

A discrete shipment size choice model with latent classes of shippers' attributes

Vortrag zur Konferenz „Verkehrsökonomik und -politik“

Raphael Piendl, Gernot Liedtke



Wissen für Morgen



Agenda

- State of the Art der Verkehrsmittelwahl
 - Total Logistic Costs - Das EOQ-Modell
 - Datengrundlage, Modellierung und empirische Schätzung
 - Integration latenter Klassen
 - Zusammenfassung und Ausblick
-



State of the Art der Verkehrsmittelwahl

- **State of the Art: Verkehrsmittelwahlmodell ohne explizite Logistikeinflüsse**
 - Anhand beobachteter Frachtaufträge wird ein Verkehrsmittelwahlmodell geschätzt
 - Keine Modellierung der besten Option eines Belieferungsprozesses
 - Unterschied zum Personenverkehr: Hohe Diversität der Entscheider und Güter
- **Logistikentscheidungen und Verkehrsmittelwahl beeinflussen sich gegenseitig**
 - Veränderungen der Verkehrsmiteleigenschaften führen zu Veränderungen auf der Logistikebene
 - Veränderungen des Logistikkonzepts beeinflussen die Wahl des Verkehrsmittels
 - Abbildung der Produktvielfalt (kombinierte Verkehre, Konsolidierungen etc.) nicht ohne Logistik möglich
- **Losgrößenentscheidung als Proxy für Logistikentscheidungen**
 - Erklärt einen Teil der Verhaltensheterogenität logistischer Entscheider
 - Weltweit wenige Modelle mit integrierter Losgrößenwahl existent
 - Zusätzlich: Cluster der Entscheidungsträger im Sinne einer Zielgruppenanalyse zur Erfassung weiterer Diversität



State of the Art der Verkehrsmittelwahl

- Problem: Konsistente Verknüpfung der Wahl der kontinuierlichen Losgröße mit der diskreten Verkehrsmittelwahl
- Hier angewendet: Losgrößenwahl als diskretes Entscheidungsmodell
- Viele Modelle zur Losgrößenwahl im Bereich des Operation Research
 - Anforderung für Güterverkehrsmodellierung: Aggregierbarkeit
 - Sogenannte „High-Level-Modelle“:
 - Einfache und korrekte Lösung
 - Veranschaulichung der Interaktion zwischen Lösung und strategischen Parametern
- Grundlage der Modelle: Minimierung der Gesamtkosten



Total Logistic Costs – Das EOQ-Modell

- Gesamtkosten pro Jahr für Entscheider n : $C_n(q_n) = \frac{Q_n}{q_n} F_n + Q_n c_n(q_n) + \frac{q_n}{2} a_n$
 - Q_n : bekannter und konstanter Güterfluss pro Jahr
 - q_n : Losgröße pro Transport um die gesamte Nachfrage Q_n zu befriedigen
 - F_n : Fixkosten des Transports (Bereitstellung und Bearbeitung)
 - $c_n(q_n)$: Variable Transportkosten (Annahme: proportional steigend zur Losgröße)
 - a_n : Lagerkostensatz für gehaltene Bestände pro Mengen- und Zeiteinheit
- Transportfixkosten pro Jahr fallen mit der Losgröße
- Lagerhaltungskosten pro Jahr steigen mit der Losgröße
- Analytisch optimale Lösung: $q_n^* = \sqrt{\frac{2F_n Q_n}{a_n}}$



Datengrundlage, Modellierung und empirische Schätzung

- Befragung der BVU GmbH und TNS Infratest GmbH im Rahmen der BVWP 2015
- Computer Assisted Personal Interviews: 474 Befragungen, 926 exemplarische Transportfälle
- Detailinformationen zum Transportfall abgefragt, u.a.
 - Art, Gewicht, Warenwert und Eigenschaften (z.B. zerbrechlich, brennbar etc.) des Transportguts
 - Häufigkeit des exemplarischen Transports
 - Position in der Logistikkette
 - Dauer, Kosten und Entfernung des Transports



Datengrundlage, Modellierung und empirische Schätzung

- Beschränkung auf das Verkehrsmittel LKW
 - 91.4% der Transportfälle mit dem LKW durchgeführt
- Eliminierung von Ausreißern
 - Unrealistische Ziele
 - Zu hohe Sendungsgewichte in Relation zu Entfernung und Kosten: Losgröße < 48t
- Diskretisierung der Losgröße
 - Willkürliche Einteilung problematisch, Störterm wächst potentiell mit Klassenbreite
 - 3 Losgrößenklassen gemäß eines Wachstumsfaktors $\tau = 4$ und in Anlehnung an die betriebliche Praxis gebildet
 - 0,75t – 3t: Stückgüter
 - 3t – 12t: Teilladungen
 - 12t – 48t: (mehrere) Ganzladungen



Datengrundlage, Modellierung und empirische Schätzung

- Zur Erinnerung: $q_n^* = \sqrt{\frac{2F_n Q_n}{a_n}}$
- Logarithmierung führt zu schätzbaren linearen Abhängigkeiten

$$\begin{aligned}\ln q_n^* &= \ln \sqrt{\frac{2F_n Q_n}{a_n}} \\ &= 0.5 \ln(2) + 0.5 \ln Q_n + 0.5 \ln F_n - 0.5 \ln a_n \\ &= \text{const.} + 0.5 \ln Q_n + 0.5 \ln F_n - 0.5 \ln a_n.\end{aligned}$$

- Logarithmierung führt zu konstanten Klassenbreiten von $\ln \tau = \ln 4$



Datengrundlage, Modellierung und empirische Schätzung

- Menge an Alternativen: $C = \{q_1, q_2, q_3\}$
- Nutzenfunktion der einzelnen Alternativen für logistischen Entscheider n :

$$U_{q_i,n} = \alpha_{q_i} + \beta_{q_i,Q} \ln Q_n + \beta_{q_i,a} \ln a_n + \Theta_{q_i,X} \mathbf{X}_n + \epsilon_{q_i,n}, \quad \forall n, i = 1, 2, 3.$$

- α_{q_i} : Alternativspezifische Konstante
 - $\beta_{q_i,Q}, \beta_{q_i,a}$: Parameter bezüglich des Güterflusses Q_n bzw. der Lagerkosten a_n
 - \mathbf{X}_n : Vektor mit zusätzlich beeinflussenden Attributen
 - $\Theta_{q_i,X}$: Vektor mit Parametern bezüglich der Attribute in \mathbf{X}_n
 - $\epsilon_{q_i,n}$: Zufallskomponente der Nutzen für unbeobachteten Teil
- Keine Fixkosten F_n
 - Nur LKW-Transporte untersucht
 - Annahme: Fixkosten konstant
 - Zusätzlich beeinflussende Attribute
 - Transportdistanz
 - Latente Klassen



Datengrundlage, Modellierung und empirische Schätzung

- Bestimmung der Gesamtnachfrage Q_n
 - Ergibt sich aus Losgröße q_n und der angegebenen Frequenz der Transporte
- Approximation des Lagerkostensatzes a_n
 - Warenwertdichte $a_\rho = \frac{\text{Warenwert [EURO]}}{\text{Sendungsgewicht [TONNE]}}$ als Proxyvariable
- Transportdistanz d
 - Höhere Unsicherheit bei längeren Transporten
 - Freie Kapazität bei längeren Transporten teurer
 - Entkopplung durch regionale Distributionszentren
 - Transportdistanz Box-Cox transformiert: $d_{BC} = \frac{d^{0,18} - 1}{0,18}$
- Anzahl an Beobachtungen durch fehlende Werte stark geschrumpft: 487



Datengrundlage, Modellierung und empirische Schätzung

	Estimate	Std. Error	t-value	Pr(> t)
3t - 12t:(intercept)	-6.3859	1.4563	-4.39	0.0000***
12t - 48t:(intercept)	-12.0507	1.8581	-6.49	0.0000***
3t - 12t:ln(Q)	1.0436	0.1486	7.03	0.0000***
12t - 48t:ln(Q)	2.1333	0.1934	11.03	0.0000***
3t - 12t:ln(a_ρ)	-0.1483	0.1093	-1.36	0.1749
12t - 48t:ln(a_ρ)	-0.7279	0.1422	-5.12	0.0000***
3t - 12t: d_{BC}	0.2061	0.0750	2.75	0.0060***
12t - 48t: d_{BC}	0.4786	0.0912	5.24	0.0000***

Log-Likelihood: -328.31

McFadden R^2 (ρ^2): 0.35707

Adjusted McFadden R^2 ($\bar{\rho}^2$): 0.33832

Likelihood ratio test: $\text{chisq} = 364.68$ (p.value = $< 2.22e-16$)

Frequencies of alternatives:	0,75t - 3t	3t - 12t	12t - 48t
	0.21355	0.31622	0.47023

. p<0.1; *p<0.05; **p<0.01; ***p<0.001



Integration latenter Klassen

▪ Motivation:

- Heterogenität der logistischen Entscheider und der transportierten Güter
- Große Anzahl an potentiell beeinflussenden logistischen Faktoren
- Relativ geringer Erklärungsgehalt einzelner klassischer Logistikvariablen
- Aufblähung der Modelle im statistischen Sinne
- Simplifikation notwendig

▪ Methode: Latente Klassenanalyse

- Ziel: Identifikation von nicht direkt beobachtbaren Gruppen in der Stichprobe
- Ausgangspunkt: Ausschließlich kategoriale Variablen (Kontingenztafel)
- Schätzung mittels Maximum-Likelihood-Methode
- Wahrscheinlichkeit der Klassenzugehörigkeit
- In Verbindung mit anderen multivariaten statistischen Verfahren einsetzbar (Multiple Regression, Faktorenanalyse etc.)



Integration latenter Klassen

- Diskrete Variablen: Gütereigenschaften liefern die beste Grundlage

Anzahl Klassen	Log-Likelihood	BIC
2	-2284.309	4710.949
3	-2232.398	4681.386
4	-2191.574	4673.995
5	-2175.075	4715.257
6	-2158.715	4756.797
7	-2146.076	4805.778

- Anzahl an latenten Klassen auf Basis des Bayes-Informationskriteriums (BIC) gewählt
 - Minimum bei 4 Klassen



Integration latenter Klassen

	Klasse 1	Klasse 2	Klasse 3	Klasse 4
γ_c	0.107	0.36	0.3475	0.1854
	Wahrscheinlichkeiten einer "JA"-Antwort			
Gut ist zerbrechlich:	0.1727	0.0634	0.3302	0.0000
Gut ist hochwertig:	0.2631	0.2395	0.6373	0.1266
Gut ist sperrig:	0.0000	0.2376	0.4942	0.2734
Bestimmte Temperatur:	0.6331	0.0855	0.0000	0.0000
Lebensmittel:	1.0000	0.0000	0.0000	0.0651
Brennbar, explosiv etc:	0.0000	0.2079	0.0599	0.1707
Greifer- oder Schüttgut:	0.0398	0.0000	0.0000	0.3647
Unverpacktes Flüssiggut:	0.0555	0.0000	0.0000	0.1800
Standardladungseinheiten:	0.5550	0.8113	0.2308	0.0305
Einzelanfertigung:	0.0384	0.0841	0.3974	0.0000
Zusammenfassung Artikel:	0.4781	0.4134	0.3992	0.0557
Beobachtungen: 487				
Geschätzte Parameter: 47				



Integration latenter Klassen

- Zuordnung der logistischen Entscheider zu einzelnen Gruppen: Maximale Wahrscheinlichkeit der Klassenzugehörigkeit
- Problem: Keine strikte Zuordnung
- APP_c : Mittelwerte der maximalen Wahrscheinlichkeiten für Klasse c
- Gütekriterium „Odds of correct classification“: $OCC_c = \frac{APP_c/(1-APP_c)}{\gamma_c/(1-\gamma_c)}$
- Gute Trennung der Klassen: $OCC > 5$

Klasse	Mittelwerte	Varianzen	OCC
1	0.9292	0.0199	109.4464
2	0.8492	0.0204	10.0114
3	0.8435	0.0222	10.1221
4	0.7489	0.0384	13.1055



Integration der latenten Klassen in das Logit-Modell

	Estimate	Std. Error	t-value	Pr(> t)
3t - 12t:(intercept)	-6.7827	1.5170	-4.47	0.0000***
12t - 48t:(intercept)	-13.8539	2.0166	-6.87	0.0000***
3t - 12t:ln(Q)	1.1092	0.1562	7.10	0.0000***
12t - 48t:ln(Q)	2.2525	0.2041	11.04	0.0000***
3t - 12t:ln(a_ρ)	-0.2244	0.1171	-1.92	0.0554 .
12t - 48t:ln(a_ρ)	-0.7696	0.1524	-5.05	0.0000***
3t - 12t: d_{BC}	0.1804	0.0761	2.37	0.0178*
12t - 48t: d_{BC}	0.4657	0.0942	4.94	0.0000***
3t-12t:Klasse 2	0.9026	0.4891	1.85	0.0650 .
12t-48t:Klasse 2	1.3325	0.5880	2.27	0.0234**
3t-12t:Klasse 3	1.2509	0.5026	2.49	0.0128**
12t-48t:Klasse 3	1.9546	0.6089	3.21	0.0013**
3t-12t:Klasse 4	1.1099	0.5794	1.92	0.0554 .
12t-48t:Klasse 4	2.4058	0.6832	3.52	0.0004***

Log-Likelihood: -319.58

McFadden R^2 (ρ^2): 0.37418

Adjusted McFadden R^2 ($\bar{\rho}^2$): 0.34676

Likelihood ratio test: chisq = 382.15 (p.value = < 2.22e-16)

. p<0.1; *p<0.05; **p<0.01; ***p<0.001



Zusammenfassung und Ausblick

- Die geschätzten Modelle weisen durchgehend hohen Erklärungsgehalt auf
- Richtige Vorzeichen der Einflussgrößen
- Vor allem die Gesamtnachfrage Q_n erklärt einen Großteil der Losgrößenwahl
- Warenwertdichte als Approximation der Bestandskosten geht gedämpft in die Entscheidung ein
 - Koeffizienten zu gering
 - Platzverbrauch auch Teil der Lagerhaltungskosten
 - Wichtigkeit innerhalb der Supply-Chain nicht beschrieben
- Transportdistanz erhöht die Losgröße
- Integration von latenten Klassen verbessert den Erklärungsgehalt
 - Latentes Klassenmodell auf suboptimaler Basis
 - Anzahl an Beobachtungen (487)
 - Erfassung von logistischen Einflussvariablen ausbaufähig
 - Hochwertige Güter haben wider den Erwartungen einen positiven Effekt



Zusammenfassung und Ausblick

- Eine Einbettung der Losgrößenwahl in die Verkehrsmittelwahl erhöht den Erklärungsgehalt
- Sensitivität gegenüber Maßnahmen wird exakter abgebildet
- Identifizierung latenter Klassen und deren Einbettung eröffnet weitere Möglichkeiten in der Analyse von Entscheidungsverhalten
- Integriertes Losgrößen- und Verkehrsmittelwahlmodell auf Basis der Revealed Preference (RP)-Daten hier nicht zu schätzen
 - Zu großer Anteil an LKW-Transporten
 - Hoher Verlust an Beobachtungen bei der Berechnung von Q_n
- Großes Potential der Modelle, Datengrundlage mit RP/SP Daten unabdingbar
→ univariate Faktenaufnahme in bisherige Statistiken unzureichend



Vielen Dank für Ihre Aufmerksamkeit!

