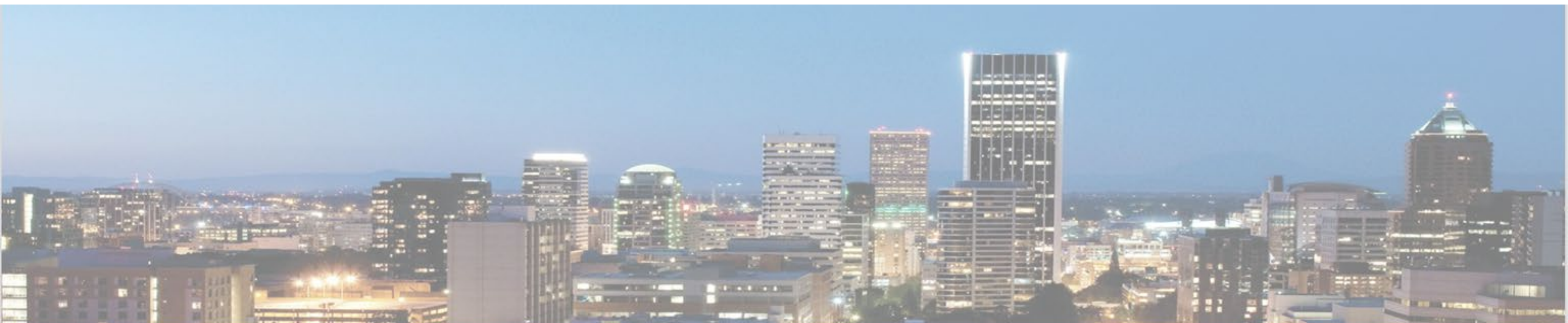
Stats meeting

Interaction analysis for continuous / survey data

Dr. Frederik Vos



The School of Business
PORTLAND STATE UNIVERSITY

My background in psychology trained me a bit deeper in statistical approaches – this presentation will cover some of my (limited) insights

Introduction

- **Background**:
  - Psychology,
  - Communication Science,
  - Business Administration / Supply Chain Management

Current:

- Assistant Professor in Supply Chain Management at the School of Business @ PSU

<u>Few comments before we start:</u>

- This will be a broad overview of different ways to analyze your data
- I am not a full-sledged statistician by training, so I will try to not anger the real statisticians too much ;)
- The focus will be on application and usefulness, less on the underlying requirements and steps to take, so I will short-cut some explanations. I can go in future presentations deeper into some of the approaches presented here

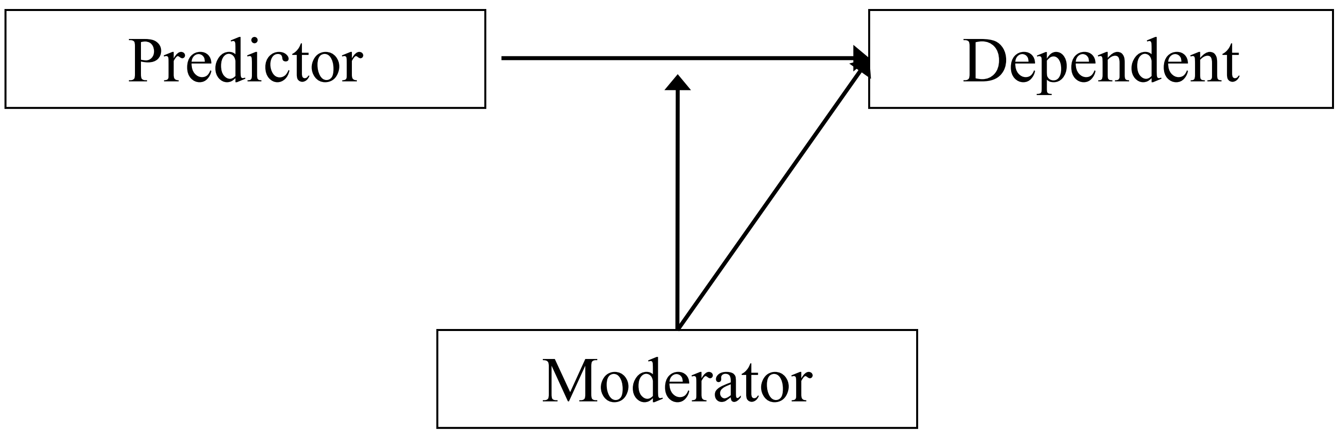
We will focus on 4 different ways to analyze interactions for continuous data – I will not take into account more complex designs, such as time-lagged designs etc.

Intro

Part	Focus on
1	“Standard” multiple regression interaction analysis
2	Polynomial Regression with Response Surface Analysis
3	SEM Multi-group analysis (MGA)
4	Qualitative comparative analysis (QCA)
5	Discussion – other ideas?

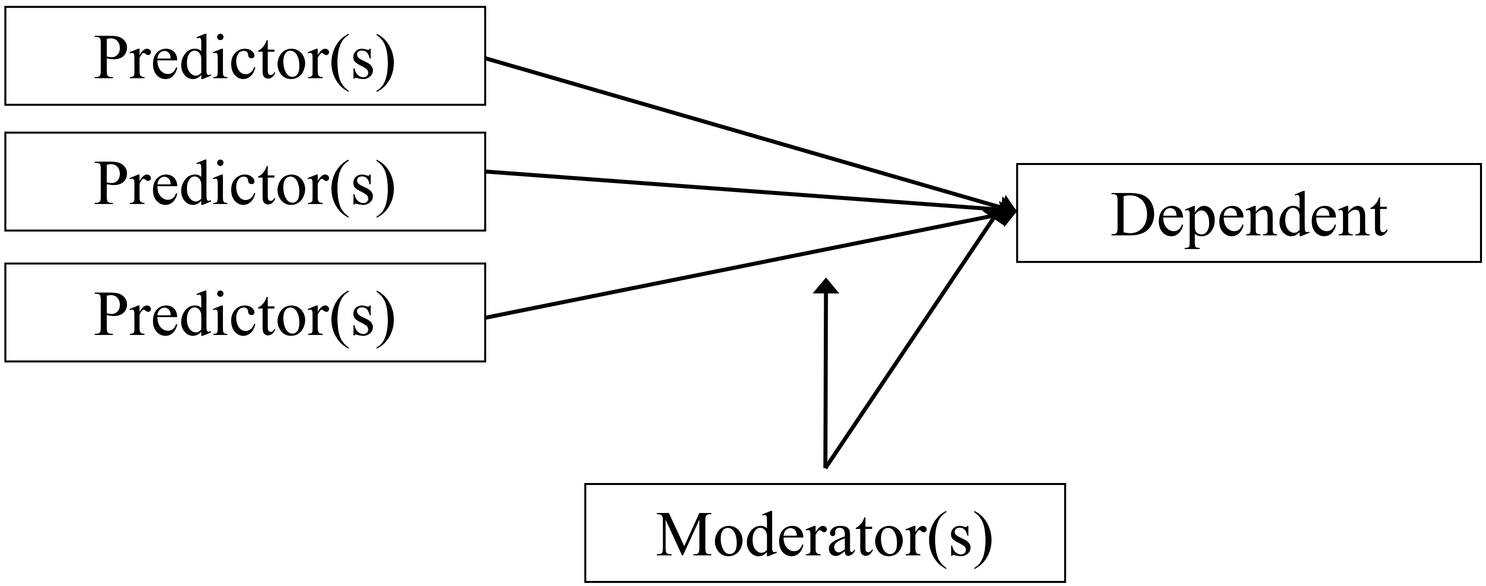
This is a simple representation of an interaction – in one way or the other, we are discussing variations of this today

Model for moderator effect



Also, more complex models are possible, including also more predictors and moderators – I try to not make it too complex

Model for interaction effect



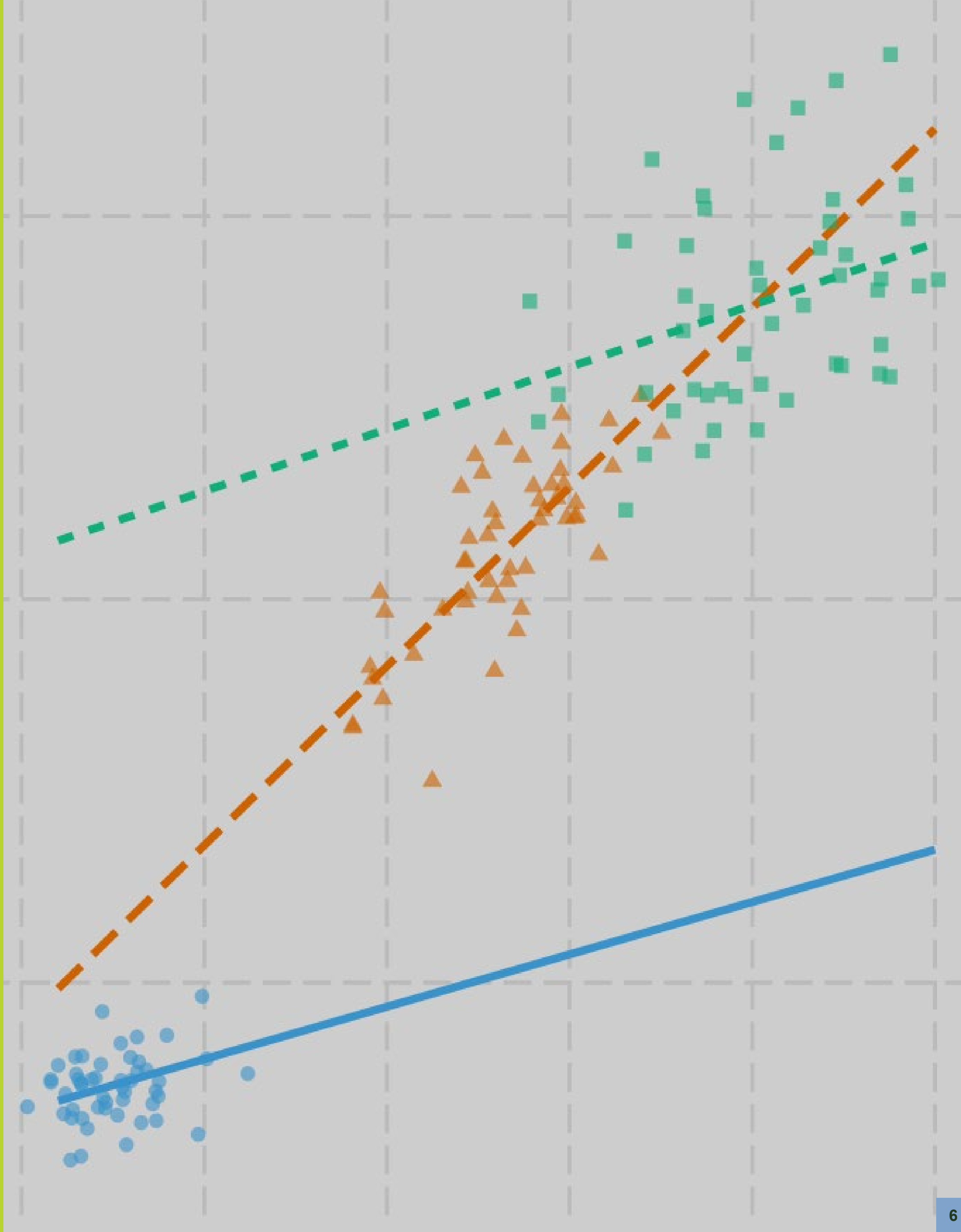
The Moderator affects the relationship between the predictor(s) and the dependent

–

Yet also combination of interactions between multiple variables are possible (more about this later)

1

Interactions in “standard” multiple regression



Moderator in a regression

Moderator effect =

- variable **Z** affects the (direction and/or strength) of the effect of **X** on **Y**, or

- Effect of **X** on **Y** depends on the level of **Z**.

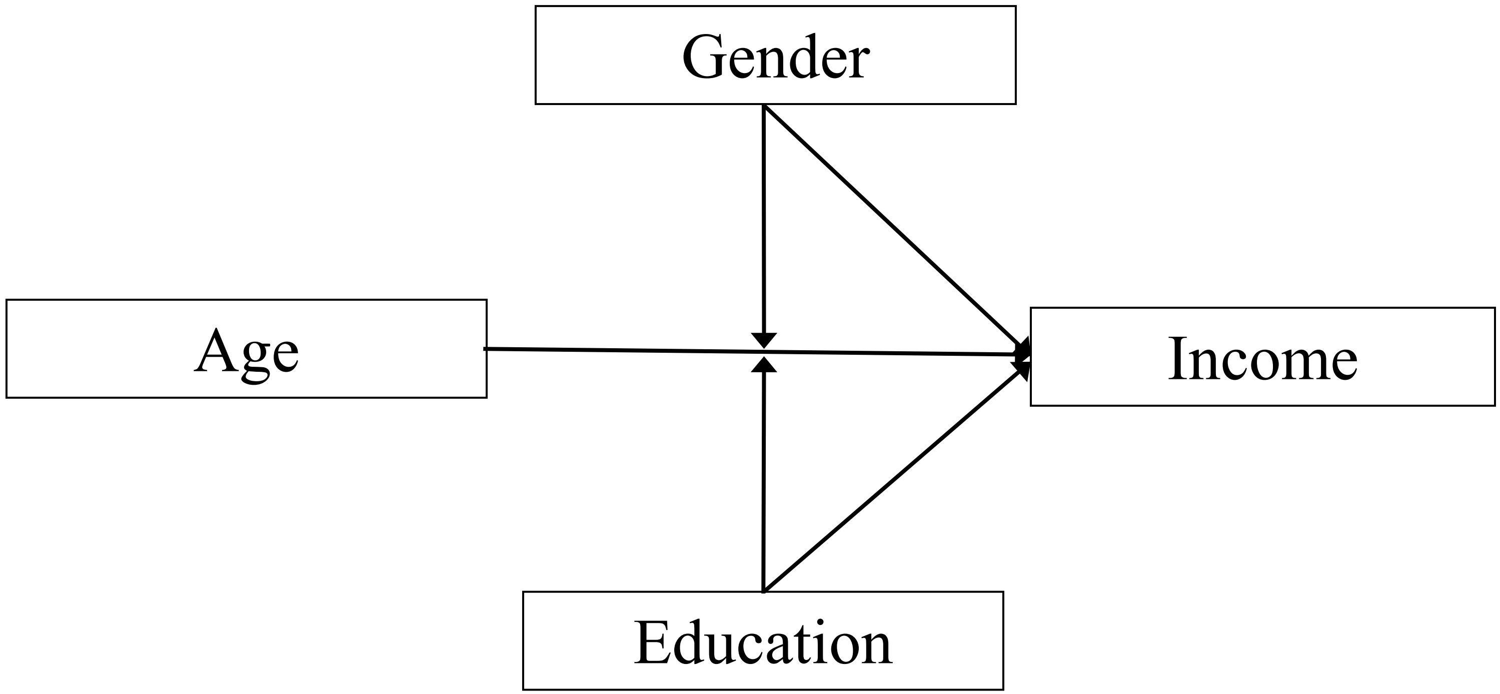
- Corresponds to an interaction effect of **X** and **Z**
  (= **X** * **Z**) on **Y**

In a standard OLS regression, can be represented as:

- $Y = a + bX + cZ + d\, X{*}Z$

A simple example of a standard OLS regression analysis involving interactions – variations of this are also possible for path modelling (like Lisrel)

Regression



This is an example of Age and Education influencing Income

Example (numbers are made-up)

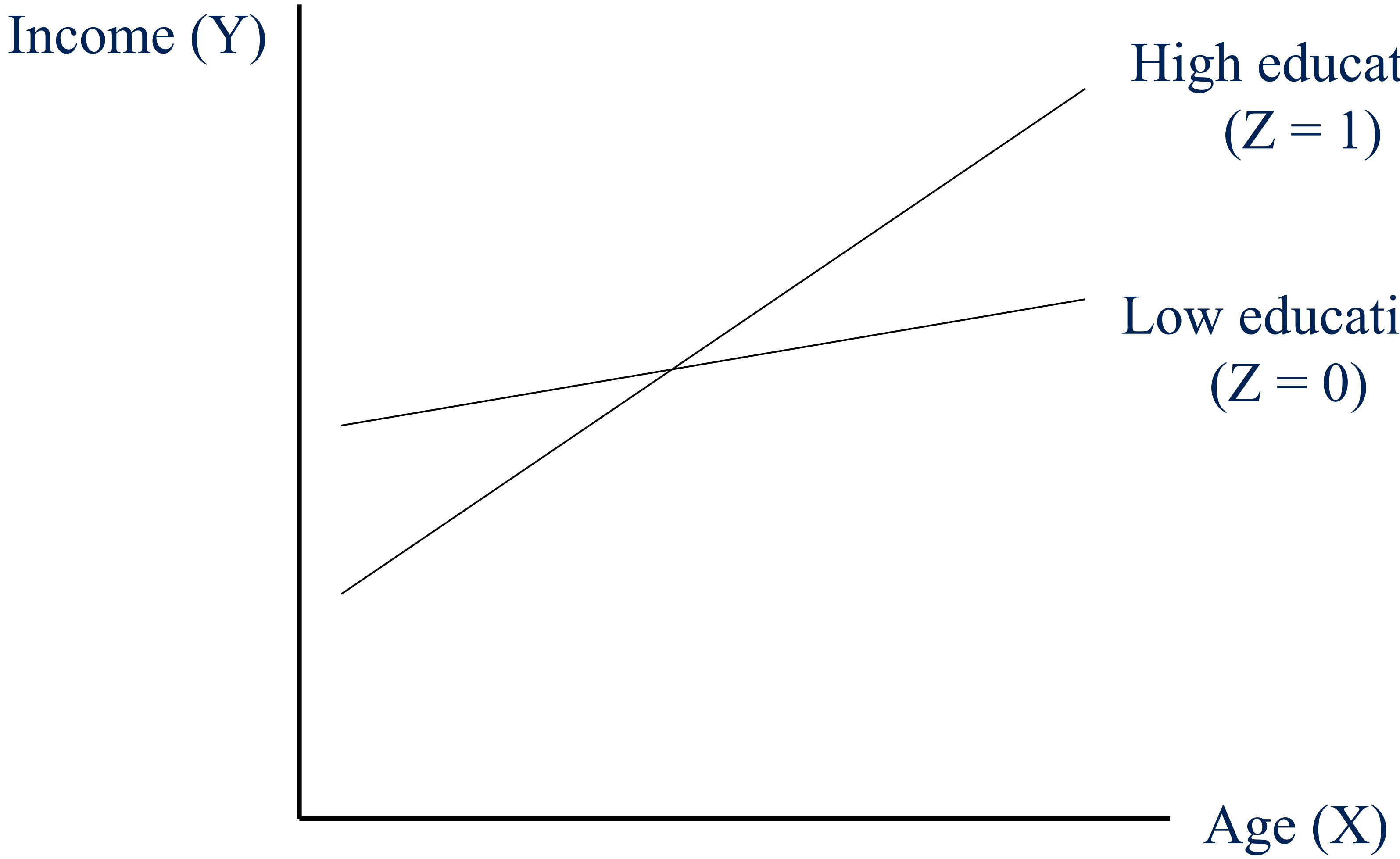
Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	25,123	,235		106,885	,000
	Age	6,850	,817	,160	8,383	,000
	Gender	-,236	,817	-,006	-,289	,772
	Education	,671	,030	,461	22,221	,000
	Age x Gender	-,021	,111	-,004	-,188	,851
	Age x Education	,322	,096	,067	3,353	,001

a. Dependent Variable: income

When visualizing an interaction, a Figure like this can be created

Visualization of the result



There are plenty of options for performing a “standard” OLS interaction analysis

Software packages and remarks

Software solutions:

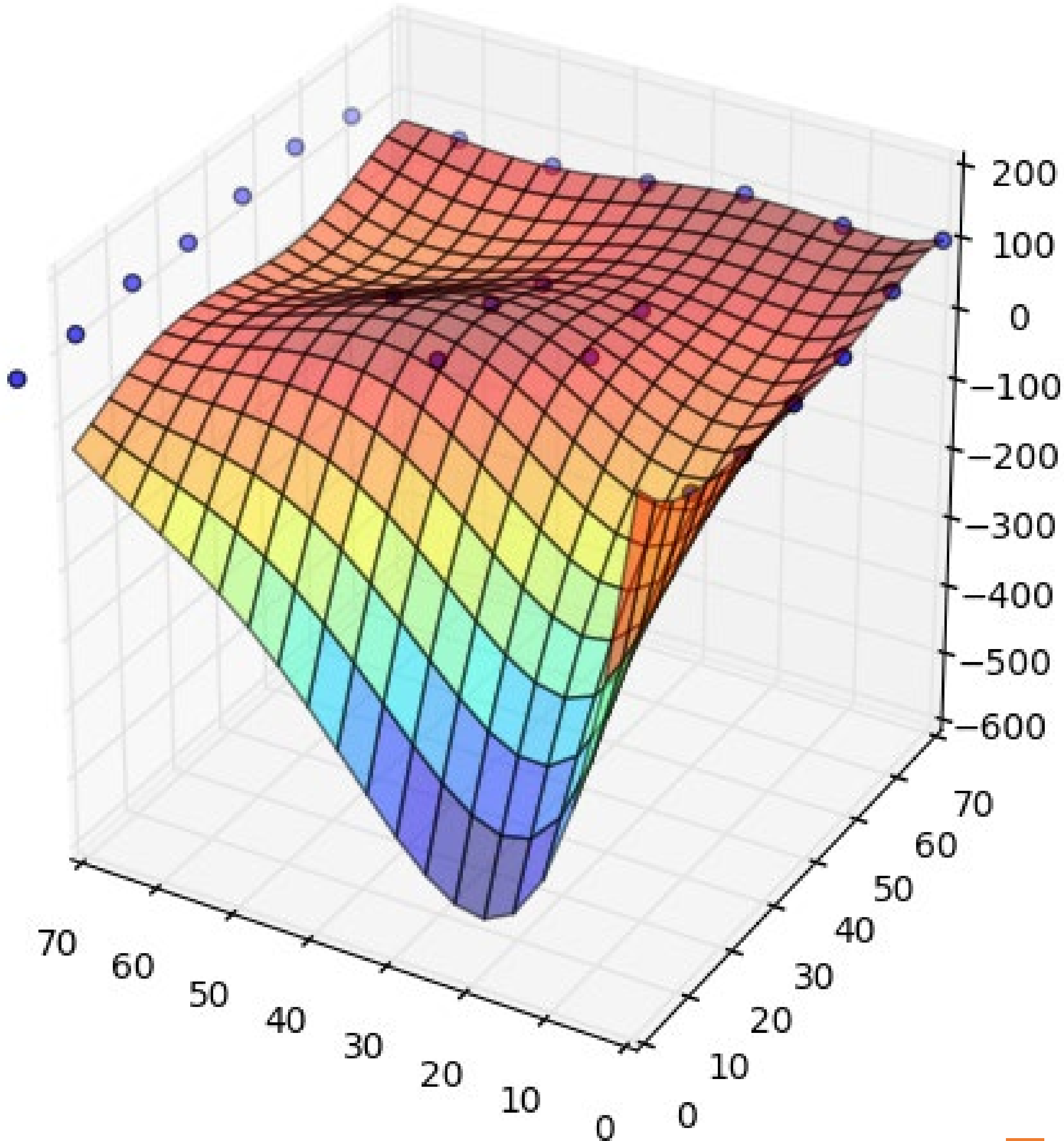
- Excel
- Almost any statistical software

Remarks:

- I presented this to get get everyone on-board with the standard analysis (most of us should know)
- It gives some meaningful insights
- More advanced methods are available

2

Polynomial Regression and Response Surface analysis



Polynomial analyses with response surface analyses helps us to assess interactions in new ways. This is one of the applications

Beyond simple interactions

Journal of Purchasing and Supply Management 24 (2018) 343–351

Contents lists available at ScienceDirect

Journal of Purchasing and Supply Management

journal homepage: www.elsevier.com/locate/pursup

ELSEVIER

The effects of balanced and asymmetric dependence on supplier satisfaction: Identifying positive effects of dependency

Marjolein C.J. Caniëls[a,*], Frederik G.S. Vos[b], Holger Schiele[b], Niels J. Pulles[c]

[a] *Faculty of Management, Science and technology, Open University of the Netherlands, P.O. box 2960, NL-6401 DL Heerlen, The Netherlands*
[b] *Faculty of Behavioural, Management and Social sciences, University of Twente, P.O. box 217, NL-7500 AE Enschede, The Netherlands*
[c] *Department of Operations, Faculty of Economics and Business, University of Groningen, P.O. Box 72, NL-9700 AB Groningen, The Netherlands*

A R T I C L E I N F O

Keywords:
Buyer-supplier dependence
Supplier satisfaction
Polynomial regression

A B S T R A C T

Studies argue that balance in dependence is critical to supplier satisfaction in buyer-supplier relationships. We examine whether asymmetric relationships can also lead to supplier satisfaction, arguing that traditional analysis methods are unsuitable for thoroughly analyzing this issue. With polynomial regression and response surface analysis combined with dyadic data, we test the relationship between (1) balanced dependence (i.e., the buyer and supplier are equally dependent on each other) and supplier satisfaction and (2) asymmetric dependence (i.e., either the supplier or buyer is the dominant party) on supplier satisfaction. The results indicate that mutual dependence is positively related to supplier satisfaction, but surprisingly, asymmetric dependence can be related to higher levels of supplier satisfaction.

In this paper, we wanted to assess whether dependency is always bad – are suppliers unhappy when being dependent on a buying firm?

The problem – how to study dependency?

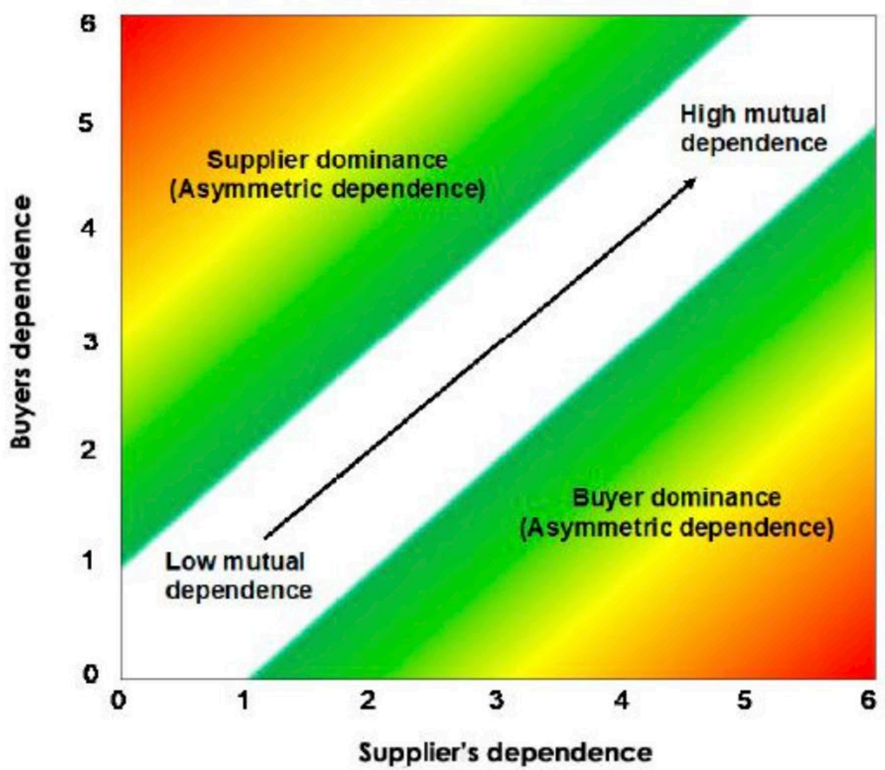


Figure 3
Buyer-supplier dependence.
Source: Adapted from Caniëls, Vos, Schiele, and Pulles (2018).

Traditional approaches to measure / model dependence between parties used to be:

- the **algebraic difference** between dependencies (Joshi, 1998; Yilmaz and Kabadayi, 2006)
- the **average or the sum** of these measures (Gundlach and Cadotte, 1994)
- or use **spline scores** (Gulati and Sytch, 2007; Kumar et al., 1995)

→ Each of these approaches reduces variation or does not show the full picture

- ➢ Solution: polynomial regression with response surface analysis (Edwards, 1994; Shanock et al., 2010)
- ➢ It also includes two non-linear effects (X^2 and Y^2) rather than only an interaction term (cross product XY)

We collected dyadic data and performed a polynomial regression to model the effect of dependence – all variables were continuous

Multiple regression as basis

Table 4: Polynomial regression examining the impact of buyer dependence and supplier dependence on supplier satisfaction

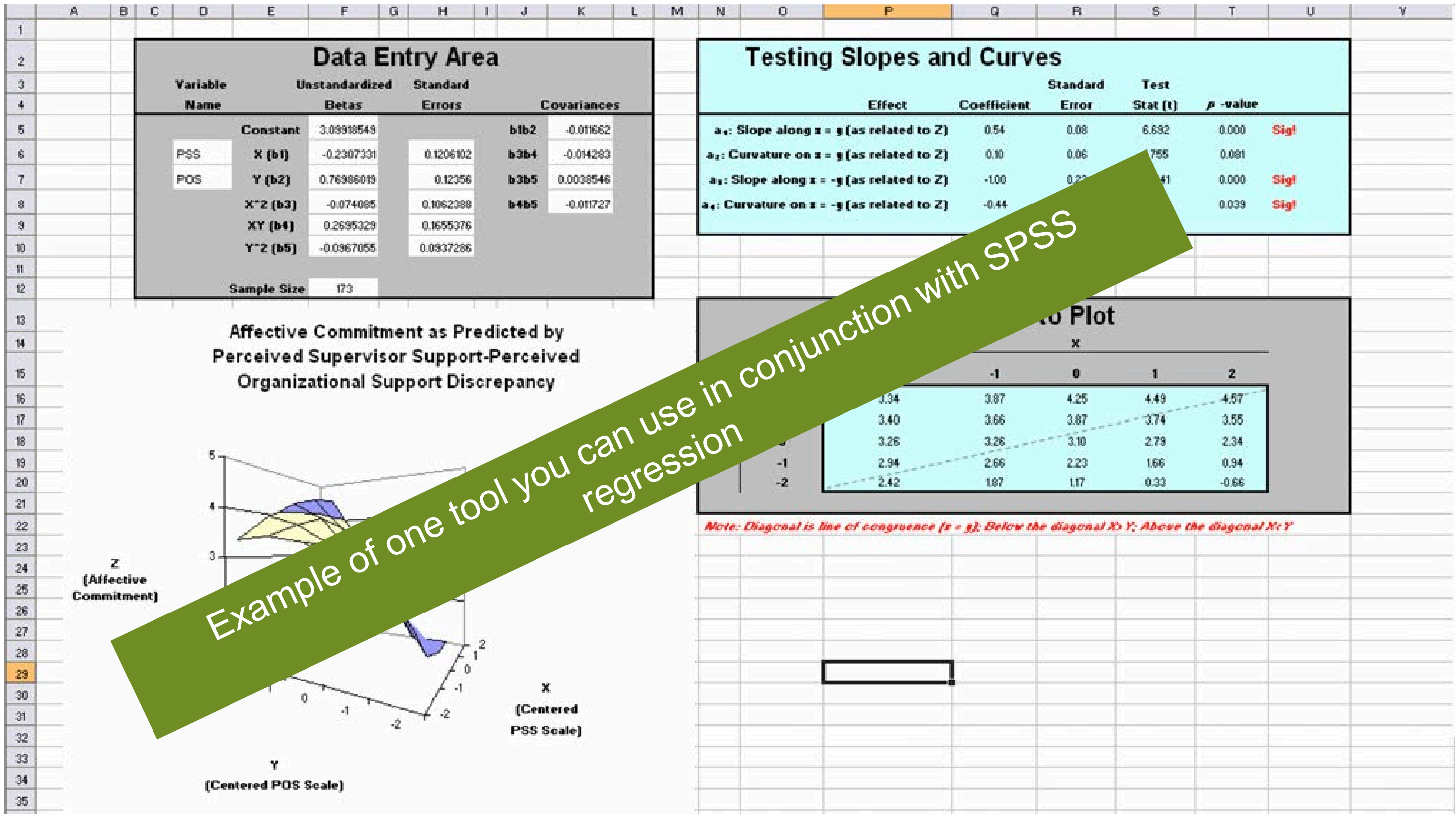
Variables	**Dependent: Supplier satisfaction**					
	Model 1		Model 2		Model 3	
	B	SE	B	SE	B	SE
Step 1						
Length of relationship	.00	.00	.00	.00	.00	.00
Step 2						
Buyer dependence (X)			.21**	.05	.05**	.36
Supplier dependence (Y)			.08*	.04	.12*	.07
Step 3						
X^2					.04	.04
Y^2					.04	.04
X * Y					-.06*	.04
Adjusted R^2	-.01		.18		.20	
R^2 change	.00		.20**		.04	

*Notes: *=p<0.05; **=p<0.01; B=unstandardized regression coefficient; SE= Standard error; N=109; Bootstrap samples=5,000.*

What would be the conclusion here?

The data from the regression and additional covariance data was then imputed into a special excel file (from Shanock et al., 2014)

Step 2 of the analyses



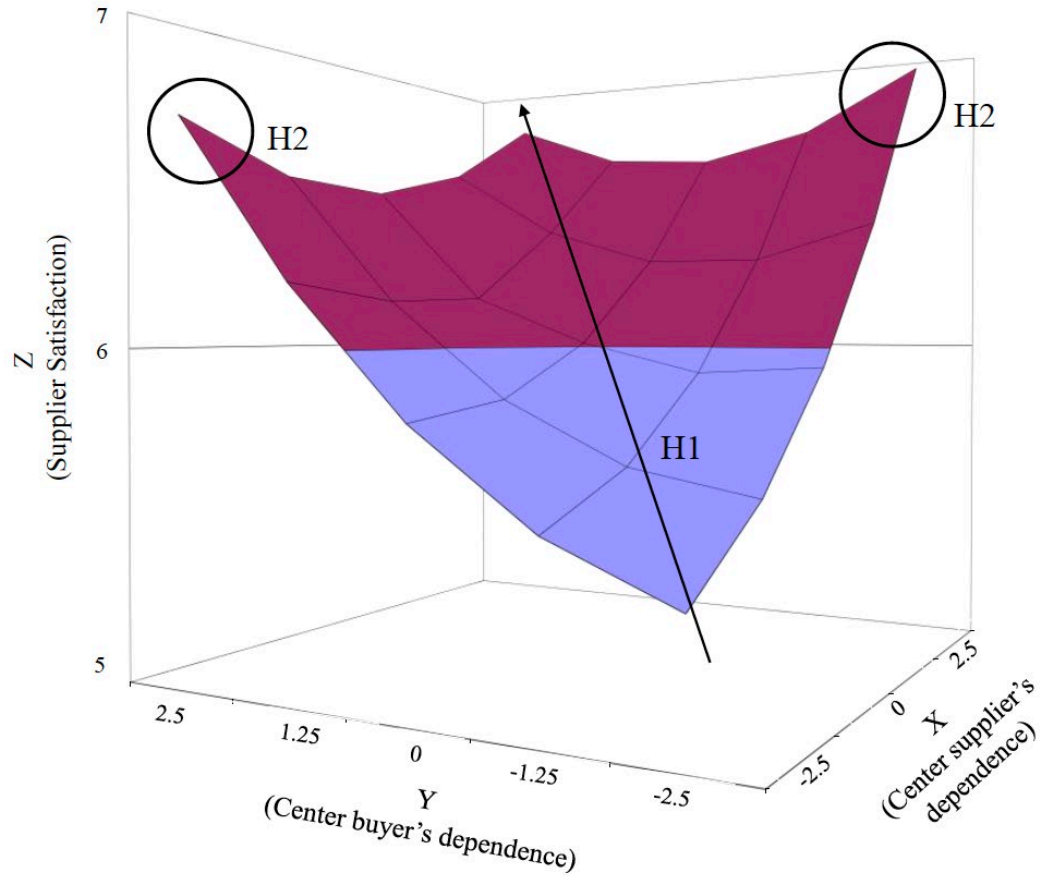
I recorded two videos to guide these analyses steps in SPSS – however, take care that the tools in R might be more sophisticated:

Discrepancy test / Step 1: https://youtu.be/_TbeXxhz8Kc

Analyses steps / Steps 2-5: https://youtu.be/jeqVy7PdhUw

The final result showed a 3D model of the interaction

Results of the analyses



- The results are very different from what we would have expected from the simple negative interaction of X and Y

Table 5 Analysis of Slopes and Curvatures, effects as related to supplier satisfaction

	Shape along balance line; Supplier dependence = buyer dependence (X=Y)		Shape along asymmetry line; Supplier dependence = - buyer dependence (X=-Y)	
Slope	a1 = b1 + b2	.30**	a3 = b1 – b2	.06
Curvature	a2 = b3 + b4 + b5	.02	a4 = b3 – b4 + b5	.13*

Notes: * $p < .05$, ** $p < .01$. *a1 and a2 represent the slope of each surface along the X=Y line, while a3 and a4 represent the slope of each surface along the X=-Y line, where b1, b2, b3, b4, and b5 are the unstandardized coefficients on X, Y, X^2, XxY, and Y^2, respectively.*

SPSS and R are quite commonly used for this type of analyses – yet the question remains whether more complex analyses are possible

The response surface analyses analysis I presented is one version of it performed via SPSS.
Common software solutions

- SPSS regression in combination with the Excel of Shanock et al. (2014).
- R-Package - Response Surface Analysis (RSM)
- R-Package - Response surface analysis (RSA)

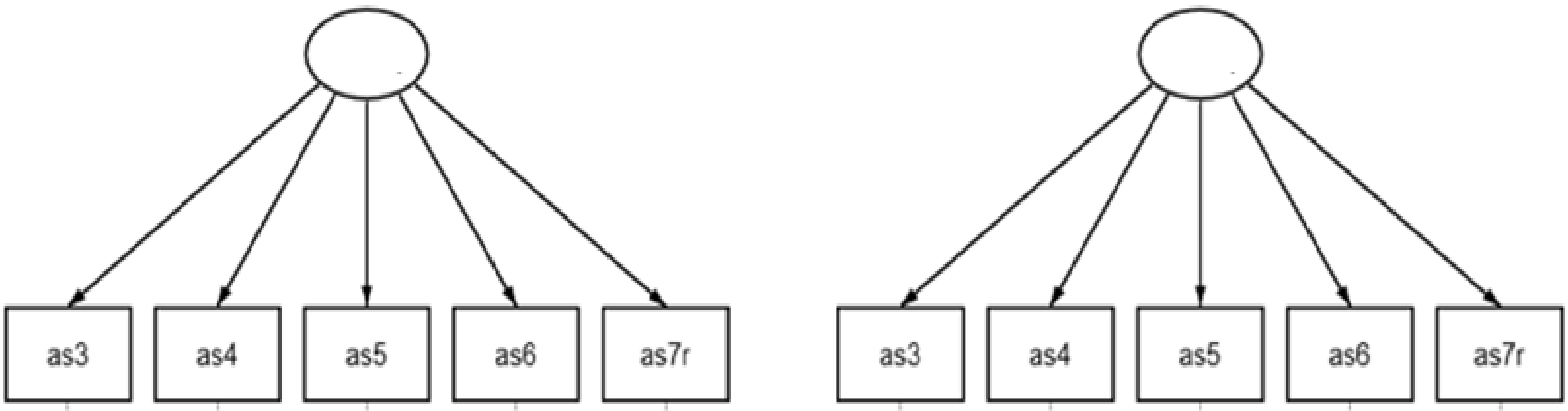
Potential other platforms that seem to support RSA (I did not assess them in detail)

- SAS/STAT (RSREG Procedure)
- Matlab
- Stata
- Python

Critical remark:

- It gives more insights than "standard" OLS interaction analysis
- A methodological reviewer at Journal of Management criticized the method – asking for more accurate modelling of the response surface, being able to identify areas of significance.
- I am not yet aware of any more sophisticated analyses already existing "as is" without substantial manual modelling/coding.
- What about complex models, such as SEM? We will cover this next.

③ SEM Multi-group Analysis



This is a paper we published with an MGA in it

Journal of Business Research 69 (2016) 4613–4623

Contents lists available at ScienceDirect

Journal of Business Research

Supplier satisfaction: Explanation and out-of-sample prediction

Frederik G.S. Vos *, Holger Schiele, Lisa Hüttinger

Faculty of Behavioral, Management and Social Sciences (BMS), University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 1 November 2015
Received in revised form 1 February 2016
Accepted 1 March 2016
Available online 26 April 2016

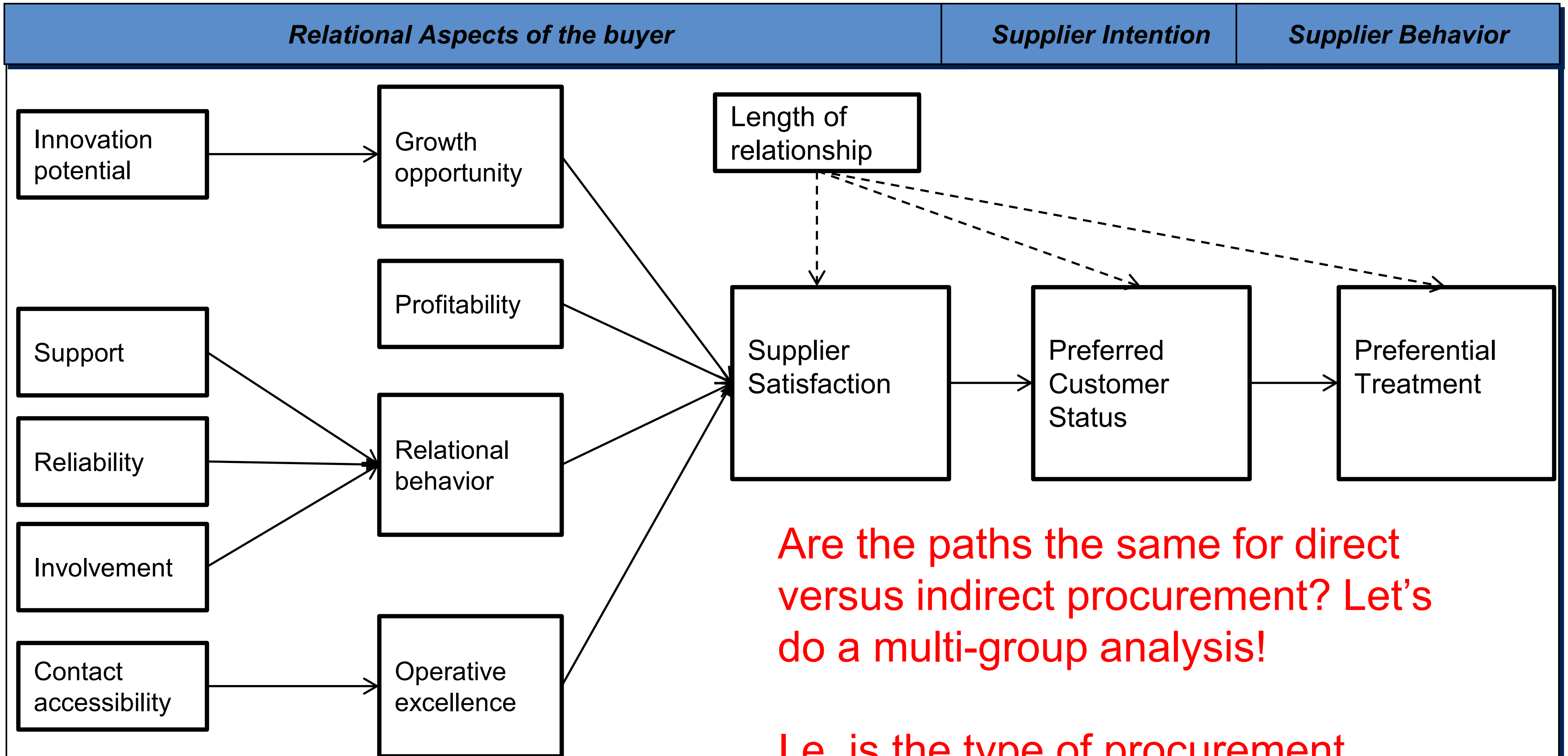
Keywords:
Supplier satisfaction
Preferred customer
Resource allocation
Cross-validation
Prediction
Replication

ABSTRACT

Many firms not only compete for customers, but increasingly compete for suppliers. Supplier satisfaction is a necessary condition for gaining and maintaining access to capable suppliers and their resources in this new competitive environment. This research replicates and extends the previous empirical research on supplier satisfaction. Additionally, this study tests an extended model for direct and indirect procurement, which assesses antecedents as well as consequences of supplier satisfaction. The findings indicate that next to growth opportunities and reliability, profitability of the relationship has a major impact on supplier satisfaction for both direct and indirect procurement. The results also show that supplier satisfaction has a positive impact on awarding the buyer preferred status, ultimately leading to preferential treatment. An additional exploratory analysis suggests the possibility for a hierarchical model consisting of first- and second-tier antecedents of satisfaction, which are particularly useful in direct procurement. Ultimately, the study provides a guide for purchasers to identify the dimensions of satisfaction to manage for satisfactory buyer–supplier relationships, namely perceived growth opportunity, relational behavior, operative excellence and profitability. The application of the new procedure for creating cross-validated, out-of-sample point predictions reinforces the practical relevance of these findings, which indicates a satisfactory prediction of cases outside the modeling sample.

We wanted to assess whether all relationships in this model are similar for different types of procurement situations (direct versus indirect procurement)

Research model



Are the paths the same for direct versus indirect procurement? Let’s do a multi-group analysis!

I.e. is the type of procurement influencing the the relationships differently?

This analysis was performed in SmartPLS

Application

- In a multi-group analysis, we estimate each group separately
- Then, the two models and findings need be compared
- Is it a significant difference?

Table 6

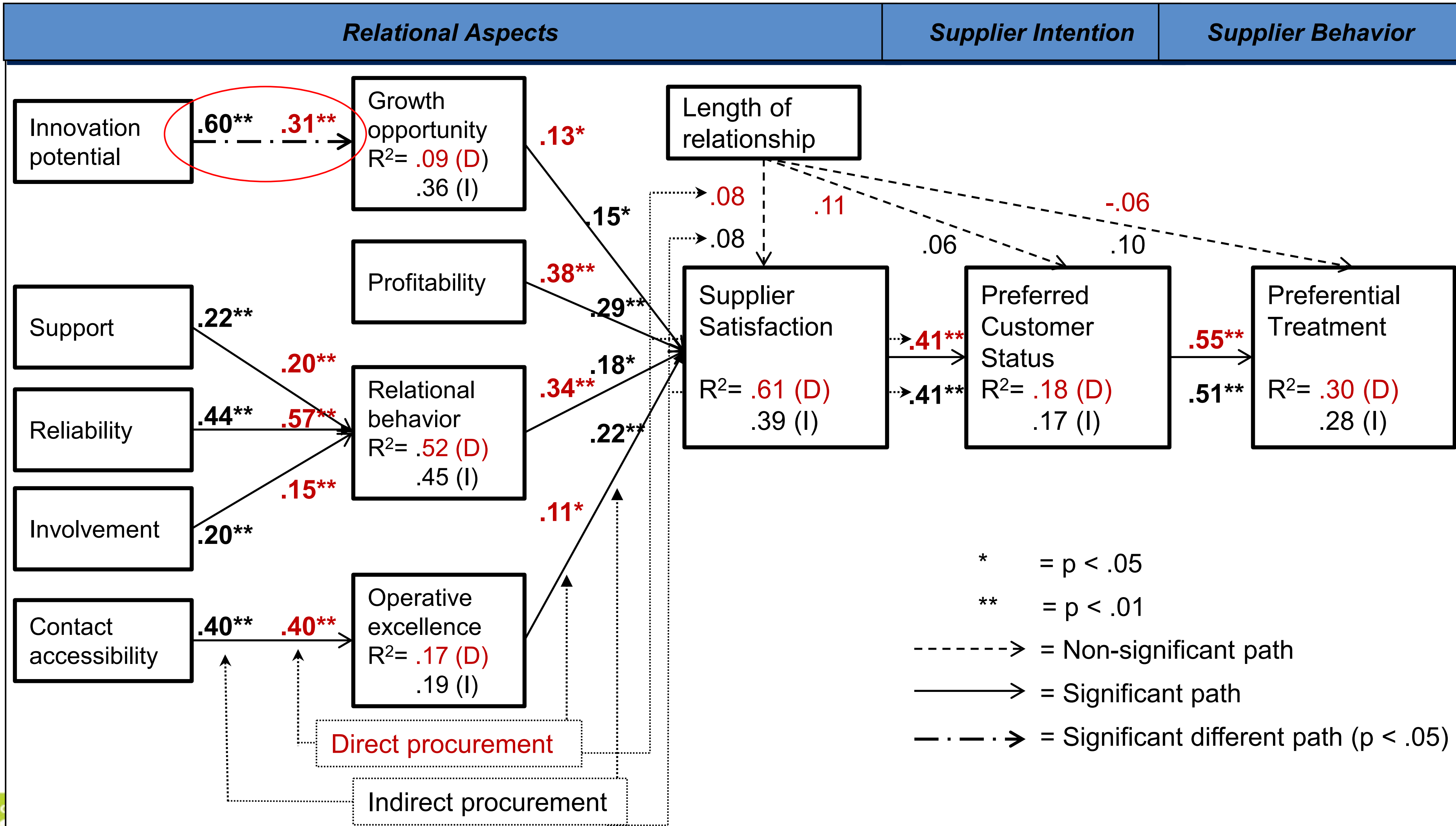
Bootstrap and effect statistics of the revised models (bootstrap samples = 5000).

Paths	β	SE	t	f^2	β	SE	t	f^2	DIFFMGA
	D	D	D	D	I	I	I	I	D|I
IP ≥ GP	.31	.07	4.64**	.10	.60	.06	1.81**	.56	.29*
S ≥ RB	.20	.06	3.32**	.06	.22	.08	2.64**	.06	.02
R ≥ RB	.57	.06	10.00**	.58	.44	.08	5.19**	.29	.13
I ≥ RB	.15	.06	2.68**	.04	.20	.08	2.52**	.06	.05
CA ≥ O	.40	.07	6.20**	.20	.40	.08	5.10**	.20	.00
G ≥ SS	.13	.05	2.33*	.03	.15	.09	1.71*	.03	.03
P ≥ SS	.38	.06	6.14**	.19	.29	.06	4.68**	.11	.08
RB ≥ SS	.34	.07	5.00**	.15	.18	.10	1.83*	.03	.16
O ≥ SS	.11	.06	1.79*	.02	.22	.08	2.74**	.05	.10
DR ≥ O	.07	.07	.94	.01	.13	.08	1.63	.02	.06
DR ≥ SS	−.05	.05	.98	.01	−.01	.07	.09	.00	.05
L ≥ PT	−.06	.07	.79	.00	.10	.08	1.33	.01	.16
L ≥ PC	.11	.07	1.54	.01	.06	.07	.87	.00	.05
L ≥ SS	.08	.05	1.46	.01	.08	.06	1.38	.01	.01
SS ≥ PC	.41	.07	5.46**	.20	.41	.08	5.22**	.20	.00
PC ≥ PT	.55	.06	9.68**	.42	.51	.06	8.53**	.36	.04

Notes: D = direct procurement; I = indirect procurement; β = standardized coefficient beta; t = t-statistic; SE = standard error of β; f2 = effect size of variance explained by predictor; DIFFMGA = difference in the multi-group analyses between direct and indirect procurement; * = $p < .05$ (one-sided); ** = $p < .01$ (one-sided); CA = contact accessibility; G = growth opportunity; I = involvement; IP = innovative potential; DR = days to respond to the questionnaire (Control); O = operational excellence; P = profitability; RL = reliability; treatment RB = relational behavior; S = support; L = length of relationship (Control); SS = supplier satisfaction; PC = preferred customer status; and PT = preferential treatment.

In the end, we found no major differences and this allowed us to say that the model seems to be quite robust for different procurement situations

Results



There are plenty of options for performing multi-group analyses (MGAs)

Software packages and remarks

Software solutions:

- PLS-based SEM (very convenient, but less accepted in top journals)
  - SmartPLS (MGA algorithm)
  - Adanco (MGA)
  - …
- Covariance based SEM
  - Amos (SPSS) (PSU example)
  - Mplus (PSU example)
  - SAS (Multiple-Group Analysis)
  - Stata (Multiple-group generalized SEM)
  - R-package Lavaan (Multiple Groups)
  - …

Remarks:

- PLS implementation quite easy to perform, in covariance-based SEM a bit more complex
- Possibilities for insignificant – significant findings exist (Sig. is different in the groups, but difference in betas not significant different between groups)
- Challenge how to split continues data into groups (Mean, Median, percentiles, middle of a scale?)
- There are several guidelines on how to perform MGAs, check the most recent ones!
- As alternative, Mplus offers latent factor SEM interaction analysis, but usually they are inconclusive / unstable

Portland State
UNIVERSITY

4 Qualitative comparative analysis (QCA)

Solution								
1	2	3	4	5	6	7	8	9
●	●	●	●	●	●			⊗
⊗	●	●	●	●	⊗		•	⊗
	●	●	●			●	●	⊗
●	●	●		⊗	⊗	⊗	•	⊗
	⊗	•		⊗	⊗	⊗	•	●
⊗		●	⊗	⊗	●	⊗	⊗	⊗
⊗	⊗	•	⊗	⊗	⊗	⊗	⊗	⊗

There have been calls for more configurational thinking in my research area

QCA call

ORIGINAL ARTICLE

Journal of Supply Chain Management WILEY

Configurational approaches to theory development in supply chain management: Leveraging underexplored opportunities

David J. Ketchen Jr.[1] | Lutz Kaufmann[2] | Craig R. Carter[3]

2021

[1]Department of Management, Harbert College of Business, Auburn University, Auburn, Alabama, USA

[2]SCM Group, WHU—Otto Beisheim School of Management, Vallendar, Germany

[3]W. P. Carey School of Business, Arizona State University, Tempe, Arizona, USA

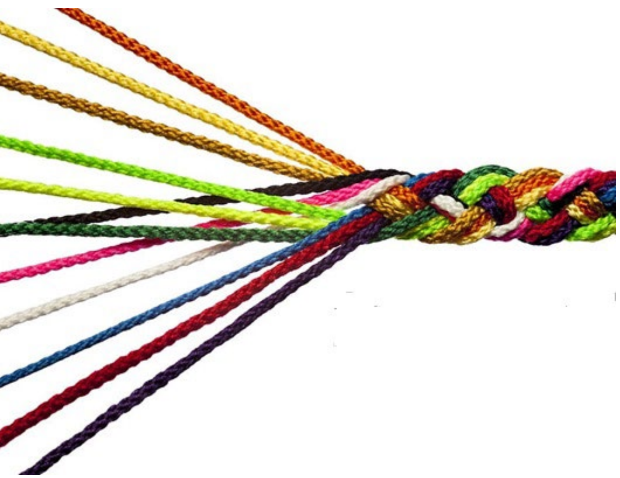
Correspondence
Craig R. Carter, W. P. Carey School of Business, Arizona State University, Tempe, Arizona, USA.
Email: craig.carter@asu.edu

Abstract

In introducing the 2020 Emerging Discourse Incubator, Flynn et al. (2020, https://doi.org/10.1111/jscm.12227) urged supply chain scholars to leverage fresh approaches in order to develop supply chain-specific theory, including approaches that are underutilized within the discipline. In response, we explain how more examination of configurations—meaningful sets of observations within a sample—can enhance theory development and, in particular, fuel the construction of supply chain-specific theory. First, we describe the value of configurational theorizing while contrasting it with two more popular approaches: one that centers on linear relationships and one that spotlights the unique features of individual observations. Second, we explain the main configurational approaches available to scholars. Here, we pay special attention to qualitative comparative analysis (QCA)—an approach to configurational theorizing that is relatively new to organizational research. Third, we offer examples of how configurational theorizing via the use of QCA can be used to develop supply chain management theory. Although QCA is employed regularly in neighboring fields, QCA remains something of a conceptual curiosity within supply chain management research. This state of affairs represents an important opportunity because QCA's emphasis on causal complexity fits well with the fact that supply chain outcomes usually arise from an array of variables—often at different levels of analysis—and

The general idea behind Qualitative Comparative analysis (QCA) is to uncover configurations (i.e. interactions) that lead to a certain outcome

General idea



Common situations

- **Causal factors combine with each other** to lead to the occurrence of an event or phenomenon.
- Different **combinations of causal factors** can lead to the occurrence phenomenon.
- Causal factors can have **opposing effects** depending on the combinations of factors

How QCA helps:

- QCA was developed for small-to-intermediate-N research designs (e.g., 5-50). In this range, there are often too many cases for researchers to keep all the case knowledge "in their heads," but too few cases for most conventional statistical techniques.

- However, scholars realized it also helps with bigger (quantitative) data and helps to assess necessity and sufficiency of conditions

- QCA may detect multiple paths, i.e. alternative causal combinations that can lead to high/low levels of the same outcome.

To differentiate from common statistical analyses, QCA researchers use a different set of terminology

Differences

Alternatives to Key Elements of the Conventional Template

(Ragin 2008, *Redesigning Social Inquiry)*

Conventional	Redesigned
1. variables	sets
2. measurement	calibration
3. dependent variables	qualitative outcomes
4. given populations	constructed populations
5. correlations	set theoretic relations
6. correlation matrix	truth table (shows kinds of cases)
7. net effects	causal recipes (INUS conditions)
8. counterfactual estimation	counterfactual analysis

In QCA, we differentiate between crisp set and fuzzy set approaches – with fuzzy set, you can capture also continuous data

Coding of the variables

- In its inception, QCA was Boolean/crisp (yes/no) – but now, it can be also "fuzzy" as well
- It works through set membership – how often are certain combinations of variables emerge to together to lead to a result – However, the "calibration" matters quite a bit

CRISP VERSUS FUZZY SETS

Crisp set	Three-value fuzzy set	Four-value fuzzy set	Six-value fuzzy set	"Continuous" fuzzy set
1 = fully in	1 = fully in	1 = fully in	1 = fully in	1 = fully in
			.8 = mostly but not fully in	Degree of membership is more "in" than "out": .5 < x_i < 1
		.75 = more in than out	.6 = more or less in	
	.5 = neither fully in nor fully out			.5 = cross-over: neither in nor out
		.25 = more out than in	.4 = more or less out	
			.2 = mostly but not fully out	Degree of membership is more "out" than "in": 0 < x_i < .5
0 = fully out	0 = fully out	0 = fully out	0 = fully out	0 = fully out

THE SCHOOL OF BUSINESS
PORTLAND STATE UNIVERSITY

As an example, imagine you perform an innovation workshop and want to know which factors lead participants to create high value and novel ideas

Example of a colleague

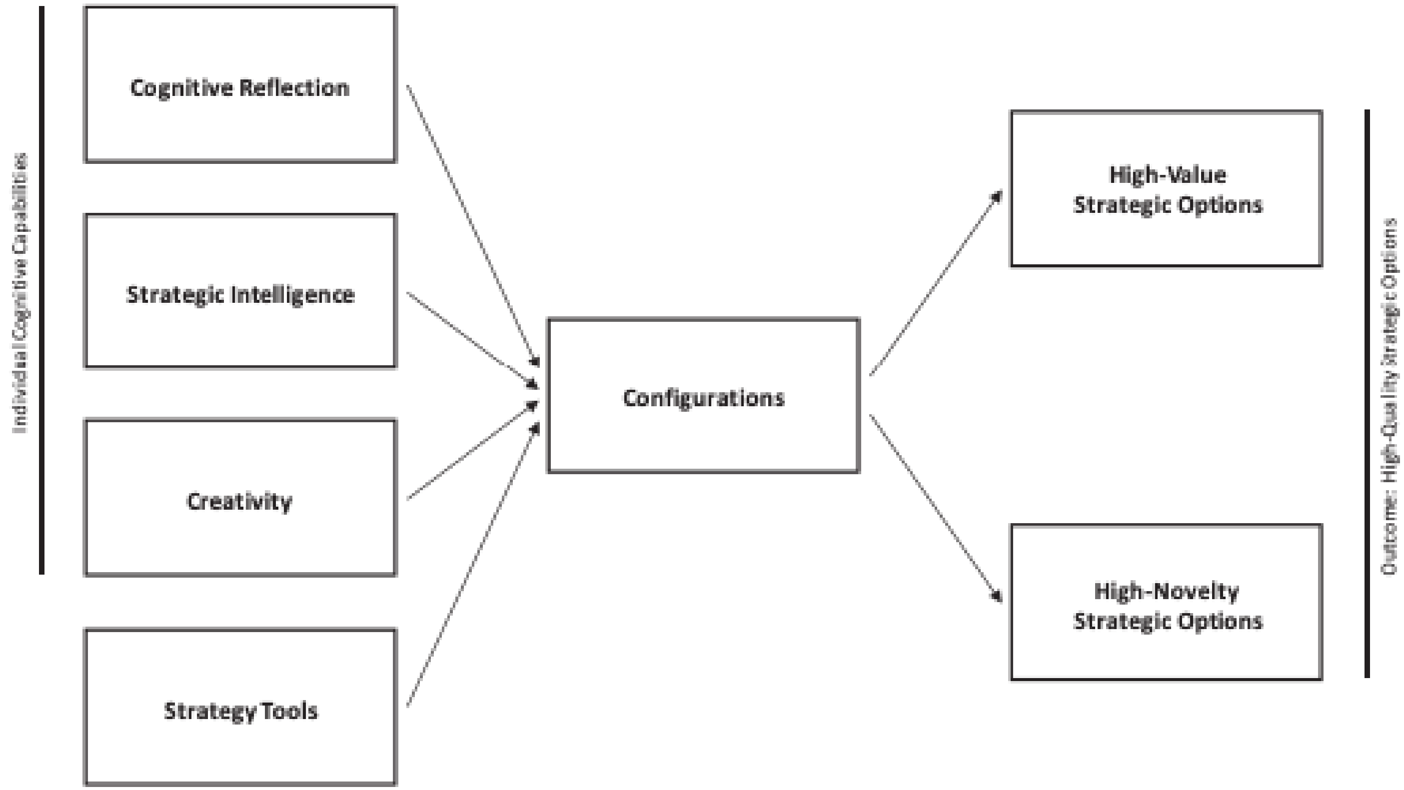


Figure 15: Configurational model of individual cognitive capabilities and strategy tools and their effect on high-quality strategic options.

The results show different recipes/combinations leading to different outcomes

Table 15: Configurations sufficient for the absence of excellence in high-value and high-novelty strategic-option generation.

	Not High-Value Strategic Options				***Not High-Novelty Strategic Options***
	3	***4***	***5***	***6***	***7***
Cognitive reflection	●	⨂		•	⨂
Strategic intelligence	●		⨂		
Creativity	•	⨂	⨂	⊗	⨂
Scenario analysis		⨂	●	•	⨂
Consistency	0.89	0.83	0.84	0.89	0.87
Raw coverage	0.20	0.21	0.17	0.23	0.23
Unique coverage	0.11	0.19	0.05	0.06	0.23
Number of cases	7	6	6	10	6
Overall solution consistency		*0.85*			*0.87*
Overall solution coverage		*0.60*			*0.23*

Low outcome

high outcome

Table 16: Configurations sufficient for excellence in high-value and high-novelty strategic-option generation after scenario analysis.

	High-Value Strategic Options		***High-Novelty Strategic Options***
	8	***9***	***10***
Cognitive reflection	⨂	●	⨂
Strategic intelligence	●	⨂	●
Creativity	●	●	
Consistency	1	0.81	0.89
Raw coverage	0.19	0.39	0.23
Unique coverage	0.14	0.34	0.23
Number of cases	2	10	3
Overall solution consistency	*0.85*		*0.89*
Overall solution coverage	*0.53*		*0.23*

Note: large black circles (●) are core present conditions, small black circles (•) are peripheral present conditions, large circles with a cross (⨂) are core absent conditions, small circles with a cross (⊗) are peripheral absent conditions; blank spaces indicate a 'don't care' condition.

Many software solutions exist to perform QCA analysis, however, the number of variables that can be included is limited

Software packages and remarks

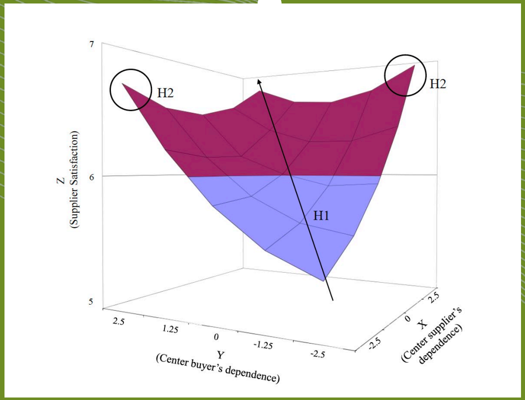
Software solutions:
- fsQCA Software (Ragin's fsQCA)
- R-package QCA
- Stata Routine (fuzzy)
- … (see website)

Remarks:
- Only one “dependent” and maximum 8-10 “independent” variables
- QCA / fsQCA is constantly evolving – check new updates
- Research design should follow a configurational perspective (not just as add-on) to be successful

Short Recap before discussion

“Standard” interaction

Response Surface



Multi-group

fsQCA

Table 16: Configurations sufficient for excellence in high-value and high-novelty strategic-option generation after scenario analysis.

	High-Value Strategic Options		***High-Novelty Strategic Options***
	8	***9***	***10***
Cognitive reflection	⨂	●	⨂
Strategic intelligence	●	⨂	●
Creativity	●	●	
Consistency	1	0.81	0.89
Raw coverage	0.19	0.39	0.23
Unique coverage	0.14	0.34	0.23
Number of cases	2	10	3
Overall solution consistency	*0.85*		*0.89*
Overall solution coverage	*0.53*		*0.23*

Note: large black circles (●) are core present conditions, small black circles (•) are peripheral present conditions, large circles with a cross (⨂) are core absent conditions, small circles with a cross (⊗) are peripheral absent conditions; blank spaces indicate a ‘don’t care’ condition.

Please share your experiences, ideas and remarks!

Open round

- Which of the methods did you until now and what are your experiences?
- Are there other methods worthwhile to explore? Any other ideas?
- Someone already looked into Bayesian approaches?
- …
- …

THE SCHOOL OF
BUSINESS
PORTLAND STATE UNIVERSITY

Dr. Frederik G. S. Vos, MSc. MSc. MSc.
Pronouns (he, him, his)
Assistant Professor of Supply Chain Management
PSU KMC 510H, 615 SW Harrison St., Portland, OR 97201
tel: (+1)6693128605, email: fvos@pdx.edu
Connect via: https://www.linkedin.com/in/frederikvos