 27:58
picture in your mind's eye a path diagram for a latent curve model. So you have five indicators that are repeated measures, you have two latent factors. How do you get from age 15, to age 16. So you start at age 15, you go down to your intercept factor up to age 16. That's part of it. And then depending on how you code time you go down to your intercept, you go over the covariance, you go up to age 16. The model is building in the covariance between age 15 and 16. The residual of age 15. And 16 Rarely is correlated in the LCM. That part that's leftover of age 15, correlating with that part that's leftover at age 16. Not because we can't do it, but because often, that residual covariance has decayed away over a year, any little bit of relation that's not related to the underlying growth process, those ripples in the ponds have just died away, and it's gone flat again. But if you have 10:22am, and 11:43am, we can still have a random intercept on those. But there's going to be a little bit of serial correlation leftover. Maybe you had a shot a coffee in the morning, maybe you've gotten an argument with your boss, and your anxiety went up. And it's going to be higher at several times, maybe before it returns back down to baseline, these intensive longitudinal designs introduce the serial correlations that we have to build into the model. Because if we don't you know what starts tomorrow, North Carolina State Fair, you know, there are two reasons to go to the North Carolina State Fair turkey legs and whack a mole.

Greg 29:38
That's all there is.

Patrick 29:40
If you have the serial correlations in your intensive design that want to be there and you don't let them those are going to whack a mole down into other parts of the model and can seriously bias other aspects of your model estimation. So

Greg 29:56
if I make the tie back to the Leighton curve, Otto was structured residuals in a panel design. Do you call it lick'em-sir?

Patrick 30:05
Let's leave that alone. Okay. rickle-pim is a special case of lick'em-sir. Where were we doesn't matter, okay.

Greg 30:18
In that model, you could connect to the residuals of a variable to the residuals of that variable at a new time point. But isn't it the case that in the LCMS, or you often don't have the residual to residual time point connections within a variable, you more often have them across variables whose relation you're trying to study? Yeah.

Patrick 30:39
And it's a testable hypothesis. But all of those, all of those are still based on this panel design. Yeah, if you're looking at a school based end of grade testing, you want to follow it across elementary school, that kind of framework may be really well suited for that. But if you're talking about daily anxiety, and nightly alcohol use, assessed morning and night over a 14 day period, the LCM taps out.

Greg 31:09
So one complication then is the serial correlation that we have. Are there any others that you want to mention while we're here? If you're

Patrick 31:15
thinking about how does something unfold throughout the day or throughout the week, often, you have to also start thinking about either cycles, or transition points. Dan, and I have some data we work with where there's 61, individuals measured with heart rate throughout the day, up to 60 times. You can cycle throughout the day, but you also go to sleep where there's a transition point. If you're studying alcohol use in a university setting, you may need to consider weekdays and weekends as cyclical. And by the way, and this is official weekend on college campuses starts on Thursday night. We did some studies in our labs, not only do we have to put in a weekend effect, we had to look up online and put in impacts of home basketball games. But think about it, you have to do that stuff. If you are on the campus of the University of North Carolina and you are studying weekday weekend drinking patterns. You darn well better put in home basketball games,

Insert 32:25
and the fairy tale Ryan put the Tar Heels to 10 News Head Coach K's legendary career has come to a close

Patrick 32:35
because there's a spike in drinking on Franklin Street. Dude, I'm doing my thing here is your alarm going off?

Greg 32:42
It is hang on just a second. Our whole episode is about the importance of getting timing right. I have a Southwest Airlines flight 24 hours from now and I have to make sure that I'm checked in. Otherwise I'm going to get a crappy middle seat.

Patrick 32:56
You don't pay the 20 bucks to buy yourself to the front of the line. How freakin cheap Are you? Okay, I want people to realize this is not a gag. When I was mid stream of doing my thing he is right now trying to get at best you're gonna get a B he thinks so boring, right? Oh, I guarantee you because you've been talking to me for the last 30 seconds all of A is gone already.

Greg 33:21
Okay, here we go. I gotta be 14. So boom, timing is everything. Okay? I honestly just didn't expect that you would still be talking

Patrick 33:30
really don't mind me at all. Okay, I'm back. So now the latent curve model is tapped out. We can't do it. Right. If you have 20 3040 measures, the SEM has said thanks so much. I had a great time tonight. I'll be here all week and as walked off the stage, you know who's going to come on and I have to admit, they get marginalized a little bit from us SEM KoolAid drinkers, right. We can do multiple indicator latent factors. We can do reciprocal effects. We can do formal tests and mediation ever look at us all buttered up in front of the mirror and flexing my pecs look at my pecs. Look at them. You can't look away from my pecs. Oh, you have 10 repeated measures. Pardon me. I have to go see a man about a horse. Seriously, is the multi level model gets dished on a lot? Yeah. And you know what the SEM was the warm up act for the multilevel model who's going to come out with this little bottle of water and it's stool.

Insert 34:37
Boy, what a crowd what a crowd.

Greg 34:40
People are throwing beer bottles at SEM to get off the stage. Is that what you're saying?

Patrick 34:44
Just like the Blues Brothers were they performed behind the chicken wire, rollin, rollin, rollin, the screens are swollen because you When patrons liked it, they threw beer bottles. When they didn't like it, they threw beer bottles. MLM is sitting back there smoking a cigarette in the back of the bar. And it's all by itself. And it's like, I knew you'd come back. Because what MLM does is it brings in the numerical measure of time as a predictor variable. And you know, what, 1018 1145 It don't make no nevermind,

Greg 35:29
you and I had our episode where we talked about the relative benefits and detriments of MLM approach versus an SEM approach to growth. And I think we did a pretty even handed treatment of that, with the help of Dr. Michelle. We landed on the latent growth curve models as having so much versatility and things that the multi level model couldn't do. But one of the big Achilles heels of the structural equation

modeling approach is the rigidity when it comes to dealing with time. So I gotta say that MLM has got some things going on here it is looking pretty fine back there in the corner of the bar,

Patrick 36:05
very briefly, MLM approaches any longitudinal design as an issue of dependent data, it is repeated measures nested within individuals. Now we also know from prior episodes, and a whole lot of great work that is out in the world, that many, many growth modeling applications are isomorphic between MLM and SEM to the 16th decimal, but it starts to separate when you get into these intensive longitudinal designs, what the MLM allows for the intensive longitudinal data are a couple of big things right off the bat, one, you can, without hyperbole, have an observation at a time point that is not shared by anyone else in the sample, you can have one person out of 1000 observations that were assessed at 9:11am. And MLM don't care. Yeah, this is a huge advantage. And what it's doing is imagine that you have a predictor, some measure, that is height, we never think about, oh my gosh, we need to have two people who share the same height to build a very, now it's just, it's on a number line. And we have this continuous number line of time. And person one was assessed here, here, here, and here, person two here, here, here. And here. Don't make no nevermind. The second advantage is there are a whole host of level one error structures that we can build in those serial correlations. And they do different things, we can do an auto regressive structure, a moving average, we can do banded toeplitz. Ooh, one of my favorite is spatial power, I use proc mixed in SAS for my own work. And they have something like 20 different structures of level one residuals. And what that is, is what we alluded to a few minutes ago, in the latent curve model, we assume that part of age 15 is uncorrelated with that part of age 16, because any relation that was there has decayed away over that 12 month window. Well, here we're saying, holy cow, I got a boatload of these structures that I can build in one to give myself some theoretical insight into how my data unfolds over time, but to it protects against whack a mole, if there is this auto regressive lag one structure at my level one residual, and it wants to be there dollars to donuts, that we got that AR one in, so that I can interpret my other covariance estimates with greater confidence that they're not bias because a whack a mole.

Greg 38:42
Okay, I got a question for you. So you've got these error structures that are in the can, right, you can check this box and we've got banded toeplitz. If you don't really have any theoretical questions around the error structure, but you just sort of view the infrastructure as a place where you want an insurance policy to help prevent whack a mole in the areas that you actually care about, what kind of error structure do you gravitate toward just to make you sleep better at night?

Patrick 39:07
I hang them all and let God sort them out. You're spot on? Not often, is there a theoretically motivated question that would distinguish an AR one from abandoned templates? You estimate 10 different structures and pick the one with the lowest Bayesian information criteria? You let the data speak to you. It's not a theoretical inference. Don't yell at me for being the reason for the replication crisis. What we're trying to do is to optimally represent this complex level one error structure so that then we can go on about our business. Yeah,

Greg 39:42
and our businesses elsewhere and our businesses elsewhere. All right. So assuming that you've wrestled with your error structure, daemons, you got something at least in place that's going to keep you out of trouble. As far as Whack a Mole goes. Tell us about some of the kinds of cool things that you can do where the action is in your model

Patrick 39:58
where you know what's neat, what you would often do in models like these even with panel data, or their time trends, does anxiety increase throughout the day and then go down? And then does it increase throughout the day and go down? Does that differ between individuals? Do you have time varying predictors? This is a big thing in ILD. Does my anxiety during the day predict my alcohol use that night? And so if I'm more anxious than I usually am, do I drink more than I usually do? Those are those kind of dynamic process questions that we want to ask. And so a lot of it is doing what we wanted to do in the first place. But being able to do it properly, do we have a predictor of an outcome that we have earlier in the day predicting later in the day, or at an hour increment or whatever, right? There are no rules on how to use these, you might have an intensive longitudinal design where you take weekly measures for 50 weeks, it doesn't matter that it's minutes or hours, it's the density of the observations, I would say it's in a way a little anticlimactic. It allows us to look at lead lag relations in the way that we really want. But doing it in a way that's proper in terms of the characteristics of our observed data. And we

Greg 41:15
get cool things like splines, right, we can have nots where there are interesting developmental junctures that things might occur, we have all kinds of flexibilities that are very, very interesting to us. And in many cases, we can get at the individual differences in those kinds of things where they're just fascinating questions here.

Patrick 41:32
And we do that with the heart rate data, we have a linear spline where we fit an individual trajectory to the heart rate data throughout the day. And then we have a knot when the person goes to sleep, and then fit a second spline of their heart rate when they're asleep. And the super cool thing is when you go to sleep can vary over individual right? And so you can literally start doing things like cosines, we've established all of us owe our high school math teachers apology,

Greg 42:06
huge apology, right?

Patrick 42:08
Am I ever gonna have to figure the area under a curve between two fixed points is so stupid? Oh, yeah, that's hypothesis testing. And another one is, when am I ever going to use a cosine function? Well, instead of fixed transition points, the smooth cycles that go up and down and up and down, you can build those. And it's fascinating. So one of the things

Greg 42:31

that you and I talked about, in a previous episode, our Lego episode, right, where we talked about re parameterization, one of the key messages was trying to get as much fidelity between the model and your research question and the things that you're describing. It's not just like, we could put in a cosine. That's awesome. It's that you think that there's a particular process that subscribes to that cycle, or you think that there's a particular process that's reasonably well characterized by two lines that meet at a particular point, I am just as you know, a huge fan of using something that has the flexibility to try to build what you need for what it is you're studying. And this does, I have to say, have an awful lot of that

Patrick 43:10
it's the laser cannon on the flower shop,

Greg 43:12
Annie's flower shop,

Patrick 43:14
there are a couple of pretty significant limitations. So this approach, not limitations in what we're able to do, because everything that we've talked about it does it really, really well, it's a limitation and what we're not able to do, imagine that we have anxiety during the day and alcohol use that evening. And we follow that for 14 days. That would be a very reasonable application for this. Alright, well, let's think about some of the limitations that the MLM naturally has first, in a single model, we're only able to look at one direction of that effect. Yeah, we can look at anxiety to alcohol use, but we cannot simultaneously look at alcohol use back to anxiety.

Greg 43:59
So drinking might make me more or less anxious, right? That's

Patrick 44:03
right. Yeah. So I did a project a number of years ago with a wonderful guy named Tom wrote a BA and his dissertation, he did an early days diary like paper pencil. He was really interested in panic attacks, and what were risk factors for that. And in the morning, he would have the individual report how fearful Are you that you will experience a panic attack during the day and then that night he assessed how many panic attacks did you have during the day? He was interested in was your fearfulness in the morning associated with experiencing more panic attacks during the day, but did your experience of panic during the day predict your fearfulness the next morning? So there are some really important questions that we have about reciprocal relations.

Greg 44:55
And the MLM approach really just isn't suited for that. That's exactly right.

Patrick 44:59
Yeah. And now the LCM is exceptionally well suited for that. Whether it be the random intercept cross like panel model or the LCMS, or however you approach it, it's really well suited to saying earlier panic predicts later fear earlier fear predicts later panic, there's just that little problem that we can't use that

model with intensive longitudinal data, right? So we do not have reciprocal relations within the MLM the other drunken punch to the face, I know what's coming. Well, then you tell me,

Greg 45:32
there's this thing that you said back when the earth was cooling. When we started this podcast, you said something early and often. And that was if you blow measurement, everything else is a problem. What I'm not seeing in this model is the thing that you and I tout all the time as being so crazy important. And that is the ability to accommodate potential measurement error here. I mean, I can say, I have outstanding measures of everything that I have in my model. And I might be able to work in the background with some very fancy, very modern ways of trying to get the best possible scores that technology and money can buy. But I can't actually model the latent variable directly in this system.

Patrick 46:15
And as Bauer said, in the MNLF, a episode, and I can make it worse yet, which is we're pushing to cell phones. Do you feel anxious? Yes or no? Do you feel fearful? Yes or no? Do you feel uncomfortable? Yes, no, we got to go back to first principles. Smitty Stevens, the principle the assignment of numbers to observations, there are a lot of dissertations out there to be done to look at not only measurement error in these models, but the other characteristics, are these continuous, are these normal are these ordinal, or these Likert? Are these counts of binary, which means we need to be thinking about these as Poissons or negative binomials. The MLM as we've wax poetic up to this point assumes the dependent variable is continuous, normal and error free. Yeah. So funny story. None of those hold.

Greg 47:15
Just like when we talked about the pluses and minuses with regard to growth modeling. We're in the situation here where we have pluses and minuses with regard to the intensive longitudinal data. We've talked about all the wonders of an MLM approach, but can we buy something back? That's more in the SEM family? We can,

Patrick 47:31
but we've never met him before. So it turns out SEM has an older brother. He's gonna come out for open mic night. Now we came out because we've seen SEM before and he's hilarious. He's funny as heck. You know,

47:45
we have a lot of expressions in the English language. And we never seem to examine them very closely, like legally drunk. leave the guy alone. He's legally drunk.

Patrick 47:56
We saw MLM before it oh man kills every time. Who's this DSM guy? Nobody's ever heard of him. He's gonna come out now. What's the deal with lampshade? I mean, if it's a lamp, why do you want shade? This is dynamic structural equation models. This is the brainchild of a lot of very smart people. Tihomir Asperhouv, Ellen Hamaker, Bengt Muthen, and many others have contributed to this. Can we take some of the basic principles of the structural equation model? The two big ones are Yeah, you want to

come back the other way with a reciprocal effect? Sure. Doesn't bother me. You can go anywhere you want. And second, you want to multiple indicator latent factor. Yeah, I can do that.

Greg 48:43
Be careful what you wish.

Patrick 48:46
There we have squared the circle, which is remember maximum likelihood who had baby roiled himself up and was flexing in front of the mirror, maximum likelihood heard the police sirens outside and they're like I'm out of here, and maximum likelihood is gone. right out the window and down the alleyway as the cops pull up. And what I mean is DSEM, which we're not going to talk about here because it's an entire episode in and of itself DSEM, I think conceptually brings in LCM SR like approach within the intensive longitudinal design. You can have reciprocal relations. You can have multiple indicator latent factors. You can do SEM-like things that we want to do, but we can do them within the ILD kind of high density framework. But what we do is ML broke out the window and ran away when he heard the sirens. It's a whole new ballgame in estimation. There are a lot of unresolved issues, we have to approach it from a Bayesian perspective. I have no concerns about that. This is just brand new, and we've got a lot To learn,

Greg 50:00
well, then I wouldn't leave it up to us to explain, explain it, I wouldn't leave anything up to us. So maybe we should get someone in the future who actually knows a little bit about DSM to help us out with that one.

Patrick 50:10
So let's seal that jar of pickles and put it on the back shelf for winter. Now I'm just making it up. I don't even know what that means. Okay, let's save DSM for a future conversation and maybe loop in some real talent to help us talk about that.

Greg 50:26
Maybe we better go take a nap on the Davenport. Don't even try No.

Patrick 50:32
Come on. You're from Seattle. Maybe we should go buy a cup of coffee and go umbrella shop and

Greg 50:42
those are my does respect my people. All right, everybody. Take care everybody.

Patrick 50:53
Thank you so much for listening. You can subscribe to Quantitude on Apple podcasts, Spotify, or wherever you distract yourself from the midterm elections. And please leave us a review. You can also follow us on Twitter. We are @quantitudepod, and check out our webpage at quantitudepod.org for past episodes, playlists, show notes, transcripts and other cool stuff. Finally, you can get Quantitude-themed merch to fulfill your obligatory upcoming secret Santa's assignment@redbubble.com Where All

proceeds go to Donors Choose to support low income schools. You have been listening to Quantitude: the podcast equivalent to Southwest boarding number C58. Quantitude has been brought to you by Elon Musk, whose career trajectory has gone from revolutionising electric vehicles to perfecting human spaceflight to becoming obsessed with blue checkmarks and will soon complete its arc when he swings by Greg's house in a few days to help out with the podcast by the proper use of the Oxford comma proud to be the number one source of marital conflict among academic couples since 1905. And by in-person and Zoom hybrid job talks. Just when you thought that a pandemic job taught couldn't possibly get any more stressful. Now it does. This is most definitely not NPR