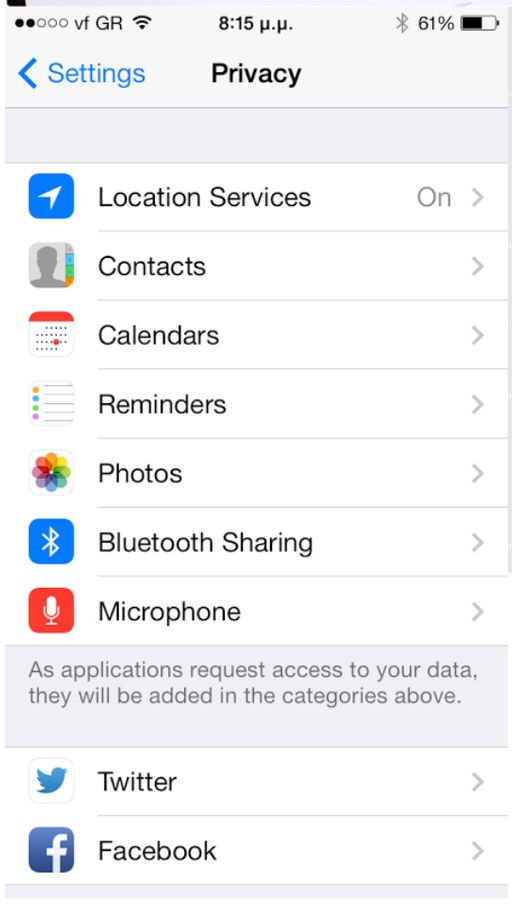
Measuring the value of *, for * in {privacy, preventing a fatality}

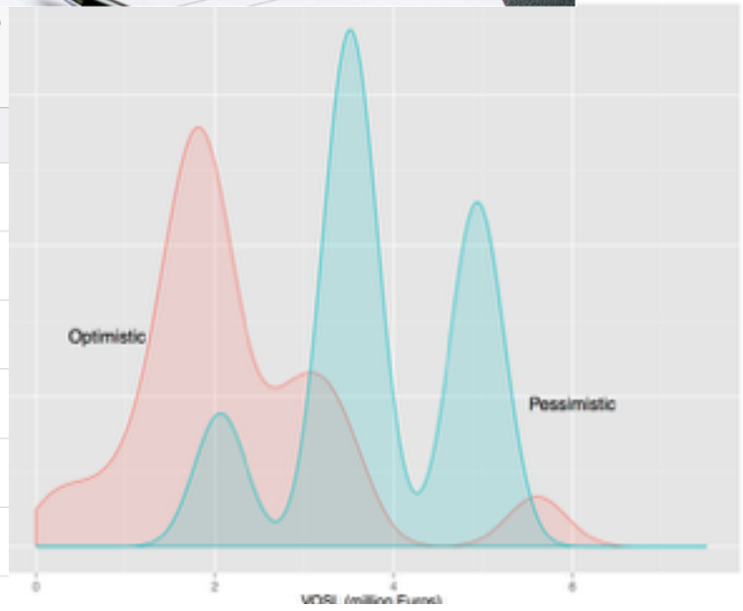
Constantinos Antoniou
National Technical University of Athens, Greece

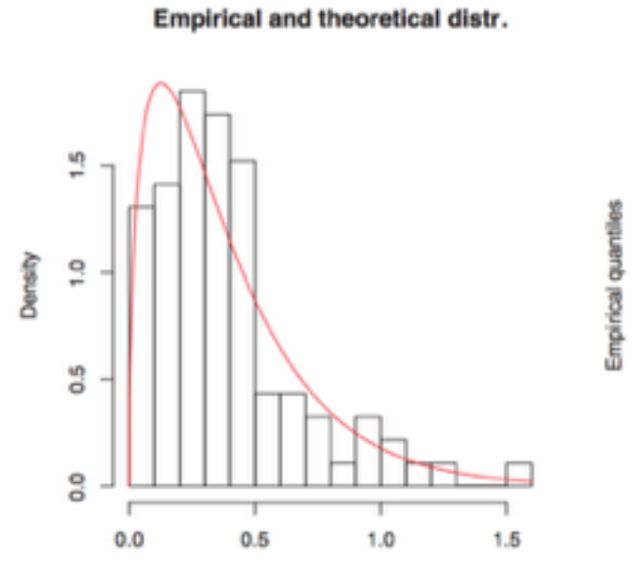
EPFL, April 21, 2016

Outline

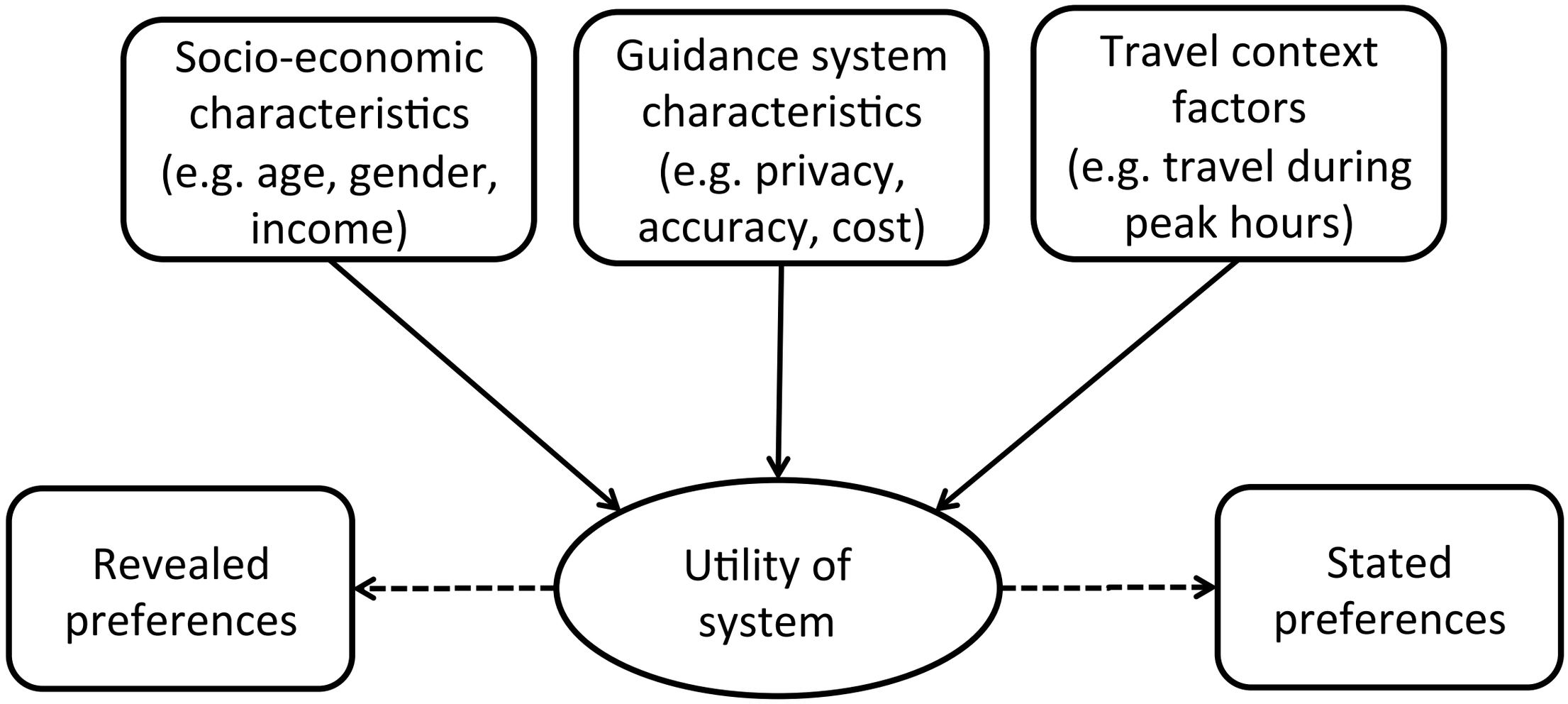
- Background
- Methodology
- Survey design and data collection
- Value of privacy
- Value of preventing a fatality





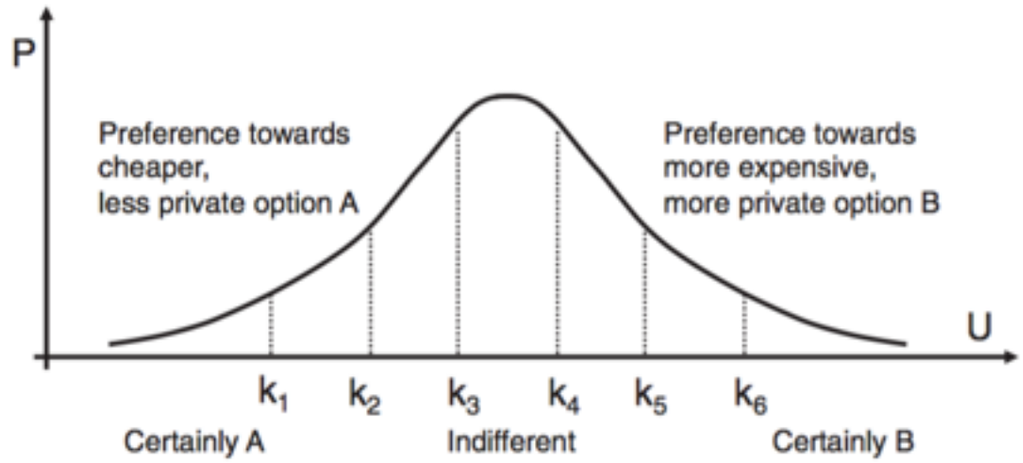


Behavioural model



Methodology

- Ordered logit/probit model
- Random effects model
- Interact *main* variables with explanatory variables
  - Allows us to estimate distributions of coefficients
- Marginal rates of substitution



$$MRS = -\frac{dY}{dX}\bigg|_{U=U_1}$$

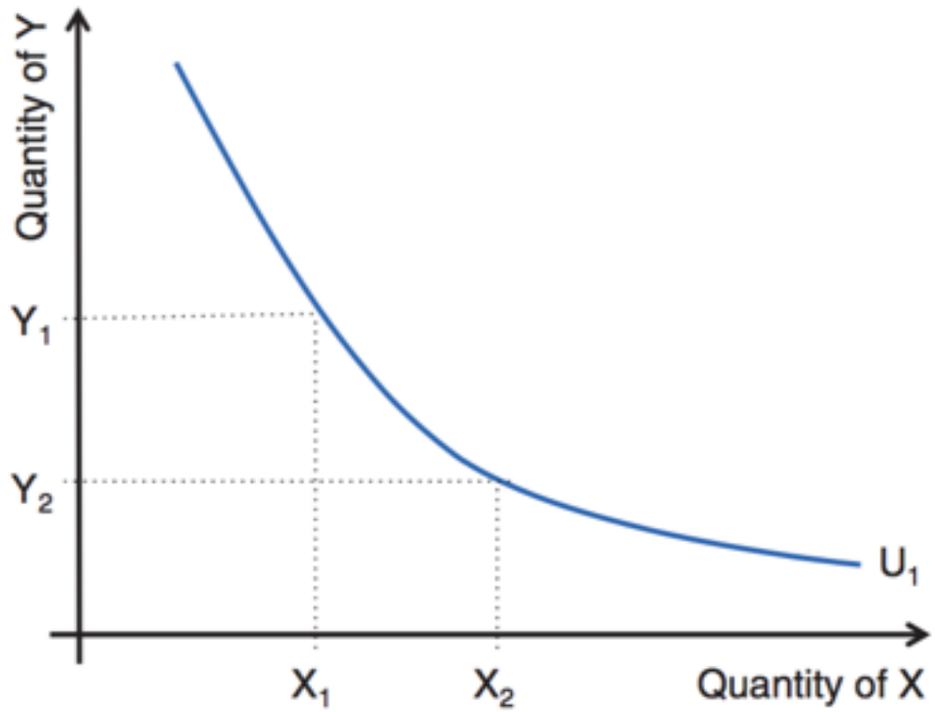


Figure 3 Indifference curve.

$$
\begin{aligned}
V_i = \quad & \Delta cost_i * \left(a_0 + \sum_k a_k \; variable_{ki} \right) \\
& + \Delta privacy_i * \left(\beta_0 + \sum_l \beta_l \; variable_{li} \right) \\
& + \Delta accuracy_i * \left(\gamma_0 + \sum_m a_m \; variable_{mi} \right) + \ldots
\end{aligned} \tag{2}
$$

where:

$\Delta cost_i, \Delta privacy_i, \Delta accuracy_i$, are the difference in cost, privacy level, and accuracy level between the two alternatives.

a_0, β_0, γ_0 are the main coefficients for cost, privacy, and accuracy.

a_k, β_l, γ_m are the additional coefficients, which measure the effect of cost, privacy, and accuracy in the other variables that enter the model specification.

"..." corresponds to additional explanatory parameters in the model.

Survey design

Which alternative route would you choose?						
		Alternative A			**Alternative B**	
Mode of operation Accuracy Cost		No location data Better 8€/month			Driving pattern Good 2€/month	
1	2	3	4	5	6	7
Strong preference towards A	Moderate preference towards A	Mild preference towards A	Indiffe-rent	Mild preference towards B	Moderate preference towards B	Strong preference towards B

have units of €/privacy level, which is the expected unit for a value-of-privacy (VOP) measure:

$$VOP = \frac{\beta_{privacy}}{\beta_{cost}} \left(\frac{utils/privacy\ \ level}{utils/€} = €/privacy\ \ level \right) \quad (3)$$

Similarly, the value of service accuracy (VOA) can be obtained by:

$$VOA = \frac{\beta_{accuracy}}{\beta_{cost}} \left(\frac{utils/accuracy\ \ level}{utils/€} = €/accuracy\ \ level \right) \quad (4)$$

Antoniou and Polydoropoulou (2014)

Initial results

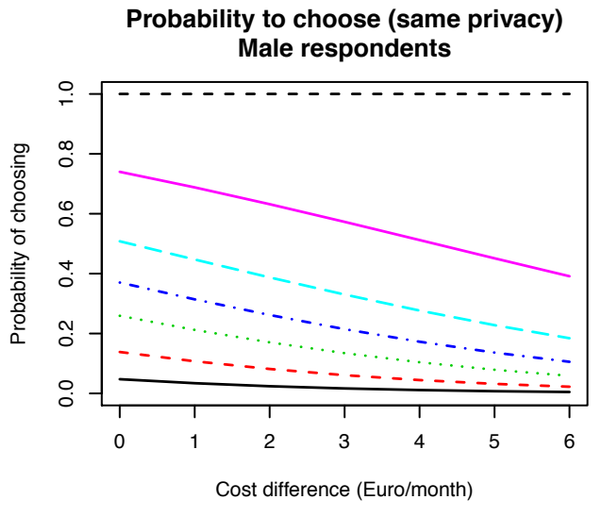
	Value of privacy €/month for each privacy level	Value of accuracy (€/month)
Ordered logit models		
All respondents	2.21	0.96
Random effects ordered logit models		
All respondents	2.19	0.95
Female respondents	2.62	1.04
Male respondents	1.80	0.83

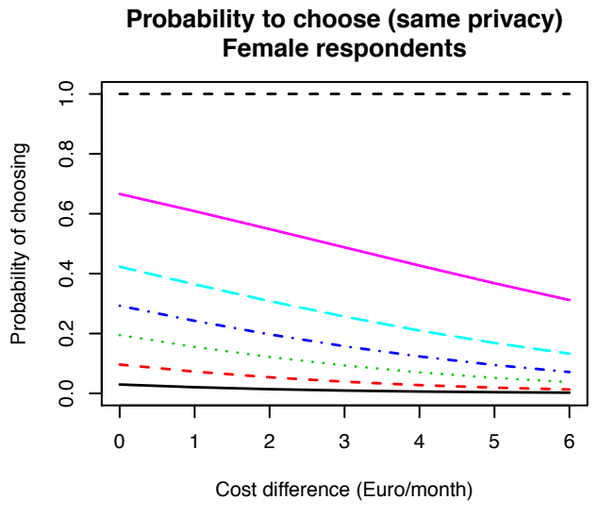
Antoniou and Polydoropoulou (2014)

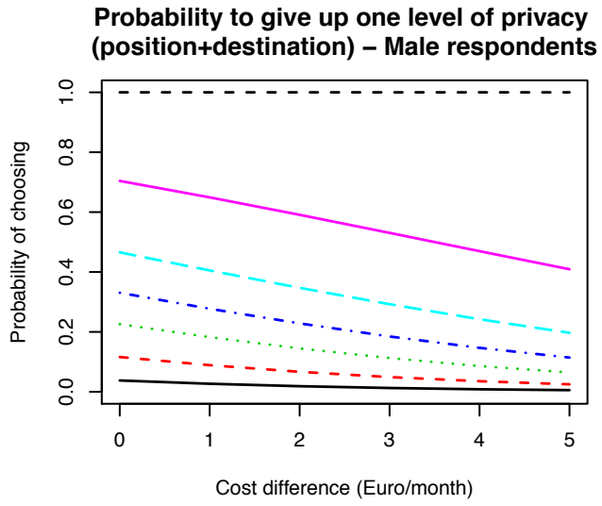
Results

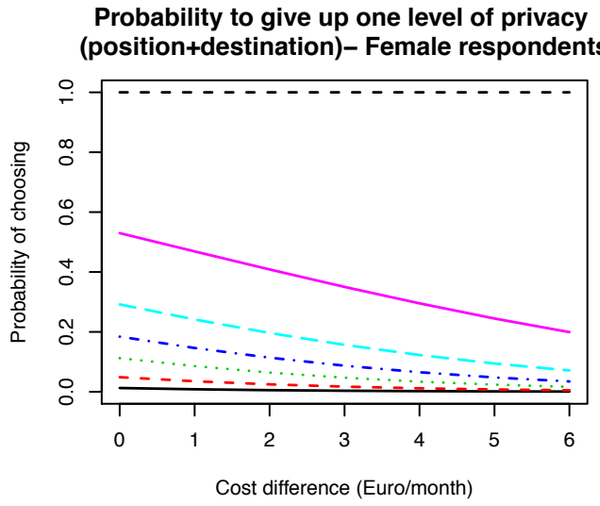
- Presented a framework for the estimation of the willingness to accept giving up privacy
- It can be inferred that women are willing to give out less personal information (relative to men)
- Average amount that the respondents would like to receive in order to give up one level of privacy is 2.2€/month
- Women would demand 2.6€/ month and men 1.8€/month

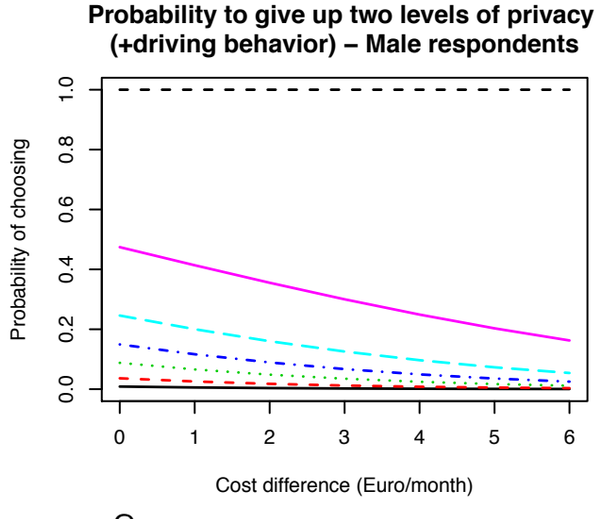
Antoniou and Polydoropoulou (2014)

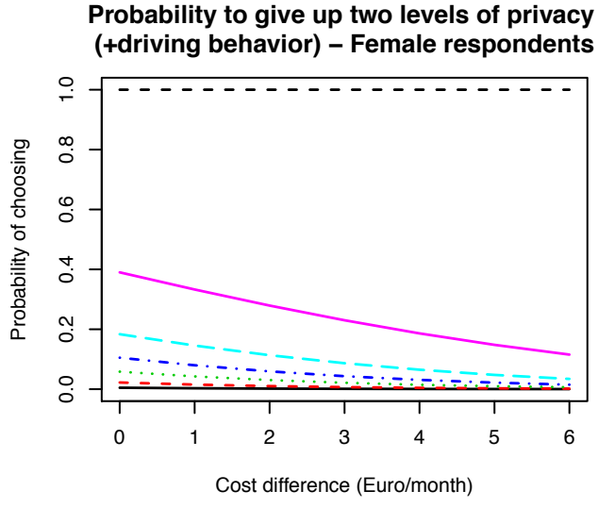












$$
\begin{aligned}
V_i = \quad & \Delta cost_i * \left(a_0 + \sum_k a_k \; variable_{ki} \right) \\
& + \Delta privacy_i * \left(\beta_0 + \sum_l \beta_l \; variable_{li} \right) \\
& + \Delta accuracy_i * \left(\gamma_0 + \sum_m a_m \; variable_{mi} \right) + \ldots
\end{aligned} \tag{2}
$$

where:

$\Delta cost_i, \Delta privacy_i, \Delta accuracy_i$, are the difference in cost, privacy level, and accuracy level between the two alternatives.

a_0, β_0, γ_0 are the main coefficients for cost, privacy, and accuracy.

a_k, β_l, γ_m are the additional coefficients, which measure the effect of cost, privacy, and accuracy in the other variables that enter the model specification.

"..." corresponds to additional explanatory parameters in the model.

Antoniou and Polydoropoulou (2015)

Table 2 Model estimation results—all respondents.

Variables	All respondents	
	Estimate	*t* Value
Cost (€/month)	−0.116	−2.000
Accuracy	1.271	3.973
Cost × Suburb	0.077	4.068
Cost × Peak	−0.147	−2.123
Cost × Own GPS	0.112	2.228
Cost × Male	−0.062	−3.248
Cost × Age less than 25 years	−0.159	−1.444
Accuracy × Own GPS	−0.920	−3.147
Accuracy × Use GPS frequently	−0.285	−3.514
Accuracy × Children	−0.130	−2.076
Operation × Peak	−0.389	−1.817
Operation × Use GPS frequently	−0.119	−2.922
Operation × Children	−0.069	−1.581
Operation × Income	0.072	4.161
Operation × Age less than 25 years	−0.625	−1.818
Operation × Age more than 50 years	−0.227	−2.616
Operation × Own GPS and drive both in urban and rural roads	−0.174	−1.922
Threshold parameters for index model		
1|2	−1.117	−14.044
2|3	−0.527	−7.132
3|4	−0.181	−2.496
4|5	0.127	1.748
5|6	0.569	7.572
6|7	1.181	14.214
Summary statistics		
Number of observations	840	
Initial log-likelihood	−1620.07	
Final log-likelihood	−1581.63	
AIC	3209.26	

Antoniou and Polydoropoulou (2015)

Distributions of estimated coefficients

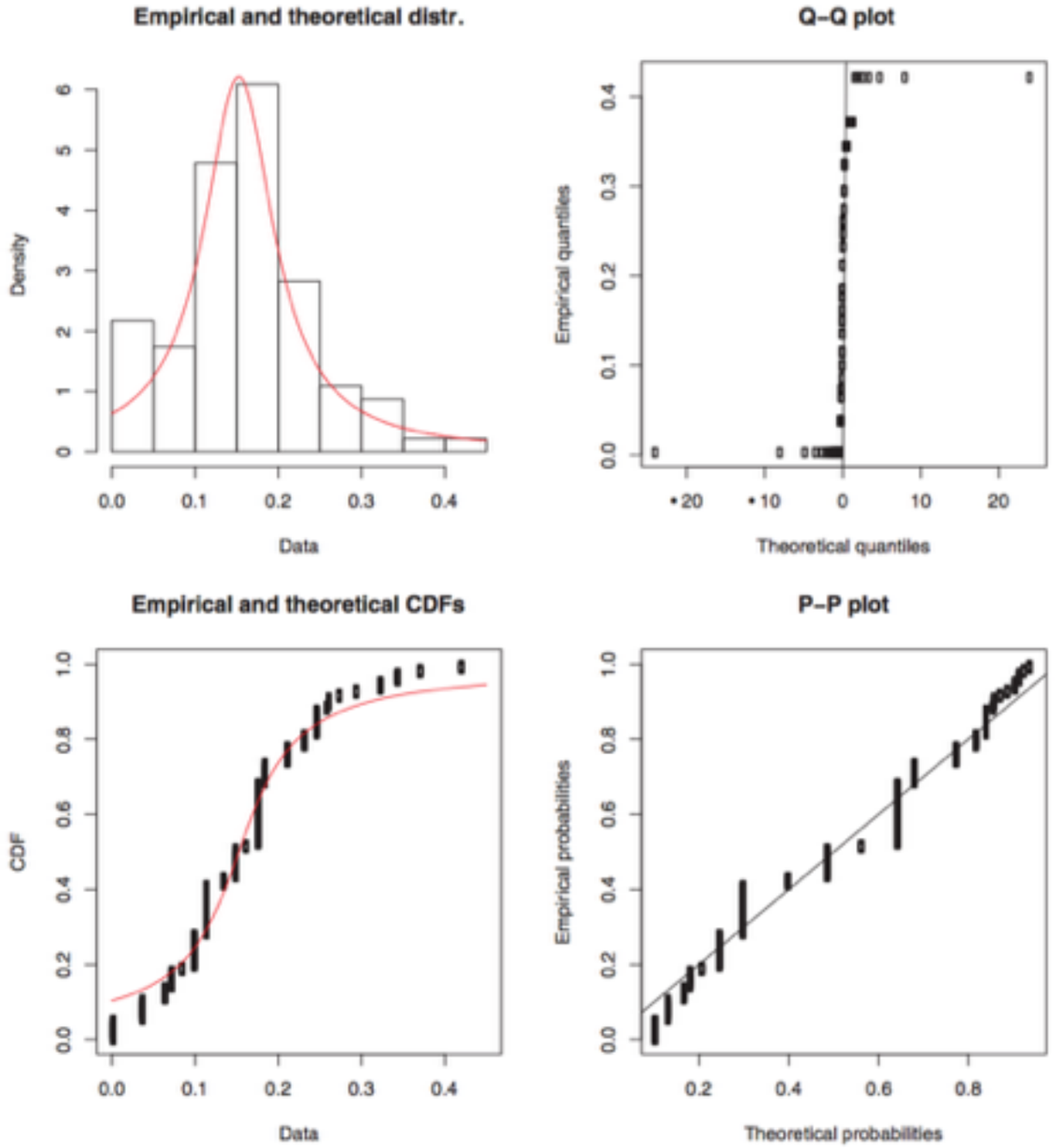


Figure 5 Estimated distribution of cost coefficient (and fit of Cauchy distribution).

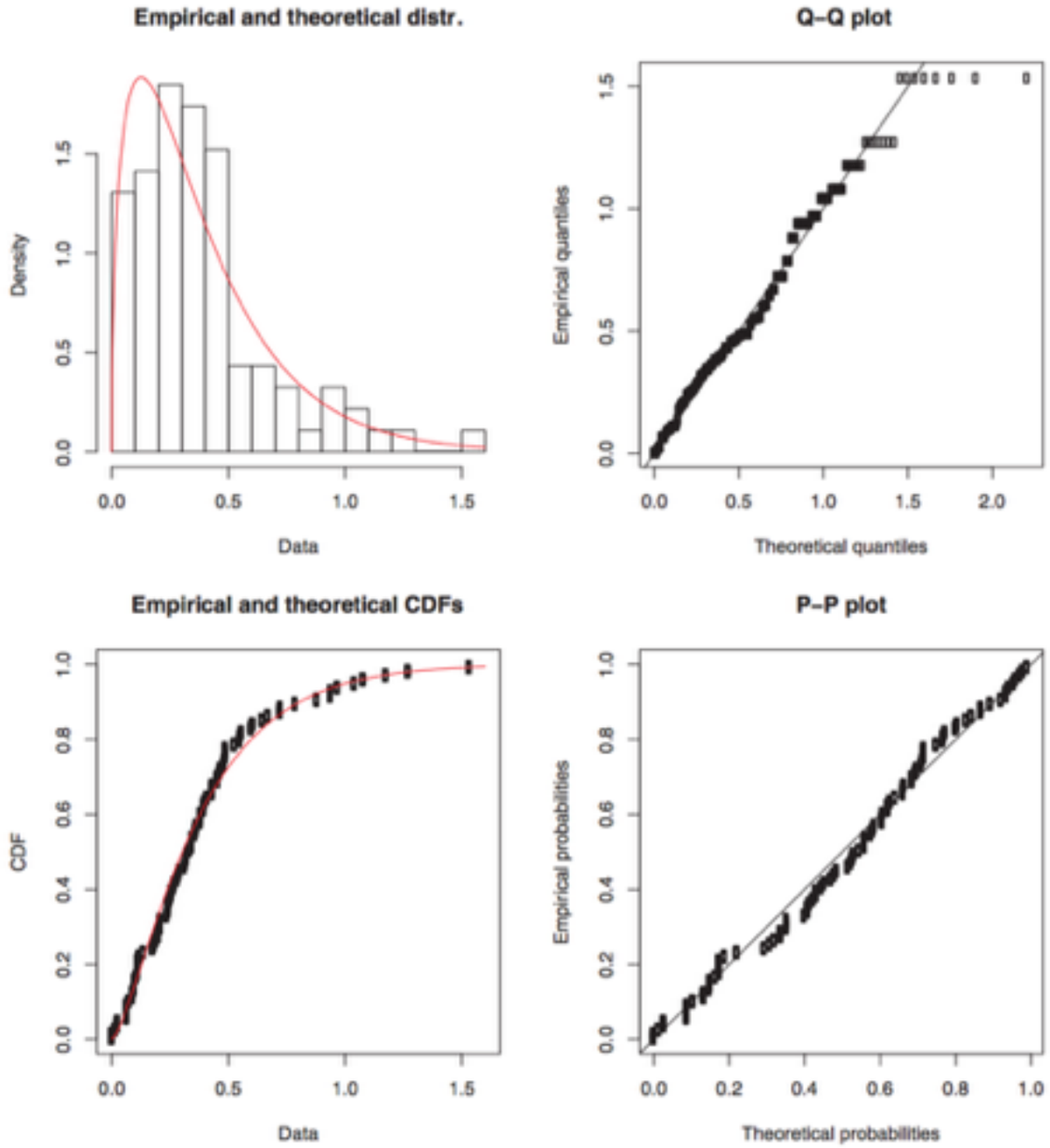


Figure 6 Empirical distribution of level of operation coefficient (and fit of gamma distribution).

Antoniou and Polydoropoulou (2015)

Questionnaire Design

Introduction

Optimistic information

Pessimistic information

Part 1
General questions on the participant's driving experience

Part 2
The choice experiment

3 Variables (Cost, Risk and Time)
+
3 Difference levels
↓
27 Hypothetical Scenarios (criteria of level balance, orthogonality, minimal overlap and utility balance)
3 (9-choice) blocks

Correspondence with a real route situation (reference values)

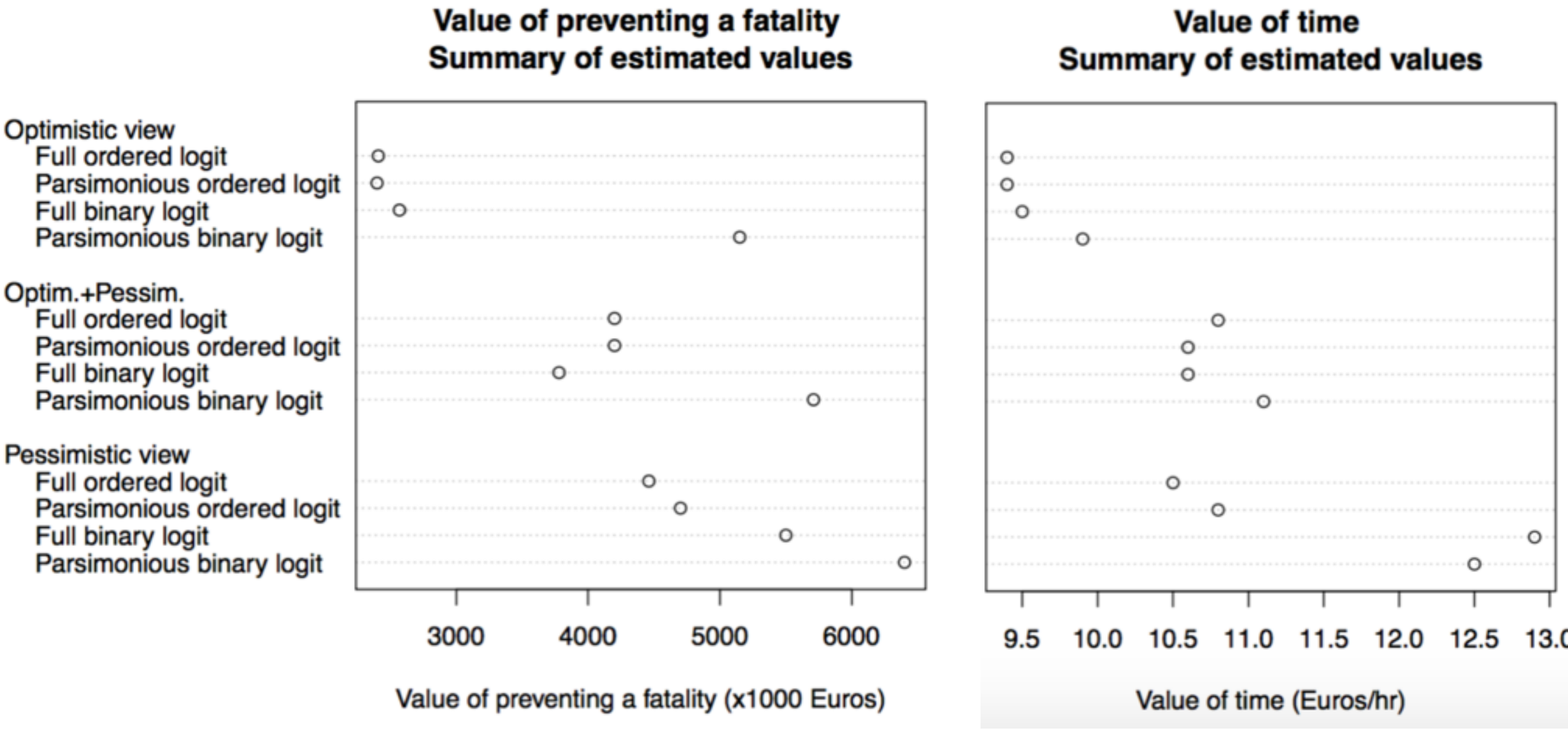
	Alternative A		Alternative B	
Cost	6 €		7 €	
Risk	24 fatalities/year		20 fatalities/year	
Time	3 hours		2 hours and 15'	
Certainly A	Probably A	Indifferent	Probably B	Certainly B

Part 3
Questions on the participant's perceptions about road safety and socio-economic data

Antoniou and Kostovasilis (2012)

ORDERED LOGIT MODELS	Pooled		Optimistic		Pessimistic	
Variable	**Est.**	**z-value**	**Est.**	**z-value**	**Est.**	**z-value**
Travel cost	-0.176	-3.710	-0.161	-2.378	-0.147	-2.179
Travel time	-1.900	-12.712	-1.518	-7.645	-1.546	-7.672
Risk	-0.136	-4.824	-0.071	-1.896	-0.12	-3.082
Self-reported driving ability: better than average	0.505	3.577	-0.238	-1.16		
Data: face-to-face interview	0.854	5.850			1,105	5.067
Road safety measures perception: insufficient	-0.277	-1.892	-0.379	-1.92		
Crashes as passenger	-0.209	-2.805	0.277	2.578		
Kilometers traveled per year (x1000)			-0.015	-1.654	-0.036	-2.656
Young (Age < 30 years)					0.565	2.601
Gender: Male	-0.823	-5.353	0.658	2.801		
Road: other	-0.686	-3.629			-1.764	-6.203
Road: Athens-Thessaloniki National Highway	-0.573	-3.542			-0.305	-1.343
Intercept						
k 1 | 2	-2.217	-7.039	-1.912	-4.401	-1,799	-4.298
k 2 | 3	-1.171	-3.854	-0.925	-2.179	-0.806	-1.994
k 3 | 4	-0.946	-3.129	-0.713	-1.683	-0.61	-1.516
k 4 | 5	0.468	1.565	0.446	1.056	0.87	2.158
Observations	846		423		423	
AIC	1912.1		1011.4		978.4	

Antoniou and Kostovasilis (2012)



Antoniou and Kostovasilis (2012)

3.4 Value of preventing a fatality

Antoniou and Kostovasilis (under review)

Suppose the following general formulation for the systematic component of the utility function is used (to get the full utility specification one needs to add the error term):

$$V_i = \Delta cost_i * \left(a_0 + \sum_k a_k \, variable_{ki}\right) +$$

$$+\Delta time_i * \left(\beta_0 + \sum_l \beta_l \, variable_{li}\right) + \quad (1)$$

$$+\Delta risk_i * \left(\gamma_0 + \sum_m a_m \, variable_{mi}\right) + \ldots$$

where:

$\Delta cost_i, \Delta time_i, \Delta risk_i$ are the difference in travel cost, travel time and travel risk level between the two alternatives;

a_0, β_0, γ_0 are the main coefficients for cost, time and risk;

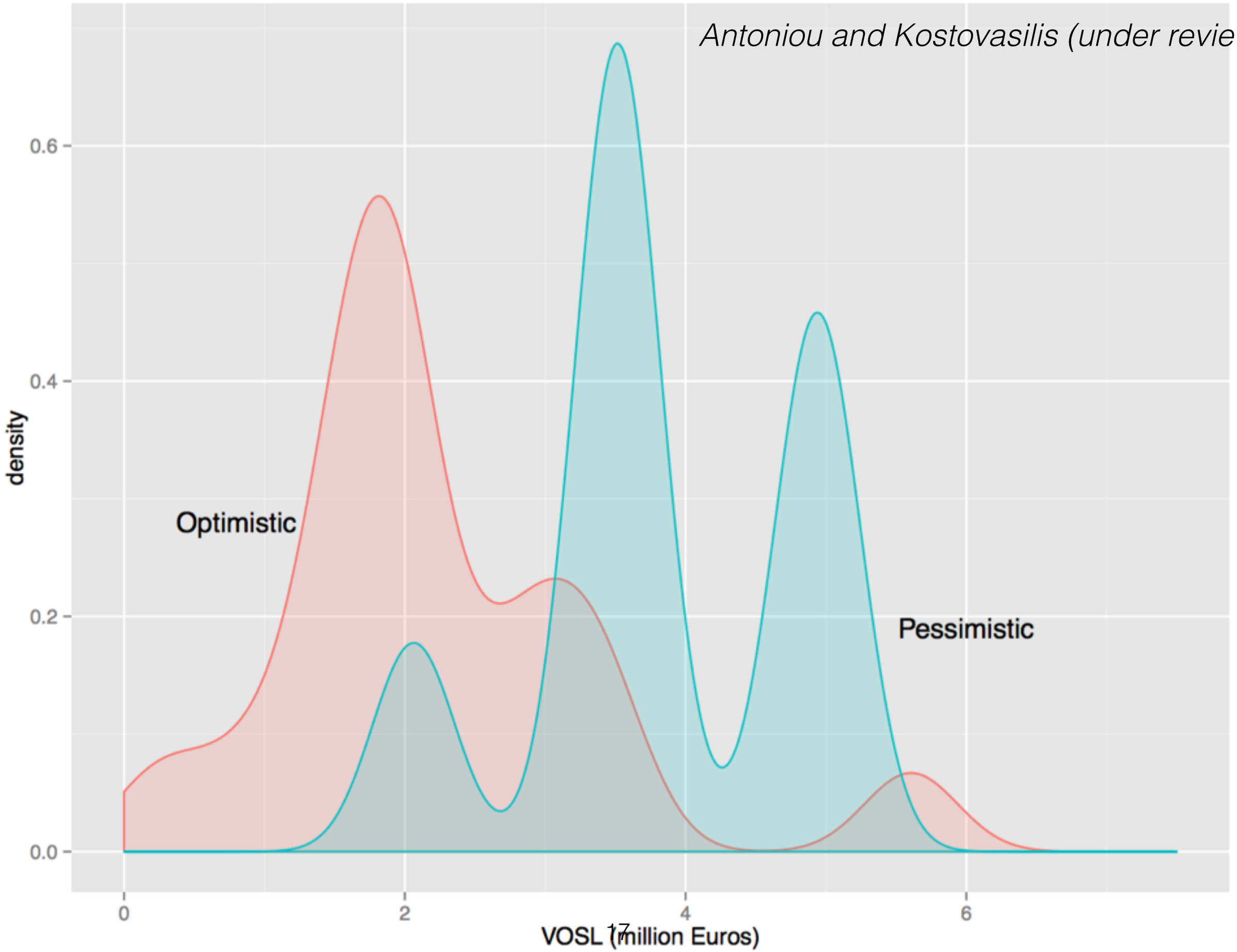
a_k, β_l, γ_m are the additional coefficients, which measure the interaction of cost, time and risk in the other variables that enter the model specification; and

"…" corresponds to additional explanatory parameters in the model.

Antoniou and Kostovasilis (under review)

	Optimistic		Pessimistic	
Variables	*Estimate*	*t-value*	*Estimate*	*t-value*
Intercept	1.152	8.335	1.108	7.835
Travel time	-0.585	-2.616	-0.606	-2.802
Travel risk	-0.311	-3.585	-0.163	-6.556
Cost*(Age <30 years)	-0.258	-4.572	-0.244	-4.542
Cost*(45<Age<65)	-0.308	-3.590		
Cost*(Km driven per year x1000)	0.009	3.514		
Time*(Drive mostly on other main roads)	-0.686	-2.821	-0.598	-2.126
Time*(Drive mostly on Athens-Thessaloniki Motorway)	-0.529	-2.404		
Time*Email response			-0.829	-4.109
Time*(Road Safety Measures Considered Insufficient)	-0.594	-3.072		
Time*(Had Accident as Driver)			-0.214	-3.188
Time*(Had Accident as Passenger)			-0.380	-3.136
Risk*Female	-0.089	-4.976		
Risk*(Drive mostly on other main roads)			0.052	2.050
Risk*(Drive mostly on Athens-Thessaloniki Motorway)			0.051	2.891
Risk*Email respondent	0.122	6.872		
Risk* (Age < 30 years)	0.173	1.970	0.035	1.582

Antoniou and Kostovasilis (under review)



References

- Antoniou, C. and K. Kostovasilis (2012). Can external stimuli affect the perceived value of statistical life? Proceedings of the 91st Annual Meeting of the Transportation Research Board, January 2012, Washington, D.C.

- Antoniou, C. and A. Polydoropoulou (2014). How likely are travelers to give up information in exchange for better user information services? Proceedings of the 93rd Annual Meeting of the Transportation Research Board, January 2014, Washington, D.C.

- Antoniou, C. and A. Polydoropoulou (2015). The value of privacy. Evidence from the use of mobile devices for traveler information systems. Journal of Intelligent Transportation Systems: Technology, Planning and Operations, 19(2), pp. 167-180.

- Antoniou, C., and K. Kostovasilis (under review). How may external information affect traffic risk perception?

Measuring the value of *, for * in {privacy, preventing a fatality}

Constantinos Antoniou
National Technical University of Athens, Greece

EPFL, April 21, 2016