Preferred citation style

Schmid, B. (2016) Explaining the Choice between In-Store and Online Shopping - A Behavioral Experiment, presentation at the *Verkehrsingenieurtag 2016*, Zurich, March 2016.

Explaining the Choice between In-Store and Online Shopping - A Behavioral Experiment

Basil Schmid

IVT
ETH Zurich

Alumni Verkehrsingenieurtag
17^{th} of March 2016

Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Acknowledgments

Kay W. Axhausen: Supervision and support

Simon Schmutz: Data processing and personal advice

Milos Balac: Routing in MATSim

Maxim Janzen: LaTeX beamer template

SNSF: Project funding

"Wie der Onlinehandel die Strassen verstopft"

Source: Tagesanzeiger, 23rd of February 2016

Switzerland's digital shopping revolution?

- Market share of online and mail order sector reached 10 % mark, with growth rates in the double-digit range (Rudolph et al., 2015)
- 115 million packages distributed by "Die Post"
- Over 4 billion kilometers traveled by light goods vehicles, tendency increasing (BfS, 2014)
- Online shopping of books and electronic gadgets account for over 25 % of total market shares, while food products account for roughly 5 %
- ICT usage in Switzerland: Over 30 % of the Swiss population is online at least once per hour; even higher for mobile phone users

Online vs. in-store shopping in Switzerland

- Barriers to online shopping: Swiss study reveals substantial differences in age, gender and income between online and in-store shoppers (Rudolph et al., 2004)
  - Usage: Changes in current shopping routines required
  - Value of online shopping: Missing sales personnel; delivery time lag
  - Risk: Product uncertainty, information asymmetry or security
  - Psychological barriers: Tradition, beliefs and experience (early 1990's: Rapid increase of Swiss households with computer and internet access)
- Perceptions and attitudes of online vs. in-store shoppers differ significantly

Shopping in a Post-Car World

- Omitting private motorized vehicle justified by car-less policy developments; availability of innovative modes
- Hypotheses:
  - Substitution effect from in-store towards online shopping, especially for larger and heavy shopping baskets
  - Taste heterogeneity mainly determined by attitudes towards online shopping
- How sensitive are individuals towards different attributes related to their choice btw. online vs. in-store shopping?
- How do income and attitudes affect attribute sensitivities?
- What is the distribution of attitudes, and which socio-demographic characteristics are affecting them?

Shopping in a Post-Car World ...

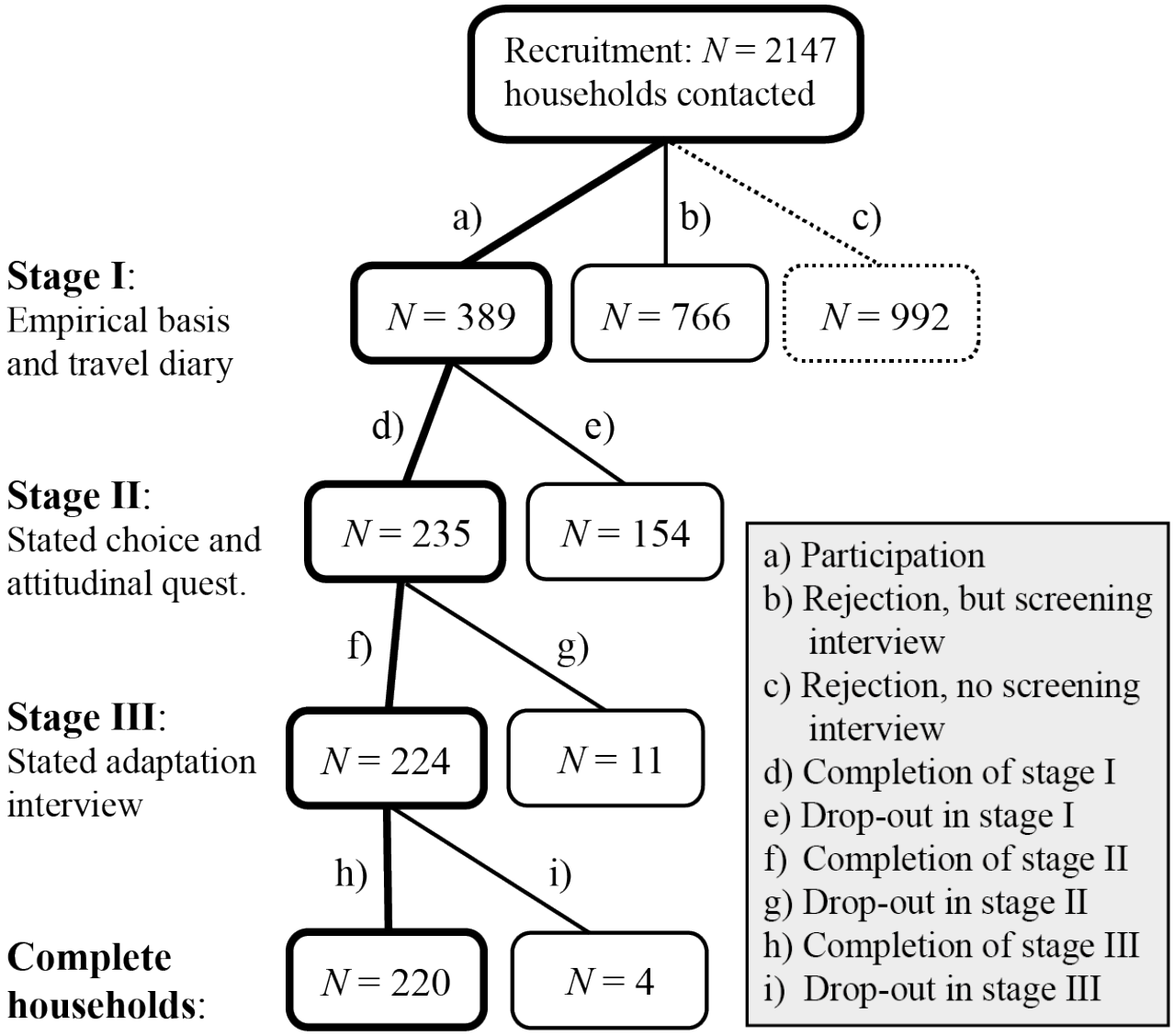
Source: www.focus.de, 3rd of December 2015

Related literature on shopping behavior

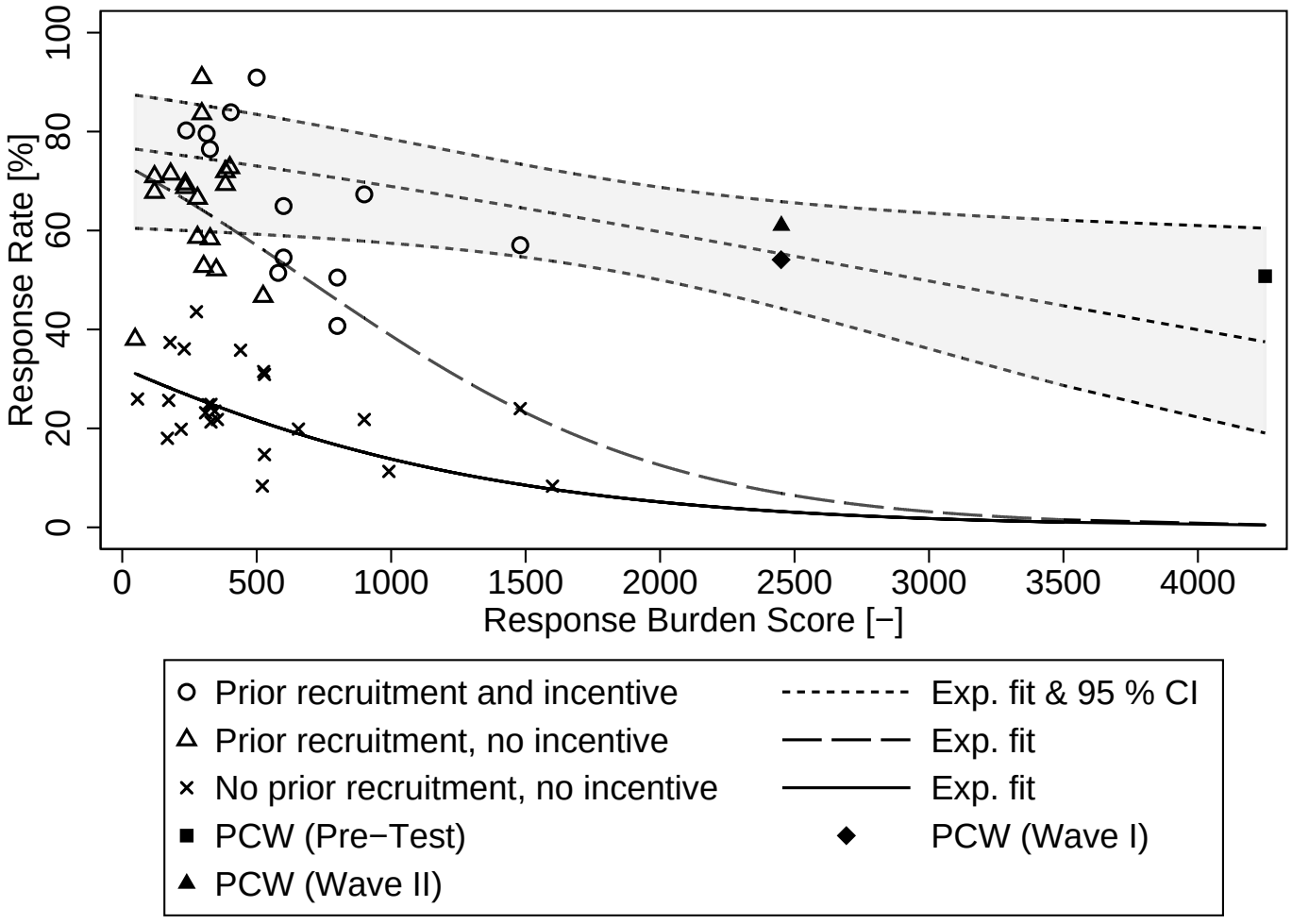
- Rotem-Mindali and Salomon (2007): Product price is indicated as main reason for online shopping (ISR)
- Dijst, Farag and Schwanen (2008): Choice model on in-store vs. online shopping, but no alternative-specific attributes. Attitudes play major role in explaining preferences (NL)
- Mokhtarian and Tang (2012): Joint choice (strong dependency) of different purchase/pre-purchase channels when ordering/buying clothes, including attitudes (US)
- Zhai et al. (2016): Shopping behavior for search/experience goods differs between shopping channels and for different stages, i.e. information and product trial (US)

$\Longrightarrow$ Post-Car World: First alternative-specific hybrid choice model in this research field

Post-Car World: A multi-stage travel survey



Response behavior @ IVT, ETH Zurich



Source: Axhausen, Schmid and Weis, 2015

Data (220 households; 339 participants)

Variable	Value	MZ2010 [%]	PCW15 [%]
Household income	Not reported	24.1	5.7
	$<$ 12'000 CHF	61.0	27.6
	$\geq$ 12'000 CHF	18.4	61.8
Residential location	City centre	38.9	50.0
	Agglomeration	54.8	43.1
	Rural	6.3	6.9
Household type	Single-person household	31.6	18.7
	Couple without kids	33.0	25.2
	Couple with kids	26.6	48.0
	Single-parent household	5.8	4.5
	Living community	3.1	3.7
Education	Low	21.0	14.7
	Medium	54.9	22.3
	High	24.1	63.0
Season tickets	None	37.3	11.0
	Half-fare card	51.8	72.9
	GA	10.9	16.1

Experimental conditions

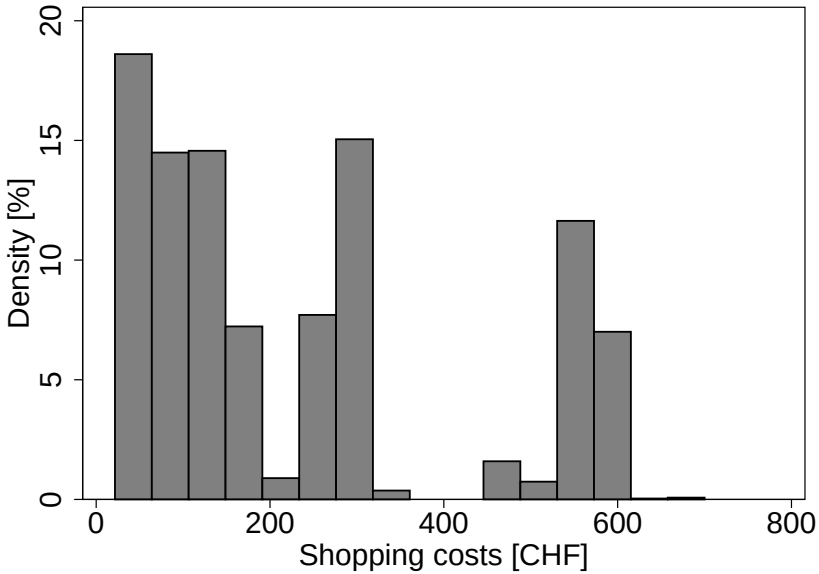
- Coherent choice situations:
  - home based round trip for in-store alternative
  - no social motives; buying goods is the one and only purpose
  - groceries and durable goods experiment: "Daily or weekly grocery shopping" and "multimedia, HiFi and electronic (household) appliances"
  - quality of the goods is assumed to be identical between the two shopping channels
  - in-store alternative either with public transport, carsharing or carpooling (no private cars)
- Pivot approach: If a shopping trip has been reported, reference values depend on expenditures, time use for shopping and traveled distance

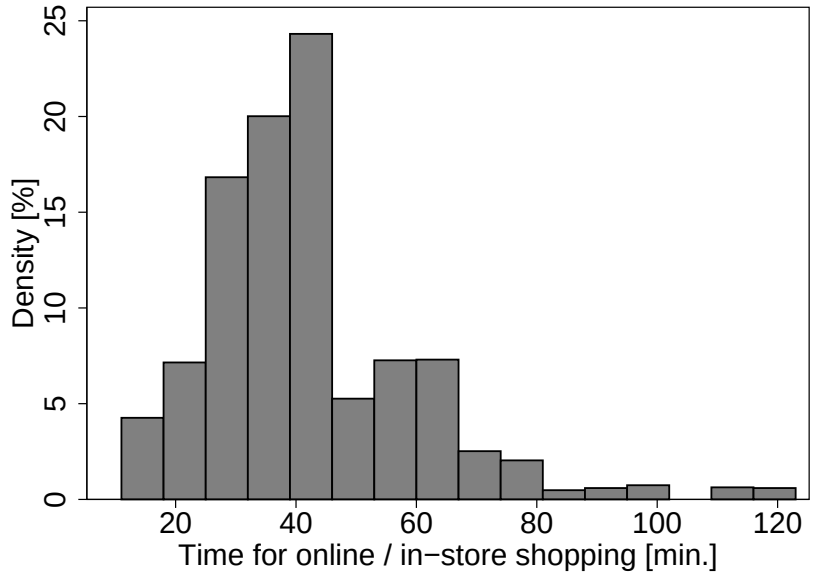
Experimental Design: Attribute levels

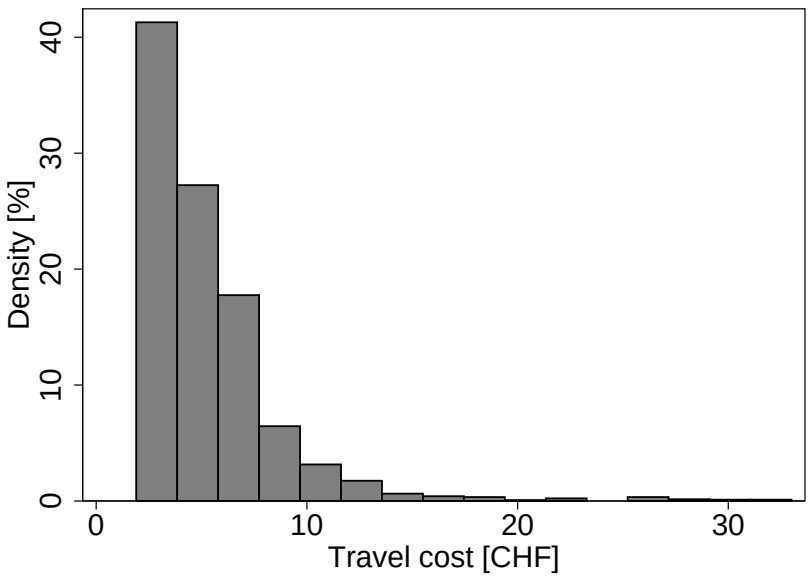
Attributes	Online	In-store	Levels
Shopping costs ONL	$\surd$	–	$-10\%, -5\%, +/-0\%$
Shopping costs IS	–	$\surd$	$-5\%, +/-0\%, +5\%$
Time for shop. ONL	$\surd$	–	$-20\%, -10\%, +5\%$
Time for shop. IS	–	$\surd$	$-10\%, +/-0\%, +10\%$
Delivery cost and duty	$\surd$	–	0, 5, 10, 15 CHF
Travel cost	–	$\surd$	$-20\%, +10\%, +40\%$
Delivery time groceries	$\surd$	–	$<$ 1 day / 1-2 days / $>$ 2 days
Delivery time durables	$\surd$	–	2-4 days / 4-7 day / $>$ 1 week
Travel time	–	$\surd$	$-30\%, +/-0\%, +30\%, \geq$ 3 min.
Size / weight of the goods basket	$\surd$	$\surd$	Low / medium / high (same for both alternatives)

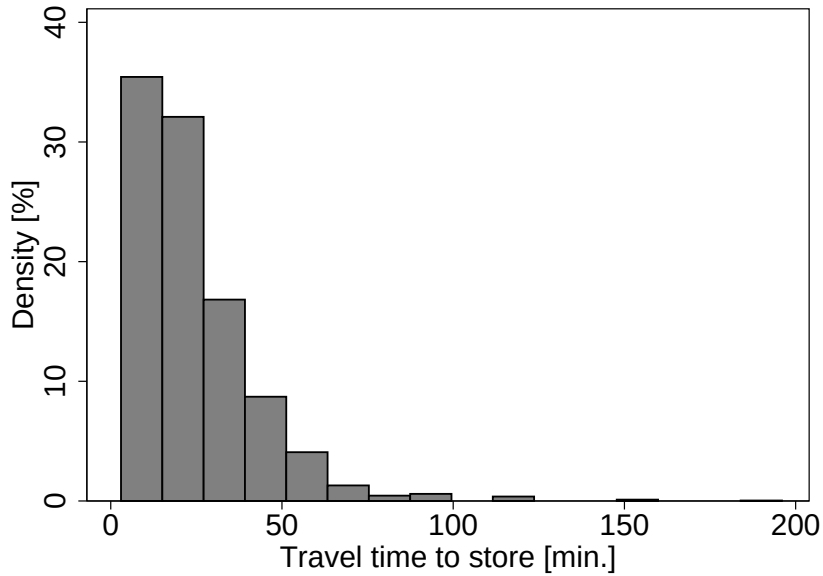
- Efficient design; 3 blocks with 8 choice sets
- Participants are assigned to the "groceries" (38 %) or "durable goods" (62 %) experiment based on reported shopping trips

Attribute distributions

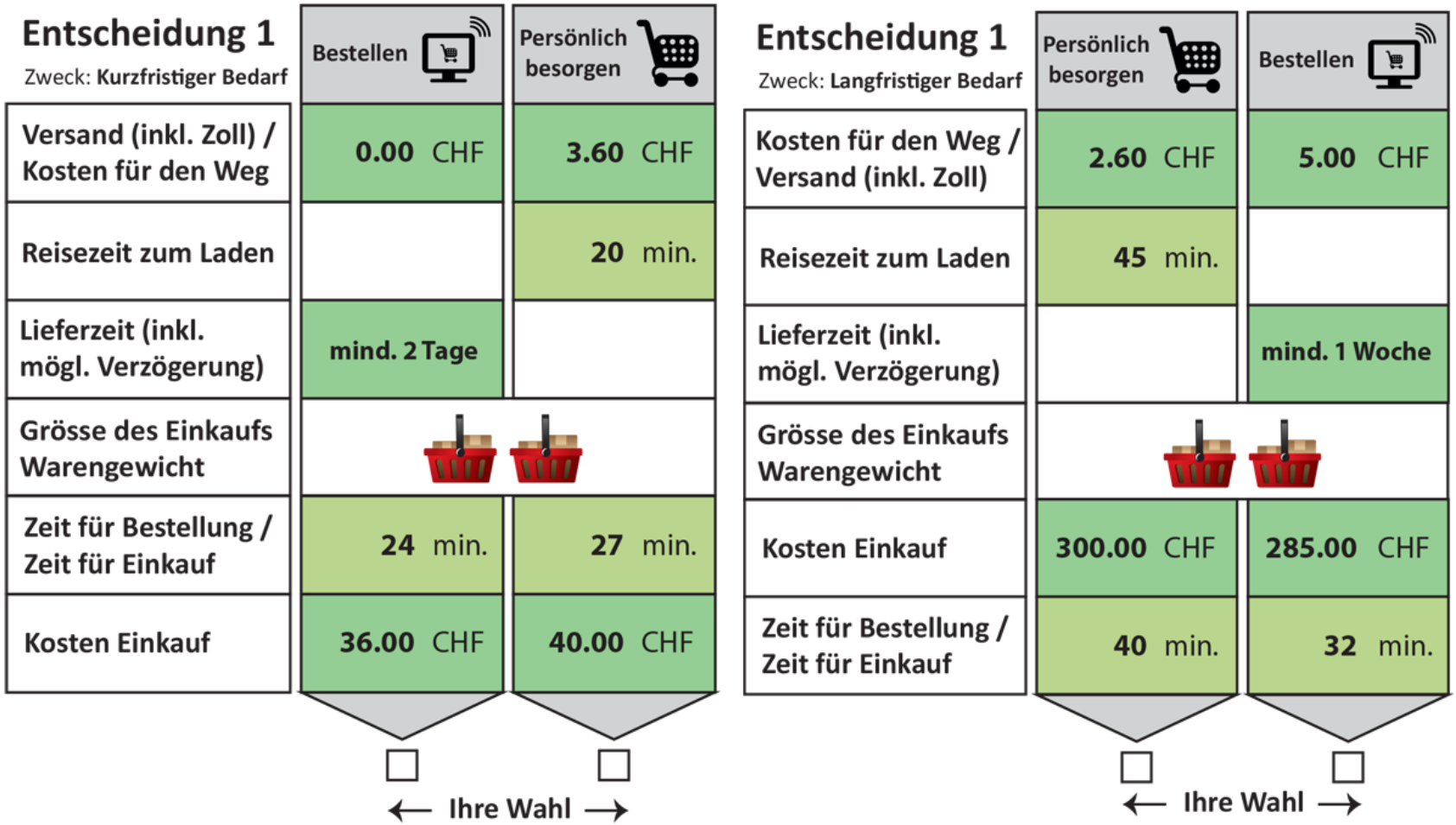








Example of choice situations



Attitudes towards online shopping

Ihre Einstellung zum Einkaufen und Online-Shopping

Trifft...

		ganz genau zu	eher zu	eher nicht zu	überhaupt nicht zu
1.	Ich bestelle oft Produkte im Internet.	☐	☐	☐	☐
2.	Online-Shopping ist mit Risiken verbunden.	☐	☐	☐	☐
3.	Einer der Gründe, weshalb ich kein Online-Shopping betreibe, ist, dass ich nicht gerne meine Kreditkarten-nummer über das Internet weitergebe.	☐	☐	☐	☐
4.	Das Internet ist mehr ein Fluch als ein Segen.	☐	☐	☐	☐
5.	Der Nachteil beim Online-Shopping ist, dass ich die Produkte nicht physisch begutachten kann.	☐	☐	☐	☐
6.	Online-Shopping erleichtert mir das Vergleichen von Preisen und verschiedenen Produkten.	☐	☐	☐	☐
7.	Das Risiko, ein falsches Produkt zugeliefert zu be-kommen, ist einer der Hauptgründe, weshalb ich kein Online-Shopping betreibe.	☐	☐	☐	☐

Attitudes towards online shopping

- Measures of different statements regarding
  - attitudes towards online shopping and internet usage in general
  - risks and credit card fraud
  - pros and cons of online shopping
- Exploratory factor analysis to ...
  - reduce the dimensionality of data to the most essential elements (general attitudes)
  - remove sources of covariance and measurement noise
  - estimate uncorrelated factor scores with $\mu \approx 0$ and $\sigma \approx 1$ as a first step for developing the hybrid choice model

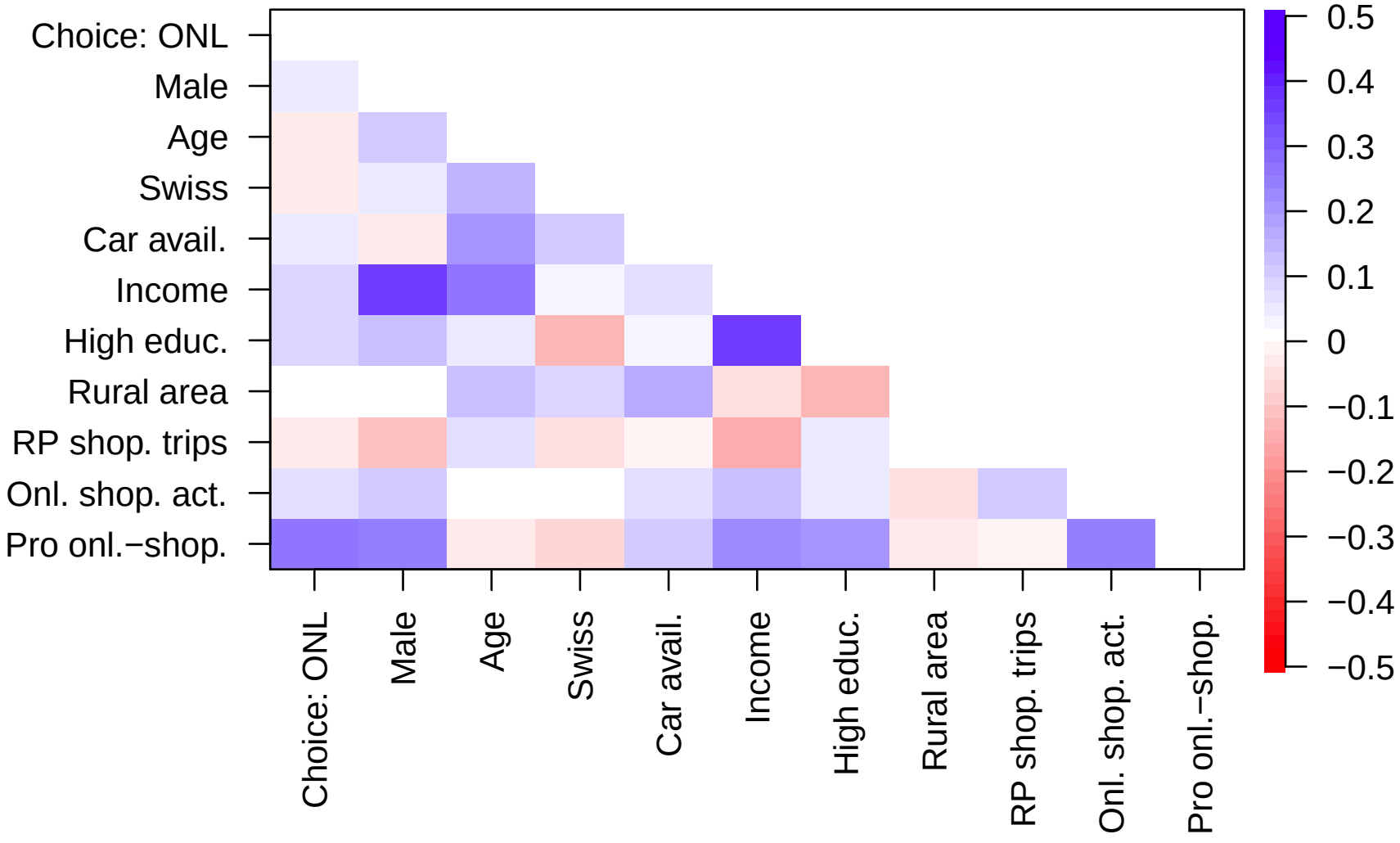
Attitudes towards online shopping

Questionnaire item	Factor loading
sh1: I often order products in the internet	+0.68
sh2: Online shopping is associated with risks	–0.49
sh3: Credit card fraud is one the reasons why I don't like online shopping	–0.66
sh4: The internet has more cons than pros	–0.54
sh5: A disadvantage of online shopping is that I cannot physically examine the products	–0.30
sh6: Online shopping facilitates the comparison of prices	+0.53
sh7: The risk of receiving a wrong product is one the main reasons why I don't like online shopping	–0.60

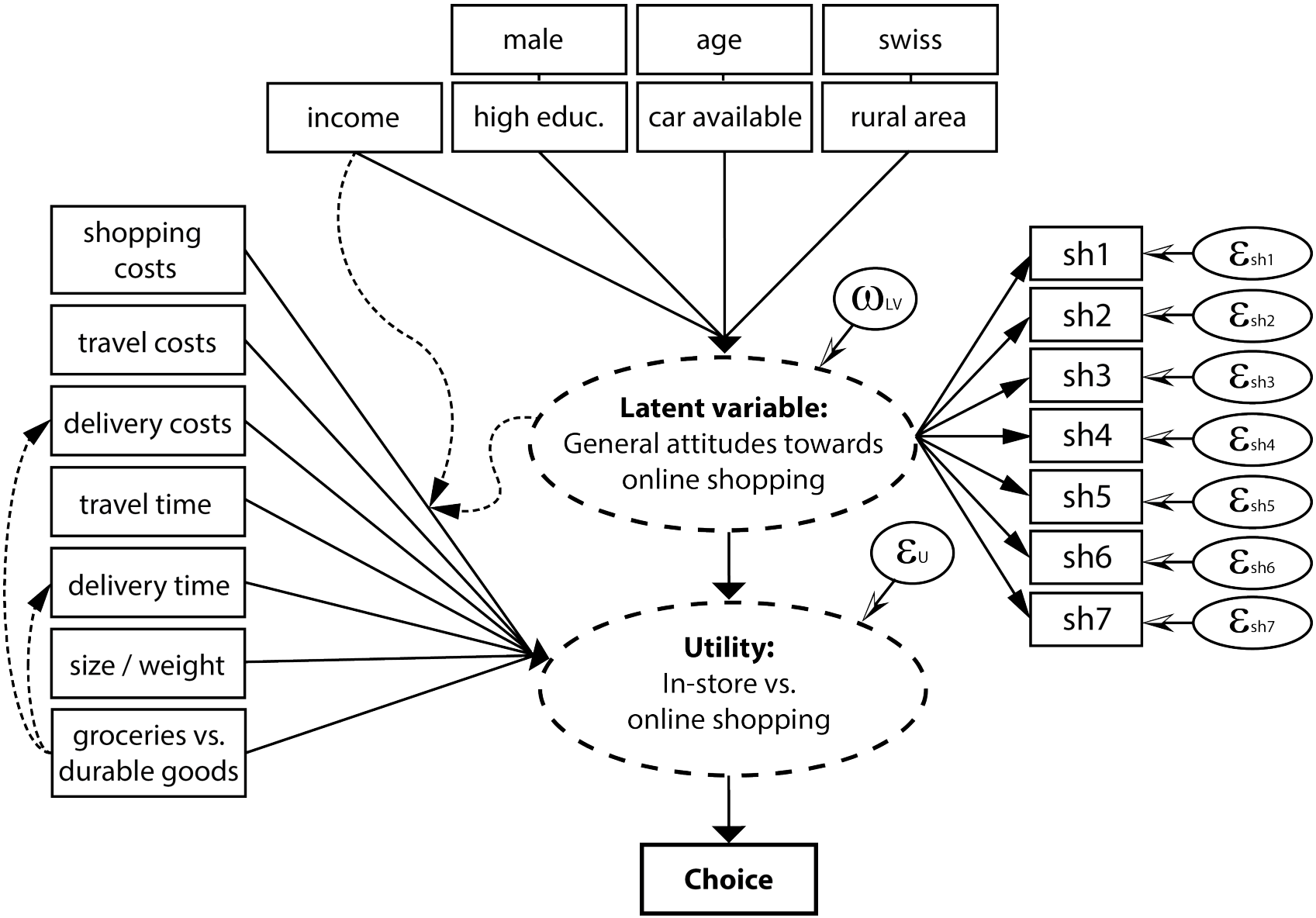
Attitudes and socio-economic characteristics

?

Correlogram: The facts



Modeling Framework: Hybrid choice



Modeling Framework: Structural model

Utility equation for shopping channel i with choice attributes X_{i_n} and the latent online shopping variable LV_n:

$$\begin{aligned} U_{O_n} = \beta_{C_O} + \beta_{O_n} \cdot X_{O_n} + \beta_{sc,O} * sc_O * \left(\frac{inc}{\overline{inc}}\right)^{\lambda_{inc}} + \\ \mu_{LV} \cdot (LV_n - \overline{LV_n}) + \\ \mu_{sc,LV} \cdot sc_{O_n} \cdot (LV_n - \overline{LV_n}) + \epsilon_{O_n} \end{aligned} \tag{1}$$

$$\begin{aligned} U_{IS_n} = \beta_{IS} \cdot X_{IS_n} + \beta_{sc,IS} * sc_{IS} * \left(\frac{inc}{\overline{inc}}\right)^{\lambda_{inc}} + \\ \mu_{sc,LV} \cdot sc_{IS_n} \cdot (LV_n - \overline{LV_n}) + \epsilon_{IS_n} \end{aligned} \tag{2}$$

Latent variable equation with socio-economic characteristics X_n:

$$LV_n = \overline{LV_n} + \kappa_X \cdot X_n + \omega_{LV_n} \tag{3}$$

Modeling Framework: Structural model

Relative importance of choice attribute X_{i_n} compared to shopping costs as a function of income and the latent variable LV_n:

$$f(inc_n, LV_n) = \frac{\beta_{X_{i_n}}}{\beta_{sc} \cdot \left(\frac{inc_n}{\overline{inc_n}}\right)^{\lambda_{inc}} + \mu_{sc,LV} \cdot (LV_n - \overline{LV_n})} \tag{4}$$

- If $\lambda_{inc} < 0$ and $\mu_{cost,LV} < 0$: Shopping cost sensitivity increases with lower income and a more positive attitude towards online shopping
- For the "average" respondent, the equation collapses to

$$f(\overline{inc}, \overline{LV_n}) = \frac{\beta_{X_{i_n}}}{\beta_{sc}} \tag{5}$$

Modeling Framework: Measurement model

Latent variable measurement equations with responses to the 7 online shopping items I_{sh}:

$$I_{sh_n} = \overline{I_{sh}} + \tau_{LV_{I_{sh}}} \cdot LV_n + \epsilon_{I_{sh_n}} \tag{6}$$

Choice equation: Choice of individual n for shopping channel i by maximizing utility U_i:

$$if \quad U_{O,n} > U_{IS,n} : choice_{i,n} = \begin{cases} \text{Online shopping} \\ else \text{ In-store shopping} \end{cases} \tag{7}$$

β_j, μ_j, λ_{inc}, $\overline{LV_n}$, κ_j, $\sigma_{\omega_{LV}}$, $\overline{I_{sh}}$, τ_{sh} and $\sigma_{I_{sh}}$ are the parameters to be estimated (42 in total)

Estimation

Likelihood of individual n choosing alternative i is the joint probability of observing the choice and the 7 online shopping items I_{sh_n}, given choice attributes and socio-economic characteristics $X_{i,n}$:

$$P(ch_{i,n}, I_{sh_n}|X_{i,n}) = \int_{\omega_{LV_n}} P(ch_{i,n}|X_{i,n}, \omega_{LV_n}) \prod_{sh=1}^{7} f_{sh_n}(I_{sh_n}, \omega_{LV_n})\phi(\omega_{LV})d\omega_{LV_n} \quad (8)$$

$$\omega_{LV} \sim N(0, \sigma_{\omega_{LV}}) \quad (9)$$

$$P(ch_{i,n}|X_{i,n}, \omega_{LV_n}) = \frac{exp(U(X_{i,n}))}{\sum_j^2 exp(U(X_{j,n}))} \quad (10)$$

$$f_{sh_n}(I_{sh_n}, \omega_{LV_n}) = \frac{1}{\sigma_{I_{sh}}}\phi\left(\frac{I_{sh_n} - \overline{I_{sh}} - \tau_{I_{sh}} \cdot LV_n}{\sigma_{I_{sh}}}\right) \quad (11)$$

Maximum likelihood estimation with PythonBiogeme version 2.4 on Euler (HPCC, 2 cores, runtime 36 min., 101 iterations)

Choice and (non-)trading behavior

- Market shares (2698 choice observations; 339 respondents): Groceries = 65 % in-store shopping and 35 % ordering; durable goods = 39 % in-store shopping and 61 % ordering
- Non-Trading behavior: Respondents choosing the same alternative (e.g. in-store shopping) in all 8 choice situations
- Almost 80 % of respondents are traders, with about 83 % for durable goods and 68 % for groceries ($p_{difference} < 0.01$)
- "Labeled" choice experiment: Non-Trading behavior is still consistent with random utility theory (too small trade-off variations with respect to non-traders underlying preferences)

(Non-)Trading behavior, by purpose

- Trading vs. non-trading between shopping channels differs by shopping purpose ($p_{difference} < 0.01$)
- Almost 30 % of respondents that are assigned to the grocery experiment are always choosing the in-store option
- Respondents with pro-online shopping attitudes have a higher probability to be online non-traders for durable goods ($p < 0.01$); opposite is not true for in-store non-traders

Estimation results: Choice models

Variable	Base model	Factor model	Hybrid model
Shopping costs	−0.021 ***	−0.024 ***	−0.022 ***
Income elasticity of shopping cost	0.041	−0.033	−0.059
Factor score x shopping costs	–	−0.007 ***	–
LV x shopping costs	–	–	−0.019 ***
Travel time (IS)	−0.022 ***	−0.024 ***	−0.025 ***
Travel cost (IS)	−0.036 **	−0.035 **	−0.036 **
Delivery time (ONL)	−0.560 ***	−0.600 ***	−0.614 ***
Delivery cost (ONL)	−0.091 ***	−0.098 ***	−0.099 ***
Delivery time x durable (ONL)	0.466 ***	0.494 ***	0.504 ***
Delivery cost x durable (ONL)	0.055 ***	0.054 ***	0.053 **
ASC (ONL)	−2.080 ***	−2.120 ***	−2.370 ***
Purpose durable (ONL)	0.152	0.047	0.065
Size	1.120 ***	1.200 ***	1.220 ***
Factor score	–	0.466 ***	–
LV	–	–	1.210 ***
Number of estimated parameters	11	13	42
McFadden ρ^2	0.20	0.25	0.70

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Estimation results: LV model

Variable	Dep. variable: LV_n
$\overline{LV_n}$	3.05 ***
Age	0.01 **
Age^2	−0.02 ***
Car always avail.	0.06 ***
High education	0.11 ***
Income	0.08 ***
Rural resid. area	−0.11 **
Male	0.25 ***
Swiss	−0.11 ***

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

- Female and Swiss non-car users with low education and income living in rural residential locations have the most negative attitudes towards online shopping
- Maximal pro-online shopping attitudes with 31 years of age
- Measurement model of latent variable (LV_n) confirms results of the factor analysis

Value of time for shopping trips/delivery

Coefficient ratios	Base model	Factor model	Hybrid model
VTTS shopping trips [CHF/h]	37.3 (62.9)	42.0 (60.0)	41.0 (68.2)
VODT groceries [CHF/day]	6.1 (26.7)	6.2 (25.0)	6.2 (27.9)
VODT durable goods [CHF/day]	2.6 (4.5)	2.4 (4.4)	2.4 (5.0)
VTTS [CHF/h]: Erath, 2006		52.90-128.85	N = 110
VTTS [US$/h]: Hsiao, 2009		5.30	N = 300
VODT [US$/day]: Hsiao, 2009		0.44-0.76	N = 300
VTTS [CHF/h]: VSS norm, 2009		12.32-20.72	N = 649
VTTS [CHF/h]: Fröhlich et al., 2014		5.90-9.10	N = 282

- Current study: Value of travel time savings (VTTS) of 40 CHF/h; about 50 % higher if considering shopping instead of travel costs as reference (values in brackets)
- Hsiao, 2009: Similar study in Thaiwan for online/in-store shopping of books
- Comparison to other Swiss studies: VTTS for shopping trips differ enormously

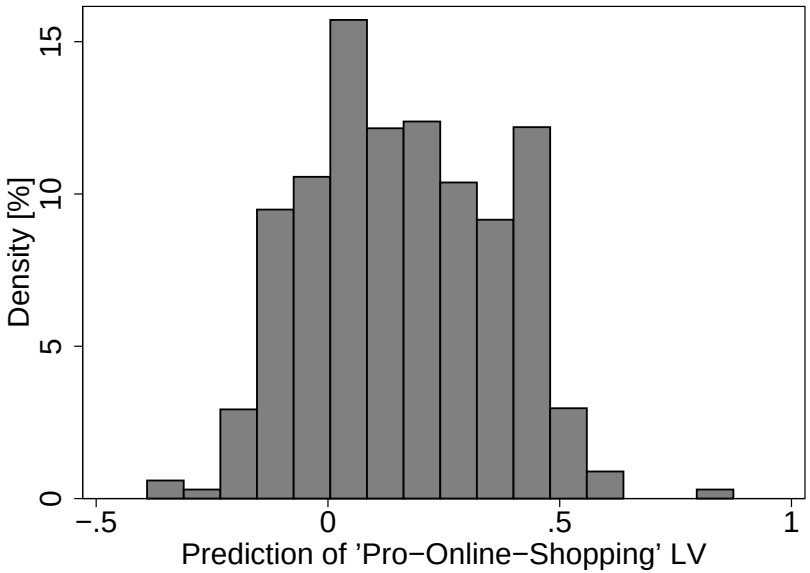
Point (cross) elasticities

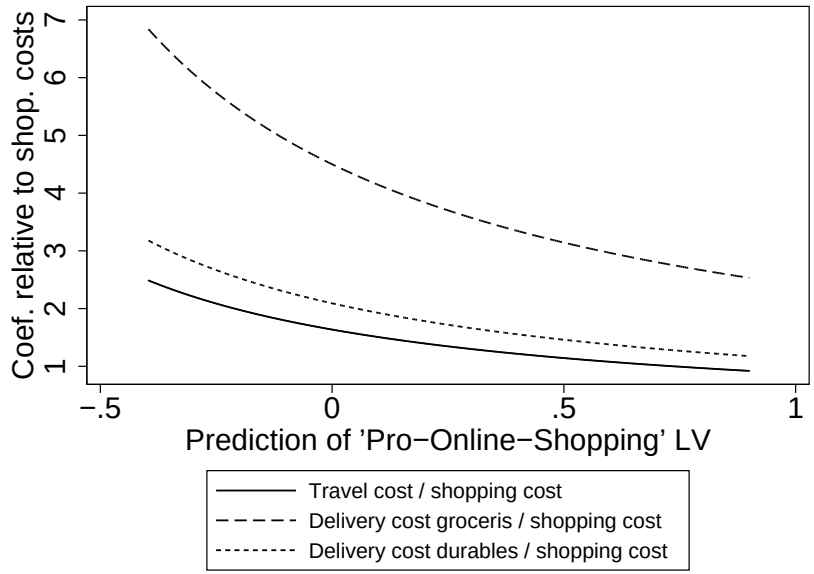
Point (cross) elasticities of hybrid model	Ordering	Trip-making
Shopping costs	–2.48 (2.60)	–2.74 (2.62)
Travel time	–	–0.31 (0.30)
Travel cost	–	–0.10 (0.09)
Delivery time, groceries	–1.20 (1.25)	–
Delivery time, durables	–0.21 (1.14)	–
Delivery cost, groceries	–0.37 (0.38)	–
Delivery cost, durables	–0.17 (0.18)	–
Size (Ordering)	1.15 (–1.20)	–

$\Longrightarrow$ relatively high elasticity of shopping costs (mean = 240 CHF), independent of shopping purpose (no sign. interaction)

$\Longrightarrow$ ceteris paribus, on average, a 1% increase in shopping costs decreases the predicted market share of online shopping by 2.5 percentage points, for pro-online shoppers (e.g. $\widetilde{LV_n} = 0.5$) by 3.6 percentage points

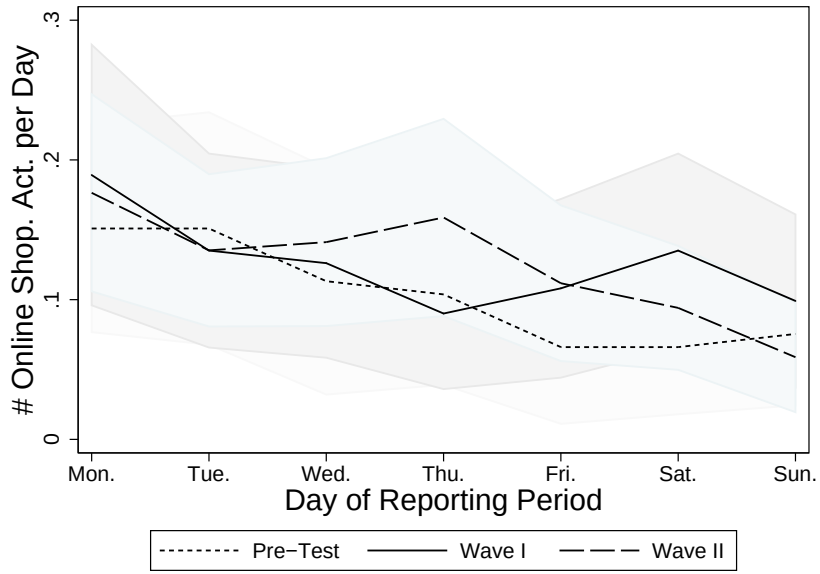
Prediction of latent variable

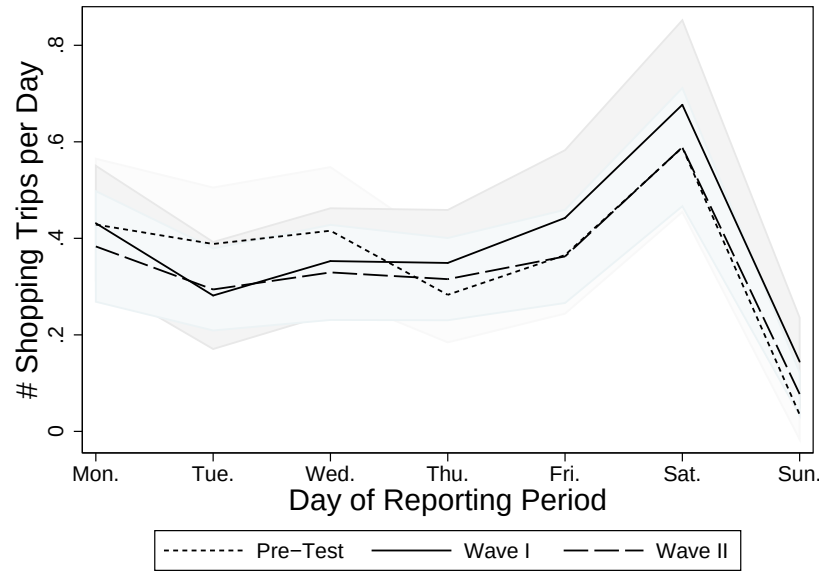




- Positive attitudes towards online shopping $\widetilde{LV_n}$ is approx. normally distributed with $\mu = 0.16$ and $\sigma = 0.20$
- Attribute sensitivities relative to shopping costs are decreasing for higher pro-online shopping attitudes $\Longrightarrow$ price-sensitive trade-off behavior by considering both alternatives as possible shopping channels

Validation with RP data





- Revealed preference data from travel and online diaries for shopping activities (N = 339 participants, 2709 persondays)
- Weekly pattern: In-store shopping trips are mostly conducted on Saturdays, while online shopping activities show a decreasing pattern over weekdays

Validation with RP data

Variable	# shop. trips per day	# onl. shop. per day
Const.	-1.143 ***	-2.360 ***
Weekday	0.000	-0.086 ***
Saturday	0.532 ***	-0.004
Sunday	-1.469 ***	-0.282
$\widetilde{LV_n}$	-0.388	1.242 ***
SE($\widetilde{LV_n}$)	(0.25)	(0.44)
σ_ϵ	0.623 ***	1.049 ***
Prob. $> \chi^2$	0.00	0.00

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

- Random-effects Poisson regressions: Strong within-subject error term correlation
- Expected effect of predicted pro-online shopping attitude $\widetilde{LV_n}$ on the number of online shopping activities per day
- $\widetilde{LV_n}$ shows a weak negative effect on the number of reported shopping trips $\Longrightarrow$ substitution effect?

Conclusions

Conclusions

- Behavioral richness and estimation efficiency increase substantially when including latent variables
- Structural model reveals distribution of LV in the population based on fundamental socio-demographic characteristics
- VTTS vs. VODT: Large potential of online shopping given the relatively high value of travel time savings for shopping trips
- Pro-online shopping attitudes lead to a sign. increase in shopping cost sensitivity $\Longrightarrow$ larger choice set when considering both online and in-store shopping as possible shopping channels
- 1 CHF $\neq$ 1 CHF: Delivery costs are perceived as more negative than travel and shopping costs (avoidability hypothesis) $\Longrightarrow$ online retailers better incorporate delivery costs in shopping prices

Questions?

Project website:

http://postcarworld.epfl.ch/