

The Podcast *Quantitude*

Greg Hancock & Patrick Curran

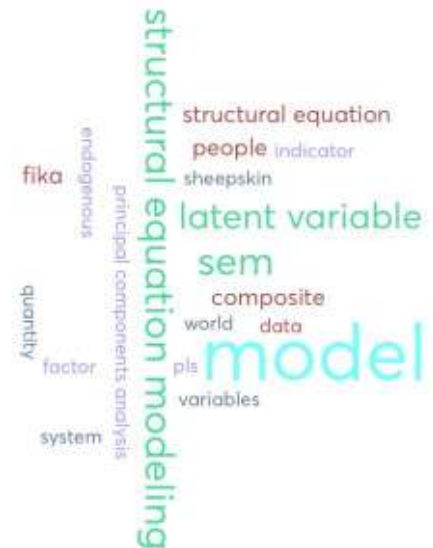
Season 4, Episode 4:

Partial Least Squares: Straight Outta Uppsala

Published Tuesday October 4, 2022 • 55:59

SUMMARY KEYWORDS

model, constructs, sem, structural equation modeling, latent variable, people, fika, composite, structural equation, data, pls, indicator, endogenous, factor, principal components analysis, sheepskin, world, quantity, variables, system



constructs

Greg 00:04

Hi everybody. My name is Greg Hancock and along with my formative Mode B friend Patrick Curran, we make a quantity food. We're a podcast dedicated to all things quantitative, ranging from the irrelevant to the completely irrelevant. In today's episode, we talk about partial least squares, a technique that resembles structural equation modeling, but with a lot of flexibility, including but not limited to its ability to accommodate both reflective and formative constructs. Along the way, we also mentioned dark red bubble, sheepskin, snorting, SEM, the Coors Light beer bong hotel rings ghost, Larry, the Cable Guy, the Marvel Metaverse, the evil eye, bad luck, schlep rock fika, Abba, and pieces left on the table. We hope you enjoyed today's episode. Have you happened to check RedBubble and see all the cool stuff that we have? I have not, there's so much cool stuff that we have. I just thought I should take this opportunity to let you know, obviously, you know, we have stickers and notebooks and all of that. I would like to thank the people out there, not just for posting your really cool pictures of you with the merch. But as we mentioned in the tag, all of the proceeds go to help donorschoose.org. And what that means is that the stuff that folks out there have bought have helped dozens and dozens of classrooms through things like calculators for math classes, hands on materials for teaching statistics, a whole variety of things. And that's all due to the folks who are listening out there. So thank you very much to everybody.

Patrick 01:36

I may have misunderstood something. I thought we were doing this whole podcast to become rich. I've been waiting for three years for the check from you. And I've just been trying to be patient. I don't want to be a wiener about it. Did I misunderstand how this works?

Greg 01:54

I think we can solve this if you just Google academia. Sorry.

Patrick 02:02

But thank you everybody. And it is so fun. Then my kid drew a couple of things that are up there in the corner, toyed and whatnot.

Greg 02:09

Now, there's some things up there that your daughter didn't draw, and that my daughter didn't draw and that I didn't draw and that none of us Drew.

Patrick 02:18

I'm sorry. Is there like a dark red bubble that we're part of that I don't know about yet? Here's

Greg 02:23

bootleg merch Can you not? There's more fake quantity and stuff than there is real quantity and stuff on there.

Patrick 02:33

This is a real question. Why I yeah, I mean, there's no money involved. We have a market share of negative \$1,000 a year. I mean, it's a real question. Why on earth like bots, are they just trying to copy stuff? I don't

Greg 02:48

know the answer. I was wondering if it was bots also. But there's a bunch of stuff that says quantity food on it and different fonts and then there's a ton of stuff that says Jiffy on it which Jiffy by the way is extremely

Patrick 03:00

happy about. Hi guys. Hi, Jiffy

Greg 03:03

jiffy. Is a your side hustle. No, I just kicked your very tastefully done. I bought five. All right, little buddy. Thanks. Bye, guys. Have a great episode. I will make clear to folks out there that you should be able to get whatever you like. But it's only the authentic quantity pod merch that actually makes it back to donorschoose.org. And how do you know if it's authentic? It has the name of who posted it beneath it. And it says by quantity pod or something? Anyway, so I thought that was kind of interesting that we have fake stuff.

Patrick 03:39

I am fascinated by that. I don't mean to like belabor this, but I totally get making a fake account and selling things with like a Nike swoosh or like a Ferrari logo. We're nothing. We own nothing. I am perplexed by this.

Greg 03:57

Maybe this is actually one of the best indicators of where the economy is. Currently. I remember when I had my first car what was your first car

Patrick 04:06

my dad gave me his Volkswagen Beetle 1969

Greg 04:11

The one that you raced in our control episode I said Yeah, right up. I had a 1976 Mercury Capri to it was a hatchback. When I got it used. My mom got me these three seat covers, and she was so happy to do that. So she got me these and she was going there sheepskin there sheepskin. Like, you know, this was some big deal and like, okay, Mom, thanks. And so I put them on my car and they were very comfortable. And I looked at the label, and it said sheepskin so I never had the heart to tell her that it wasn't actual sheepskin, but I thought that was like the best fake name ever. sheepskin.

Patrick 04:56

It's so funny. You say that because just a couple of weeks ago I tried to do my own car repairs when I can. And with YouTube now and things that you can order online is within reason. I can do a fair amount of my car repair, I needed to get this particular part for the record. I searched online and I found genuine Honda replacement parts. Because I didn't want off brand, right I wanted it actually made by Honda. And I was about to check out and it kind of struck me that it was half the price of what I'd found in other places. And so I jumped in another tab and I searched genuine Honda parts. The company name is genuine. It was genuine 100 replacement parts. Right. And I thought that was brilliant. I almost wanted to buy it just because I so admired that it was a sheep's can. Yeah,

Greg 05:52

exactly. Well, so our topic for today is maybe in this spirit, I don't know, that remains to be seen whether or not what we're going to talk about is sheepskin, or sheepskin. What we're going to talk about today is something called partial least squares originally, when you were up visiting me over the summer, and we were talking about topics that we were going to cover, we had said oh yeah, we can go through these different estimators like ordinary least squares and maximum likelihood and to stage least squares and partial least squares, and it was just sort of uttered in the same breath.

Patrick 06:23

It was gonna be in one episode. Okay, well, that's a whole other thing.

Greg 06:27

But it turns out partial least squares is not actually an estimator. Did you know this?

Patrick 06:32

I do now.

Greg 06:35

Yeah. So partially squares, which we're going to unpack a little bit here. The funny thing is, is that there's a whole segment of the world that is just tied to partial least squares like it is their jam in the early 2000s. If you go to the management information systems, literature, or other fields, like chemo

metrics, this is it right? We did a PLS, we did a PLS, we did a PLS in our world being primarily a social science world. This doesn't cross our plate a whole lot. Exactly. And

Patrick 07:07

the reflection of that, at least to me is for you and me. Our day jobs are kind of centered around structural equation modeling. Yes, I've been in the game for 25 years. And I thought it was a method of estimation. I really did. And that is not it at all.

Greg 07:24

That's right. And in fact, you mentioned structural equation modeling. You and I we didn't just drink the structural equation modeling Kool Aid. We went to Studio 54. And snorted structural equation modeling off the bathroom counter and came out rubbing our nose so we're deep in

Patrick 07:42

Nope, you did I Coors Light beer bonded that's only half a joke, which is that's all there was right? It wasn't like I could pick from this or this or this or this when you and I came up through the system. Bollen's book came out in 89. And I use that book in my first sem class in 1990. Right, this was the only game in town when I came up through the system. Now ironically, the foundation for partially squares was developed years earlier. Right, it had not permeated at least my instructional system that I was part of. And what I find fascinating is I think that's impart their literally continental preferences over this. Much of the work in the applications are dominated in Europe.

Greg 08:43

That's absolutely true. Structural equation modeling and this thing called partial least squares that we're gonna start getting into. Both of those weren't just born in Europe. They weren't even just born in the same country, which is Sweden. They were born at the same university, who Uppsala University, as coming out of a guy named Herman Wold Uppsala. Herman Wold was the advisor of Who do you remember?

Patrick 09:15

Okay, I remember but you're gonna mock me for saying his name. So I'm just gonna say Carl, and then you can say his last name because Tova taught you how to say yes.

Tove 09:26

That's right. Hey, this Toby Larson, Greg, Swedish coach and occasional quantitative linguistics consultant.

Greg 09:32

So it's first name isn't even Carl, come on. Have you? Throw me a bone. All right, fine. Fine. His American name is Carl. His stripper name is Carl. His actual name is Colgate que je and last name is Jada scug. Yeah, the pronunciation is slightly off on that one.

Greg 09:49

So Karl Joreskog as we call him in the US was the student of Herman Wold and Karl Joreskog is you know the father of the structural equation modeling that you and I practice that really He had it seeds planted back in confirmatory factor analysis in the mid to late 60s, where structure was imposed upon that latent variable system, and then started carrying off into the 70s. And it didn't just carry off on a theoretical level. And I think this is really, really important. It carried off with software to, it's great when there's the mathematics of these kinds of methods. But if you can't put it in people's hands, and there's a problem, there was LISREL. And there were punch cards that went with LISREL, and mainframes and all of that kind of thing. But when we think about how structural equation modelling got a head start, the pieces were in place, not just in terms of the mathematics and the theoretical foundations, but actually the software to be able to do it. And I think that's a critical thing. But you know, after structural equation modeling started taking off Herman Gould sort of watching his academic child go off and do great things, he sort of scratched his head and said, oh, boy, cool, gay. There's a lot of assumptions wrapped up in your structural equation modeling thing in your list rel thing, and some of those things make me uneasy, and we should probably just rattle off some of the assumptions that exist, you're gonna pop quiz me on this, I don't mean it to be a pop quiz, but let's go 30 seconds

Patrick 11:17

within the SEM and what my understanding is of what Volt's reaction was kind of twofold. One is it goes back to something that we've talked about on prior episodes, which is one of the key advantages of the SEM is that it's a priori. And one of the key limitations of the SEM is that it's a priori. So we have exploratory methods, we have DFA, we have Eigen values and Eigen vectors and rotation and all of these things. And then if you've had that in a class people say, yeah, yeah, but you're letting the data drive your decisions. don't saturate your factor learning matrix structured in a way that's consistent with theory. And that's a huge advantage. I mean, massive, huge, huge, massive advantage of confirmatory factor analysis in the SEM in general, except one, it's not. Some paraphrasing of volte is there are often situations where we are data rich, but theory poor, and that we need a methodology that allows us to do sem light things, but without being so shackled to a strict a priori parameterization of our model before we ever begin.

Greg 12:29

So you'd said that there were two primary things, the first one then has to do with feeling locked into this very confirmatory very odd, priori structured kind of endeavor. What's the other one that you had rolling around some

Patrick 12:41

of the asymptotic regularity conditions for maximum likelihood estimation? So what he seemed to be more concerned about was a rather strong assumption about multivariate normality, and that you have a sufficiently large sample size where these asymptotic conditions of consistency, unbiasedness and efficiency and asymptotic normal sampling distributions, maybe those don't come online, given the characteristics of the data that we have in our sample? That's absolutely right.

Greg 13:08

And so his thinking was, could we do something that is, and this is where I really want to be careful of my language. Can we do something that is an approximation to that? This is where some people's hackles will get raised? Not PLS, people necessarily, but sem people, right? You and I had a whole episode on principal components analysis and how like one of your first disclosures after the spider pig theme, and

Patrick 13:34

spider fig spider fig.

Greg 13:43

After those disclosures, I mean, the whole episode really was about principal components analysis what it is, and that it is not a factor model. And some people just completely get their underwear and a bunch over that for sure. And there's a whole segment of literature from the 80s and early 90s, where people were just butting heads over. Is it a latent variable model? Is it not a latent variable model? Can it be used to estimate latent variable models, but at its core principal component analysis is just a composite model. And it's not the only way that we have to form composites. And so Wold was sort of wondering, can we bring to bear some of the ideas of compositing rather than the latent variable methods that we have, but use it as an approximation to a system that otherwise looks a lot like a structural equation model?

Patrick 14:29

And in that spirit, you know, whose ghost that I saw on my back deck as I was reading some of this material was Harold Hotelling. Ah, nice member. He's local. He's in Chapel Hill, but he was at UNC back in the 30s, but

Greg 14:45

not buried in your backyard.

Patrick 14:48

Nevermind Hotelling, who was the original developer of principal components analysis, never addressed factorial rotation because he did not care about the numeric Recall values of the weights from a substantive standpoint, he had a very practical motivation, which is he had a large amount of data. And he wanted to reduce it to a smaller amount of data. And whatever those optimal weights were, which came out of Eigen values and Eigen vectors, is that allowed him to get the composites, get a cup of coffee, say hard to merge out front, and then go do whatever he was going to do with those composites. Obviously, this is not principal components in PLS, but it has that spirit of Hotelling, which is we have a problem to solve existing methods work really, really well, if we meet those assumptions very often, we don't meet those assumptions. And this is a really nice alternative that allows us to do some things that we wouldn't otherwise be able to

Greg 15:49

imagine that we had what you and I would think of as a structural equation model with some latent variables. And we don't need to dig too deep into labeling them yet. But let's imagine that we had to

exogenous factors or constructs or call them whatever you want. And let's say one endogenous, that depends on both of those. So in our head, right, we're asking you to visualize use your mind's eye. Both of those have paths coming into the dependent or endogenous latent variable. And let's just say that those two exogenous factors covary for you, in our practical people, and maybe worried about some of the things that we're starting to allude to, we might say, why don't you just put a score in there for that first factor? Why don't you just put a score in there for that second factor, if just put a score in there for that third factor? And go do yourself a regression? I mean, why not? Right,

Patrick 16:44

exactly. It's a GTR done kind of thing.

Greg 16:49

Larry, the cable guy here,

Patrick 16:51

but there's an element to that. And I actually find you and me quintessentially practical. Oh, nice. We're trying to achieve a goal we got an end game. We know what the high road is. What I mean by that is, if we are able to use a multiple indicator latent factor if that latent factor is properly defined, if we have adequate sample size, if we have adequate distribution, if the model is properly specified, that's kind of the gold standard. But as the pls people talk about with the data rich theory poor and violations of these asymptotic conditions, we need a way to achieve what our goal is. And this is a promising alternative for doing that. Yeah, we got an endgame right. It's all Marvel Universe. This is like the STMS endgame.

Greg 17:38

So is this some other part of the metaverse multiverse? metaverse? Which is

Patrick 17:42

it? You don't know?

Greg 17:43

I just remember that Dr. Strange was like trying to patch some stuff up and it just got kind of weird.

Patrick 17:53

What the hell was wrong with you?

Greg 17:55

It was so long, I had to pee. So I left came back. And it's like, what? Hey, could

Patrick 17:58

we get back on task here?

Greg 18:00

I'm sorry, what were we talking about? All right. So let me actually label some of these hypothetical factors, right. Now. Let's imagine that we had to exogenous factors, and one of them was exposure to

discrimination. And by exposure to discrimination, I might ask, let's say kids, to what extent have you felt that you have been discriminated against will say, on the basis of race ethnicity, on the playground, in your classroom, walking home from school, when you are outside of school, in the social media that you encounter, in the sports that you play outside of school, etc. So imagine we had this construct that is exposure to discrimination in the world where you and I live the structural equation modeling world that operates with a very specific structure associated with our factors, right, the latent variable has a structure that it inherited from the confirmatory factor world, where we assume that the latent variable actually causes its indicators. But in this case, I think a pretty good argument could be made that your exposure to discrimination is like a bucket. And each one of these kinds of experiences that you have fills your bucket up a little bit more, or fills your bucket up a little bit more, meaning that the arrows might actually go from the variables into this construct rather than the other way. And if that's the case, if we would agree that something like that is a reasonable description of what's going on in the relationship between the variables in the construct, sem gets really uneasy

Patrick 19:33

about that. That reminds me of a good friend I had in grad school, literally linguae Lily is at the University of Washington at Seattle, your alma mater, Go Huskies. She and I were out for coffee and I was showing her a new latent factor that I was working on. And it was uncontrollable life stressful events and it was in children. It was a series of things that the child had no control over. But that happened to Now, so their cat died, their grandmother ran away. Wait, maybe it was the other way. But there were this series of things. And she was really funny. She said, but does that factor like an evil lie? Like somehow somebody, this hex on you? And I was like, no, no, it's lamda aside lambda prime, it took me 10 years to figure out what she was talking about. And it's this issue that you're saying I love

Greg 20:26

the evil eye. I have referred to it as an incredibly old reference that maybe you won't even get. In the cartoon, the Flintstones there was a character named bad luck, schlep rock. Oh, lousy. Me. And bad luck Schleper act, just anything bad happened to bad luck, schlep rock. And so that construct can be if you've tried to frame it as latent, it would have to imply that everybody has a certain amount of bad luck schlep rock in them quick, what

Patrick 20:58

is your wedding anniversary? It is June. But you remember bad

Greg 21:08

luck? My dad, let me watch five hours on school days. So of course, I remember that. Worry. If you model it is a latent variable, you have got to own it.

Patrick 21:17

Yeah. And we've talked on episodes before that. If you draw a single headed arrow, you go into the saloon, you order two fingers of whiskey, you throw it back, slam the glass on the bar top. And to everyone in the saloon, you say I believe my latent factor causes my indicator.

Movie Clip 21:37

I'm Captain Augustus McCray. This is Captain Woodrow F call. I like a shot of whiskey. And so in my companion, i Besides whiskey, I think will require a little respect.

Patrick 21:48

If you have a series of math items, it is not unrealistic to say you have an underlying math ability, that in part determines the probability you're gonna get an item correct. I can sleep at night thinking that the reason you got a 92% on your test score is because your underlying math ability caused those responses. But as you say, you start thinking about Wait a minute, there's some underlying latent propensity that causes you to be discriminated against in social media or on the playground. And the SEM with a traditional multiple indicator latent variable is very poorly suited to deal with that to the point that you have misstated your causal process.

Greg 22:35

Exactly. Right. So now imagine that we have a model where we really are at odds from a theoretical standpoint, with a construct being represented in that traditional latent framework, it takes a lot of trickery to be able to get the standard structural equation modeling framework to do that. And even when we can, there are all kinds of limitations on our ability to do that. That's true. Well, let's imagine now that someone has a system where there to exogenous constructs are things that no they look at and say there's no way that's latent in the traditional sense, it really seems to make much more sense that it is as we call formative that the variables come together to form that particular construct. In the pls world, the traditional latent constructs where the factor itself is influencing its measured indicators, we typically call that a reflective system. In the pls world, they call that mode A, when the variables are coming into the construct influencing the construct, they refer to that as a Mode B system. One of the beauty of the partial least squares modeling framework is that it doesn't care. I don't mean that you don't specify it. But it says, you get those however you want. And we will take it from here. And I have to say, I find that kind of attractive,

Patrick 23:50

and I find it kind of attractive in the ghost of Hotelling. That is a very similar distinction between principal components analysis and the common factor model. And this is where people threw stuff at each other through the 70s and 80s. And even in the early 90s, right is principal components analysis, a factor model, blah, blah, blah. But what you just described in the different modes really does distinguish PCA and the common factor model. And that is the common factor model is believed to have given rise to the set of items. And the reason we observed the correlation among the items in the way that we did is because they have a shared underlying cause. One might say a common cause, while might say a common factor underlies those, the principal components we can very pragmatically think about reversing those arrows and the items that we have induce a composite that we can compute directly, we don't have to estimate that as a factor score. So that notion of do the arrows radiate out from some latent factor or construct all or are we optimally weighting the items that then induce a composite, that is principal components and common factor analysis

Greg 25:08

in the composite variable system that is PLS, it's not going to be principal component analysis, although there are techniques that try to merge principal component analysis into this. This is not that. So what I thought we would do is talk a little bit about how it works, not necessarily getting into all the weeds, but just generally how this works. And to start, I'm going to imagine that system that I described earlier with two exogenous factors and one endogenous factor, and I'm using the word factor, honestly, a little bit uncomfortably. If I say the word latent variable, I have to say, I'm a little bit uncomfortable, the pls community might go, what's your problem? It's still, we still think of it as something that you didn't see. But the SEM that I started at Studio 54. It's just so deep inside me that I have a hard time even referring to these as latent variables, but I'm the weird one, as far as the pls community would consider. Yeah,

Patrick 26:01

that's why they would consider you the weird. What's the deal with that Hancock.

Greg 26:07

That's it. So let's imagine in that system that we have our two exogenous constructs. And both of those are Mode B. What that means to us coming in from the outside is that those are both formative systems where the variables are coming into form that particular construct. And then let's imagine that our outcome or dependent factor in this model is a traditional latent variable, a reflective system mode A as it would be called in this world, where the factor itself is influencing is responsible for the relation among the variables, an overview of the way pls operates, if we had to come up with something without the benefit of a whole lot of computational horsepower, we could just take the indicators, and honestly, whether it's reflective or formative, whether it's mode A or Mode B, we could just take those indicators and get some sort of proxy score for each of those factors. And that is, what happens in pls is that you start by getting a score to take the place of each of those factors. And it is a simple sum, typically, of the standardized variables that serve as indicators. So all of your indicators for whether it's mode A or mode B, convert them into z scores, you sum them up taking into account direction of things, of course, and you go, boom, I have at least a start or proxy for each of the three constructs that I care about.

Patrick 27:28

And if you just stop there, that's a measured variable path analysis. For those of you who have taken an SEM class, very often, the arc of a story that is told is you do some review of multiple regression. And then you expand the multiple regression to a path analysis, where you have multiple dependent variables, and you have mediation and you have chi square tests, and all of those things. And then you move to a multiple indicator latent factor. Well, what I like is that first step of taking just a measured composite of the items on your construct is a measured variable path analysis. But pls adds this really clever iterative procedure. Yeah.

Greg 28:10

So once we get these proxies for each of our three constructs, what do we do with them now? Well, if that was the end of it, we would just go ahead and do the regression using the to exogenous composites, and have them predict and just a regular old, ordinary least squares regression have them predict that endogenous composite, and we would be done. But when we do that, we do get some estimates for the structural relations in that particular model. Once we have those structural relations in

there, we now actually have a series of relations among all three of those constructs. Between the two exogenous constructs, we have some estimate of correlation between the dependent and the independent constructs, we have some structural relations, some standardized beta weight kinds of things, what we can actually do in this system is we can now get predicted scores for each of these based on the structural relations that exist there. And so in an iterative process, the information that we got from this first pass at estimating the structural connections allows us to get new estimates for those proxies.

Patrick 29:15

So when we think about how we do business, as usual, we're done right? If we're doing the least squares, or if we have a second dependent variable, that's a mediating model, and we're using maximum likelihood. That's it right? What you walk through is we take our items, we make a composite, we take the composites, we fit a model. But here what you're saying is, is Wait a minute, we've got these model implied predicted values of the composite, could we use those in some way to update how we're computing the composite itself?

Greg 29:47

Exactly right. And so once we do that, based on the estimated relations that we have among the constructs, and there are different ways of doing that, but it involves whatever a construct is attached to Whether it's another exogenous construct and endogenous construct, again, there are different ways of doing it. But you try to use the information from that model to get updated scores, just as you said, once you do that, what happens to those updated scores? Well, you take them from that structural model in this world, it's referred to as the inner model, and then you carry it back out to the measurement model, or what here is called the outer model. And how you do that is going to depend whether or not you have a construct that is mode A, which is reflective, traditional the way you and I think about things, or Mode B, if you have a mode a system, then you can get predicted scores for each of those measured indicators. By doing just a simple regression, I can use the proxy to predict indicator one, I can use the proxy to predict indicator two, etc, I can go through one by one and do that in that reflective or mode a kind of system, when I have something that is Mode B that is a formative kind of system. That's where all the variables are actually coming into the construct. That's just a multiple regression, because I now have updated scores for the construct, and I have all of the scores for the indicator variables, I can run a multiple regression. And now I have updated relations between the indicator variables. And the particular construct. Once I have those, what I can actually do now is get new predictive values for that inner model. And there is this iterative process. And you alluded to that earlier, where we go intermodal outer model, intermodal outer model until things stabilize. And when things converge satisfactorily, then we have our final estimates of those constructs. And we just do an OLS regression boom, done and done with the idea that we have created a system whose goal in the end is to try to explain variance. And that's different from what you and I are used to.

Patrick 31:56

That's right, and you start moving into the topic that's getting increasing appreciation in the field, often linked to machine learning, which is distinguishing prediction and explanation, we can have a prediction model, that is, look, we're going to roll up our sleeves, and we're going to do whatever we need to do to maximize our ability to make y and \hat{y} as close together as possible. That is not as much of what we

might think of as an explanation, where we keep Karl Popper's corpse happy. And we impose restrictions on the system and make those testable hypotheses. And that's all in model fit chi squares and RMSEA and things like that, which do not exist in the pls

Greg 32:43

totally, absolutely right about fit. And I want to make sure that we're very, very clear that the system that I described was just an example where we had two exogenous constructs and one endogenous construct, that was just an example. And that's an example that very much mirrors a regression model with two predictors and one outcome. But this extends to all different types of models. If you think about all of the crazy connections that we might have, we might have exogenous constructs going to an endogenous construct going to more endogenous constructs, we might have mediated kinds of systems. So draw the hairy structural model that you're accustomed to drawing, and then attach your indicator variables in a way that is consistent with the way you believe they should be attached for each of those constructs. And some of them might be mode A the traditional latent variable way we're accustomed to thinking about things that reflective system, some of them might be Mode B, which is the formative system where the variables actually pour into the construct, and pls says, whatever, you get it and then hand it to me and I will turn the crank.

Patrick 33:48

So I think we've got a pretty good 30,000 foot face pressed to the window on the aeroplane looking over the Grand Canyon. We've got our head around mode, a Mode B inner, outer. So where do we go next?

Greg 34:02

Well, no more yet, because it would be a Swedish tradition that at this point, we have a fika Do you know what a fika is?

Patrick 34:08

I already went before we started recording. So no, I'm good. You took a fika but I'm fine. He told her Can you help

Tove 34:15

us out on the whole fika thing? So I think it can mean slightly different things but it tends to be a break when we drink coffee, eat cake or something and just relax and talk to

Greg 34:26

people. Thank you. So Patrick, we definitely need to have a fika I have you been to Sweden by the way

Patrick 34:31

I have I love Scandinavia is I've been to Stockholm and I've been to Uppsala.

Greg 34:37

I have never been to Sweden. I just want to say that. So if anybody out there wants to invite me to Sweden. It is on my list. There aren't very many more places on my list. But if you want to invite me to

Sweden, DM me, I totally want to go all right. This has been a lovely fika and we're back. But I have no idea what your question was before we feel good. What What was your question actually?

Patrick 35:00

I really am gonna go to the bathroom.

Greg 35:02

Okay, well, instead of our usual elevator music, we should probably play some ABA

Patrick 35:06

play big red car again, I got a profanity laced text from you a couple of days ago about how big red car got stuck in your head, it was

Greg 35:16

brutal, don't even say big red car anyway, in fact, I'm going to insert some ABA right here anyway, just to clear that from my head.

Patrick 35:33

Hi, everyone. Alright, so I came back from my break, and Greg took a break as well, and he's not here, in postprocessing, I'm going to put a little bit more of the grid car here, just to mess with him to

Greg 35:46

the big man. Traveling there, and we'll travel.

Patrick 35:52

Okay, we're both back. Let's pan back and think about some of the broad characteristics of where this can and can't be applied. And then how we might use this thoughtfully in just pursuing our science. We could

Greg 36:08

break this up in terms of different aspects of the model and modeling process. For example, we could talk about the characteristics of the data and what we need to be in place for pls to be a viable option we talked about and this was one of Volt's concerns, we talked about sem because it's based on maximum likelihood or variations on maximum likelihood, we talk about it being a large sample technique without ever being able to actually define that. But pls in the end is based on an iterative least squares kind of process. And so it puts fewer demands in terms of sample size. That's kind of a nice thing. That is

Patrick 36:47

a nice thing. We've talked about before how so much of the field focuses on sample size in terms of power, which of course it should, you need to know whether to use a poop emoji or an eggplant emoji in your grant application. But much less attention is paid to model stability. And also having a sense of whether you have a sufficiently large sample for those asymptotic properties to come online in the way that we think that they are.

Greg 37:17

Exactly. We already mentioned that it doesn't have the heavy distributional reliance on normality. But as you pointed out, we've come a long way since the 1960s and 1970s. As you and I had a whole episode I think about non normality, right. So in SEM, we've had ways to deal with

Patrick 37:32

this. Yeah. And so it's interesting, because Where Wold was really worried about things like normality, assumption, independence, assumption, things like that. We actually have really good ways now to deal with that. So we have robust standard errors, we have corrected test statistics, we have very well developed ways for handling ordinal items. It's not to say that pls is not still advantageous, but ml isn't quite as handcuffed as it used to be

Greg 38:03

totally agree with that. And MPLS. You know, formally, it's tied to the same distributional assumptions that regression is, and we've talked about that quite a bit, especially back when we were doing OLS. But it also will use a bootstrapping technique to get the standard errors that it needs, which helps to work around some of the other issues, including dependence, right, to some extent,

Patrick 38:23

as long as dependence is a nuisance variable. And this is what we argued in a court ordered setting with McNeish, was about the unnecessary ubiquity of multi level models is Dan is exactly right in everything that he said in there is that there are very good ways of correcting for violations of independent residuals and a whole broad class of models. As long as you don't have to disaggregate effects, as long as there aren't within group in between group effects. Because then what a corrected standard error does if you don't disaggregate as it gives you a proper standard error for the incorrect effect. Yeah. And so it's just another stealing Bollen's line, fine print, that pls is really well suited for addressing violations of independence by using the Bootstrap. But it's still assuming that there's an overall effect that's of interest and not an effect that needs to be disaggregated. In some structural way.

Greg 39:22

I think that's not only a really important point, but a point that gets glossed over in the pls stuff that I have seen.

Patrick 39:29

And I think it gets equally glossed over in the cluster corrected SEM. Oh, you have kids within schools and multiple schools will just say cluster equal schooling you're fine. So I don't think that's unique to pls

Greg 39:42

missing data. Now, I feel like in the structural equation, modeling world we have kind of nailed missing data. This is a regression based world though. What do you tell your students when you teach regression, what do you tell them about missing data?

Patrick 39:54

Well, in standard Oh, well, less, it's less wise deletion. Anybody who's missing anything We'll see you. Thanks for coming out tonight. You don't have to go home, but you can't stay here. Yep. What I do say though, is well, you can move even with the regression framework to a maximum likelihood estimator and useful information, maximum likelihood under certain assumptions that is missing at random, we have an episode on this, we won't get into the details of that. You can also use multiple imputation. And you can do that within a regression framework as well. And that is you build a model for missingness impute values, estimate your regression model, you do that 10 or 20 times gather all the things together and combine them using formulas that exist. What I almost always tell my students in my teaching, though, is you should probably move to a full information maximum likelihood setting within an SEM framework and incorporate partially missing data there,

Greg 40:51

well, then it wouldn't be P L. S, would it? Would be what PML? I'm sorry, by definition. You can't do that in here. That's exactly right. But yeah, I mean, to the extent that I have found any information about it, usually it sort of revolves around, we'll just don't get too much missing data. And, you know, like, try to keep it below 5%. And if you really have to maybe throw some values there in the holes, and it shouldn't matter. All right, you're good to go. Godspeed. And in other worlds that we inhabit, we would go. I'm not so sure about that.

Patrick 41:22

Yep. Although what you just described is what sem did for half a century.

Greg 41:27

It's true.

Patrick 41:30

A lot of this stuff is not unique to PLS, right? ML had a Don't Ask Don't Tell policy on missing data in non normal data and dependent data for a lot of decades, not just a lot of years, a lot of decades. But yes, in the pls in its current. And we want to stress this because this is another area that is ripe for a lot of dissertations or postdoctoral research projects, in its current form. To my understanding, there is not a widely available method to handle missing data in the way that we would with maximum likelihood estimation. That

Greg 42:08

is my understanding as well, for sure. This can get at model complexity in a variety of other ways, but not actually that way. You know, there are things that the original version of structural equation modeling has expanded into, you know, things like higher order constructs, or latent class kinds of things, or mediation, all of those are kinds of things that fit very, very nicely under the structural equation modeling umbrella. A lot of those things represent the methodological developments that have been going on in the last 10 years or so within the pls community as well, trying to expand the type of models that you can address through pls. So you can still have that wonderful benefit of having different types of variable systems mode A Mode B, while having the structural models that are of interest, these inner models that really represent the theory that you care about.

Patrick 42:55

And that's a really good point you're making is there a lot of ongoing developments in pls as we speak? So historically, and you indicated this earlier in the conversation, you just standardized everything? Yeah. And again, that's what we did an EF a for 100 years, right? You standardize your variables, you have the correlation matrix. That's exactly right. But what is the byproduct of that? What you have means have zeros and variances of one? Well, who cares? Nobody cares? What means and variances are? In less you want to study change over time? Yeah, unless you want to do a multiple group analysis in which you examine weak or strong or strict or partial invariance, unless you want to do an MNLF a like thing. And so but I think a lot of people are saying, Okay, we've really figured out the core parts of this. Now, how can we build this out in ways that generalize the applicability in practice?

Greg 43:54

Yeah, it's a very exciting thing, right? And there's the whole pls methodological community is developing these methods, trying to get them incorporated into existing software, whether it's proprietary software or AR based packages. So there's a lot of cool stuff that's going on. And what I hope that we're getting out of this conversation is that pls isn't meant to be sheepskin, although Wold really wanted it to be this thing that does some of the same stuff as SEM, but maybe under different restrictions. I think it complements sem really, really nicely. I don't think that people necessarily should say, Well, I'm doing SVM unless I can't, and then I'm gonna go over to pls. I think it's nice to be able to think about, alright, what's the model? Let me get everything laid out on the table. What's the model in terms of the inner model, that structural portion that we care about the outer model, the measurement part that we care about? What characteristics do I have in my data, and then what might be the appropriate technique to try to draw from and even though we sort of joked about sheepskin, I would like us to think about this as just another option that people have.

Patrick 45:03

One thing I love about this is it's very clever in saying what let's take a deep breath. And let's think about what is another way that we could approach this. So maximum likelihood has these chips that we believe are over the horizon, but we can't see them. And so based upon the characteristics of our sample data, we want to get the best estimates possible. This is not again, to reiterate an earlier point, this is not a method of estimation, this is a fundamentally different way of approaching the entire modeling process. It is not without limitations in the kind of structures that we can do. There's a difference in path models, it can be a full sem or a path analysis. But there are models that are called recursive or non recursive. It is one of the few things in quant that is the opposite of what you would think it is. And that's how you remember it. And that's how you remember a recursive model Colloquially speaking, the influences move from left to right, and there are no correlated residuals, that's a recursive model. All right, there are no feedback loops. There are no bi directional effects. And there are no correlated disturbances in your dependent variable, a non recursive model, you can have either or correlated disturbances and feedback loops, right? Pls is currently limited to recursive models. I don't feel like that's a death knell, I think a lot of our models are recursive, not all of them. I think in cross sectional data, we don't often have feedback loops, I gotta tell you, it's very common in my own kind of work, where I do have correlated disturbances. So say, I have two mediators. And so I have several exogenous predictors. I have two mediators that are like separate mediators is that there are

two specific indirect effects. It's very common to correlate the disturbance of those two mediators. And currently, we're not able to factor that into the pls approach. And so

Greg 47:12

that might be one of those branch points for you that says, I better keep it in the SEM world, as Eric covariances in your measurement model might be as extreme cross loadings, ones that are really non ignorable, that's harder in this particular system. Or honestly, just if you want an assessment of fit, right, the world here doesn't tend to emphasize global fit, because the whole goal is to maximize variance not to explain the variances and covariances behavior as part of a larger system, it's driving at trying to explain R squared for things that are endogenous within the part of your model that you actually care about. So if you say, and I think this is part of where Wold was getting at, you know, if you have a model that you have a strong theory about, but you really want to put it to the test, pls isn't necessarily this global testing framework. It emphasizes more about the estimation of things rather than the actual explanation, per se.

Patrick 48:07

I really like the approach. I mean, I like conceptually what it's trying to do. What it's conceding, right. We've talked before about how you have to concede certain battles in a war so that you can marshal resources in another part of what you want to achieve. I like all of that. One thing to keep in mind is that one of Volt's original motivations was to operate in more of an exploratory perspective, that we don't always have a really strong a priori sense of what we're trying to do with the data and paraphrasing his term as we're often in a data rich theory, poor setting. I very much see that. But it's important to realize we're still locking things down in this model, in a confirmatory kind of way. That is, pls is not going to tell us the optimal number of constructs we need, we have to define the constructs and define the items that go with the constructs, pls is not going to be kind of like a regularization method where smaller parameters leave the model and bigger parameters stay in the model, we still have to state what the structure is that we are interested in among our constructs. And so we just have to keep in mind that this is not exploratory in an AFA kind of way, but it is less rigid in a cause indicator versus effect indicator perspective on what leads to that composite in your model.

Greg 49:41

Yeah. And you know, the world that you and I tend to live in that structural equation, modeling world is all about constraints, constraints in the measurement portion of the model constraints in the structural portion. And we might have constraints in the structural portion of this model, but we're still only estimating relations in a view Very traditional OLS kind of way that doesn't necessarily feel those constraints because things are so very partial in the way that we do things. One thing that Volvo had said in many of his writings was the equivalent of I don't know what it would be in Swedish, but it'll all come out in the wash, whatever the Swedish version of that would be.

Tove 50:19

So there's no literal translation. But in Swedish, we say something like just gonna last to say,

Greg 50:25

as we get larger sample sizes, as we get more indicator variables. In the end, you latent variable structural equation modelers and US PLS, people. In the end, we all just sort of come together and reach the same inferential conclusions, which he would argue and I think, maybe reasonably so is the goal, in the end is to try to understand what the relations are, or at least approximate them reasonably, even if not exactly.

Patrick 50:50

It's all about how can we take the data that we have available to us and make a valid and reliable inference about the nature of the relations among the constructs in a way that helps us understand something that we didn't know before? This is just another arrow in our quiver of saying we have a theoretical question, we have data available to us. And we want to make a probabilistically based inference about the nature of the relations among our measures. And this is just another way of approaching that problem. And

Greg 51:23

for those of you who are steeped in a PLS tradition, maybe you will think about some of the comparative benefits of structural equation modeling, if you haven't thought about that before. And for those of us who are in this SEM kind of world, there is this other option out there pls that might be tailored very, very well to our model, our purpose our data, and I think we should consider that, as you said, another arrow in our quiver, this is

Patrick 51:47

target rich for dissertations and Master's in grant applications, things like missing data, how would we extend this longitudinally? What are diagnostic measures that might be available? And then one that I'm really interested in myself, and we've alluded to this before, when we were talking about to stage least squares is, might we, for a given model, have a principled way of estimating our model using full information, maximum likelihood using two stage least squares, using partial least squares, and use those results jointly to try to triangulate on a set of relations that we have the greatest confidence in? And if those three methods converge on a discussion section outstanding? If they don't, well, then it's an intellectual goose to say, we've got to better understand why are these different from one another? And in which of these do I have the greatest confidence, because if you only rock back and forth and say, maximum likelihood is consistent, efficient, unbiased, and asymptotically normally distributed, and that's all I'm going to do. You are blinding yourself to other insights into your data that might help you make a better data informed decision about the nature of your constructs.

Greg 53:08

I like that point very, very much. So as per usual, you and I put a tremendous amount of planning into this episode, when I'm texting each other.

Patrick 53:18

It's 930. Last

Greg 53:19

night, 1130 last night, but you know, also you and I are not experts in pls. It's something that you and I are gaining familiarity at. And as you said, if someone out there is listening who has expertise in this, they'll say oh, but they didn't talk about this. They didn't talk about that. Yeah, that's absolutely true. At the end of this episode, there are going to be a lot of pieces leftover on the table.

Patrick 53:38

Just like when you buy something from IKEA. Team first dish team Dorkin FIRFER.

Tove 53:47

That's not a real word

Greg 53:48

Good night, everybody.

Patrick 53:50

Thank you, everybody. Take care. Bye bye. Thank you so much for listening. You can subscribe to corner today on Apple podcasts, Spotify, or wherever you download your cacophonous noise to drown out midterm political ads. And please leave us a review. You can also follow us on Twitter we're at quantity food pod and check out our webpage at quantity and pod.org for past episodes, played live show notes, transcripts and other cool stuff. Finally, you can get quantitative themed merch -- get the real stuff not the fake -- redbubble.com We're all proceeds go to Donors Choose to support low income schools. You have been listening to Quantitude, the only official pumpkin spice podcast for Fall corner today has been brought to you by the double asteroid redirection test in which NASA launched a rocket 7 million miles to intercept a lump of rock 500 feet across at 14,000 miles per hour with the sole intention of making the rest of us feel bad about our own contributions to science by corner tunes Tova Nadir 3000 A newly available download that allows everyone to have their very own personal Swedish interpreter, and by the House of Windsor who proudly anoint Quantitude the Royal podcast consort. This is most definitely not NPR.

Greg 55:21

In fact, I'm going to insert some ABA right here anyway just to clear that from my head

Patrick 55:36

Hi, everyone. Alright, so I came back from my break and Greg took a break as well and he's not here. And so, in post processing, I'm going to put a little bit more of the grid corner just to make no

Greg 55:49

no no