Homogamy in risk aversion: Evidence from the travel survey MIMETTIC

Nathalie Picard (THEMA, UCP)
Sophie Dantan (CES, ENS Cachan)
André de Palma (CES, ENS Cachan)

EPFL, April 21, 2016

Outline

- Travel uncertainty and scheduling costs in the literature
- Measuring Risk Aversion: Mimettic Survey
- Modelling answers to scenarii to estimate risk aversion
- Estimation with individual heterogeneity
- Estimation with individual and couple-specific heterogeneity

Travel uncertainty and scheduling costs in the literature

Valuation of variability in the literature

- Several approaches to model individual utility:
  - Scheduling model (Small 1982)
  - Mean-variance (Bates et al. 2001; Fosgerau & Karlstrom 2011)
  - Value of reliability (Lam & Small 2003; de Palma, Lindsey & Picard 2005)
  - Extension of scheduling model (Li, Tirachini, Hensher 2010, Tseng et Verhoef 2008)
  - ➔ Need to model and estimate risk aversion in relation to scheduling cost
- Correlation between spouses' VOT and different VOT with/without spouse (de Palma, Lindsey, Picard 2015)
  - ➔ To the best of our knowledge, no results on the relation between spouses' risk aversions

A few references

- Bates, J., Polak, J., Jones, P., Cook, A., 2001. The valuation of reliability for personal travel, *Transportation Research E*, 37-2/3, 191-229
- Coulombel, N. de Palma, A. (2014) Variability of Travel Time, Congestion, and the Cost of Travel, *Mathematical Population Studies*, 21:4, 220-242
- de Palma, A., Picard, N., (2005). Route choice decision under travel time uncertainty, *Transportation Research A*, 394, 295-324
- de Palma A., N. Picard (2006), Route choice behavior with risk averse users, in *Spatial Evolution and Modelling*, P. Nijkamp & A. Reggiani eds, Ch. 7, 139-178
- de Palma A., R. Lindsey, N. Picard (2005). Urban Passenger Travel Demand, in *The Blackwell Companion to Urban Economics*, R. Arnott & D. McMillen eds
- de Palma A., R. Lindsey, N. Picard (2015). Trip-timing decisions and congestion with household scheduling preferences. *Economics of Transportation,* 4, 118-131
- Fosgerau, M., Karlstrom, A. (2010). The value of reliability, *Transportation Research B,* 44, 38–49
- Lam, T., Small, K., (2003). The Value of Time and Reliability: Measurement from a Value Pricing Experiment. University of California Transportation Center.
- Li, Z. Tirachini, A., Hensher, D. (2012) Embedding Risk Attitudes in a Scheduling Model: Application to the Study of Commuting Departure Time, *Transportation Science,* 46, 170-188
- Polak, J., S. Hess, X. Liu (2008). Characterising Heterogeneity in Attitudes to Risk in Expected Utility Models of Mode and Departure Time Choice, The Transportation Research Board

Measuring Risk Aversion: Mimettic Survey

Data collection

- Stated Preferences web-based survey on a panel maintained by TNS-SOFRES
- Financial incentives (coupon; value independent from answers)
- Data collection period: late 2011 – early 2012
- Sample: 5210 respondents including 1047 couples in which both spouses answered

Questions to measure risk aversion

0% 25% 50% 75% 100%

s2tcq1

Supposons encore que vous allez refaire ce déplacement en RER ou en train de banlieue pour motif Travail à 09H05 , toujours avec exactement les mêmes conditions, sauf cette fois, pour le temps de trajet.

Il y a 2 options qui s'offrent à vous. Dans l'une le temps de trajet est fixe et dans l'autre le temps de trajet peut varier. Laquelle préférez-vous ?

- le trajet dure 30 minutes
- le trajet a une chance sur deux de durer 10 minutes et une chance sur deux de durer 45 minutes

0% 25% 50% 75% 100%

s2vpq1

Supposons encore que vous allez refaire ce déplacement en voiture en tant que conducteur pour motif Travail à 09H05 , toujours avec exactement les mêmes conditions, sauf cette fois, pour le temps de trajet.

Il y a 2 options qui s'offrent à vous. Dans l'une le temps de trajet est fixe et dans l'autre le temps de trajet peut varier. Laquelle préférez-vous ?

- le trajet dure 45 minutes
- le trajet a une chance sur deux de durer 20 minutes et une chance sur deux de durer 60 minutes

Sequence of question

1st question: choice between 2 routes: S / R

Il y a 2 options qui s'offrent à vous. Dans l'une le temps de trajet est fixe et dans l'autre le temps de trajet peut varier. Laquelle préférez-vous ?

- le trajet dure 30 minutes
- le trajet a une chance sur deux de durer 10 minutes et une chance sur deux de durer 45 minutes

→ *S : safe travel time*

→ *R : risky travel time*

If S chosen: R more attractive in 2nd question

et si maintenant, toujours pour le même déplacement, vous avez 2 autres options, laquelle préférez-vous ?

L'hypothèse est toujours que vous refaites ce déplacement, dans exactement les mêmes conditions, sauf en termes de temps de trajet, fixe ou variable;

- le trajet dure 30 minutes
- le trajet a une chance sur deux de durer 5 minutes et une chance sur deux de durer 40 minutes

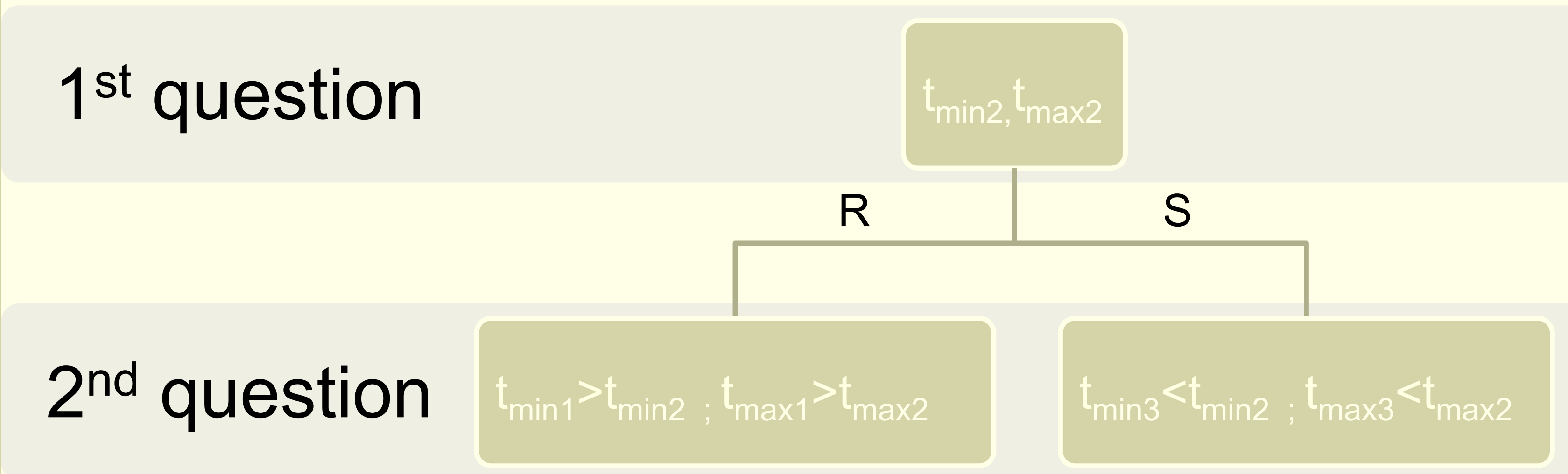
If R chosen: 2nd question

et si maintenant, toujours pour le même déplacement, vous avez 2 autres options, laquelle préférez-vous ?

L'hypothèse est toujours que vous refaites ce déplacement, dans exactement les mêmes conditions, sauf en termes de temps de trajet, fixe ou variable;

- le trajet dure 30 minutes
- le trajet a une chance sur deux de durer 16 minutes et une chance sur deux de durer 51 minutes

Question tree & risk aversion interval



Using revealed preference to customize stated preference questions

- Anchoring questions…
  - Same purpose, same mode same departure time
    ➔ Potential endogeneity problem, not corrected yet
  - Safe tt "close" to revealed one
  - Simple risky travel time: binary distribution
    - 50% chance ttinf < tt0
    - 50% chance ttsup > tt0
- Anchoring makes questions more realistic and easier to answer but a perfect anchoring induces severe drawbacks
  - De Palma and Picard, 2005 showed it makes travel times proposed in the scenarii endogenous to VOT and RA and severely biases the associated coefficients
  - Solution: random selection of travel times proposed in scenarii
  - 10 series of scenarii for each mode;
  - A series “close" to revealed tt is randomly selected for each respondent

Modelling answers to scenarii to estimate risk aversion

PRELIMINARY SPECIFICATION

Risk aversion threshold computation

- Threshold θ^* defined by indifference condition:

$$U\left(S;\theta^*\right) = \mathrm{E}\left[U\left(R;\theta^*\right)\right] \Leftrightarrow U\left(tt_0;\theta^*\right) = 1/2U\left(tt_{\inf};\theta^*\right) + 1/2U\left(tt_{\sup};\theta^*\right)$$

- Individual prefers Safe Option iff s/he is more risk averse:

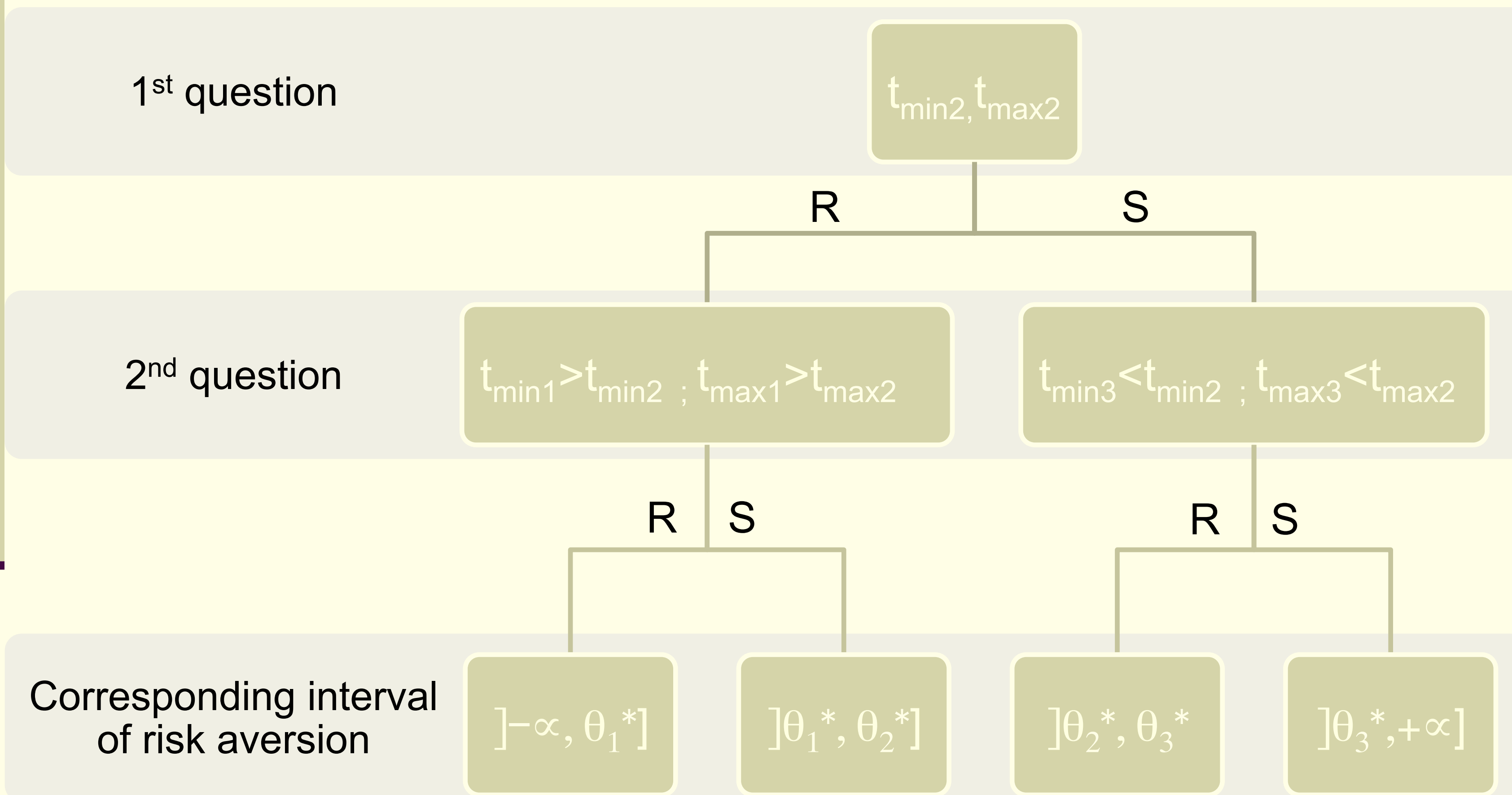
$$U\left(S;\theta\right) \mathrm{f}\ \mathrm{E}\left[U\left(R;\theta\right)\right] \Leftrightarrow \theta > \theta^*$$

- Individual prefers Risky Option iff s/he is less risk averse:

$$U\left(S;\theta\right) \mathrm{p}\ \mathrm{E}\left[U\left(R;\theta\right)\right] \Leftrightarrow \theta < \theta^*$$

- Each scenario designed such that ?? threshold for each question, and thresholds are correctly ordered

Question tree & risk aversion interval



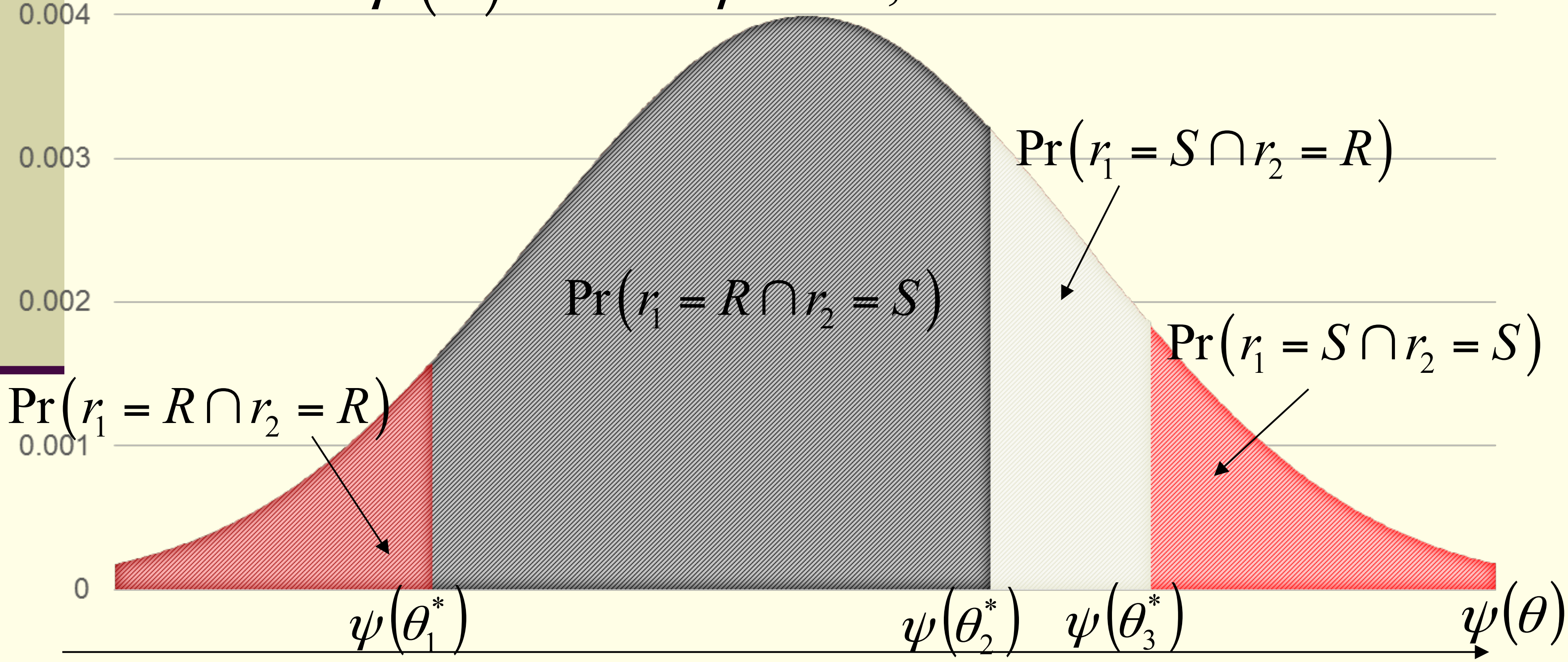
Testing utility function: CARA vs CRRA

- CARA = Constant Absolute Risk Aversion
  - →Risk aversion (and thus answer) identical if 1h (or any other value) added to all tt in scenario
- CRRA = Constant Relative Risk Aversion
  - →Risk aversion (and thus answer) identical if all tt in scenario multiplied by 2 (or any other value)
- Design to test CARA / CRRA
  - A constant **added** to all tt in half scénarii
  - All tt **multiplied** by a constant in the other half scénarii

Econometric model: ordered probit

- Extension of methodology developed by de Palma and Picard TRA, 2005
- Assume that there exists an increasing function $\psi(.)$ such that

$$\psi(\theta) = \nu = X\beta + \sigma\varepsilon,\ \varepsilon \text{ standard normal}$$

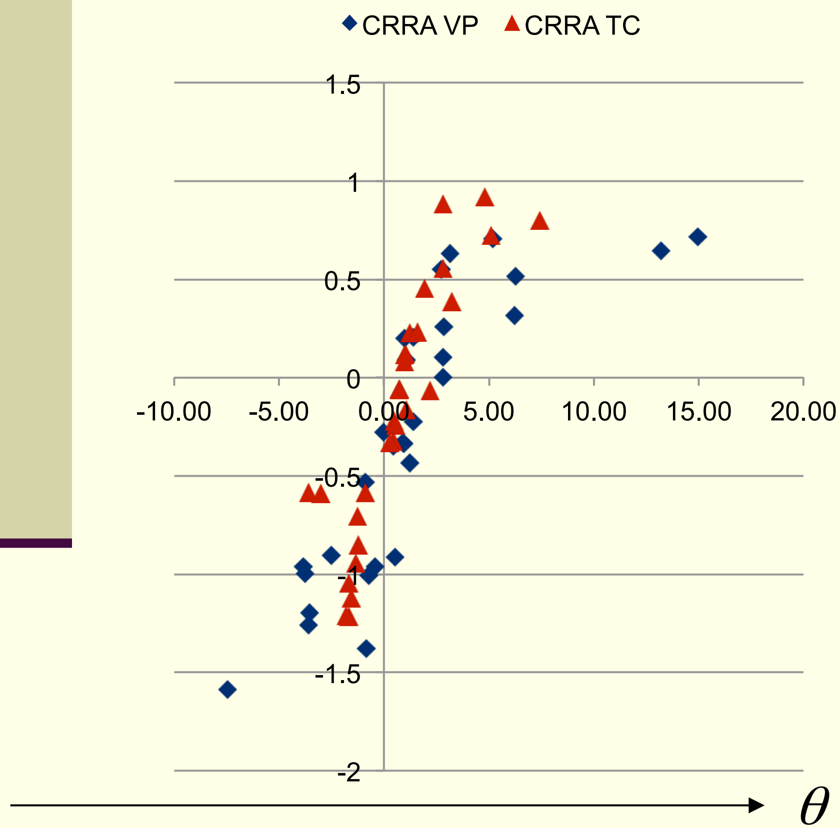


Econometric model: ordered probit without explanatory variables

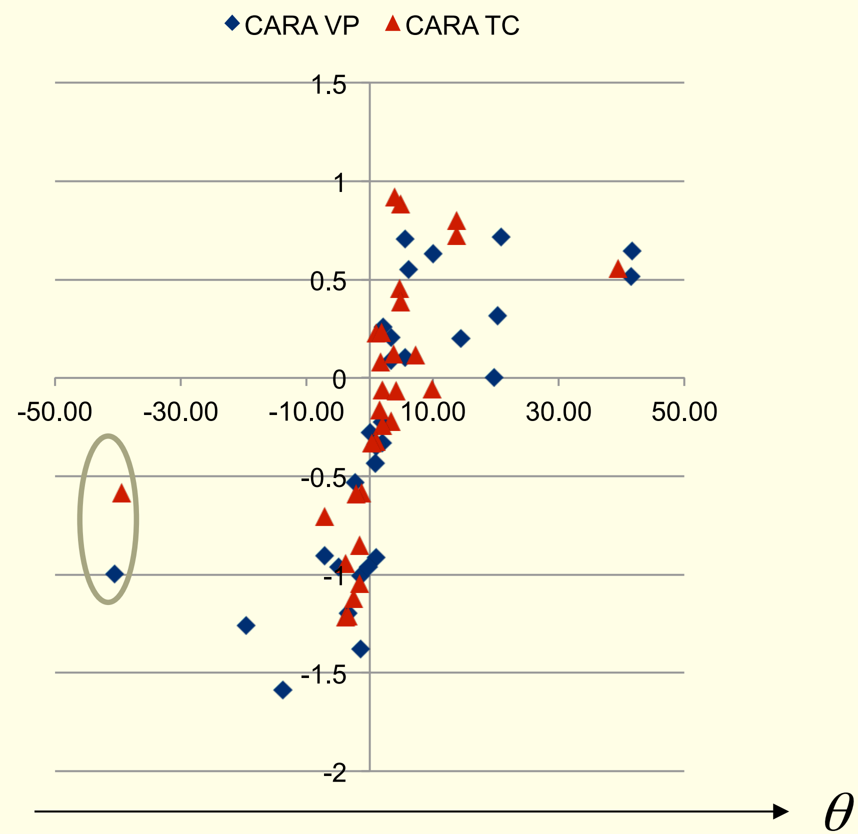
- Assumes no individual observed heterogeneity
- Imposes no restriction, except for:
  - the support of the distribution of θ
  - very little structure on the link between series of scenarii
- For a given series i=1,…,10 and question j=1,2,3
  - thresholds ν_{ij}^{*} estimated from % in each of the 4 cells in the tree, by inverting the standard normal CDF
  - thresholds θ_{ij}^{*} given by indifference condition
  - Correspondence: $\psi\left(\theta_{ij}^{*}\right)=\nu_{ij}^{*}$ gives the value of ψ(.) at 3*10=30 points for each mode

Test in favor of CRRA rather than CARA

CRRA



CARA



3-threshold ordered probit

$$P(S,S) = P\left(\varepsilon_i > \psi\left(\theta_2^*\right) - X_i\beta\right)$$

$$= 1 - F\left[\frac{1}{\sigma} \times \left(\psi\left(\theta_2^*\right) - X_i\beta - \mu\right)\right]$$

$$P(S,R) = P\left(\varepsilon_i < \psi\left(\theta_2^*\right) - X_i\beta\right) - P\left(\varepsilon_i < \psi\left(\theta_1^*\right) - X_i\beta\right)$$

$$= F\left[\frac{1}{\sigma} \times \left(\psi\left(\theta_2^*\right) - X_i\beta - \mu\right)\right] - F\left[\frac{1}{\sigma} \times \left(\psi\left(\theta_1^*\right) - X_i\beta - \mu\right)\right]$$

$$P(R,S) = P\left(\varepsilon_i < \psi\left(\theta_1^*\right) - X_i\beta\right) - P\left(\varepsilon_i < \psi\left(\theta_3^*\right) - X_i\beta\right)$$

$$= F\left[\frac{1}{\sigma} \times \left(\psi\left(\theta_1^*\right) - X_i\beta - \mu\right)\right] - F\left[\frac{1}{\sigma} \times \left(\psi\left(\theta_3^*\right) - X_i\beta - \mu\right)\right]$$

$$P(R,R) = P\left(\varepsilon_i < \psi\left(\theta_3^*\right) - X_i\beta\right)$$

$$= F\left[\frac{1}{\sigma} \times \left(\psi\left(\theta_3^*\right) - X_i\beta - \mu\right)\right]$$

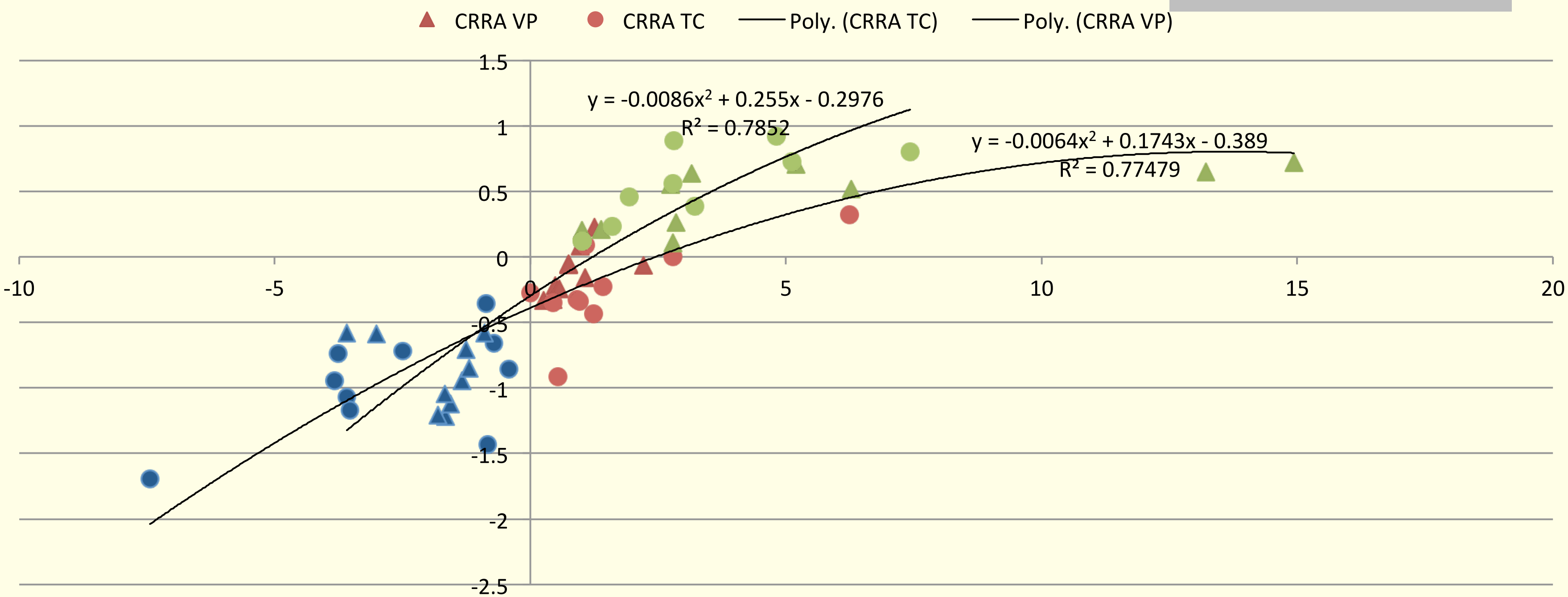
Thresholds for public transport

Scenario	CRRA			CARA			CDF			ν		
1	-7.44	0.91	2.74	-13.80	1.97	6.20	5.6%	37.1%	71.0%	-1.59	-0.33	0.55
2	-3.83	1.42	14.94	-4.89	1.97	20.79	16.8%	41.0%	76.4%	-0.96	-0.23	0.72
3	-3.76	0.00	1.00	-40.56	0.00	14.44	15.9%	39.1%	58.0%	-1.00	-0.28	0.20
4	-3.59	2.80	6.27	-19.69	19.69	41.53	10.4%	50.0%	69.8%	-1.26	0.00	0.52
5	-3.53	0.96	5.19	-3.37	0.99	5.63	11.6%	36.8%	76.1%	-1.20	-0.34	0.71
6	-2.49	6.24	13.21	-7.22	20.28	41.57	18.3%	62.5%	74.0%	-0.91	0.32	0.64
7	-0.87	1.08	3.15	-2.34	3.39	10.09	29.8%	53.5%	73.7%	-0.53	0.09	0.63
8	-0.83	0.53	2.80	-1.41	0.99	5.63	8.4%	18.1%	54.2%	-1.38	-0.91	0.11
9	-0.71	0.44	1.39	-1.41	0.99	3.37	15.8%	36.4%	58.2%	-1.00	-0.35	0.21
10	-0.42	1.24	2.85	-0.29	0.88	2.10	16.8%	33.2%	60.2%	-0.96	-0.43	0.26

Thresholds for private car

scenario	CRRA			CARA			CDF			ν		
1	-3.59	0.74	2.80	-39.38	9.87	39.38	27.9%	47.7%	71.2%	-0.58	-0.06	0.56
2	-3.01	1.25	4.81	-2.16	0.99	3.94	27.7%	59.0%	82.1%	-0.59	0.23	0.92
3	-1.81	0.44	1.93	-3.37	0.99	4.74	11.4%	37.3%	67.6%	-1.21	-0.32	0.46
4	-1.67	0.26	1.61	-1.58	0.29	1.92	14.8%	36.9%	59.1%	-1.05	-0.33	0.23
5	-1.66	0.74	5.12	-3.80	1.97	13.80	11.2%	47.7%	76.6%	-1.22	-0.06	0.73
6	-1.56	2.20	7.43	-2.59	4.22	13.83	13.1%	47.5%	78.8%	-1.12	-0.06	0.80
7	-1.33	0.56	1.03	-3.80	1.97	3.80	17.3%	40.5%	54.8%	-0.94	-0.24	0.12
8	-1.26	0.50	1.00	-7.22	3.40	7.22	24.0%	41.3%	54.7%	-0.71	-0.22	0.12
9	-1.20	1.07	3.22	-1.57	1.57	4.92	19.8%	43.6%	65.0%	-0.85	-0.16	0.39
10	-0.90	0.97	2.80	-1.33	1.65	4.95	28.0%	53.2%	81.2%	-0.58	0.08	0.88

Parametric specification for ψ(θ)



- A good parsimonious parametric specification for individual ***i*** using mode ***m***

$$\psi\left(\theta_i^m\right) = \log(\theta_i^m - \theta_{\inf}^m) = \mu_m + X_i \beta_m + \sigma_m \varepsilon_i^m , \; m = PT, PC$$

Ordered probit vs interval regression

- If $\theta_{\inf}^{m}$ were known, and if all parameters were estimated independently for PC and PT,
- then imposing this parametric specification would amount to estimate an interval regression model on $Y_i^m = \log(\theta_i^m - \theta_{\inf}^m) = \mu_m + X_i\,\beta_m + \sigma_m \varepsilon_i^m$
- Maximum likelihood to jointly estimate

$$\mu_m, \beta_m, \sigma_m,\ m = PT, PC \text{ and } \theta_{\inf}^{m},\ m = PT, PC$$

- Results not reported here show that:

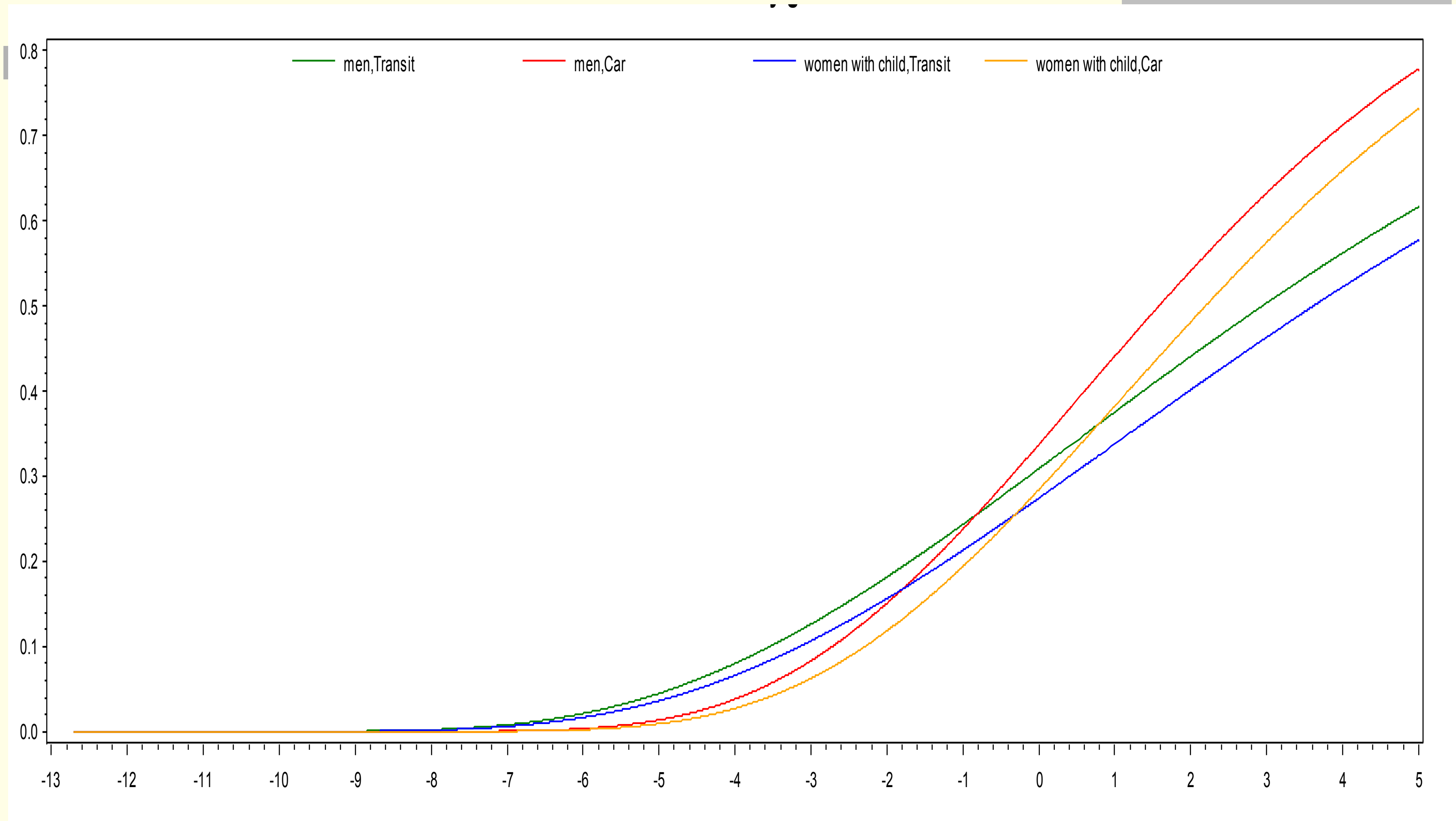
$$\mu_{PT} \neq \mu_{PC}, \beta_{PT} \approx \beta_{PC}, \sigma_m \neq \sigma_m \text{ and } \theta_{\inf}^{PT} \approx \theta_{\inf}^{PC}$$

- Non-significant differences omitted to improve efficiency of estimates

Preliminary estimation results, *generalized* interval regression model

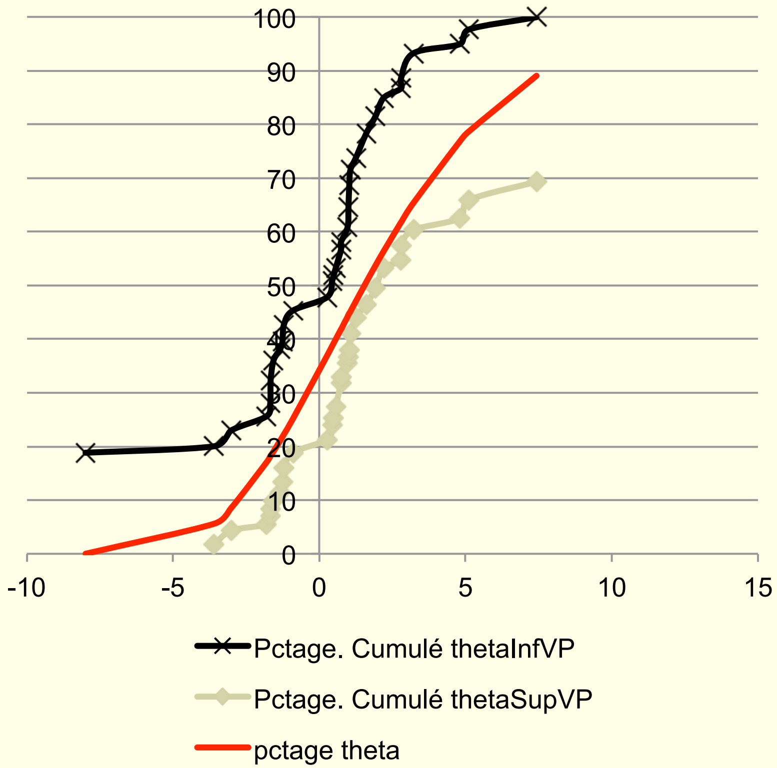
Mode-specific θ_{inf}	no		yes
σ^{PC}	0.279	***	0.248
σ^{PT}	0.417	***	0.429
μ^{PC}	2.66	***	2.772
$\mu^{PT}-\mu^{PC}$	0.091	***	-0.048
θ_{inf}^{PC}	**-12.715**	***	**-14.366**
θ_{inf}^{PT}	**-**		**-12.327**
Woman with no /old children (ref=man)	0.014		0.012
Woman with young children	0.042	**	0.040
Motive="work/professional appointment", reference	-		
*(occup="self-employed/farmer")	-0.016		-0.014
*(occup="white-collar")	-0.038	**	-0.035
*(occup="none") (student, unemployed)	-0.097	**	-0.092
Motive="visit to someone"	-0.06	**	-0.054
Motive="appointment"	-0.04	*	-0.038
Motive="sport. leisure"	-0.064	**	-0.061
Motive="shopping/going back home"	-0.094	***	-0.085
Number of observations	3814		3814
Log-Likelihood	-5405.5		-5405.5

An irrealistic proportion of risk-lovers

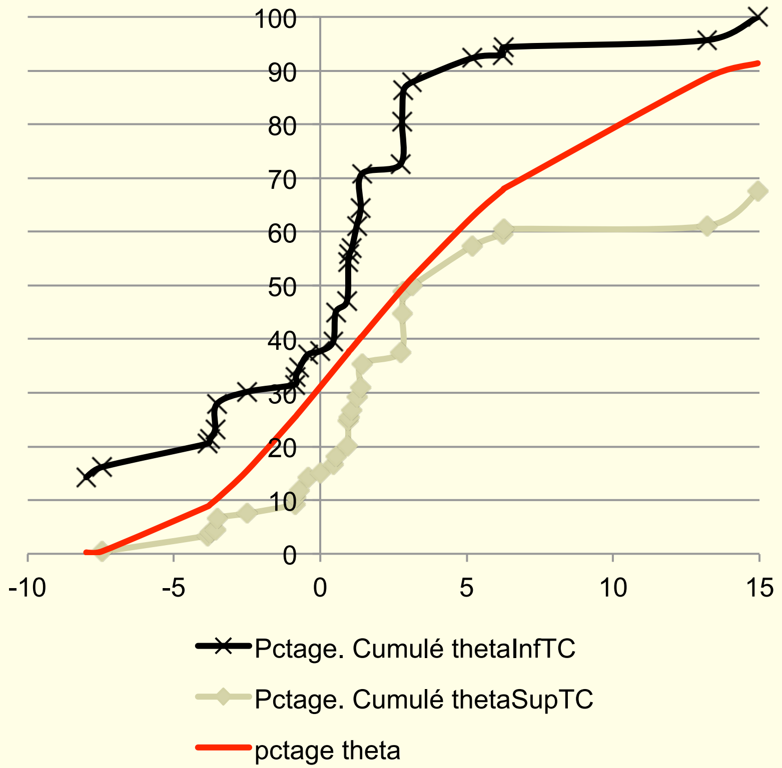


…which is not a related to the parametric specification

CDF of estimated, min and max θ, PC



CDF of estimated, min and max θ, PT



Modelling answers to scenarii to estimate risk aversion

INTRODUCING RANDOM ANSWERS

Risk lovers or random answers?

- Wrong incentives may induce a large number of random answers for complex scenarii
- Assumption: Probability of random answer at question *j* depends on
  - the corresponding threshold θ_j^* (polynomial specification)
  - individual characteristics Y_i

$$\pi_{ij} = \frac{1}{1+\exp\left(-\left[a+b\cdot\theta_j^{*}+c\cdot\theta_j^{*2}+d\cdot\theta_j^{*3}+Y_i\cdot\gamma\right]\right)}$$

- Joint estimation of risk-aversion and probability of random answers
- Similar to a latent variable model

Simultaneous estimation of random answer and risk choice model

$$\tilde{P}_i(S,S) = (1-\pi_1)\cdot(1-\pi_2)\cdot\left\{1-F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_2^*)-X_i\beta-\mu\right)\right]\right\}+\pi_1\cdot\pi_2\cdot\frac{1}{4}$$

$$+\pi_1\cdot(1-\pi_2)\cdot\frac{1}{2}\cdot\left\{1-F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_2^*)-X_i\beta-\mu\right)\right]\right\}+(1-\pi_1)\cdot\pi_2\cdot\frac{1}{2}\cdot\left\{1-F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_1^*)-X_i\beta-\mu\right)\right]\right\}$$

$$\tilde{P}_i(S,R) = (1-\pi_1)\cdot(1-\pi_2)\cdot\left\{F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_2^*)-X_i\beta-\mu\right)\right]-F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_1^*)-X_i\beta-\mu\right)\right]\right\}+\pi_1\cdot\pi_2\cdot\frac{1}{4}$$

$$+\pi_1\cdot(1-\pi_2)\cdot\left\{F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_2^*)-X_i\beta-\mu\right)\right]\right\}+(1-\pi_1)\cdot\pi_2\cdot\frac{1}{2}\cdot\left\{1-F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_1^*)-X_i\beta-\mu\right)\right]\right\}$$

$$\tilde{P}_i(R,S) = (1-\pi_1)\cdot(1-\pi_2)\cdot\left\{F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_1^*)-X_i\beta-\mu\right)\right]-F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_3^*)-X_i\beta-\mu\right)\right]\right\}+\pi_1\cdot\pi_2\cdot\frac{1}{4}$$

$$+\pi_1\cdot(1-\pi_2)\cdot\left\{1-F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_3^*)-X_i\beta-\mu\right)\right]\right\}+(1-\pi_1)\cdot\pi_2\cdot\frac{1}{2}\cdot\left\{F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_1^*)-X_i\beta-\mu\right)\right]\right\}$$

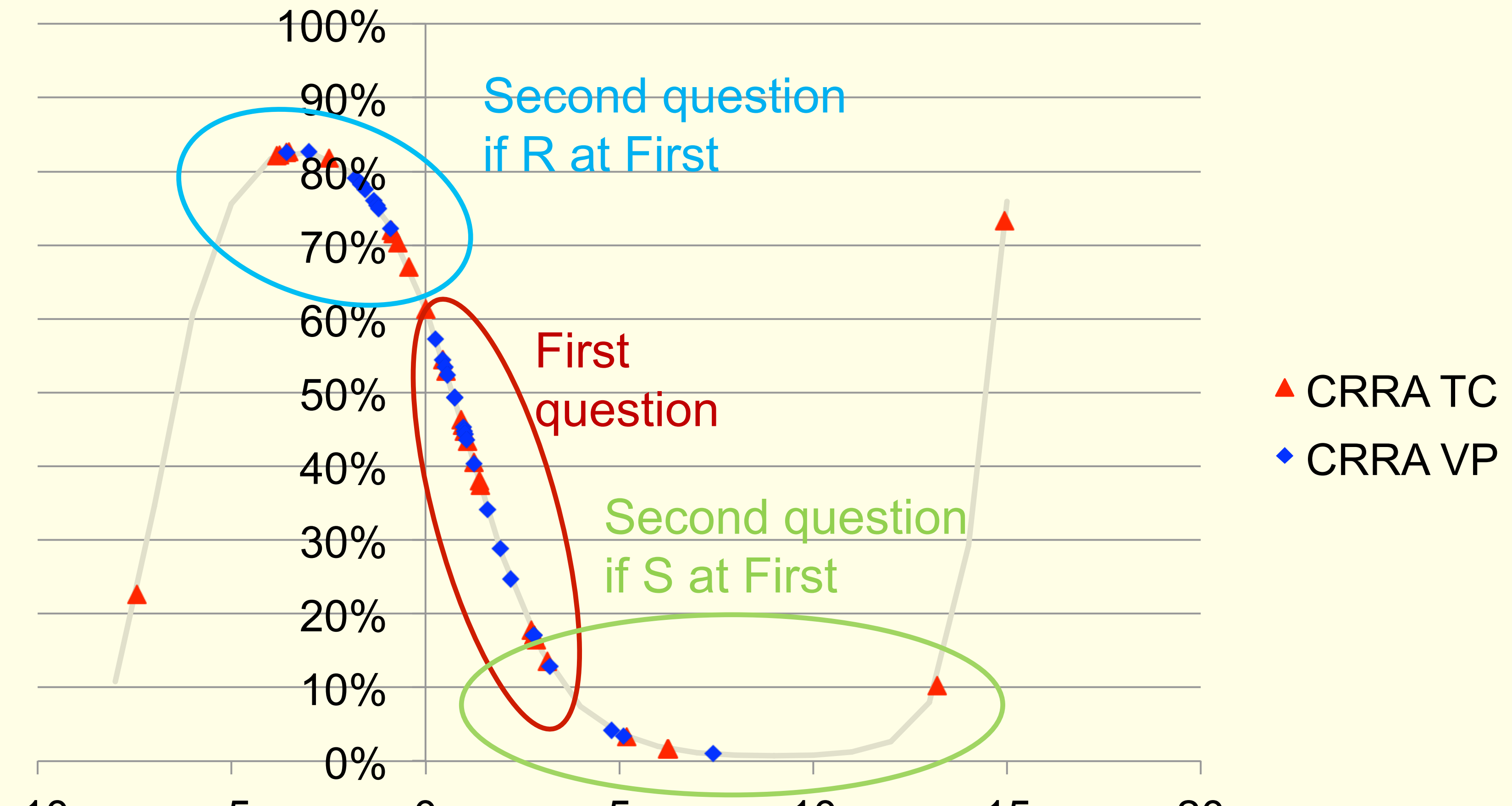
$$\tilde{P}_i(R,R) = (1-\pi_1)\cdot(1-\pi_2)\cdot\left\{F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_3^*)-X_i\beta-\mu\right)\right]\right\}+\pi_1\cdot\pi_2\cdot\frac{1}{4}$$

$$+\pi_1\cdot(1-\pi_2)\cdot\left\{F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_3^*)-X_i\beta-\mu\right)\right]\right\}+(1-\pi_1)\cdot\pi_2\cdot\frac{1}{2}\cdot\left\{F\left[\frac{1}{\sigma}\times\left(\tilde{\psi}(\theta_1^*)-X_i\beta-\mu\right)\right]\right\}$$

Estimation results with and without random answers

	Without		With	
σ^{PC}	0.279	***	2.193	***
σ^{PT}	0.417	***	1.522	***
μ^{PC}	2.66	***	0.697	***
$\mu^{PT}-\mu^{PC}$	0.091	***	0.502	***
θ_{inf}	**-12.715**	***	**0.031**	
Woman with no /old children (ref.=man)	0.014		-0.012	
Woman with young children	0.042	**	0.151	
Motive="work/professional appointment", ref	-			
*(occup="self-employed/farmer")	-0.016		1.034	**
*(occup="white-collar")	-0.038	**	-0.392	***
*(occup="none") (student, unemployed)	-0.097	**	-0.221	
Motive="visit to someone"	-0.06	**	-0.346	
Motive="appointment"	-0.04	*	-0.104	
Motive="sport. leisure"	-0.064	**	-0.284	
Motive="shopping/going back home"	-0.094	***	-0.703	***
Probability of random answer				
Woman (ref=man)			-0.362	*
self-employed/farmer (ref=wage-earner)			1.357	***
inactive or unemployed (ref=wage-earner)			0.685	***
Number of observations	3814		3814	
Log-Likelihood	-5405.5		-5087.0	

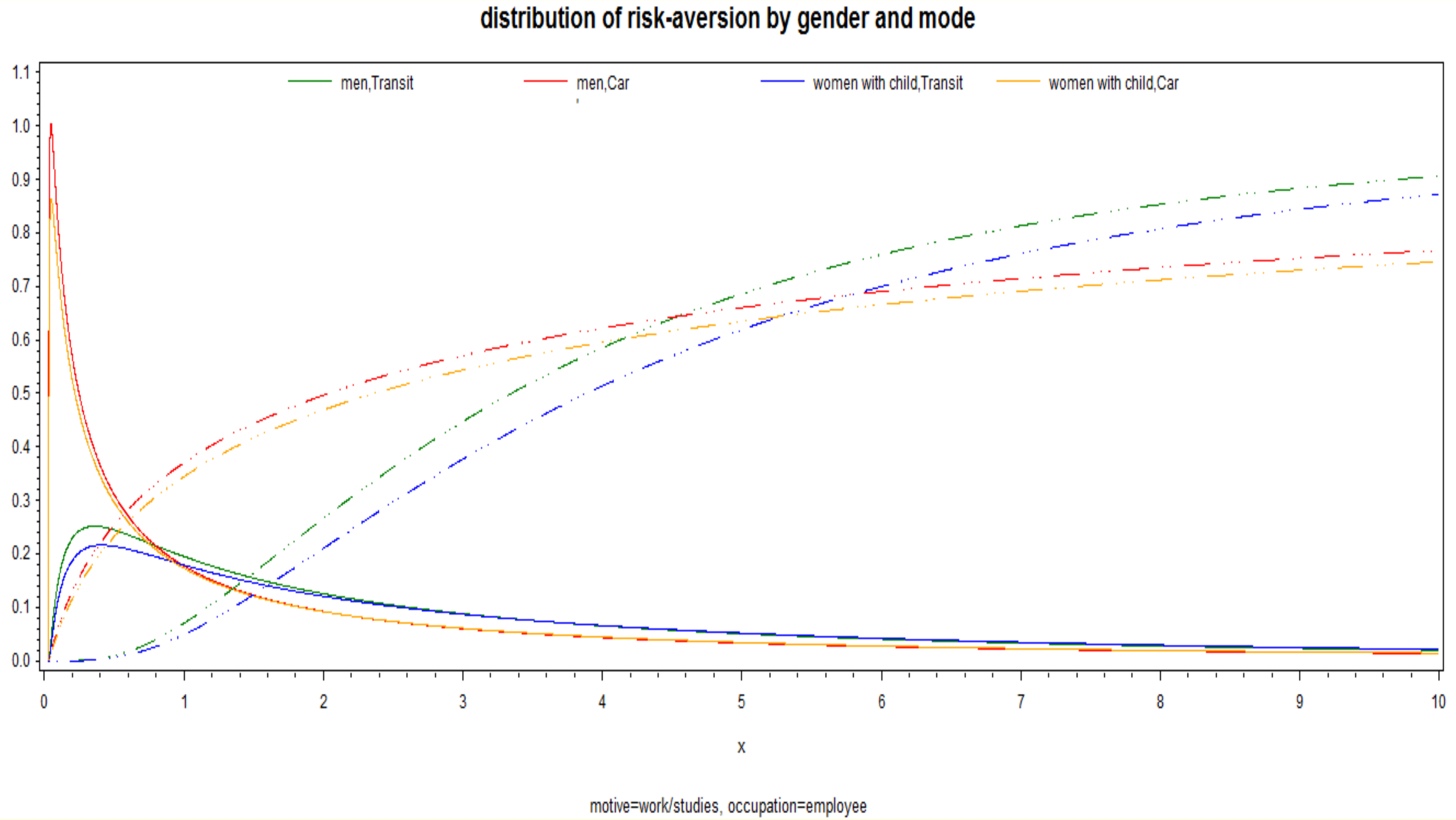
Probability of random answer by threshold



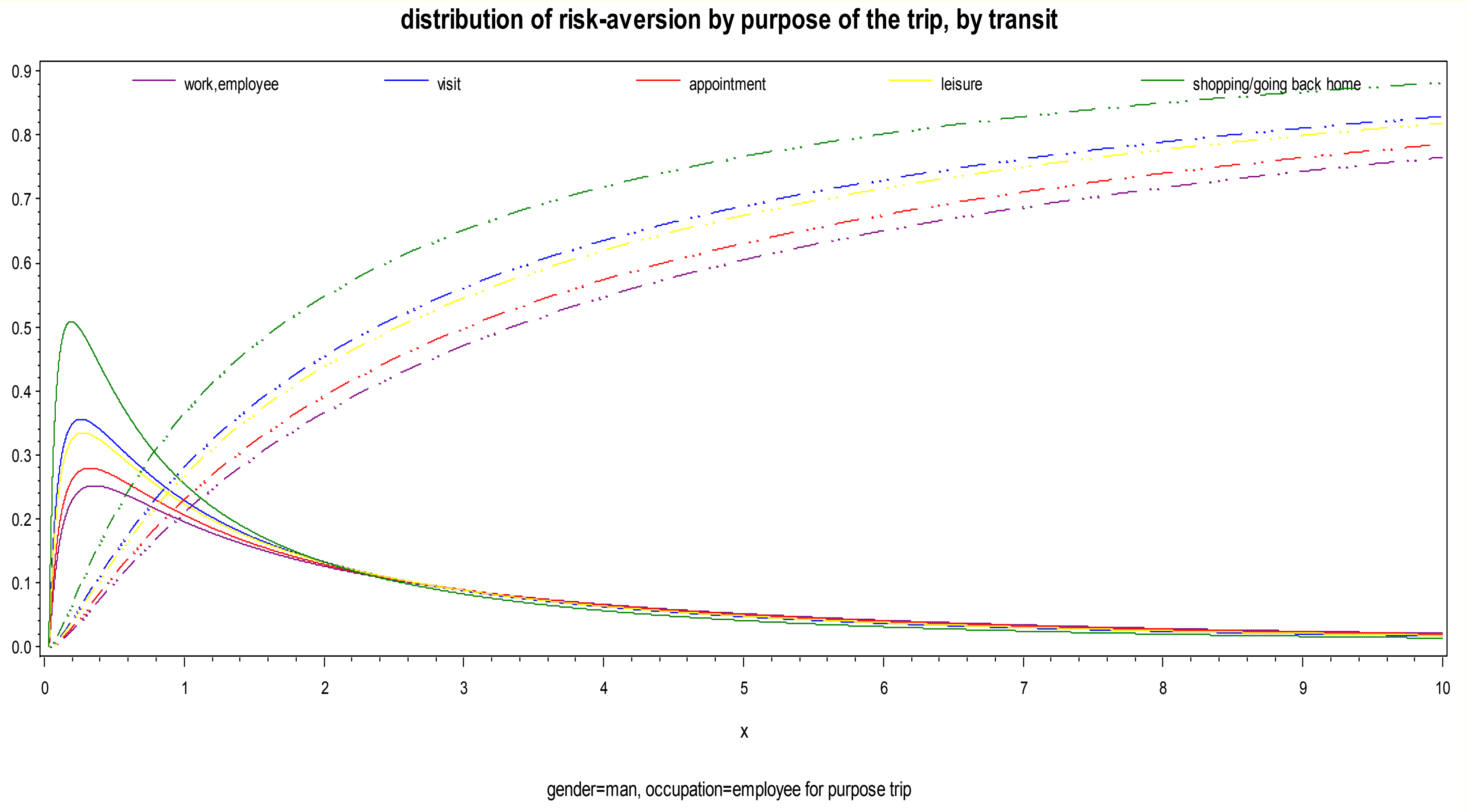
Interval regression with 2 or 3 thresholds

- Probability or random answer at 2nd question
  - close to 1 if R at first
  - close to 0 if S at first
- Probability or random answer at 1st question
  - (less) close to 0
- Imposing limiting probabilities amounts to
  - forget the 2nd answer if R at 1st
  - estimate a 2-threshold *generalized* interval regression model
  - … and solves the % "risk lover puzzle" equally well as the joint estimation of probability of random answer

Distribution (density & CDF) of RA by gender and mode, purpose=work, employee



Distribution of RA by trip purpose, man



Modelling answers to scenarii to estimate risk aversion

HOMOGAMY IN RISK AVERSION

Estimation with couple-specific effect

- The model we wanted to estimate

$$\psi\left(\theta_{GC}^{m_G}\right) = X_{GC}\beta + \gamma_G\sigma_C \cdot \varepsilon_C + \sigma_{G,m_G}\varepsilon_{GC}, G = Male, Female, m_G = PC, PT$$

- But only the sign, not the value of γ_G is identified
- Positive assortative mating ⇔ same sign for γ_M and γ_F ➔ $\gamma_F=\gamma_M=1$ wlog

$$\psi\left(\theta_{GC}^{m_G}\right) = X_{GC}\beta + \sigma_C \cdot \varepsilon_C + \sigma_{G,m_G}\varepsilon_{GC}$$

- Negative assortative mating ⇔ different sign for γ_M and γ_F ➔ $\gamma_F=-\gamma_M=1$ wlog

$$\psi\left(\theta_{FC}^{m_F}\right) = X_{FC}\beta + \sigma_C \cdot \varepsilon_C + \sigma_{F,m_F}\varepsilon_{FC}$$

$$\psi\left(\theta_{MC}^{m_G}\right) = X_{MC}\beta - \sigma_C \cdot \varepsilon_C + \sigma_{M,m_M}\varepsilon_{MC}$$

Results with couple-specific effect

- Positive assortative mating
- No gender differences of σ by mode
- Significant heteroskedasticity in couple effect
  - marital status
  - residence (Paris/inner ring/outer ring)
  - homeownership (owner/renter)

		estimate
σ^{PC}	1.6124***	0.7370***
σ^{PT}	1.7498***	0.6350***
μ^{PC}	0.5364***	0.6039***
μ^{PT}-μ^{PC}	0.3924**	0.3317**
Man, reference		
Woman with no /old children	0.0787	0.0963
Woman with young children	0.1676	0.1993*
living in couple		-0.0335
θ^{PC}_{inf}	-0.3268***	-0.4802***
θ^{PTii}_{inf}	-0.4298**	-0.3343***
Motive="work/professional appointment", worker, reference		
*(occup="self-employed/farmer")	0.0141	0.0923
*(occup="white-collar")	-0.2358***	-0.2183**
*(occup="none") (student, unemployed)	-0.4497**	-0.3936**
Motive="visit to s.o"	-0.3285**	-0.3234**
Motive="appointment"	-0.1644	-0.1784
Motive="sport. leisure"	-0.2967**	-0.2750*
Motive="shopping/going back home"	-0.5740***	-0.5815***
σ		
*residence in Paris		-0.9916***
*residence in Inner Ring		1.0012***
*residence in Outer Ring		1.0808***
*residence in Paris		0.2935
*residence in Inner Ring		0.0080
*residence in Outer Ring		-0.2242
*married couple		0.1968
couple with children <=3 yr old		-0.1291
couple with children >3 yr old		-0.0224
homeownership		-0.2661*
#obs (#couples observed)	3814	3814 (1047)
Log-Likelihood	-4078	-4067

Conclusions and extensions

- Ongoing research
- Preliminary results
  - Probably different tomorrow!
- Still a lot of work to do
- Comment and suggestions welcome