Accepted refereed manuscript of: Sandorf ED, Persson L & Broberg T (2020) Using an Integrated Choice and Latent Variable Model to Understand the Impact of "Professional" Respondents in a Stated Preference Survey. *Resource and Energy Economics*, 61, Art. No.: 101178.
https://doi.org/10.1016/j.reseneeco.2020.101178

Using an Integrated Choice and Latent Variable Model to Understand the Impact of "Professional" Respondents in a Stated Preference Survey

Erlend Dancke Sandorf [†a], Lars Persson[b], and Thomas Broberg[b]

[a]University of Stirling, Stirling Management School, Economics Division

[b]Umeå University, School of Business & Economics, Department of Economics, Centre for Environmental and Resource Economics

Abstract: Internet panels are increasingly used for stated preference research, and members of such panels receive compensation for each completed survey. One concern is that over time this creates professional respondents who answer surveys to receive the monetary compensation. We identify professional respondents using data on panel tenure, survey response frequency, completion rates and total number of completed surveys. We find evidence of two types of professional respondents: "hyperactives" who answer surveys frequently and "experienced" who have long panel tenure and a large number of completed surveys. Using an integrated choice and latent variable model in a stated preference survey, we find that "hyperactive" respondents are less likely to choose the 'status quo' and have a more stochastic choice process as seen from the econometrician's point of view, whereas "experienced" respondents have a relatively more deterministic choice process. Our results show that "hyperactive" respondents significantly impact estimated values.

Keywords: Professional respondents, Internet panels, Discrete Choice Experiments, Integrated Choice and Latent Variable Model

JEL codes: C35, Q51

†University of Stirling, Stirling Management School, Economics Division, FK9 4LA, Stirling, United Kingdom
Email: e.d.sandorf@stir.ac.uk
Acknowledgements:
The first author acknowledges funding from the Jan Wallander and Tom Hedelius foundation, and the Tore Browaldhfoundation grant no. B2015-0490:1; and funding from the European Union's Horizon 2020 research and innovation programmeunder the Marie Sklodowska-Curie grant agreement No 793163 (INSPiRE). Erlend Dancke Sandorf did the majority ofthe work while he was working at the Swedish University of Agricultural Sciences. In additiona, financial support fromEnergiforsk is gratefully acknowledged. Any remaining errors are the sole responsibility of the authors.Journal Pre-proof

1 Introduction

Internet panels have become a popular way to gather stated preference data (Lindhjem and Navrud, 2011). As these panels have matured, population coverage has improved to the point where the vast majority of the population have access to the internet (Baker et al., 2010; Internet World Stats, 2017)[1]. While sampling concerns with respect to internet panels have diminished over time - concerns with respect to response quality have increased (see e.g. Baker et al., 2010; Hess and Stathopoulos, 2013; Hillygus et al., 2014; Gao et al., 2016; Campbell et al., 2017). Internet panel members have agreed to receive and answer surveys for compensation. Granted, other survey modes such as face-to-face interviews and mail-out surveys also offer compensation, but in internet panels - respondents receive surveys on a wide range of topics at regular intervals. A concern is that this creates "professional" respondents who elect to participate in multiple panels and surveys to obtain the economic incentive offered to them (Baker et al., 2010)[2]. If this type of rent seeking behavior is indeed their true motivation for participating in any given survey, then how far can we trust that the answers they provide are a reflection of their underlying preferences?

Much of the research on panel and response quality is done in the context of non-probability panels, and research with respect to "professional" respondents is mostly published in white papers by industry or in conference proceedings (Baker et al., 2010; Whitsett, 2013; Hillygus et al., 2014). In the non-probability panels, respondents are often recruited using banner-ads on web-pages, e-mail blasts or recruitment campaigns. This poses two major issues for the use of results. First, there is a large problem with self-selection into the panel. This is most often cited as one of the root causes for the creation of professional respondents (Baker et al., 2010; Baker and Downes-Le Guin, 2007). This self-selection is reduced in probability based panels, but depending on how often panel members are changed, certain members can have significant experience. Second, the non-probability panel itself constitutes a non-random sample. This is particularly problematic if the purpose of the study is to extrapolate out of sample and say something about the general population.

Previous studies exploring the impact of professional respondents have often relied on a single indicator to capture "professionalism" (see e.g. Hillygus et al., 2014, for an overview). We believe that any one indicator is unable to fully capture the latent construct "professionalism". To overcome this, we include several of the indicators characterizing "professional" respondents used in the literature in an Integrated Choice and Latent Variable (ICLV) model (see e.g. Ben-Akiva et al., 2002; Fosgerau and Bjørner, 2006;

[1] In Sweden, for example, 92.9 percent of the population has access to the internet (Internet World Stats, 2017).
[2] Sometimes respondents who are members of multiple panel and frequently participate in surveys are referred to as "hyperactives", "frequent survey takers" or "survey savvy".

Hess and Beharry-Borg, 2012). To help us identify proper indicators, we used an exploratory factor analysis to see which indicators correlated with which latent constructs. Most notably, we use exact measures of panel tenure, average number of completed surveys per month, survey completion rates and total number of completed surveys (Whitsett, 2013). Furthermore, to understand who these "professional" respondents are, we included socio-demographic characteristics such as education, income and employment status in the structural equations for the latent variables. The ICLV modeling framework, allows us to use multiple indicators to identify professional respondents, and at the same time avoid some of the issues related to measurement error and potential endogeneity bias from including such measures directly in the choice model (Ben-Akiva et al., 2002).

The purpose of this paper is to understand better how "professional" respondents affect elicited preferences and ultimately measures of interest such as willingness-to-pay or willingness-to-accept. Given the prevailing ideas that "professional" respondents are "in it for the money", it is reasonable to suspect that they seek to maximize their income from answering surveys rather than provide thoughtful and accurate answers. It is possible that this translates into a form of simplifying behavior where a respondent will take the low effort approach of consistently choosing the 'status quo' or just randomly choosing alternatives without regard to what they contain. As such we test three hypotheses: That professional respondents have i) a higher propensity to choose the status quo, or ii) a relatively more stochastic choice process as seen from the econometrician's point of view, and iii) different willingness-to-accept estimates compared to less professional respondents. In doing so, this paper is the first to explore the role of professional respondents in the context of a stated preference survey, and specifically how such respondents impact measures such as willingness-to-accept. This paper is also one of the few to explore this in the context of a probability-based internet panel. Lastly, we use recorded indicators of "professionalism" rather than self-reported measures.

Our results show that there are two main types of professional respondents: i) "hyperactives" who have a high completion rate and survey frequency, and ii) "experienced" who are long term members of the panel and has answered a large number of surveys. We conjecture that the problematic respondents are those identified as "hyperactive". They are less likely to choose the 'status quo' alternative and have a more stochastic choice process as seen from the econometrician's point of view. This translates into a marked downward shift in compensation sensitivity in the ICLV model. Respondents identified as "experienced" are not significantly less likely to choose the 'status quo', but they do have a more deterministic choice process. Failing to consider professional respondents leads to deflated and less precise, i.e. wider distribution, conditional WTA estimates.

The rest of the paper is organized as follows: In Section 2 we provide a brief overview of the relevant strands of the literature dealing with professional respondents and "symptoms" of professional respondents; in Section 3 we introduce the modeling framework; in Section 4 we provide information on the case study and summary statistics of the relevant variables; Section 5 we present and discuss our results; and in Section 6 we conclude the paper, discuss the implications of our results for stated preference practitioners and provide a few avenues for future research.

2 Background

A professional respondent, or a frequent survey responder, is thought to be (i) motivated by economic incentives, (ii) a member of multiple panels, (iii) answering many surveys, and (iv) a long-term member of the panel (Baker et al., 2010; Whitsett, 2013; Hillygus et al., 2014; Matthijsse et al., 2015). The main concern with professional respondents is that answering multiple different surveys often may alter the way they respond (Whitsett, 2013). However, the question remains open exactly how it affects their response patterns and ultimately measures of interest (Hillygus et al., 2014; Matthijsse et al., 2015). Given the nature of these respondents - that they are motivated by the economic incentive - we would expect that they are drawn to the non-probability panels because it enables them to answer more surveys and earn more money. Indeed, the majority of research on professional respondents has been done in that context (Hillygus et al., 2014). The focus in this section, and indeed this paper, is on internet panels, and we refer the reader to Sturgis et al. (2009) for an overview of work done in the context of more traditional panels such as mail-out panels. Whitsett (2013) provides a review of 16 articles exploring the impact of professional survey respondents on data quality. He finds that the three most common definitions, or metrics, used to identify these respondents, are: i) the number of panels to which they belong, ii) the number of surveys they take per week/month/year, and iii) panel tenure, i.e. how long they have been members of any given panel. However, the results with respect to how professional respondents affect data quality remained inconclusive. Hillygus et al. (2014) use an opt-in panel and test hypotheses with respect to professional respondents along two dimensions: i) the number of panels they belong to, and ii) the number of surveys they have completed. They put forth the argument that professional respondents are more likely to be motivated by the economic incentive, whereas other respondents are more likely to be motivated by the topic of the survey. This is reflected in their results, which show that professional respondents knew less about the topic. The authors argue that this reduced the "bias" in their sample, which did comprise a large group of topic-knowledgeable respondents. Furthermore, contrary to expectation, they found that professional respondents spent more time on the survey, were less likely to engage in "straight-lining" behavior, and reported higher levels of effort. Matthijsse et al. (2015) studies professional respondents using a dataset from the Netherlands Online Panels Comparison study, which includes 19 panels and covers roughly 90% of the online market research respondents in the country. They find that roughly 17% of respondents can be classified as professional and that they are more likely to be members of multiple panels, answer more surveys, are more likely to be motivated by the economic incentive and have a slightly faster completion time (contrary to the findings of Hillygus et al. (2014)). Furthermore, they find that professional respondents are more likely to be women, unemployed, report worse health and lower life satisfaction. However, they find no differences in terms of age, nationality, household size, religion or education. Lastly, they find that these respondents are no more likely to engage in survey satisficing.

A related body of literature focuses on panel conditioned respondents. Unlike professional respondents - these are not recruited, but trained (Hillygus et al., 2014). A conditioned respondent is a member of a panel (in the true sense), where they receive the same (or very similar) survey at regular intervals. This allows respondents to think, reflect and refine their answers between waves of the survey (Toepoel et al., 2008). As such, this type of survey experience may affect responses because attitudes, knowledge and behavior may be influenced by previous exposure to the same survey or topic (Toepoel et al., 2009). In an early study using the Knowledge Network panel in the US, Dennis (2001) argues that panel conditioning does not necessarily pose a problem for research conducted on online samples because, while a respondent may answer a large number of surveys, the surveys are on diverse topics, which would help mitigate the panel conditioning effect. However, he does acknowledge the possibility that highly specialized samples, e.g. economics professors, might suffer from panel conditioning effects. Toepoel et al. (2008) and Toepoel et al. (2009) compare "fresh" and "trained" respondents using two Dutch internet panels. They find no difference between the two types of respondents in terms of how they processed the different design elements of the survey, e.g. number of items on the screen or different ranges of Likert scale questions. However, they find that panel conditioned respondents are more likely to engage in survey satisficing; take short-cuts and not pay full attention to each question; and complete the survey significantly faster. Interestingly, they find no significant difference in completion time for respondents when answering the key questions of the survey, in this case: attitudes, behaviors and facts, but that the difference in response time can be found on the "knowledge questions", i.e. those that require prior knowledge on the topic. This suggests that it is not survey experience in general (e.g. answering multiple surveys), but direct experience with those type of questions that drives the observed difference between respondents (Toepoel et al., 2009). A similar result was found by Das et al. (2007). Sturgis et al. (2009) put forth and test the cognitive stimulus hypothesis, which states that people who are repeatedly exposed to the same topic over multiple survey implementations deliberate more carefully, gather information and discuss with friends and family between different waves of the survey, and that this leads to more reliable and stable answers to attitudinal questions over time. Testing the hypothesis they find that this deliberation and refinement of answers is not a function of panel membership length, but rather how many times they have encountered a particular question. This is in line with the findings of e.g. Toepoel et al. (2009). Chandler et al. (2015) conduct a test-retest experiment to explore panel conditioning effects using Amazon's opt-in panel MTurk. In the first wave, respondents completed several tasks and experiments common in psychology, e.g. anchoring, gains versus loss framing, and low- versus high-category scales. Upon completing the first wave, respondents were randomly allocated to one of three conditions, where the difference was the time before being invited to the second wave, which was the same experiment as in the first wave. They found that the conditioning effect was quite strong and led to a 25 % reduction in overall effect size. They argue that this is likely caused by respondents deliberating more carefully between surveys and that it removes "intuitive" answers in the second wave, which can be more prone to error. This implies that larger samples, or weighting, might be necessary to detect actual effects if large parts of the sample are panel-conditioned.

It is certainly possible that the prevalence of professional respondents is lower in probability based internet panels, such as the one used in this paper, simply because opting to become a member is impossible. However, given that internet panels have been around for more than a decade, and some panels invite members to respond to surveys at a high rate, many respondents may have substantial experience answering surveys. In normal internet panels, panel conditioning is likely to play a minor role since respondents are surveyed on a wide range of topics and rarely respond to the same survey more than once. That said, given the relative rarity of stated preference surveys, we believe that substantial previous experience with such surveys is unlikely. Panel conditioning might play a role in test-retest studies, but exploring this in more detail is left for future research.

3 Theoretical Framework

3.1 Integrated Choice and Latent Variable (ICLV) Model

In this section we describe the integrated choice and latent variable (ICLV) model and its constituent parts. The ICLV model comprises three main components: i) a structural model, which describes the latent variable as a function of observable socio-demographic characteristics; ii) the choice model, which links the latent variable with the observed choice in each choice task; and iii) a measurement model, which explains observed indicators of professional respondents as a function of the latent variable. In Figure 1 we provide an overview of the modeling framework. The rectangles depict observable variables, the ellipses unobservable (latent) variables, and the solid and dashed arrows represent structural and measurement equations respectively (Ben-Akiva et al., 2002). The ICLV models have a relatively long history in transport research (see e.g. Fosgerau and Bjørner, 2006; Bolduc and Alvarez-Daziano, 2010; Hess and Rose, 2012; Hess and Stathopoulos, 2013), but only recently have they been introduced to other domains of economics (Hess and Beharry-Borg, 2012; Dekker et al., 2016; Czajkowski et al., 2017).

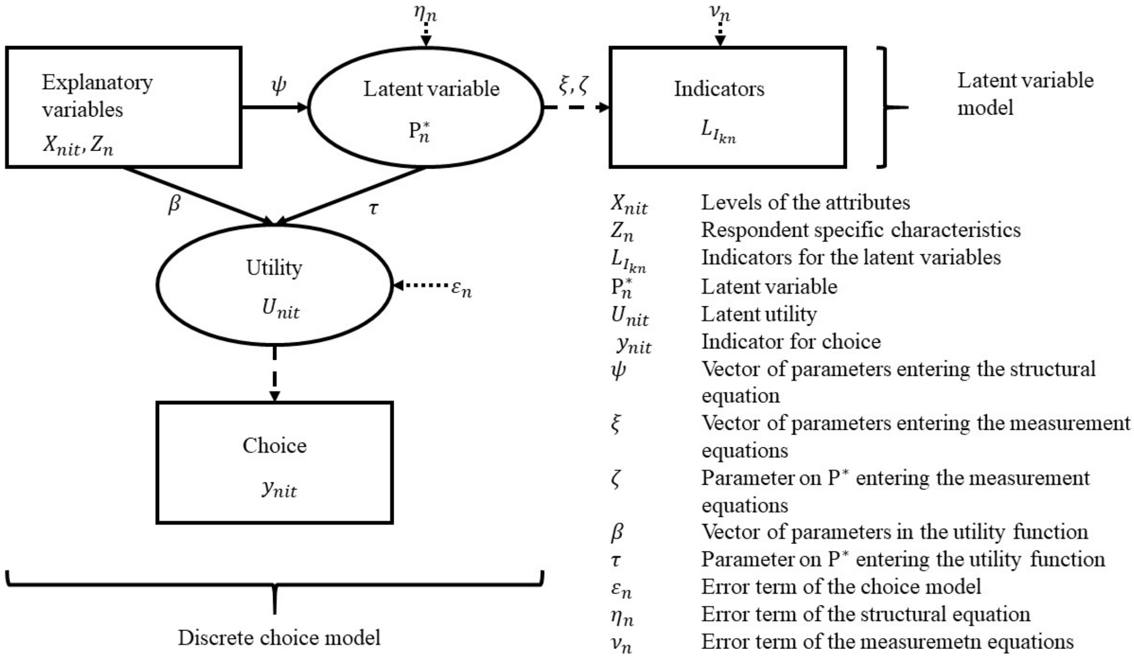


Figure 1 - The ICLV Framework (adapted from Ben-Akiva et al. (2002)

3.1.1 The structural model

We are interested in how professional respondents impact elicited preferences from a stated choice experiment. Since we cannot observe whether a given respondent is professional, we treat this as a latent variable. In Equation 1 we let a respondent's professionalism P_n^* be a linear-in-the-parameters additive function of observable characteristics Z_n, where ψ is a vector of parameters to be estimated and gives the marginal effect of each Z_n. The literature review revealed that whether or not a respondent was identified as professional varied across socio-demographic characteristics. η_n is a normally distributed error term with zero mean and unit standard deviation. This normalization is necessary for identification (Hess and Beharry-Borg, 2012). Specifying the latent variable in this manner implies that it is continuous and normally distributed. We can interpret the latent variable P_n^* as the degree of professionalism (ordinal scale), i.e. a higher value of P_n^* implies a more professional respondent.

$$P_n^* = \psi Z_n + \eta_n \tag{1}$$

3.1.2 The choice model

The analysis of stated preference data is based on the assumption that people maximize utility, and that their choices can be described by a random utility model (McFadden, 1974). The actual utility that an individual n experiences from choosing alternative i in any choice situation t is unobserved (latent) and can be written using a linear in the parameters and additive structural equation (Equation 2a), where β is a vector of parameters, X_{nit} a vector of attributes and ε_{nit} an error-term. While we cannot observe the level of utility directly, we do observe which alternative the respondent chose y_{nit}. We can then use Equation 2b to link unobserved utility to observed choice, where i and j indexes the chosen and non-chosen alternatives respectively, and C_{nt} is the complete set of available alternatives for individual n in choice situation t. Different assumptions about the error-term ε_{nit} lead to different choice models such as e.g. the multinomial logit and the mixed logit.

$$U_{nit} = \beta X_{nit} + \varepsilon_{nit} \tag{2a}$$

$$y_{nit} = \begin{cases} 1 & U_{nit} > U_{njt}, \forall j \in C_{nt}, j \neq i \\ 0 & \text{otherwise} \end{cases} \tag{2b}$$

We assume that the error-term ε_{nit} is *i.i.d.* type I extreme value distributed (Gumbel) with variance $\frac{\pi^2 s^2}{6}$. Then, the probability that individual n chooses alternative i in choice situation t can be described by the multinomial logit model in Equation 3 (McFadden, 1974).

$$\Pr(y_{nit}|\beta, X_{nit}) = \prod_{t=1}^{T} \frac{\exp(\lambda_n\left[\beta_0 \mathrm{ASC}_{nit} + \beta X_{nit}\right])}{\sum_{j \in C} \exp(\lambda_n\left[\beta_0 \mathrm{ASC}_{nit} + \beta X_{njt}\right])} \tag{3}$$

where λ_n is a scale parameter that is inversely related to the error-term, ASC_{nit} an alternative-specific constant (ASC) for alternative i, and β and X_{nit} as defined above. In most applications the scale parameter is assumed constant and equal to unity. However, in the present paper we combine two datasets. In that case we need to explicitly consider the possibility that unobserved factors vary between them, i.e. that they have different error variances and by extension, different scale parameters. We accommodate this in a straight forward manner by estimating a relative scale parameter (Swait and Louviere, 1993). This implies setting the scale of one dataset to unity and let the scale parameter for the other be freely estimated. Specifically, we specify λ to be: $[\lambda_b \mathrm{d}_b + \lambda_g \mathrm{d}_g] / \lambda_b$, where λ_b is set to unity for identification; d_b and d_g are indicators for whether a respondent belongs to the "brown" or "green" dataset; and λ_g is the relative scale parameter for the "green" dataset. We clarify this distinction in section 4. This specification ensures that scale is positive and can be interpreted as scale relative to unity.

The latent variable for professionalism is assumed to follow a normal distribution, and when included in a multinomial logit model, it is the only source of heterogeneity between respondents. Such a specification runs the risk of confounding heterogeneity arising from professionalism and that from preferences. To reduce this confounding, we allow for heterogeneous preferences by assuming that the preference parameters follow pre-specified distributions. Let Θ be a vector of random parameters, including the alternative specific constant for the status quo alternative, and Ω denote the moments of the parameter distributions. Then we can denote the joint density of the parameters β by $f(\Theta|\Omega)$. This then leads to the unconditional mixed logit probability in Equation 4:

$$\Pr(y_{nit}|\Omega, X_{nit}) = \int \prod_{t=1}^{T} \frac{\exp(\lambda_n [\beta_0 \text{ASC}_{nit} + \beta X_{nit}])}{\sum_{j \in C} \exp(\lambda_n [\beta_0 \text{ASC}_{nit} + \beta X_{njt}])} f(\Theta|\Omega) d(\Theta) \tag{4}$$

This integral does not have a closed form solution, but is approximated through simulation. To test our hypotheses related to professional respondents, we interact the latent variables with the scale and the mean of the distribution for the alternative specific constant respectively. Since P_n^* is normally distributed with zero mean, it means that when we take the exponent scale is log-normally distributed. This specification ensures that scale is positive. We are able to separately identify the impact of professional respondents on scale because it also enters the measurement equations[3] (defined below)(Hess and Rose, 2012; Hess and

[3]We do acknowledge that we cannot be sure that what we are detecting is scale and not just correlation across distributions, but for convenience - we refer to the effect as a scale-effect. To reduce the possibility that our scale simply picks up correlation across distributions, we specify a full correlation matrix in our mixed logit model.

Stathopoulos, 2013; Hess and Train, 2017). We show the full specification in Equation 5.

$$\Pr(y_{nit}|\Omega, X_{nit}) = \int \prod_{t=1}^{T} \frac{\exp(\lambda_n \exp(\tau P_n^*) \left[(\tau P_n^* + \beta_0)\text{ASC}_{nit} + \beta X_{nit}\right])}{\sum_{j \in C} \exp(\lambda_n \exp(\tau P_n^*) \left[(\tau P_n^* + \beta_0)\text{ASC}_{nit} + \beta X_{njt}\right])} f(\Theta|\Omega) d(\Theta) \tag{5}$$

3.1.3 The measurement equation for the latent variable

To help identify the latent variable P_n^* we link it to several observable indicators hypothesized to be associated with professionalism, such as panel tenure, survey frequency and completion rates. To identify suitable indicators we started with those already used in the literature reviewed in section 2. To refine our choice of indicators and to increase our certainty that they correlate to the same latent constructs, we conducted an exploratory factor analysis (see section 5). All of our measurements are continuous variables and we estimate the probability of the outcome using the normal density as depicted in Equation 6:

$$L_{I_{kn}|P_n^*} = \frac{1}{\sigma_{I_k}\sqrt{2\pi}} e^{-\frac{(I_{kn} - \bar{I_{kn}} - \zeta_k P_n^*)^2}{2\sigma_{I_k}^2}} \tag{6}$$

where I_{kn} is the indicator and $\bar{I_{kn}}$ the mean of the indicator, ζ_k the effect of the latent variable on the k^{th} indicator and P_n^* the latent variable. Notice that we have centered the indicator on zero, i.e. zero mean, and estimate the empirical standard deviation σ_{I_k} (Hess and Beharry-Borg, 2012; Hess and Stathopoulos, 2013).

3.2 The joint likelihood function

The structural-, choice- and measurement models are jointly estimated for efficiency reasons (Daly et al., 2012) and the corresponding likelihood function is defined as

$$\text{L}(\beta, \psi, \xi, \zeta, \sigma, \gamma) = \int \int \prod_{k=1}^{K} \Pr(i_{nit}|\beta, X_{nit}, P_n^*) L_{I_{nk}|P_n^*} f(P_n^*|\psi) d(P_n^*) f(\Theta|\Omega) d(\Theta). \tag{7}$$

We take the product over the likelihood of each measurement equation k and integrate over the latent variable P_n^* to obtain the joint likelihood function linking the latent variable and the choice model. This function does not have a closed form solution, but is approximated through simulation.

3.3 Model estimation

All models were coded in R (R Core Team, 2016) and we used 2500 MLHS draws (Hess et al., 2006) per individual to simulate the distributions of preferences and latent variables. We allowed the preference distributions to be correlated, and let the compensation be log-normally distributed while all non-monetary attributes were normally distributed. To test for local optima we generated a large number of starting values at random and ran several of the best fitting models to completion.

4 Empirical Case Study and Data

To test our hypotheses we use data on hypothetical electricity contract choice published in Broberg et al. (2017). The purpose of the survey was to investigate the potential demand response through behavioral change by exploring people's willingness-to-accept compensation to opt in to a demand-side management program characterized by partial and temporary load control. Respondents were asked to choose between different hypothetical contracts that differed in (i) maximum load allowed, (ii) duration and frequency of the load constraint, (iii) whether the households were free to decide how to adapt to the load control in terms of appliance choice, and (iv) a compensation for accepting the contract. A sample choice task is shown in Figure 2. The five attributes and the corresponding levels were combined into 16 choice tasks using an efficient design based on minimizing the D_b-error and generated in Ngene. We conducted 2 pilot studies with 100 respondents to obtain more precise priors to use in updating the design. The 16 choice tasks were blocked into 2 blocks of 8 choice tasks. The order of the choice tasks was randomized between respondents.

The study was administered in July and August of 2017. Respondents were randomly recruited, but restricted to be respondents residing in one- and two dwelling buildings. Recruited respondents were then randomly allocated into a "green" treatment group or a "brown" control group. The difference was that respondents in the "green" treatment were exposed to a simple "green cheap talk script" focusing on climate change prior to answering the choice tasks. In total, 2014 completed surveys were obtained equally split between the two treatments. In addition to the choice question part, the questionnaire contained questions related to socioeconomic variables and behavioral aspects related to the electricity market and energy consumption. For more details on the survey and data see Broberg et al. (2017).

Which of the following A, B or C contracts would you choose if offered to you? Unless otherwise stated in the agreement, everything else works as today, for example, the electricity price you pay and how often it changes.

	Contract A	Contract B	Contract C – as today
Load control	5000 watt	3500 watt	As today
Choice of appliances	Pre-determined given the load	Flexible given the load	As today
Duration	4.30pm-7.30pm	5pm-6.30pm	-
Number of days	5 days	20 days	-
Compensation	2500	750	-
My choice	[ ]	[ ]	[ ]

Figure 2 - Sample choice card

5 Results

5.1 Data

In Table 1 we provide descriptive information on the variables entering our models. The top part of Table 1 shows the variables entering the structural equations, the middle part those entering the measurement equations, and the bottom part those that are only part of the exploratory factor analysis. To create indicators we used data on panel tenure, and the number of invited, started and completed surveys. The variables of particular interest are those entering the measurement equations, i.e. those we believe are indicators of "professionalism". We see that average panel tenure is 5.73 years, with a standard deviation of 3.55, which indicates that some respondents have considerable experience as panel members, but that there is large variation in the sample. To disentangle panel tenure from activity, we construct an average survey per month measure. On average, panel members answer 3.56 surveys per month, with a standard deviation of 2.24. We also include measures of survey completion rates and number of completed surveys. In addition to answering several surveys per month, it is possible that professional respondents spend less (or more) time on the choice tasks. On average, respondents spent just over 2 minutes answering all 8 choice tasks, but there is substantial variation. Lastly, we use a self-reported effort measure elicited on a Likert scale. The majority of our respondents reported medium to high levels of effort.

Table 1 - Summary statistics of select variables entering the models

Variable	Type	Mean	SD	Min	Max
Male	Indicator	0.52	0.50	0.00	1.00
Higher education	Indicator	0.52	0.50	0.00	1.00
Age 30 - 39	Dummy	0.12	0.32	0.00	1.00
Age 40 - 49	Dummy	0.17	0.37	0.00	1.00
Age 50 +	Dummy	0.60	0.49	0.00	1.00
Household income (median)	Indicator	0.54	0.50	0.00	1.00
Permanent employed	Indicator	0.51	0.50	0.00	1.00
Total number of people in household	Continuous	2.68	1.25	1.00	12.00
Avg. nr. of surveys per month	Continuous	3.56	2.24	0.20	16.33
Survey completion rate	Continuous	0.54	0.15	0.05	1.00
Panel tenure (years)	Continuous	5.73	3.55	0.01	12.19
Nr. of completed surveys	Continuous	199.29	140.81	1.00	823.00
Effort	Categorical	3.37	1.10	1.00	5.00
Total time on choice tasks†	Continuous	2.11	1.07	0.15	3.58
Total time on survey†	Continuous	28.79	16.11	3.00	56.00

† Winsorized at the 80th percentile

5.2 Exploratory factor analysis

To understand how the different measures of professionalism correlates and thereby feed into the measurement model, we conducted an exploratory factor analysis. In Table 2, we show the results of a model with four factors using principal axis factoring with promax rotation. The purpose of an exploratory factory analysis is to understand which variables correlate with which unobserved latent constructs, i.e. finding patterns between the variables in our data. A larger absolute factor loading indicates a stronger correlation with the latent construct. Hoyos et al. (2015) suggests using an exploratory factor analysis to understand better which variables makes for good candidates to use in the measurement equations. To ease interpretation of the factor analysis we use an axis rotation, here: promax rotation. The benefit of using an oblique rotation, such as promax, is that it allows for correlations between the factors, which we return to below. The exploratory factor analysis suggests two types of professional respondents, a) those that answer a large number of surveys per month and have a relatively high survey completion rate, i.e. "hyperactives" (Baker et al., 2010), and b) those that have a lot of experience measured by the number of years they have been members of the panel and the total number of surveys they have answered, i.e. "experienced". These variables are clearly correlated with different latent constructs. Moreover, it suggests that the time you spend on the survey and choice tasks, and the self-reported effort, correlates with other latent constructs than the two "professional" latent variables. That said, we do not focus on effort or speeding in this paper[4].

[4]Such behavior has been extensively explored in other studies (see e.g. Hess and Stathopoulos, 2013; Börger, 2016; Dekker et al., 2016; Campbell et al., 2017, and references therein)

Table 2 - Exploratory factor analysis

Variable	Factor 1	Factor 2	Factor 3	Factor 4
	“Hyperactives”	“Experienced”	“Speeders”	“Effort”
Effort	0.10		0.17	0.61
Panel tenure (years)	−0.43	0.85		
Survey completion rate	0.65			−0.11
Number of completed surveys	0.34	0.84		
Avg. number of surveys per month	0.90	−0.11		
Total time on choice tasks (minutes)†			0.72	0.20
Total time on survey (minutes)†			0.49	−0.11
Proportional variance	0.22	0.21	0.12	0.07
Cumulative Variance	0.22	0.43	0.55	0.62

All factor loadings less than 0.1 are removed
† Winsorized at the 80th percentile

Table 3 - Exploratory factor analysis

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1.00	0.03	−0.11	0.02
Factor 2	0.03	1.00	−0.10	0.08
Factor 3	−0.11	−0.10	1.00	0.25
Factor 4	0.02	0.08	0.25	1.00

In Table 3 we show the correlation between the four factors identified above. As is evident by the off-diagonal elements, all four factors are practically uncorrelated. This also suggests that effort or speeding is not a function of professionalism as defined by the panel meta data variables used in this paper. Given this, and our interest in professional respondents, we decided to include two latent variables in our integrated choice and latent variable model: 1) LV1 - “hyperactives” and 2) LV2 - “experienced”. In the next section, we explore how these two types of professional respondents affect elicited preferences.

5.3 The ICLV model - the structural equations

In Table 4 we show the results of the structural equations that link observable socio-demographic characteristics with the latent variables of professionalism. We chose the variables to include in the model based on what has been used in the literature previously (see section 2). ’Male’ is a gender dummy, ’higher education’ is a dummy equal to one if a respondent has university level education, the age dummies are relative to the omitted category ’18-29’, ’permanent employment’ is a dummy equal to one if a respondent is permanently employed, ’household income - median’ is a dummy equal to one if a respondent is a member of a household with above median income, and ’number of people in the household’ is a count variable. Since the structural equation is additive and linear-in-the-parameters, the interpretation of the parameters is straight forward. We see that "hyperactive" respondents are more likely to be younger men with permanent employment living in larger households, whereas "experienced" respondents are more likely to be older men with permanent employment living in smaller households. This characterization of respondents stands in contrast to those reported by Matthijsse et al. (2015), who found that "professional" respondents were more likely unemployed women.

Table 4 - Structural equations

	LV1 - "Hyperactive"		LV2 - "Experienced"	
	Est.	s.e.	Est.	s.e.
ψ_{Male}	0.4435***	0.0559	0.3171***	0.0513
$\psi_{\text{Higher education}}$	−0.0727	0.0526	−0.0016	0.0530
$\psi_{\text{Age 30-39}}$	−0.1708	0.1247	−0.0814	0.1107
$\psi_{\text{Age 40-49}}$	−0.5064**	0.1197	0.3955**	0.1006
$\psi_{\text{Age 50+}}$	−1.2126***	0.0662	−0.1716***	0.0679
$\psi_{\text{Permanent employment}}$	0.6066***	0.0669	0.5379***	0.0680
$\psi_{\text{Household income - median}}$	0.0887	0.0614	−0.0422	0.0559
$\psi_{\text{Number of people in household}}$	0.1089***	0.0234	−0.1120***	0.0193

'***' - 1 % level, '**' - 5 % level, '*' - 10 % level
Adjusted robust standard errors

5.4 The ICLV model - the measurement equations

Given our interest in professional respondents, we use the indicators identified in the literature as pertaining to these respondents (see e.g. Hillygus et al., 2014). Our exploratory factor analysis revealed that these variables correlated with two different latent constructs, which we labeled "hyperactive" and "experienced". The chosen indicators for "hyperactiveness" were average number of surveys per month and survey completion rate, while the chosen indicators for "experience" were panel tenure and total number of completed surveys. These indicators enter the measurement models for the related latent variable. Interestingly, time spent on the survey and choice tasks, and self reported effort correlated with different latent variables. As such neither would be good to include in the measurement models. Given our interest in professional respondents, and that speeding and effort has been studied in a discrete choice experiment context before (see e.g. Hess and Stathopoulos, 2013; Börger, 2016; Dekker et al., 2016; Campbell et al., 2017, and references therein), we do not pursue investigations in that direction in the present paper.

The top half of Table 5 shows the estimated empirical standard deviations, and the bottom half shows the marginal effect of the latent variable on the observed indicator. Conforming to expectations, we see that "hyperactive" respondents answer significantly more surveys per month and have a higher completion rate, and that "experienced" respondents have longer panel tenure and a higher total number of completed surveys. Taken together with the exploratory factor analysis, this provides a strong indication that our latent variables are indeed capturing the hypothesized behavior.

Table 5 - Measurement equations

	LV1 - "Hyperactive"		LV2 - "Experienced"	
	Est.	s.e.	Est.	s.e.
$\xi_{\text{Avg survey per month}}$	0.8972***	0.0391		
$\xi_{\text{Survey completion rate}}$	0.1209***	0.0021		
$\xi_{\text{Panel tenure (years)}}$			2.8534***	0.0513
$\xi_{\text{Nr of completed surveys}}$			0.0554***	0.0039
$\zeta_{\text{Avg survey per month}}$	1.6874***	0.0471		
$\zeta_{\text{Survey completion rate}}$	0.0696***	0.0031		
$\zeta_{\text{Panel tenure (years)}}$			2.0010***	0.0837
$\zeta_{\text{Nr of completed surveys}}$			0.1227***	0.0033

'***' - 1 % level, '**' - 5 % level, '*' - 10 % level
Adjusted robust standard errors

5.5 The choice model

In Table 6 we report the results from a mixed logit model and the integrated choice and latent variable model. We report the coefficients of the lower Cholesky matrix and correlation matrix in Table A1 and Table A2. We see that, on average, utility is increasing in the level of compensation and decreasing with increased restrictions. We see that a looser restriction is preferred to a stricter one, a shorter duration to a longer and that fewer days are preferred to more days affected. We see that the alternative specific constant for the 'status quo' alternative is positive and significant, which indicates that on average, people are unwilling to enter into contracts that would mean a restriction on their use of appliances. Finally, we note that the relative scale parameter is not different from unity.

Table 6 - The mixed logit model and the integrated choice and latent variable model

	MIXL				ICLV - MIXL			
	Mean		SD		Mean		SD	
	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.
Compensation	1.0777***	0.0576	1.1987***	0.0643	−0.3022***	0.0785	1.1674***	0.0967
Max 3500	−0.2306***	0.0572	0.9937***	0.1248	−0.1655***	0.0594	0.8981***	0.1395
Max 2000	−0.9897***	0.0889	1.6442***	0.1137	−0.9754***	0.0950	1.5659***	0.1024
Flexible use	0.1604***	0.0486	0.3298***	0.1614	0.1438**	0.0598	0.5876***	0.1059
Duration 90 mins	−0.2552***	0.0690	0.7508***	0.1345	−0.2194***	0.0716	0.6814***	0.0837
Duration 180 mins	−0.9800***	0.0959	1.6788***	0.1448	−0.9819***	0.1026	1.4818***	0.1161
Days 10	−0.5576***	0.0630	0.6842***	0.2099	−0.5780***	0.0648	0.4908***	0.1461
Days 20	−0.8534***	0.0856	1.0425***	0.1412	−0.9134***	0.0930	1.0064***	0.1964
ASC	0.6330***	0.1507	5.4023***	0.5073	0.9403***	0.1544	5.5106***	0.1612
Scale relative to brown	0.9501	0.0558			0.9652	0.0630		
τ_{LV1} – SQ					−0.2137*	0.1098		
τ_{LV2} – SQ					0.1392	0.1308		
τ_{LV1} – Scale					−0.1176***	0.0386		
τ_{LV2} – Scale					0.1230***	0.0391		
LL	-11019.1				-17584.990			
AIC	22148.2				35335.979			
BIC	22571.0				35974.027			
K	55				83			
N	16112				16112			

'***' - 1 % level, '**' - 5 % level, '*' - 10 % level
Adjusted robust standard errors
The relative scale is not significantly different from unity

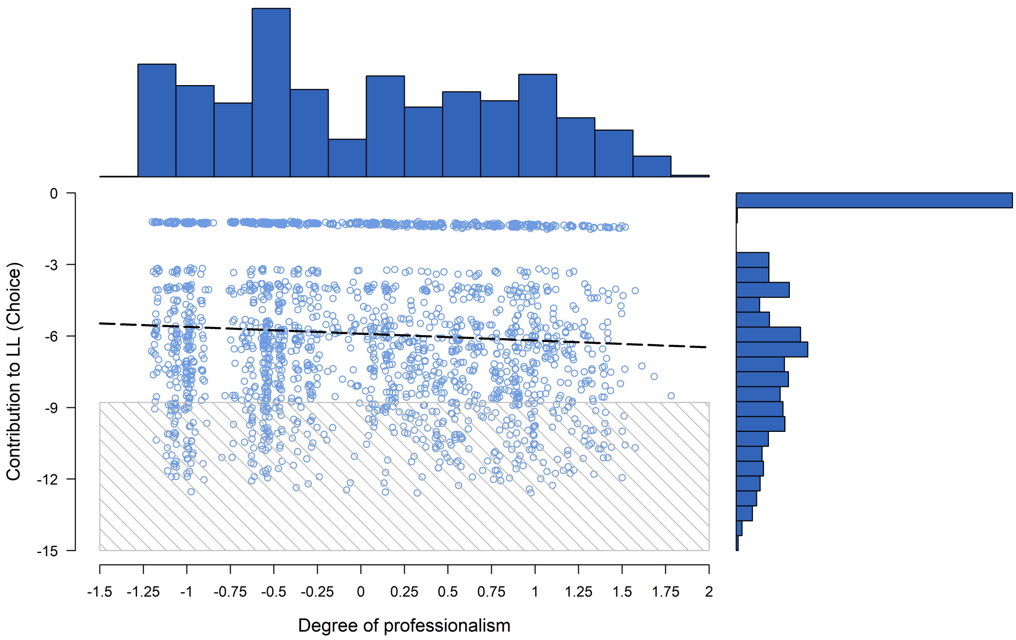
Next, we take a look at the results from the integrated choice and latent variable model. We note that the parameter distributions are relatively similar, except for a shift in the marginal utility of money - a point which we return to in detail below, but remark that direct comparison is not appropriate since the models are subject to different scaling. First, let us consider the first latent variable, i.e. the "hyperactive" respondents. We see from Table 6 that the estimate on the interaction with the mean of the preference distribution for the 'status quo' alternative is negative and significant at the 10 percent level. This suggests that "hyperactive" respondents are less likely to choose the 'status quo' alternative. If we look at the effect of the latent variable on scale, we see that the more "hyperactive" a respondent is, i.e. the higher the value of the latent variable, the smaller is the scale parameter, and the more stochastic the choice process appears from the econometrician's point of view. This result is consistent with fewer 'status quo' choices. It is possible that these respondents want to give the impression that they are making choices and trade-offs rather than always choosing the 'status quo'.

In Figure 3a we have plotted the value of the latent variable against a respondent's contribution to the log-likelihood function of the choice model. A higher absolute log-likelihood value means a poorer fitting model. By extension, an individual's contribution to the log-likelihood is an indication of how well the model explains his or her choices. What we see is that as the value of the latent variable increases, so does the contribution to the log-likelihood function, i.e. the model is worse at explaining these respondents' choices. This is consistent with what the parameter estimates showed: "hyperactive" respondents have a more stochastic choice process as seen from the econometrician's point of view [5].

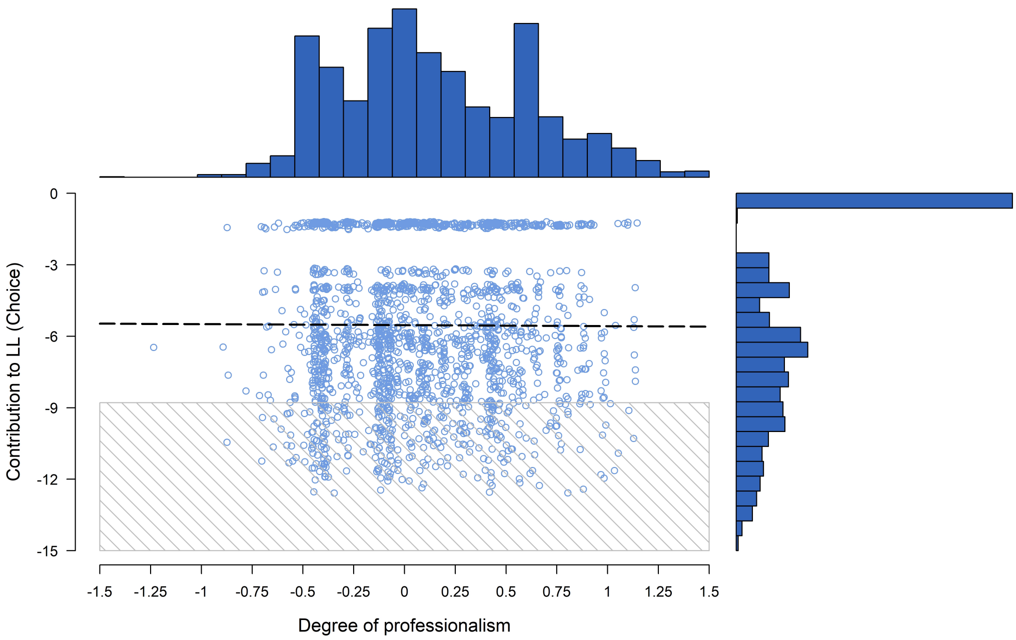
Now, let us take closer look at the professional respondents classified as "experienced". From Table 6 we see that they are more likely to choose the 'status quo' alternative, but not significantly so as evident by the parameter on the interaction with the mean of the preference distribution. However, the positive and significant term for the scale expression indicates that they have a more deterministic choice process as seen from the econometrician's point of view. Taking a look at Figure 3b we see that there is no discernible relationship between the latent variable and the contribution of the log-likelihood value of the choice model, which suggests that the model predicts these choices equally well.

Taken together, these results suggest that the type of professional respondent that could be problematic for the estimation of preferences is the one that has a high survey completion rate and frequency of answering. The total experience of answering surveys measured as the total number of survey and panel tenure appear to not have a detrimental impact on elicited preferences. This is a potentially important result with respect to the type of respondents we sample in internet panels, and interesting in the sense

[5] There is also a set of respondents for which the model explains choices really well, these respondents are almost exclusively those that consistently chose the 'status quo' alternative.



(a) Latent variable 1 – "Hyperactive"



(b) Latent variable 2 – "Experienced"

Figure 3 - The latent variable plotted against the contribution to the log likelihood value of the choice model

that there appeared to be no correlation between these respondents and response time or self-reported effort. We believe this is an interesting avenue for future research to explore.

5.6 Welfare estimates

In Table 7, we report the conditional willingness-to-accept (WTA) distributions broken down by model and attributes. The distributions are Winsorized at the 2.5^{th} and 97.5^{th} to remove the influence of outlying WTA estimates on the mean (Tukey, 1962). In practice, any value larger (smaller) than the 97.5^{th} (2.5^{th}) percentile is set to the value of the 97.5^{th} (2.5^{th}) percentile. As such, we retain all individuals in the following and report the max and the min of the Winsorized distributions. As expected, people require larger compensations for more severe restrictions. For example, the mean (of the conditional means) WTA is SEK 209.00 to accept a restriction of 2000 kWh compared to SEK 75.40 for 3500 kWh. While willingness-to-accept is lower for the remaining restrictions, compensation needs to be higher for the more severe ones. We notice that the median WTA is relatively close to zero. This is likely connected to the large number of 'status quo choices' and respondents always choosing the SQ. Considering the WTA estimates from the ICLC-MIXL model, we see the same pattern in WTA across attributes, but we note that they are substantially larger, although the means are still within the ranges of the compensation attribute. This upward shift in mean WTA is likely caused by the change in estimated compensation sensitivity between models. When we consider how the degree of professionalism affects the mean utility of choosing the 'status quo' alternative we are explaining a larger degree of the choice between alternatives as the function of this alternative. In the ICLV-MIXL model we observe that people, on average, are less sensitive to changes in compensation. Consider a 'hyperactive' respondent. He is unlikely to choose the 'status quo' alternative as evident by the negative and significant interaction term. Under the MIXL model, this behavior was captured by a combination of the estimated compensation sensitivity and the utility of choosing the 'status quo' alternative. Under the ICLV-MIXL model, this behavior is to a much larger extent captured by the interaction term with the 'status quo' alternative, which results in a downward shift in compensation sensitivity. Said another way, if you are more likely to choose either of the hypothetical contracts at a larger range of offered compensation levels, your compensation sensitivity is likely lower. This suggests a possible confounding between the 'status quo' alternative and estimated compensation sensitivity, and furthermore that professional "hyperactive" respondents, if their somewhat erratic choice behavior is not explicitly considered, we might under-estimate WTA. Another interesting point concerns the spread of the distribution of conditional means. What we see is that once we consider the behavior of the two types of professional respondents, the WTA distributions become much tighter, which indicates more precise WTA estimates. This is likely a function of the somewhat larger coefficients of variation, i.e. mean divided by standard deviation, for the non-compensation attributes in the ICLV-MIXL model.

Table 7 - Distribution of mean conditional estimates

Model	Attribute	Min	Median	Mean	Max
MIXL	Max 3500	−2410.00	0.04	75.40	3020.00
MIXL	Max 2000	−4770.00	1.03	209.00	5770.00
MIXL	Flexible use	−1140.00	−0.06	−30.90	954.00
MIXL	Duration 90 mins	−1590.00	0.61	57.80	1940.00
MIXL	Duration 180 mins	−3370.00	2.46	164.00	4470.00
MIXL	Days 10	−1100.00	2.32	135.00	1910.00
MIXL	Days 20	−2360.00	2.38	196.00	3960.00
MIXL	ASC - SQ	−18800.00	0.10	−531.00	18400.00
ICLV MIXL	Max 3500	−61.40	0.25	146.00	744.00
ICLV MIXL	Max 2000	−69.00	4.16	831.00	3370.00
ICLV MIXL	Flexible use	−1010.00	−0.92	−240.00	12.10
ICLV MIXL	Duration 90 mins	−212.00	0.02	−10.20	120.00
ICLV MIXL	Duration 180 mins	−12.10	9.45	318.00	1440.00
ICLV MIXL	Days 10	−3.85	3.49	372.00	1420.00
ICLV MIXL	Days 20	−14.70	4.39	553.00	2190.00
ICLV MIXL	ASC - SQ	−21900.00	−0.17	−5590.00	333.00

The distributions are Winsorized at the 0.025^{th} and 0.975^{th} percentile because of large outliers.

6 Conclusion

In this paper, we are concerned with a group of respondents in internet panels that can be classified as professional. It is believed that these respondents are motivated by the monetary incentives and as a consequence, they are more likely to speed, use simplifying strategies and spend less effort. Such behavior draws into question the validity of responses provided by professional respondents and subsequently the reliability of elicited preferences and welfare estimates used to inform policy. We identify professional respondents using actual measures of panel tenure, total number of surveys, survey frequency and completion rates. An exploratory factor analysis revealed that there are two types of professional respondents: i) “hyperactives” who has high completion rates and answer surveys frequently, and ii) “experienced” who have long panel tenure and have answered a large number of surveys. Interestingly and importantly, these two variables are unrelated to time spent on the choice tasks and survey, and self-reported levels of effort, as well as the fact that the latter two are independent of each other. It suggests that the results reported in this paper should be seen as additional to results hypothesizing connections between professionalism, speeding and effort. Furthermore, future research should focus on the connection between the three and to what extent they affect survey response quality.

We explore the impact of “hyperactive” and “experienced” respondents using an integrated choice and latent variable model, where the degree of hyperactiveness and experience is treated as a latent variable and identified by a set of structural and measurement equations. Our results show that “hyperactive” respondents are less likely to choose the ‘status quo’ alternative and have a more stochastic choice process as seen from the practitioner’s point of view. A potential explanation for this behavior could be an “afraid to be kicked-out” effect, meaning that “hyperactive” respondents try to behave good by (i) not choosing the status quo alternative and (ii) not “straight lining”, i.e. randomize the choice to a larger extent. In all cases, our result translates into a downward shift in the marginal compensation sensitivity. Furthermore, the model does worse in describing choices as evident by the contribution to the log-likelihood function. “Experienced” respondents on the other hand are more likely to choose the ‘status quo’ alternative, but not significantly so, and have a more deterministic choice process as seen from the practitioner’s point of view. For these respondents there is no connection between the level of experience and contribution to the log-likelihood function, which suggests that the model does equally well (poorly) at describing these choices. Taken together this implies that professional respondents can be split into two groups, but that only one of these – the “hyperactives” – are problematic for preference elicitation. Turning our attention to estimated willingness-to-accept, we find that the model failing to take professional respondents into account under-estimates WTA. This is most likely caused by the downward shift in the distribution of the marginal utility of income. Taking into account the influence of professional respondents leads to an upward shift in WTA, but the narrower range of the distribution of conditional means suggests that these WTA estimates are more precise. It is possible that future research should look into the impacts of addressing these respondents at the sampling stage of the survey. Should we sample to avoid them?

Going forward, we believe that further research with respect to how professional respondents affect preferences is warranted. Our study is limited in that we do not know they type of surveys respondents have answered previously. It could be that the overall effect of experience answering surveys is different from the specific effect of answering stated preference surveys. Exploring this issue would likely blur the lines between professional and panel conditioned respondents. A first step to exploring this could be to conduct a test-retest study where the sample in the retest study is split between the original and the new sample. Any difference can be attributed to panel conditioning. Furthermore, if these samples are stratified based on general survey experience and experience with stated preference surveys in particular, we would be able to disentangle the two effects. This would also allow researchers to test the cognitive stimulus hypothesis (Sturgis et al., 2009) in the context of a stated preference survey.

References

Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., Dillman, D., Frankel, M. R., Garland, P., Groves, R. M., Kennedy, C., Krosnick, J. A., Lavrakas, P. J., Lee, S., Link, M., Piekarski, L., Rao, K., Thomas, R. K., Zahs, D., 2010. AAPOR Report on online panels. The Public Opinion Quarterly 74 (4), 711–781.

Baker, R., Downes-Le Guin, T., 2007. Separating the wheat from the chaff: Ensuring data quality in internet samples. In: Proceedings of the Fifth International Conference of the Association for Survey Computing: The Challenges of a Changing World. pp. 157–166.

Ben-Akiva, M., Walker, J., Bernardino, A. T., Gopinath, D., Morikawa, T., Polydoropoulou, A., 2002. Integration of choice and latent variable models. In: Mahmassani, H. S. (Ed.), In perpetual motion: Travel behavior, research opportunities and application challenges. Elsevier, Amsterdam, pp. 431–470.

Bolduc, D., Alvarez-Daziano, R., 2010. On estimation of hybrid choice models. In: Hess, S., Daly, A. (Eds.), Choice Modelling: The State-of-the-Art and the State-of-Practice. Emerald Group Publishing Limited, pp. 259–287.

Börger, T., 2016. Are fast responses more random? Testing the effect of response time on scale in an online choice experiment. Environmental and Resource Economics 65 (2), 389–413.

Broberg, T., Brannlund, R., Persson, L., 2017. Consumer preferences and soft load control on the Swedish electricity market.
URL http://www.cere.se/documents/wp/2017/CERE{_}WP2017-9.pdf

Campbell, D., Mørkbak, M. R., Olsen, S. B., 2017. Response time in online stated choice experiments: the non-triviality of identifying fast and slow respondents. Journal of Environmental Economics and Policy 6 (1), 17–35.

Chandler, J., Paolacci, G., Peer, E., Mueller, P., Ratliff, K. A., 2015. Using nonnaive participants can reduce effect sizes. Psychological Science 26 (7), 1131–1139.

Czajkowski, M., Vossler, C. A., Budziński, W., Wiśniewska, A., Zawojska, E., 2017. Addressing empirical challenges related to the incentive compatibility of stated preferences methods. Journal of Economic Behavior & Organization 142, 47–63.

Daly, A., Hess, S., Patruni, B., Potoglou, D., Rohr, C., 2012. Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. Transportation 39 (2), 267–297.

Das, M., Toepoel, V., van Soest, A. H. O., 2007. Can I Use a Panel? Panel Conditioning and Attrition Bias in Panel Surveys.
URL http://www.ssrn.com/abstract=1012252

Dekker, T., Hess, S., Brouwer, R., Hofkes, M., 2016. Decision uncertainty in multi-attribute stated preference studies. Resource and Energy Economics 43, 57–73.

Dennis, J. M., 2001. Are internet panels creating professional respondents? Marketing Research 13 (2), 34–38.

Fosgerau, M., Bjørner, T. B., 2006. Joint models for noise annoyance and willingness to pay for road noise reduction. Transportation Research Part B: Methodological 40 (2), 164–178.

Gao, Z., House, L. A., Xie, J., 2016. Online Survey Data Quality and Its Implication for Willingness-to-Pay: A Cross-Country Comparison. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 64 (2), 199–221.

Hess, S., Beharry-Borg, N., 2012. Accounting for Latent Attitudes in Willingness-to-Pay Studies: The Case of Coastal Water Quality Improvements in Tobago. Environmental and Resource Economics 52 (1), 109–131.

Hess, S., Rose, J. M., 2012. Can scale and coefficient heterogeneity be separated in random coefficients models? Transportation 39 (6), 1225–1239.

Hess, S., Stathopoulos, A., 2013. Linking response quality to survey engagement: A combined random scale and latent variable approach. Journal of Choice Modelling 7, 1–12.

Hess, S., Train, K., 2017. Correlation and scale in mixed logit models. Journal of Choice Modelling 23, 1–8.

Hess, S., Train, K. E., Polak, J. W., 2006. On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a Mixed Logit Model for vehicle choice. Transportation Research Part B: Methodological 40 (2), 147–163.

Hillygus, D. S., Jackson, N., Young, M., 2014. Professional respondents in non-probability online panels. In: Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A., Lavrakas, P. J. (Eds.), Online Panel Research: A Data Quality Perspective, 1st Edition. John Wiley & Sons, Ltd., Ch. 10, pp. 219–238.

Hoyos, D., Mariel, P., Hess, S., 2015. Incorporating environmental attitudes in discrete choice models: An exploration of the utility of the awareness of consequences scale. Science of The Total Environment 505, 1100–1111.

Internet World Stats, 2017. Internet World Stats: Usage and Population Statistics.
URL http://www.internetworldstats.con/stats4.htm

Lindhjem, H., Navrud, S., 2011. Using internet in stated preference surveys: A review and comparison of survey modes. International Review of Environmental and Resource Economics 5, 309–351.

Matthijsse, S., Leeuw, E. D., Hox, J., 2015. Internet panels, professional respondents, and data quality. Methodology 11 (3), 81–88.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.

R Core Team, 2016. R: A language and environment for statistical computing.
URL https://www.r-project.org

Sturgis, P., Allum, N., Brunton-Smith, I., 2009. Attitudes over time: The psychology of panel conditioning. In: Lynn, P. (Ed.), Methodology of Longitudinal Surveys. John Wiley & Sons, Ltd., Chichester, Ch. 7, pp. 113–126.

Swait, J., Louviere, J., 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models. Journal of Marketing Research 30 (3), 305–314.

Toepoel, V., Das, M., Van Soest, A., dec 2008. Effects of Design in Web Surveys: Comparing Trained and Fresh Respondents. Public Opinion Quarterly 72 (5), 985–1007.
URL https://academic.oup.com/poq/article-lookup/doi/10.1093/poq/nfn060

Toepoel, V., Das, M., Van Soest, A., van Soest, A., jan 2009. Relating question type to panel conditioning: Comparing trained and fresh respondents. Survey Research Methods 3 (2), 73–80.

Tukey, J. W., 1962. The Future of Data Analysis. The Annals of Mathematical Statistics 33 (1), 1–67.

Whitsett, H. C., 2013. Understanding "Frequent Survey Responders" on Online Panels.
URL http://www.nera.com/content/dam/nera/publications/archive2/PUB{_}Frequent{_}Survey{_}Responders{_}0313.pdf

A Appendix A

Table A1 - Lower triangular Cholesky matrix and upper triangular correlation matrix for the MIXL model

	Comp	Max 3500	Max 2000	Flex	90 mins	180 mins	Days 10	Days 20	ASC
Compensation	1.1987***	0.0510	−0.0928**	−0.3947**	−0.3423**	−0.2398**	−0.9295***	0.0196	0.2245
Max 3500	0.0507	0.9924***	−0.1149	−0.3480***	−0.1294***	−0.2325*	−0.0484	−0.2362	0.0645
Max 2000	−0.1526	1.5445***	0.5428***	−0.5679***	0.9334***	−0.2640	0.0821	−0.1158	0.5534
Flexible use	−0.0379	−0.2602***	−0.1988*	0.0099	−0.7938**	0.1007	0.0299	−0.3468	0.3040
Duration 90 mins	−0.2963***	−0.1651*	0.2765**	0.5655***	0.2287	−0.3057***	0.3003	0.2471	0.1776**
Duration 180 mins	−0.5843***	−0.3610***	1.2801***	0.7737***	0.1993	−0.2641*	−0.2707***	0.8488***	0.7766***
Days 10	−0.3885***	−0.1610**	0.4108***	−0.1060	−0.0296	0.3314***	0.0216	0.3672	−0.2265
Days 20	−0.3569***	0.1234	0.4968***	−0.1913	0.2673	0.6555***	0.0651	0.3943***	−0.3013
ASC	−0.6990***	−1.6179***	−0.0219	−1.1583***	1.9611***	−4.3616***	0.9460	0.8009**	−0.5749

'***' - 1 % level, '**' - 5 % level, '*' - 10 % level

Table A2 - Lower triangular Cholesky matrix and upper triangular correlation matrix for the ICLV - MIXL model

	Comp	Max 3500	Max 2000	Flex	90 mins	180 mins	Days 10	Days 20	ASC
Compensation	1.1674***	−0.1285	−0.1800***	−0.2631	−0.4724***	0.0453	−0.4241**	−0.1503	0.2689**
Max 3500	−0.1154	0.8907***	−0.1234	−0.2854***	−0.0335**	0.0308	0.0855	−0.0657	0.1276
Max 2000	−0.2819*	1.5348***	0.1306	−0.6207***	0.9951***	−0.0854	0.0688	−0.0652	0.2948
Flexible use	−0.0725	−0.2742***	0.0783	−0.5086***	−0.4470***	0.2772*	−0.0216	−0.1609	0.2954**
Duration 90 mins	−0.1793*	0.0079	0.2183**	0.1732	0.5954***	−0.0613*	0.3173**	0.0810	0.1731***
Duration 180 mins	−0.4229***	−0.0088	0.4119***	0.2410	1.0526***	0.8254***	−0.1279*	0.8260***	0.7649***
Days 10	−0.3046***	−0.0817	0.1760**	0.1516*	−0.1334*	0.1408	−0.2231***	0.1173	−0.4496
Days 20	−0.4754***	0.2197**	0.2218*	0.1706	0.0328	0.0903	−0.7962***	−0.1306	−0.2265**
ASC	−0.1847	−0.3647**	−4.5609***	−0.9944***	2.7150***	0.7183***	0.3914**	−0.5756***	0.1935

'***' - 1 % level, '**' - 5 % level, '*' - 10 % level