

This is the peer reviewed version of the following article: Ryan, M, Mentzakis, E, Matheson, C, Bond, C. Survey modes comparison in contingent valuation: Internet panels and mail surveys. *Health Economics*. 2020; 29: 234– 242, which has been published in final form at <https://doi.org/10.1002/hec.3983> This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for self-archiving.

Survey modes comparison in contingent valuation: internet panels and mail surveys

Mandy Ryan, University of Aberdeen
Emmanouil Mentzakis, University of Southampton
Catriona Matheson, University of Stirling
Christine Bond, University of Aberdeen

Abstract

Stated preferences methods are extensively applied in health economics to elicit preferences. Whilst mailed surveys were commonly used to collect data, internet panel (IP) surveys are being increasingly used. This raises questions about the validity of responses and estimated willingness-to-pay (WTP) values generated from IP surveys. We conduct the first study in health to compare a contingent valuation (CV) internet panel survey with a mailed survey using the Electoral Roll. Our IP has a higher response rate and lower item missing response rate. The difference is reduced but remains when restricting comparisons to valid WTPs. Sample characteristics differ, with significant differences between modes for gender, age, income, and attitudes and knowledge. Whilst difference in WTP values exist, with the IP resulting in higher values, we find limited evidence that such differences are statistically significant. The mail survey has lower initial cost per response; however, once restricting samples to valid WTP responses with non-missing respondent information the cost per response across modes is similar. Our results, suggesting that internet panels generate valid and cost-effective values, are encouraging as researchers move increasingly to IPs to collect preference data.

1. Introduction

Stated preferences methods are extensively used in applied economics to elicit preferences. With the emergence of internet-based surveys and falling response rates to mailed surveys, it has become increasingly common to collect data using internet panels (IPs). IPs have advantages in comparison to traditional used modes, including: avoidance of transcription and manual input errors; faster administration; and removal of interviewer bias (Dillman et al., 2009). However, it remains unclear if different administration modes introduce biases.

Biases are an important consideration in health economics where stated preference methods have found numerous applications. However, literature on survey mode effects in health economics is limited in both the number of studies and the modes compared. Three studies have investigated mode effects within the health state/health outcome literature. Mulhern et al. (2013) compared a Computer Assisted Personal Interview (CAPI) survey (n=201) and an IP survey (n=221) looking at health state valuations. Preferences were elicited using a discrete choice experiment (DCE). To ensure overall comparability of the sample characteristics across administration modes samples were recruited following procedures employed in typical surveys. Whilst significant differences in respondents' characteristics were found (online respondents were educated to a higher level and CAPI respondents reported significantly better general health and health/life satisfaction), there was no significant difference in preferences across modes. Norman et al. (2010) compared online questionnaires and face-to-face interviews when eliciting health state preferences using the time-trade-off method. Data were collected from a convenience sample of adults (n=135) randomised 1:1. They found the online survey had higher variability, and more extremes, than in-person interview valuations. Rowen et al. (2016) compared online questionnaires (n=302) and face to face interviews (n=69) in a pairwise comparison study of social preferences for burden of illness. Whilst the invited sample was representative of the UK for age and gender, the mode of administration affected respondent characteristics; compared to the general population, interview respondents were older and more likely to be retired whereas online respondents were younger, less likely to be employed or retired, and had poorer health.

Whilst willingness to pay (WTP) is commonly used in health economics to assess values, only two studies have investigated how WTP is impacted by survey mode; both used the DCE method. In a study eliciting preferences for pharmacy services, Watson et al (2019) compared four modes: CAPI; mailed survey; and two UK IP surveys (Ipsos Mori and ResearchNow). Samples were recruited following standard approaches and each mode targeted 1000 respondents. Modes were compared according to objective measures (response rate; sample representativeness compared to the UK population; elicited values; theoretical validity; and cost per response) and subjective/self-reported measures (time taken to complete study; perceived study consequentiality; and stated attribute non-attendance). Sample characteristics differed across modes as did estimated WTP. On most measures CAPI were superior but were more expensive. On all measures, except response rates, IPs outperformed the mail survey and were cheaper. In a DCE concerned with patient preferences for health insurance, Determann et al. (2017) compared responses for two samples drawn from an online panel – one sample completed the questionnaire online (n=533) whilst the other received a paper version (n=365). They found no evidence that online surveys yield inferior results compared to paper-based surveys, but the price per respondent was lower

We present the first study comparing the direct contingent valuation (CV) method, using an IP survey and a mailed survey using the Electoral Roll (ER). We compare modes according to: response rates; item-response rates; WTP response classifications; cost per response; sample characteristics; and monetary values.

2. Methods

2.1 Context

The CV survey elicited willingness to pay (WTP) for treatment strategies for illicit drug users. For detailed information on the study design see Matheson et al. (2014). In summary, the questionnaire had four sections: demographics (e.g. gender, age); experience of drug misuse (e.g. experience with illegal drugs in the last year); knowledge and attitudes about drug users; and WTP for five treatment strategies (needle exchange; maintenance with oral methadone; community detoxification; residential detoxification and maintenance with prescribed heroin). Respondents were asked to imagine the Government were considering expanding treatment programmes paid from tax contributions. Respondents were then asked how much they would personally be willing to pay to expand each of the five treatments. Respondents who stated they were unwilling to contribute anything were asked to explain why; the aim was to distinguish 'protesters' (those who objected to the CV valuation method) from those who genuinely did not wish to contribute (those who did not value treatment strategies and had a 'valid' zero). A section of the questionnaire invited open comments, which were analysed qualitatively to aid classification of WTP responses.

2.2 Modes of data collection and sample frames

We designed a single master CV survey using word-processing software, and this was the basis for scripting the IP survey. The self-completion survey was administered using two different samples of the Scottish population who received either a mailed paper self-complete questionnaire or were part of an internet panel (IP) who completed the survey on line. Samples were recruited using standard approaches:

- The paper survey was mailed to a random sample of 3000 people over 18 years from the Electoral Roll (ER). The sample was provided by an independent sampling company and stratified by age and geographical location. Invitations to participate were to named individuals. Respondents entered a prize draw for a £100 shopping voucher as an incentive. Two reminders were sent at three weekly intervals; data collection was closed at eight weeks. Respondents were identified using ID numbers on reply-paid envelopes. Questionnaires were separated from envelopes prior to data entry to preserve anonymity but allow identification of non-responders.
- For the IP mode a stratified sample (age and geographical location) was drawn from a UK volunteer IP, ResearchNow (IP-RN), producing a guaranteed 300 respondents. Potential participants were selected at random from those eligible within strata. Invitations to complete the questionnaire were emailed to the sample. When individuals responded to the invitation they were screened against the quotas (for age and geographical location) until the target number of responses had been received. Following standard practice, IP respondents received the standard nominal incentive used by the company. Data collection took one week.

2.3 WTP response classifications

Combining stated WTP values and respondents' qualitative comments, WTP responses were classified as: *protest* (no value given and a protest against CV e.g. "I think the government should be paying for this"); *cost-based* (respondents who attempted to estimate the cost of services as opposed to the value e.g. "no idea of cost", see Ryan and San Miguel, 2000 for more on cost-based responses); missing (no WTP provided) or *valid*. Distinctions were drawn between 'full' and 'restricted' (i.e. responses without any missing values for covariates/sample characteristics) samples.

2.4 Mode cost-effectiveness

Costs for the mailed questionnaire included: questionnaire printing; paper; postage (outgoing and reply paid reminders); administration; data input; and electoral roll access (one-off payment). IP-RN sample costs included initial fixed set up costs for scripting the survey plus a cost per participant for completion. We calculated cost per WTP response and cost per 'valid' WTP response across the 'full' and 'restricted' samples.

2.5 Comparison of sample characteristics

We compared sample characteristics across modes for both the 'full' sample and for individuals with valid WTP responses. This distinction provides insight into whether representativeness is conditional on validity.

2.6 Comparison of valid WTPs

We compared WTP values and non-zero (positive) WTP values across modes for all five WTP questions. We present results for the 'full' and 'restricted' samples; this latter group is likely to be of most interest to researchers.

2.7 Regression framework

Two or more-part models have been proposed to analyse data where behaviour is separable in parts (Duan et al., 1983). This has found applications in CV with two-part specification modelling WTP conditional on non-zero WTPs (Hammit and Zhou, 2006). Others have incorporated higher dimensions to separate zero and missing responses (Yu and Abler, 2010). We adapted this framework and fitted a three-part model.

First, the probability of stating a valid (j=1), protest (j=2), cost-based (j=3) or missing (j=4) WTP follows a multinomial logit:

$$\Pr(WTP_{Group_i} = j | x_i) = \frac{e^{x_i \beta_j}}{\sum_{j=1}^J e^{x_i \beta_j}} \quad (1)$$

Second, the probability of positive WTPs, conditional on valid WTP is a probit:

$$\Pr(WTP_i > 0 | x_i, WTP = Valid) = \Phi(x_i \gamma) \quad (2)$$

Third, average WTP, conditional on positive, valid WTP was modelled as a generalized linear model with log gamma link that captured the potential long tail in the WTP distribution and has often been proposed for the health care cost modelling (Basu et al., 2004):

$$E(WTP_i|x_i, WTP_i > 0, WTP = Valid) = \exp(x_i\delta) \quad (3)$$

where x_i is a vector of individual characteristics.

Finally predicted WTPs were calculated as $E(WTP_i|x_i, WTP = Valid, WTP_i > 0)$ for both modes and differences were bootstrapped with 500 repetitions.

3. Results

3.1. Response rates, WTP classifications and cost effectiveness comparisons across modes

Table 1 presents response rates, item-response rates, WTP classifications and cost-effectiveness across modes for the full and restricted samples. We find:

- Mail and IP-RN surveys achieved sample sizes of 1067 (38% response rate) and 302 (guaranteed) respectively.
- Item non-response reduced sample sizes by 28% (303/1067) to 764 for mail, and 5% (15/302) to 287 for IP-RN.
- For both the full and restricted samples around two-thirds of respondents (61% and 66% respectively) provided a *valid* WTP in the mail survey, with the remainder consisting mainly of *missing* WTP responses (as opposed to *protest* or *cost-based* responses). The IP-RN mode had a higher *valid* response rate (79% in both groups) and lower *missing* item-response rate (6%).
- For valid WTPs, response rates for the mail mode were 22% (653/3000) for full and 17% (507/3000) for restricted samples with 78% (237/302) and 75% (227/302) correspondingly for the IP-RN mode.
- Mail costs were £6,070; the guaranteed IP-RN sample costs were £2,660. Cost per returned response was £5.64 for mail and £8.81 for IP-RN. For valid responses the costs were £9.30 for mail and £11.20 for IP-RN; for the restricted sample the two modes converge to just under £12 per response.

3.2. Sample characteristics comparisons

Tables 2 and 3 present sample characteristics by mode for the full and *valid* WTP samples. Differences exist for: age (IP-RN had the largest proportion of respondents (34%) in the 55 to 64 years category while mail had the largest proportion (27%) over 64 years); income (mail had a larger proportions of low income earners (29% cf 19%); alcohol drinker (IP-RN 88% cf mail 77%); and knowledge and attitude scores (IP-RN scoring higher in both), with the difference in attitude score no longer significant in the valid WTP sample. The mailed survey has a significantly higher number of missing item responses, ranging from 0.7% of the full sample for age to 13.3% for knowledge scores, compared to no missing values for the IP-RN. These percentages are reduced in the valid WTP sample (range 0%-10.6%).

3.3. WTP mode comparisons

Except for *residential detoxification*, we find no statistical difference in WTP between modes (Table 4). When considering only responses with a positive WTP, IP-RN respondents display significantly

larger values for three programs (*maintenance with oral methadone, community and residential detoxification*), albeit for two of them only at the 10% level, implying a larger proportion of zero WTP values for the IP-RN sample.

3.4. Regression based mode comparisons

Table 5 (Panel A) presents marginal effects from the three-part model. For Part 1, the IP-RN mode increases the probability of a valid response by 10 percentage points and lowers chances for missing WTPs by 15 percentage points. No difference across modes are observed for Protest and Cost-based classifications. In Part 2, except for *maintenance with heroin*, IP-RN mode increases the probability of reporting non-zero WTP values by 18 to 25 percentage points. For Part 3, IP-RN respondents report significantly higher monetary values for all five programs. However, such values are not necessarily representative of unconditional WTP for the programs.

Table 5 (Panel B) presents predicted average WTPs from the three part model, corresponding to average monetary values for the sample. Given the high number of non-valid and zero WTP values in the full sample, predicted monetary values are lower than the conditional values of Part 3 for both samples. Differences in predicted WTPs between modes remain but lack statistical significance.

4. Concluding comments

Our IP-RN survey had higher valid response rates; this may reflect quality/updates of IP mailing lists (compared to the electoral roll). The IP-RN also had a lower missing item response rates; the comparison with the mailed survey shrinks but remains when restricting comparisons to valid-WTPs. This finding is a consequence of IP-RN requiring all questions to be answered. Whilst a mail out with online link could potentially overcome this problem, in pilot work Watson et al (2013) found the response rate for mail to online surveys much lower than all other modes compared (14/1200; 1.2%) and subsequently did not take the comparison of this mode forward to their main study. Sample characteristic comparisons show statistically significant differences. Whilst differences in WTP values exist, with the IP resulting in higher values, we find limited evidence that such differences are statistically significant. Questions may be raised about the generalisability of our results. The mail survey had a larger percentage of over 64-year-olds. In this application, illicit drug use, the sample characteristics may not be driving differences in values. However, for applications where the condition is more likely to affect older people (e.g. social care), IPs and mailed surveys may result in differences in WTP. Future research should explore this. Further, over time this difference may reduce as older people become more computer literate. Our mailed survey had lower initial cost per response; however, for the restricted sample the cost per response across modes was similar and not significantly different. It is recognised for larger sample sizes the mail survey may become more cost-effective.

Our finding that IPs generate similar values to mailed survey, and are equally cost-effective, are encouraging as researchers move increasingly to IPs to collect preference data. They are also relevant for researchers collecting health-related quality of life or other health-related questionnaires. IPs have potential advantages when designing studies, including: the opportunity for: animations/interactive explanations; more complex questions/bidding games; and pivot designs where starting values are specific to the respondent. They also allow for collection of response time data (useful when exploring

data quality) and quicker data collection (useful when quantitative data from pilots is required to inform the main design). Given the ease with which IP surveys can be undertaken researchers should look to exploiting their many advantages in the design of surveys.

References

- Basu, A., Manning, W.G., Mullahy, J., 2004. Comparing alternative models: log vs Cox proportional hazard? *Health Economics* 13, 749–765.
- Determann, D., Lambooij, M.S., Steyerberg, E.W., de Bekker-Grob, E.W., de Wit, G.A., 2017. Impact of Survey Administration Mode on the Results of a Health-Related Discrete Choice Experiment: Online and Paper Comparison. *Value Health* 20, 953–960.
- Dillman D. Smyth J.D. Christian L.M. (2009). *Internet, mail and mixed-mode surveys: The tailored design method*. New Jersey, Wiley.
- Duan, N., Manning, W.G., Morris, C.N., Newhouse, J.P., 1983. A Comparison of Alternative Models for the Demand for Medical Care. *Journal of Business & Economic Statistics* 1, 115–126.
- Hammitt, J.K., Zhou, Y., 2006. The Economic Value of Air-Pollution-Related Health Risks in China: A Contingent Valuation Study. *Environ Resource Econ* 33, 399–423.
- Matheson, C., Jaffray, M., Ryan, M., Bond, C.M., Fraser, K., Kirk, M., Liddell, D., 2014. Public opinion of drug treatment policy: exploring the public's attitudes, knowledge, experience and willingness to pay for drug treatment strategies. *Int. J. Drug Policy* 25, 407–415.
- Mulhern, B., Longworth, L., Brazier, J., Rowen, D., Bansback, N., Devlin, N., Tsuchiya, A., 2013. Binary Choice Health State Valuation and Mode of Administration: Head-to-Head Comparison of Online and CAPI. *Value in Health* 16, 104–113.
- Norman, R., King, M.T., Clarke, D., Viney, R., Cronin, P., Street, D., 2010. Does mode of administration matter? Comparison of online and face-to-face administration of a time trade-off task. *Qual Life Res* 19, 499–508.
- Rowen, D., Brazier, J., Keetharuth, A., Tsuchiya, A., Mukuria, C., 2016. Comparison of Modes of Administration and Alternative Formats for Eliciting Societal Preferences for Burden of Illness. *Appl Health Econ Health Policy* 14, 89–104.
- Ryan M. and San Miguel F. Testing for consistency in willingness to pay experiments. *Journal of Economic Psychology* 21, 305-317, 2000.
- Watson V. Porteous T. Ryan M. Survey mode effects when eliciting public preferences for health care: A pilot discrete choice experiment study. Health Economists' Study Group. Warwick, June 2013.
- Watson V. Porteous T. Bolt T. Ryan M. (2019) Mode and Frame Matter: assessing the impact of survey mode and sample frame in choice experiments *Medical Decision Making* (forthcoming).
- Yu, X., Abler, D., 2010. Incorporating zero and missing responses into CVM with open-ended bidding: willingness to pay for blue skies in Beijing. *Environment and Development Economics* 15, 535–556.

Table 1. WTP response classification and costs between modes

Full Sample					
	Mail Survey (N=1,067)		Internet Panel (N=302)		Test Statistic (<i>p</i> -value)
	Frequency	Percentage	Frequency	Percentage	
Valid Response	653	61.2%	237	78.5%	$\chi^2 = 69.15 (0.001)***$
Protest	83	7.8%	34	11.3%	
Cost Based	17	1.6%	11	3.6%	
Missing	314	29.4%	20	6.6%	
Total	1,067	100.0%	302	100.0%	
Cost per response	£5.64		£8.81		
Cost per Valid response	£9.30		£11.2		
Restricted Sample					
	Mail Survey (N=764)		Internet Panel (N=287)		Test Statistic (<i>p</i> -value)
	Frequency	Percentage	Frequency	Percentage	
Valid Response	507	66.36%	227	79.09%	$\chi^2 = 44.90 (0.001)***$
Protest	60	7.85%	32	11.15%	
Cost Based	13	1.70%	10	3.48%	
Missing	184	24.08%	18	6.27%	
Total	764	100.0%	287	100.0%	
Cost per response	£7.95		£9.27		
Cost per Valid response	£11.97		£11.72		

Notes:

Mail survey costs calculated at £6,070. Internet Panel costs calculated at £2,660.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Full Sample includes all respondents; Restricted Sample includes only those who responded to all items (i.e. controls for item-response)

Table 2. Sample Characteristics for full samples: Comparison by mode

	Mail Survey		Internet Panel		Test Statistic (p -value) for proportions excluding Missing Value observations
	Frequency (N=1,067)	Percentage	Frequency (N=302)	Percentage	
Gender					
Male	488	45.74%	155	51.32%	$\chi^2 = 1.93 (0.165)$
Female	555	52.01%	147	48.68%	
Missing Values	24	2.25%	0	0.00%	
Age					
Under 35 Years	168	15.75%	49	16.23%	$\chi^2 = 48.1 (0.001) ***$
35 to 44 Years	151	14.15%	52	17.22%	
45 to 54 years	230	21.56%	65	21.52%	
55 to 64 Years	220	20.62%	104	34.44%	
Over 64 Years	291	27.27%	32	10.60%	
Missing Values	7	0.66%	0	0.00%	
Employment					
Part-time	165	15.46%	58	19.21%	$\chi^2 = 6.84 (0.145)$
Full-time	437	40.96%	114	37.75%	
Unemployed	40	3.75%	14	4.64%	
Retired	307	28.77%	76	25.17%	
Other	104	9.75%	40	13.25%	
Missing Values	14	1.31%	0	0.00%	
Income					
Less than £15,000	318	29.80%	60	19.87%	$\chi^2 = 21.05(0.001)***$
£15,000 to £24,999	177	16.59%	65	21.52%	
£25,000 to £34,999	180	16.87%	68	22.52%	
£35,000 to £49,999	136	12.75%	54	17.88%	
Greater than £49,999	173	16.21%	40	13.25%	
Missing Values	83	7.78%	15	4.97%	
Alcohol drinker					
No	234	21.93%	37	12.25%	$\chi^2 = 14.43(0.001)***$
Yes	822	77.04%	265	87.75%	
Missing Values	11	1.03%	0	0.00%	
Smoker					
Never	511	47.89%	126	41.72	$\chi^2 = 4.52 (0.104)$
Ex-smoker	367	34.40%	114	37.75	
Current	179	16.78%	62	20.53	
Missing Values	10	0.94%	0	0.00%	
Experience of drug misuse					
No	706	66.17%	192	63.58%	$\chi^2 = 2.21 (0.137)$
Yes	330	30.93%	110	36.42%	
Missing Values	31	2.91%	0	0.00%	
Knowledge					
Knowledge score	5.34	3.11	7.15	2.47	$z = -8.80 (0.001)***$
Missing values	13.31%		0.00%		
Attitude					
Combined attitude score	54.33	8.97	55.35	9.37	$z = -1.88 (0.06)*$
Missing values	12.56%		0.00%		

Table 3. Sample characteristics for Valid WTP samples: Comparison by mode

	Mail Survey		Internet Panel		Test Statistic (p -value) for proportions excluding Missing Value observations
	Frequency (N=6530)	Percentage	Frequency (N=237)	Percentage	
Gender					
Male	287	43.95%	126	53.162%	$\chi^2 = 5.06 (0.024)**$
Female	356	54.95%	111	46.84%	
Missing Values	10	1.53%	0	0.00%	
Age					
Under 35 Years	119	18.22%	39	16.46%	$\chi^2 = 27.23 (0.001) ***$
35 to 44 Years	101	15.47%	40	16.88%	
45 to 54 years	143	21.90%	49	20.68%	
55 to 64 Years	141	21.59%	84	35.44%	
Over 64 Years	149	22.82%	25	10.55%	
Missing Values	0	0.00%	0	0.00%	
Employment					
Part-time	103	15.77%	45	18.99%	$\chi^2 = 5.14 (0.272)$
Full-time	284	43.49%	87	36.71%	
Unemployed	31	4.75%	11	4.64%	
Retired	170	26.03%	63	26.58%	
Other	62	9.49%	31	13.08%	
Missing Values	3	0.46%	0	0.00%	
Income					
Less than £15,000	191	29.25%	50	21.10%	$\chi^2 = 21.05(0.001)***$
£15,000 to £24,999	110	16.85%	52	21.94%	
£25,000 to £34,999	113	17.30%	52	21.94%	
£35,000 to £49,999	95	14.55%	42	17.72%	
Greater than £49,999	113	17.30%	31	13.08%	
Missing Values	31	4.75%	10	4.22%	
Alcohol drinker					
No	131	20.06%	29	12.24%	$\chi^2 = 7.41 (0.006)***$
Yes	518	79.33%	208	87.76%	
Missing Values	4	0.61%	0	0.00%	
Smoker					
Never	323	49.46%	102	43.04%	$\chi^2 = 4.06 (0.131)$
Ex-smoker	214	32.77%	83	35.02%	
Current	111	17.00%	52	21.94%	
Missing Values	5	0.77%	0	0.00%	
Experience of drug misuse					
No	401	61.41%	147	62.03%	$\chi^2 = 0.014 (0.906)$
Yes	241	36.91%	90	37.97%	
Missing Values	11	1.68%	0	0.00%	
Knowledge					
	Mean	Std.	Mean	Std.	Mann-Whiney Test Statistic (p -value)
Knowledge score	5.81	2.98	7.15	2.48	$z = -5.83 (0.001)***$
Missing values	10.57%		0.00%		
Attitude					
Combined attitude score	54.67	9.22	54.82	9.93	$z = -0.285 (0.77)$
Missing values	8.12%		0.00%		

Table 4: Comparisons of valid WTP values between modes for the five evaluated treatment programs.

	Needle exchange		Maintenance with oral methadone		Community detoxification		Residential detoxification		Maintenance with heroin	
	Mail	Internet Panel	Mail	Internet Panel	Mail	Internet Panel	Mail	Internet Panel	Mail	Internet Panel
E(WTP)										
Mean	10.22	10.77	8.28	9.91	16.1	18.4	14.1	25.6	5.16	6.07
S.E.	(3.11)	(2.70)	(1.62)	(2.80)	(2.87)	(4.96)	(1.64)	(9.04)	(0.89)	(1.37)
N	653	237	653	237	653	237	653	237	653	237
t-test (p-value)	0.10 (0.91)		-0.51 (0.61)		0.41 (0.68)		1.87 (0.061)*		0.53 (0.59)	
E(WTP WTP >0)										
Mean	24.11	38.84	21.3	39.1	32.44	56.7	27.9	75.0	18.4	25.7
S.E.	(7.25)	(8.75)	(4.05)	(10.26)	(5.64)	(14.3)	(3.29)	(25.7)	(2.99)	(4.98)
N	277	67	254	60	324	77	329	81	183	56
t-test (p-value)	-0.90 (0.36)		-1.84 (0.067)*		1.79 (0.074)*		3.28 (0.001)***		1.198 (0.232)	

Notes: *** p<0.01, ** p<0.05, * p<0.1

Table 5: Marginal effects of mode on the three parts of the estimation model and predicted mean WTP by mode for the five evaluated treatment programs

Panel A					
Part 1: P(WTP=Valid vs Protest vs Cost-based vs Missing)					
	Valid	Protest	Cost based	Missing	
Internet Panel (=1 if from IP-RN sample)	0.106*** (0.0312)	0.0315 (0.0218)	0.0122 (0.0102)	-0.150*** (0.0236)	
	Needle exchange	Methadone maintenance with oral methadone	Community detoxification	Residential detoxification	Maintenance with heroin
Part 2: P(WTP>0 WTP=Valid)					
Internet Panel (=1 if from Internet Panel sample)	-0.198*** (0.041)	-0.184*** (0.039)	-0.248*** (0.043)	-0.224*** (0.043)	-0.049 (0.035)
Part 3: E(WTP WTP=Valid & WTP>0)					
Internet Panel (=1 if from Internet Panel sample)	29.8** (12.3)	28.8** (12.8)	38.5** (16.2)	50.5*** (17.9)	13.5** (6.80)
Panel B					
E(WTP x, WTP=Valid, WTP>0)					
Mail	8.81*** (1.44)	8.25*** (1.13)	14.30*** (2.03)	12.60*** (1.56)	6.13*** (0.95)
Internet Panel	16.70*** (5.06)	15.30*** (5.07)	23.50*** (6.78)	25.70*** (6.87)	10.60*** (2.90)
Mail vs Internet Panel difference	-7.91 (4.42)	-7.07 (4.83)	-9.23 (6.17)	-13.10 (6.58)	-4.53 (2.82)

Notes:

- Standard errors in parentheses.
- *** p<0.01, ** p<0.05, * p<0.1
- All models include covariates for age, gender, employment status, income, alcohol drinker status, smoking status, drugs experience, and attitude and knowledge scores.