Title: Using repeated choice experiments to evaluate the impact of policy changes on cervical screening

Authors: Meliyanni Johar[1*], Denzil G Fiebig[2], Marion Haas[3], Rosalie Viney[4]

[1] University of Technology Sydney

[2] School of Economics, University of New South Wales

[3] Centre for Health Economics Research and Evaluation (CHERE), Sydney

[4] Centre for Health Economics Research and Evaluation (CHERE), Sydney

Running Title: Johar et al.

Abstract:

Australia was one of the first countries to introduce a publicly funded HPV vaccine program, and its introduction coincided with a media campaign to promote regular cervical screening. One issue with HPV vaccination is how it impacts on demand for screening. This study examines changes in women's screening preferences following these two interventions, using a novel approach to policy evaluation based on repeated discrete choice experiments. The study extends our previous analysis of attitudes to screening by taking advantage of the timing of the choice experiments to examine the impact of the two policy changes on determinants of screening. We find that, unexpectedly, willingness to screen is generally lower post-interventions. The reason for this trend appears to be related to HPV vaccination. We also find that interventions have minor impacts on how women value screening attributes. Our approach allows us to examine the impact of provider behaviour. A simulation demonstrates that under certain conditions, participation rates can be increased by 40% to 50% if health providers actively encourage women to undertake a cervical screening test.

[1*] **Corresponding author**

University of Technology Sydney

Po Box 123

Broadway 2007

Email: meliyanni.johar@uts.edu.au

1. Introduction

Cervical cancer is one of the most preventable forms of cancer, but remains the second most common women's cancer worldwide (Parkin, 2005). In Australia, about 735 women are diagnosed with the cancer every year, and it is predicted that 1 in 150 women will develop the cancer by the age of 75. These numbers would be substantially lower if all women engaged in preventive behaviours. Cytology tests are available to detect pre-cancerous lesions, and under Medicare, the standard Pap test is free to Australian women. It is estimated that through regular screening, 90% of cervical cancer cases can be prevented. Despite this, there is evidence that Australian women are under-screened (Fernbach, 2001). Lack of awareness about cervical cancer and screening programs, and misunderstanding among women about their eligibility for these programs are among the leading reasons for this trend, suggesting that raising awareness is a necessary step towards successful cervical cancer prevention (Belkar et al. 2006; Mullins et al., 2008; Fernbach, 2001; Marshall et al., 2007).

The most recent cervical screening awareness campaign at a national level in Australia was launched in 2007. The campaign was led by a television advertisement aired through national networks prompting women to make a screening appointment if their last Pap test was more than two years ago. The experiences of previous health-oriented campaigns in general suggest that televised messages are a powerful means of influencing the behaviour of their audiences (Mullins et al., 2008; Dobbinson et al., 2008; White et al., 2003).

Unlike other forms of cancer, the cause of cervical cancer is known: infection through the Human Papilloma virus (HPV). Preventing HPV infection with through vaccination is therefore a potentially effective strategy for preventing cervical cancer. In April 2007, the Australian government implemented an HPV vaccination program, one of the first national programs to be launched worldwide. The vaccination program provided HPV vaccines free to females aged 12 – 26 years old, through a school based program for girls 12-18 and through a "catch up program" delivered by primary care providers for women aged 18-26 . The public was informed about the availability of the vaccine through articles, banners, posters and pamphlets, but, unlike the screening promotion campaign, the introduction of the national vaccination program was not advertised on television.

The aim of this study is to examine women's screening preferences in response to these two interventions: the screening promotion and the vaccination program. The approach taken extends a policy evaluation to include analysis of the effects of the interventions on the determinants of screening. Specifically, discrete choice experiments (DCEs), a form of stated preference (SP) technique, are used to elicit women's screening preferences as well as their valuations of various screening factors. SP data have become increasingly popular in health economic studies, providing behavioural data which are not available from revealed preference (RP) data sources (e.g., market survey) (Fiebig et al., 2009; Salkeld et al., 2000; Scott and Vick, 1998; Ryan et al., 2006; Costa-Font and Rovira, 2005; Dolan and Tsuchiya, 2011). By analysing the effects of the interventions on the screening determinants, in addition to their impact on choices, policymakers can make informed use of

these factors to shape future screening behaviour. For example, Gerard et al. (2003) use SP data to analyse determinants of breast screening participation and find that, while attributes of the screening are important, women's participation decisions are also affected by factors such as the manner in which they were informed about screening and the screening environment offered by the health provider which can be policy targets.

In our study, the identification of the effects of the interventions will be achieved through comparisons of outcomes of the 'treated' and 'control' groups, in the standard sense that only the former was exposed to any intervention. The availability of an identical DCE conducted in 2004, before both interventions were introduced (DCE1), provides one control group, allowing us to compare the preferences before and after the policy interventions. In addition, several other control groups are formed through spatial variations and randomisation in the follow-up DCE collected after both interventions (DCE2). These extra comparison groups are useful to isolate out any common (time) trend effect.

The existing literature predicts that awareness would increase women's willingness to be screened (Fernbach, 2001; Marcus and Crane, 1998; Jenkins et al., 1999; Mullins et al., 2008). A previous screening promotion campaign in New South Wales, for instance, recorded a 30% increase in screening uptake within 4 months of the promotion campaign (Shelley et al., 1991). However, never before has a screening promotion campaign been implemented at the same time as a vaccination program, another means to prevent cervical cancer. It could be hypothesised that the two interventions

could support one another, but alternatively, the vaccination program could also counteract the effectiveness of the screening campaign, for example due to misconceptions that vaccination can substitute for screening (Newall et al. 2007; Kulasingam and Myers, 2003). This study will provide the first results in relation to these issues.

2. Background

Cervical cancer is a cancer of the cervix. The cancer develops when women are infected with high-risk strains of HPV (13-18 strains) for a number of years. As the progression of the cancer to an invasive state is slow (up to 10 years) and pre-invasive stages are largely asymptomatic, regular screening is crucial. Early detection of abnormal cells is known to have high curative rates. In Australia, the current recommendation is for women to begin screening between the age of 18 and 21, or a year after commencing sexual activity, whichever is later, and continue two yearly screening until the age of 70.

For the past 40 years, the Pap test has been the main means of cervical screening, but recently alternative tests have been developed, most recently a HPV test to detect specifically high-risk strains of HPV. When used in conjunction with a Pap test, this test is almost 100% accurate. A Pap test alone is typically about 50-85% accurate (Salmeron et al., 2003; Wright et al., 2000). In Australia, women would not typically be offered the choice of a HPV test unless they first had a positive Pap test.

The causal link between HPV infection and cervical cancer has led to development of HPV vaccines to prevent an initial HPV infection. The first vaccine to receive marketing approval ad government funding in Australia was Gardasil, which targets HPV types 16 and 18 which together cause 70% of cervical cancer and HPV types 6 and 11 which cause an estimated 90% of cases of genital warts. Being vaccinated does not exempt women from screening because the vaccine does not protect against other high-risk strains of HPV or eliminate existing exposure to HPV. The vaccine is therefore most effective when received prior to sexual debut.

3. Discrete Choice Experiments

While DCEs have previously been used in a policy evaluation context, a novelty of this study is a methodological one demonstrating the use of repeated DCEs conducted before and after a policy intervention.

A DCE is one stated-preference (SP) method of producing behavioural data that asks its respondents (subjects) for their preferred choice, as opposed to observing their actual decisions in real market situations (which fits a type of revealed-preference (RP) data). The behavioural foundation of SP methods is Lancastrian consumer theory, which proposes a decomposition of utility derived from the consumption of a product into the utilities derived from its attributes. Hence, in a DCE, the product of interest is described by its attributes and their associated levels, which jointly set a scenario. Studies have found that although stated choices are made in a hypothetical setting, in which there is no real consequence of making the choice, one goes through a similar decision making process as in the real market setting (Louviere et al., 2000; Telser and Zweifel, 2007).

The DCEs in this study contain the screening determinants and test options. The first DCE (DCE1) was developed by Fiebig et al. (2009) in 2004 (pre-interventions) to quantify the role of these screening factors in women's screening decisions. In June 2007, after both government interventions were introduced, the experiment was rerun (DCE2). The content of the DCEs is described below.

A scenario is described by a combination of alternative-specific attributes, which reflect the characteristics of a given test, and common or context attributes, which aim to capture the environment in which the screening decision has to be made. Common attributes are fixed across alternatives. Table 1 summarises all the attributes and their levels.

[Insert Table 1 about here]

The combination of common and alternative-specific attributes produces over 16 million ($4^4 \times 2^4 \times 4^{3 \times 2}$) possible scenarios. Experimental design techniques are as described in Burgess and Street (2004a, 2004b) and Fiebig et al. (2009) are used to reduce the potential scenarios to a manageable fraction, while retaining the ability to identify the utility weight of each attribute independently of each other. The process leads to 512 scenarios, comprising 32 treatment combinations for the common attributes and 16 treatment combinations for the alternative-specific attributes. The scenarios are then blocked into 16 versions of 32, with each version including all 32 treatment combinations for the common attributes. A respondent is randomly allocated

to one of these 16 versions. After considering a scenario, the respondents have to choose whether to have a standard Pap test (*P*) or a liquid-based Pap test (*L*), and whether to have an additional HPV test with the selected Pap test (*PH*, *LH*). The No test (*NT*) option is also available.

The surveys also collect personal information on the respondents, such as age and income. Also included in DCE2 are four additional questions regarding their awareness about the HPV vaccine and test and personal experience with any of these measures.

4. Methodology

4.1. The econometric model

The foundation of our modelling is random utility theory. Thus, if

(1) $U_{isj} = V_{isj} + \varepsilon_{isj}$,

where U_{isj} represents the indirect utility function of respondent i in scenario s for alternative j, V_{isj} being the deterministic component of the utility and ε_{isj} capturing all other factors affecting utility that are not included in, then the respondent i will choose *j* over *l* if:

(2) $V_{isj} + \varepsilon_{isj} > V_{isl} + \varepsilon_{isl}$, $\forall l \neq j$ ($l, j = P, L, PH, LH, NT$).

The presence of the random component ε in (2) makes it a probabilistic statement. The probability that the alternative *j* is chosen over the other possible alternatives therefore can be written as:

(3) $\Pr(V_{isj} - V_{isl} > \varepsilon_{isl} - \varepsilon_{isj} \forall l \neq j)$.

For the deterministic component, let

(4) $$V_{isj} = \tilde{x}_{isj}^{'}\tilde{\beta} + z_i^{'}\delta,$$

where $\tilde{\beta}$ and δ are vectors of parameters to be estimated, measuring the utility weights of attributes $\tilde{x}_{isj}$ on screening choice and the influence of socio-demographics z_i on choice, respectively. Generally, the linear-in-parameters specification is assumed by discrete choice models. To increase flexibility, attributes with multiple levels have been represented by categorical variables that allow for non-linear relationships between them and utility. The different levels of screening attributes are effects-coded, whilst the different categories of socio-demographic variables are represented by a set of dummy variables. Effects-coding the attributes separates out their effects from the effects of the omitted categories of the socio-demographic variables on screening choice. For each attribute, the parameter for the reference group is internalised in the parameters of the included levels, and is given by the negative of their sum (see Bech and Gyrd-Hansen, 2005 for further arguments for effects-coding).

By assuming a probability distribution for $(\varepsilon_{isl} - \varepsilon_{isj})$, we can estimate the deterministic component of the utility. Given the panel nature of the data, we estimate a mixed logit (MXL) model which also allows dependence between alternatives. Let us rewrite (4) so as to separate out the alternative-specific constants (ASCs) from $\tilde{x}_{isj}$ and denote the remaining vector of attributes as x_{isj}. The model is now:

(5) $$V_{isj} = dP_{isj}\alpha_{iP} + dL_{isj}\alpha_{iL} + dPH_{isj}\alpha_{iPH} + dLH_{isj}\alpha_{iLH} + dNT_{isj}\alpha_{iNT} + x_{isj}^{'}\beta + z_i^{'}\delta,$$

where *dP*, *dL*, *dPH*, *dLH* and *dNT* are dummy variables for standard Pap, liquid-based Pap, joint standard Pap and HPV test, joint liquid-based Pap and

HPV test, and no test option, respectively and we allow for a random intercept α_{ij} :

(6) $\alpha_{ij} = \alpha_j + \omega_{ij}, \qquad \omega_{ij} \sim iid(0, \Omega),$

where α_j represents its mean and ω_{ij} denotes a random component that represents a deviation from the mean. A significant deviation around the mean would indicate the presence of inherent (individual-specific) taste heterogeneity in the sample population. The resulting composite error term consists of two parts, ε_{isj} that is *iid* (identically and independently distributed) and ω_{ij} that would follow a yet-to-be specified distribution and induce heteroskedasticity and correlation over alternatives. Notice that ω_{ij} varies over respondents, but is fixed over repeated scenarios faced by a respondent, thereby inducing serial correlation across scenarios.

The assumption of additive error term implies that the main source of heterogeneity is inherent taste heterogeneity among women. Some variations however may be related to a particular attribute, and the error component model does not capture this source of heterogeneities. Extending the model to also allow for heterogeneous preferences may improve the fit and explanatory power of the model, but the size of the resulting model can be overwhelming, as there are quite a number of attributes to be considered as random. The main source of heterogeneity could also be scale heterogeneity, instead of inherent taste heterogeneity, or a combination of inherent taste and scale heterogeneity. The application of a more flexible model, such as the generalized multinomial logit model (GMNL) proposed by Fiebig et al. (2010) to allow for alternative sources of heterogeneities is left for future

study. However, as we will see later, our results suggest that the role of taste heterogeneity is dominant.

The normal distribution is used for (6) reflecting that there are people who tend to choose a given alternative, and there are others who tend not to prefer it. The location of the mean will suggest the prevalence of each kind of preferences.

Under the normality assumption, the choice probability in the MXL model is a mixture of logits with a multivariate normal mixing distribution. Conditional on the random parameters, the probability will follow the standard logit specification. However, as respondent's taste is unobserved, the unconditional probability is an integral of the conditional probabilities over all possible values of the random parameters, weighted by its probability density function. This problem has no closed-form solution and is approximated numerically through simulation by:

(7) $$\hat{P}_{isj} = \frac{1}{R}\sum_r \left[\frac{\exp[x_{isj}'\beta + z_i'\delta + \alpha_j + \omega_{ij}^r]}{1+\sum_l \exp[x_{isl}'\beta + z_i'\delta + \alpha_l + \omega_{il}^r]} \right]$$

where $j, l = P, L, PH, LH, r$ indexes a particular draw and R is the total number of draws. As only differences in utilities matter, the no test alternative is chosen as the base with an associated utility of zero.

The mean of these simulated probabilities is then taken to the objective function to be maximised by Maximum Likelihood. Halton draws are used in the simulation instead of random draws to increase the accuracy of

estimation. The random intercepts are allowed to be freely correlated with each other; implying that Ω in (6) is given by:

$$(8) \qquad \Omega = \begin{bmatrix} \sigma_{P,P} & & & \\ \sigma_{P,L} & \sigma_{L,L} & & \\ \sigma_{P,PH} & \sigma_{L,PH} & \sigma_{PH,PH} & \\ \sigma_{P,LH} & \sigma_{L,LH} & \sigma_{PH,LH} & \sigma_{LH,LH} \end{bmatrix}.$$

4.2. Treated and control groups

To identify the effects of the interventions on screening choice and utility weights, we need definitions of the treated group, which has been affected by the interventions, and the control group. The DCE1 respondents are a clear control group, as this survey was collected pre-interventions. However, relying on time variation has the limitation of not knowing the extent of each woman's exposure to the interventions; there is no information in the survey, for example, about women's awareness of the screening promotion advertisement. In the extreme case, it is possible that all women in DCE2 are unaware of both the screening promotion materials and the vaccination program, making them fit the definition of a control group. To deal with this problem, several other definitions for the treated and control groups are proposed. There are 4 cases in total, which are summarised in Table 2.

[Insert Table 2 about here]

Essentially the other comparison samples (Cases 2–4) are split-samples of DCE2. Case 2 uses the extra questions about HPV awareness and define respondents who have ever heard of, or experienced the HPV test or vaccine prior to the experiment as treated. Meanwhile, Case 3 and 4 make use of the

randomisation exercise in DCE2, which allocated respondents into two groups, but only one of the two groups was treated. The treatment was information on HPV facts and HPV-based measures. To ensure that women in the treated group have just the amount of information set by the experiment, those with prior knowledge of HPV are excluded. Finally, women who have prior knowledge of HPV and got randomly allocated into the treatment group can also form a treated sample.

The MXL model is estimated independently for treated and control samples. By doing so, we can test if samples have different scales (the overall extent of unobserved heterogeneity). In discrete choice models, the scale factor is confounded by the utility parameters. However, the identification of the scale factor is desirable, as a larger scale implies a lower variance of the unobservables, which may result from increased awareness. Furthermore, from a policy point of view, different policy strategies are appropriate if women have the same response patterns with respect to choice attributes, but one group of women is more variable in its behaviour than others, from those that are appropriate in the situation in which the underlying behavioural parameters have genuinely changed after the interventions.

5. Results

5.1. Screening choice responses

Respondents for the survey were a stratified random sample of women living in New South Wales aged 18-69 who had previously had a Pap test. Each experiment involves a different sample. DCE1 consists of 167 previously-screened women. As each respondent provided responses to 32 scenarios,

there are 5,344 respondent-scenarios. In DCE2, there are 154 previously-screened women to make a total of 4,928 respondent-scenarios. However, one respondent is dropped because she is over 70 years of age and cervical screening is no longer recommended for women in this age group and 25 respondent-scenarios are dropped due to multiple responses in a given scenario. Compared to women in DCE1, women in DCE2 tend to be younger, more educated, born overseas and have higher incomes (Johar et al., 2009). We control for these differences in the estimation.

The distribution of responses is reported in Table 3. In all cases, the shares of women who chose no test are shown to be higher in the treated than in the control samples. The difference is statistically significant in Case 1 and 4, which both involve comparison against pre-treatment period. This pattern is inconsistent with the prior expectation that awareness would motivate screening participation. Investigating further, this overall increase in non participation is found to be driven by young women's choices. For instance, in DCE1, the share of women under 20 who chose no test was 26%, while in DCE2, this group of women has 55% non-participation rate. This age-specific phenomenon thus hints that the drop in participation rate is related to the parallel vaccination program, which is targeted to these young women. For instance, they may falsely believe that vaccination and screening are alternative strategies to prevent cervical cancer, and that getting vaccinated can substitute for screening. If so, the vaccination program actually counteracts some of the effect of the screening promotion campaign, rather than working together with it to achieve the common goal of cervical cancer

prevention. Meanwhile, the share of a given test alternative is on average stable between a pair of comparison samples.

[Insert Table 3 about here]

5.2. Preference and attribute values

The MXL models are estimated using the routine by Hole (2007) in STATA. A way to measure the gain from estimating MXL from *iid*-based models is to compare their log-likelihood values. It is found that the improvements are quite substantial, measuring 28 to 30% (22 to 28%) of the log-likelihood values of the conditional logit model without (with) alternative-specific constants across samples and 14 to 26% of the log-likelihood values of multinomial logit.

Table 4 reports the results for Case 1. First considering the random intercepts, in the treated sample (DCE2), all the mean intercepts are negative and statistically significant, predicting that in given a scenario, the reference women in the sample (i.e., young women with low education and income, born in Australia, and who were never smokers) would tend to choose not to be tested. On the other hand, the reference women in the control sample (DCE1) are indifferent between no test and a standard Pap test. The location of each of the mean intercepts in DCE2 is further to the left from its counterpart in DCE1, suggesting results that are consistent with the raw data discussed earlier, that the joint interventions have generated negative

preferences towards testing in general.[1] However, the interventions seem to reduce the extent of heterogeneity surrounding a given test alternative. The standard deviations around the means are all large and significant, but are slightly smaller in magnitude in DCE2. Moreover, these deviations are not always larger than their respective means, which is the case in DCE1. Meanwhile, both samples exhibit significant correlation in pairwise alternatives, except between *P* and *LH*. This exception is sensible, as serial correlation is personal, and women who have a taste for technology would be most likely to choose a liquid-based Pap test over the standard Pap test.

[Insert Table 4 about here]

Next, to compare attribute parameters from the two samples, Figure 1A plots the set of estimates from DCE2 against those obtained from DCE1. In non-linear models, this device isolates differences in scales (overall variance) from genuine differences in utility weights; coefficient estimates from different samples are not directly comparable due to confounded scales, which can be sample-specific. Scaling phenomenon implies a systematic difference between the parameter estimates from different samples, with estimates from the sample exhibiting a larger scale being scaled down. On a scatter plot of treated sample against control sample, these estimates will have a linear relationship with slope steeper than a 45-

[1] One can find the probability of getting a value less than 0 of a normally distributed random variable with mean and standard deviation equal to the MXL estimates for each alternative. For instance, for a standard Pap test the probability of getting a value less than 0 of a normally distributed random variable with mean 1.5 and standard deviation 2.5 is 0.73 in DCE2.

degrees line. Meanwhile, points above the 45-degree line indicate larger estimates in the treated sample.

[Insert Figure 1 about here]

For the attributes, the plot shows that the differences in utility weights are largely systematic. While there is no obvious reason why women should change their valuations of screening attributes following the interventions, one can imagine that awareness of screening importance reduces the weights on costs and/or increases the weights on accuracy. The results however suggest that screening participation has always been highly influenced by provider characteristics and recommendation, and costs and test accuracy received unchanging weights.

The different coding system for attributes and the socio-demographic variables turns out to be important, as the effects of socio-demographics, unlike the attributes' weights, vary with samples. Although most of them are not statistically significant, the reversing sign of the coefficients on education and smoking variables to positive is noteworthy. That is, in DCE2, higher education increases the propensity to test, and smokers and ex-smokers are more likely to test than non-smokers. The changing behaviour related to smoking habits in particular is a positive outcome from the perspective of women's health, as smoking increases the risk of developing cervical cancer. The coefficient on ex-smokers is statistically significant at the 5% level.

Figure 1B-D summarises the results from other cases. The underlying results can be found in Johar et al., (2009).[2] As in Case 1, the relationship between attributes' weights in the other comparison samples is largely one-to-one. On the other hand, inherent taste for screening and socio-demographics' effects are sample-specific; due to the dummy-coding of the socio-demographic variables, they are linked with the random intercepts.

Comparing the inherent preference for screening between women with prior knowledge of HPV (Prior sample) and those who were unaware of it (No Prior sample), the reference women in the Prior sample are found to be much less averse towards screening. In Figure 1B, this result is depicted by all the mean intercepts of the tests (ASCs) located above the 45-degrees line. However among women in the No Prior sample, those who have smoking history and/or have high income and education are much more likely to participate in screening. Meanwhile, from the comparison samples based on randomisation, women in the Informed sample tend to be more averse towards testing than those in the Uninformed sample. In Figure 1C, in contrast to the earlier result, all the test intercepts lie below the 45-degrees line. A similar pattern is portrayed in Figure 1D, in comparing women in DCE1 and women in the Prior, Informed sample. These last two results are somewhat puzzling given that the treatment in the randomisation exercise was information on HPV facts. A possible explanation has to do with women's changing assessment of their susceptibility to developing the cancer. For instance, the treatment mentioned the causal link of HPV infection to cervical cancer and the fact that in most cases, the HPV infection will clear by itself. It

[2] Also available as supplementary online material to this article.

could be the case that women who had ever heard of HPV previously thought cervical cancer is caused by some other factors, which might have been more acute. If so, this new information may cause them to revise their risk of contracting the cancer downwards. Some women might have encountered the HPV test or vaccine, but even test participants could be unaware of HPV facts.

5.3. Eligibility for free Gardasil

Whether it is the parallel vaccination program in particular that creates the aversion towards testing can be checked by removing young women (under 30 years old), who are eligible for the vaccination program, from the DCE2 sample. This restriction reduces the sample by nearly half.

Table 5 shows that now all test means are statistically indifferent from zero, suggesting that the targeted vaccination program that occurs at the same time as the screening promotion is a part of the story. Meanwhile the results regarding the screening attributes are largely consistent with those obtained from the unrestricted sample; the correlation coefficient between them is 0.98. As additional information on the test preference of older women, the restriction was also imposed on DCE1. We find that all test mean intercepts are negative and significant. Variations around the means are substantial in any case, but that for the joint liquid-based Pap and HPV test alternative is considerably smaller in size in the DCE2 sample.

[Insert Table 5 about here]

To sum up, despite the screening promotion effort, the majority of women (still) prefer not to be screened. Spatial comparisons suggest that this is due to a reduction in the taste for screening related to HPV events. Meanwhile, the values of screening attributes to a typical woman and the overall scale, which one may interpret as measuring the extent of uncertainty surrounding screening decision-making in general, appear to be independent of any intervention. The policy implication following these results is therefore for future screening promotion effort to integrate the HPV innovations. In particular, the relationship between cervical cancer, screening, HPV facts, and HPV-based measures must be communicated in an orderly fashion to avoid confusion and prevent women from making false self-assessments of their risk of developing the cancer. Better delivery of information may also reduce the uncertainties surrounding the screening decision.

5.4. Policy simulations

Given the significant role of providers in women's screening decisions, stimulating their involvements seems to be a plausible strategy to boost screening rates. Using the attributes related to the GP in the experiment, simulation is used to forecast the impact of this strategy. As alternative strategies, let us consider a price reduction, which is the common policy instrument to increase demand and an investment in Research and Development (R&D) that produces a more accurate Pap test. Currently, the standard Pap test is covered by Medicare, but the newer tests would involve positive out-of-pocket costs.

Consider the case for a representative woman who is on-time for screening (the last test occurred within 2 years). To reflect reality, we specify recommended screening interval at 2 years. Other attributes are selected so that the predicted screening rate in DCE1 for 20-69 years old is consistent with the actual two-year participation rate for these groups of women according to the Australian Institute of Health and Welfare's (AIHW) report, which is 58% in 2003-2004 (AIHW, 2008). The corresponding market share for 2007 onwards is not (yet) available. This alignment requires the GP to be specified as male and as the regular GP of the women. The HPV test costs $50, and is not recommended. With regards to test accuracy, we set the false positive and negative rates at 1% and a 20%, respectively

The 'price effect' (E_1), 'provider effect' (E_2) and 'R&D effect' (E_3) to screening participation are found as follows:

(9) $$\overline{E}_{1NT} = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{P_{iNT}^{L}} - \overline{P_{iNT}^{H}}\right),$$

(10) $$\overline{E}_{2NT} = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{P_{iNT}^{P}} - \overline{P_{iNT}^{NR}}\right),$$

(11) $$\overline{E}_{3NT} = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{P_{iNT}^{A}} - \overline{P_{iNT}^{L}}\right),$$

where N is the sample size, and $\overline{P_{iNT}}$ is the individual average of the probability of no test (*NT*) alternative from 100 draws, drawn from their estimated distributions. The parameters are given by the MXL estimates. The superscript H (L) denotes the case in which screening was recommended without the test type specified and the liquid-based Pap test costs an additional $40 ($10), superscript P (NR) denotes the case in which the GP recommended the standard Pap test (not recommending testing), and

superscript A denotes the case in which the accuracy of the standard Pap increases to a 0.1% false positive rate and a 5% false negative rate. The R&D effect is found using the probability of no test under the low price case, $\overline{P_{NT}^{L}}$, as the reference point. Within-sample variations are achieved by the random draws as well as variations in socio-demographics characteristics.

Table 6 reports the results for (9) – (11). For the extent of the policy change considered, all of the three policies have considerable effects on the screening rate. The price subsidisation can reduce the non-participation rate by around 20% in most samples. Meanwhile, by encouraging health providers to take a more active role in screening promotion (perhaps, in women's visits for other purposes), the non-participation rate can be reduced by 40 to 50%. A large provider effect is indeed not impractical (Myers et al., 2008). R&D spending on technology research has a similar-size effect to the provider effect, but arguably, the R&D returns take longer. For an immediate impact, the government therefore may consider the price strategy, although a generous subsidy (in the above case 75% of the costs) may be needed for large effects, and/or extending incentives for opportunistic screening (e.g., amending the current Practice Improvement Program (PIP)). For the longer run effect, resources can be allocated to R&D that improves test accuracy.

[Insert Table 6 about here]

6. Conclusion

This study has analysed changes in women's attitudes towards cervical screening following the latest (2007) screening promotion campaign and a

parallel vaccination program providing HPV vaccine, Gardasil, in Australia. The successes of previous screening promotion campaigns and other preventive health campaigns (e.g., the SunSmart program in Victoria state, Australia, which aims to lower skin cancer rate) led to the expectation that the promotion campaign would substantially increase the cervical screening rate. However, it is found that the proportion of women willing to be screened is generally lower after the joint interventions. This trend is unexpected, but at least in Australia, there is no precedent for concurrently running a screening promotion campaign and a vaccination program.

The reduction in the participation rate appears to be associated with HPV events. First, the reduction in willingness-to-screen is particularly marked among young women, who can obtain Gardasil for free under the vaccination program. Meanwhile, there is little evidence that given individual characteristics, the older women are averse towards testing. These results therefore suggest that while screening and vaccination are both preventive means for cervical cancer, the effectiveness of the screening promotion effort need not be enhanced by the vaccination program. Second, women who were newly informed about HPV facts tend to have a stronger aversion towards test alternatives than otherwise similar uninformed women. Their willingness to be screened may fall as they misinterpreted HPV facts and re-adjusted their risk of developing the cancer downwards. If so, it is clear that women require clarification about the position of screening in the face of the innovations related to HPV.

Through a simulation exercise, several potential strategies to increase future screening rates were evaluated, and the result suggests that encouraging a more active role of health providers is the most effective strategy among those considered to achieve this goal, capable of reducing the non-participation rate by close to one half. Meanwhile, R&D spending on technology that improves test accuracy can be justified on the basis of its expected sizeable impact on screening participation in the longer term.

References

Australian Institute of Health and Welfare. Cervical screening in Australia 2005-2006. Cancer Series 2008 no.41. <http://www.aihw.gov.au/publications/can/csa05-06/csa05-06.pdf>

Bech M, Gyrd-Hansen D. Effects Coding in Discrete Choice Experiments. Health Economics 2005; 14(10); 1079-1083.

Belkar R, Fiebig D, Haas M, Viney R. Why Worry About Awareness in Choice Problems? Econometric Analysis of Screening for Cervical Cancer. Health Economics 2006; 15; 33-47.

Burgess L, Street D. Optimal and Near-optimal Pairs for The Estimation of Effects in 2-level Choice Experiments. Journal of Statistical Planning and Inference 2004a; 118; 185-199.

--- Optimal Stated Preference Choice Experiments When All Choice Sets Contain A Specific Option. Statistical Methodology 2004b; 1; 37-45.

Costa-Font J, Rovira J. Eliciting Preferences for Collectively Financed Health Programmes: The 'Willingness To Assign' Approach. Applied Economics 2005; 37(14); 1571-1583.

Dobbinson S J, Wakefield M A, Jamsen K M, et al. Weekend Sun Protection and Sunburn in Australia: Trends (1987–2002) and Association with SunSmart Television Advertising. American Journal of Preventive Medicine

2008; 34(2); 94-101.

Dolan P, Tsuchiya A. Determining the Parameters in a Social Welfare Function Using Stated Preference Data: An application to Health. Applied Economics 2011; 43(18); 2241-2250.

Fernbach M. The Impact of the 1999 Media Campaign on Cervical Cancer Screening Knowledge and Self-efficacy. Anti-Cancer Council of Victoria 2001, in Fernbach M (ed.). PapScreen Victoria Evaluation Report vol.1, Melbourne, Australia, p.73-85.

Fiebig D, Haas M, Hossain I, Street D, Viney R. Decisions About Pap tests: What Influences Women and Providers? Social Science & Medicine 2009; 68(10); 1766-1774.

Fiebig D, Keane M, Louviere J, Wasi N. Accounting for the Scale and Coefficient Heterogeneity. Marketing Science 2010; 29(3); 393-421.

Gerard K, Shanahan M, Louviere J. Using Stated Preference Discrete Choice Modelling to Inform Health Care Decision Making: A Pilot Study of Breast Screening Participation. Applied Economics 2003; 35(9); 1073-1085.

Hole A. Fitting Mixed Logit Models by using maximum Simulated Likelihood. Stata Journal 2007; 3; 388-401.

Jenkins C N H, McPhee S J, Bird B A, et al. Effect of a Media-Led Education Campaign on Breast and Cervical Cancer Screening among Vietnamese-American Women. Preventive Medicine 1999; 28(4); 395-406.

Johar M, Fiebig D G, Haas M, Viney R. Evaluating Changes in Attitudes towards Cervical Cancer Screening Following a Screening Promotion Campaign and a Free Vaccination Program: Results from Discrete Choice Experiments. CHERE Working Paper 2009/3.

Kulasingam S L, Myers E R. Potential Health and Economic Impact of Adding a Human Papillomavirus Vaccine to Screening Programs. Journal of the American Medical Association 2003; 290(6); 781-789.

Louviere J, Hensher D A, Swait J D. *Stated Choice Methods: Analysis and Application*. Cambridge: Cambridge University Press; 2000.

Marcus A C, Crane L A. A Review of Cervical Cancer Screening Intervention Research: Implications for Public Health Programs and Future Research. Preventive Medicine 1998; 27; 12-31.

Marshall H, Ryan P, Roberton D, Baghurst P. A Cross-sectional Survey to Assess Community Attitudes to Introduction of Human Papillomavirus Vaccine. Australian and New Zealand Journal of Public Health 2007; 31(3); 235–242.

Mullins R, Wakefield M, Broun K. Encouraging the Right Women to Attend for Cervical Cancer Screening: Results From a Targeted Television Campaign in Victoria, Australia." Health Education Research 2008; 23(3); 477-486.

Myers E, Huh W K, Wright J D, Smith J. The Current and Future Role of Screening in the Era of HPV Vaccination. Gynecologic Oncology 2008; 109; s31-s39.

Newall A, Beutels P, Wood J, Edmunds W, MacIntyre C. Cost-effectiveness Analyses of Human Papillomavirus Vaccination. Lancet Infectious Diseases 2007; 7(4); 289–296.

Parkin D, Bray F, Ferlay J, Pisani P. Global Cancer Statistics, 2002. CA Cancer Journal Clinic 2005; 55(2); 74-108.

Ryan M, Netten A, Skåtun D, Smith P. Using Discrete Choice Experiments to Estimate a Preference-based Measure of Outcome – An Application to Social Care for Older People. Journal of Health Economics 2006; 25; 927-944.

Salkeld G, Ryan M, Short L. The Veil of Experience: Do Consumers Prefer What They Know Best? Health Economics 2000; 9(3); 267–270.

Salmeron J, Lazcano-Ponce E, Lorinez A. et al. Comparison of HPV-based Assays with Papanicolaou Smears for Cervical Cancer Screening in Morelos State, Mexico. Cancer Causes and Control 2003; 14; 505-512.

Scott A, Vick S. Agency in Health Care: Examining Patients' Preferences for Attributes of the Doctor-patient Relationship. Journal of Health Economics 1998; 17(5); 587-605.

Shelley J, Irwig L M, Simpson M, Macaskill P. Evaluation of a Mass-media-led Campaign to Increase Pap Smear Screening. Health Education Research 1991; 6; 267-277.

Telser H, Zweifel P. Validity of Discrete-Choice Experiments Evidence for Health Risk Reduction. Applied Economics 2007; 39(1); 69-78.

White V, Tan N, Wakefield M, Hill D. Do Adult Focused Anti-National Tobacco Campaign Have Impact on Adolescents? The Case of the Australian Smoking Campaigns. Tobacco Control 2003; 12(Supll.II); 23-29.

Wright T C, Denny I, Kuhn I, Polleck A, Lorinez A. HPV-DNA Testing of Self-collected Vaginal Samples Compared with Cytologic Screening to Detect Cervical Cancer. Journal of the American Medical Association 2000; 283; 81-86.

Table 1: Attributes and levels in the DCE

Attributes	Levels	
Common		
Last cervical screening appointment	1 year ago; 2 years ago; 3 years ago; 5 years ago	
The recommended screening interval	1 year; 2 years; 3 years; 5 years	
Contact with GP	Regular GP seen for most care; Never seen before	
Sex of GP	Female; Male	
Recommendation of GP	No test; Standard Pap; Liquid-based Pap; Any Pap test	
Financial incentive to GP	No; Yes	
Alternative-specific (Pap test)	**Standard Pap**	**Liquid-based Pap**
Cost of Pap test	$0; $10; $20; $30	A+$10; A+$20; A+$30; A+$40
False negative rates	1/20; 1/15; 1/10; 1/5	1/100; 1/33; 1/20; 1/10
False positive rates	1/1000; 1/250; 1/150; 1/100	1/2000; 1/500; 1/150; 1/100
Alternative-specific (HPV test)	**No HPV test**	**HPV test**
Cost of HPV test	0	$50; $100; $150; $200
Recommendation to additional test	0	No test; test

Table 2: Comparison samples

Case	Treated	Control
1.	DCE2	DCE1
2.	DCE2 with prior knowledge (Prior)	DCE2 without prior knowledge (No prior)
3.	DCE2 without prior knowledge, informed (Informed)	DCE2 without prior knowledge, uninformed (Uninformed)
4.	DCE2 with prior knowledge, informed (Prior, Informed)	DCE1

Table 3: Sample mean of choice responses

Case No.	1		2		3		4	
Choice/Sample	Treated	Control	Treated	Control	Treated	Control	Treated	Control
	DCE2	DCE1	Prior	No Prior	Informed	Uninformed	Prior, Informed	DCE1
No test	0.429*	0.373	0.431	0.419	0.432	0.410	0.444*	0.373
Standard	0.228*	0.296	0.188*	0.271	0.243*	0.292	0.197*	0.296
Standard, HPV test	0.103	0.095	0.099	0.108	0.106	0.110	0.068*	0.095
Liquid	0.118	0.118	0.126	0.111	0.148*	0.083	0.126	0.118
Liquid, HPV test	0.123	0.119	0.157*	0.090	0.073*	0.105	0.165*	0.119
N	5344	4871	2427	2444	1053	1391	1373	5344

Note: * indicates difference with control group is statistically significant at 1% level based on two-mean sample comparison test. The sample size is the number of respondents * 32 scenarios. In DCE2, the sample excludes the 70 year old women and multiple responses as described in Section 5.1.

Table 4: MXL results – Case 1

	DCE1		DCE2			DCE1		DCE2	
	Coeff.	p	Coeff.	p		Coeff.	p	Coeff.	p
Socio-demo					**Alt-spec: Pap test**				
Age	0.005	0.620	0.004	0.693	Cost: A+$20	0.318	0.000	0.268	0.000
Trade certificates	-0.222	0.618	0.312	0.478	Cost: A+$30	0.060	0.189	0.036	0.444
Some uni	-0.525	0.173	0.889	0.139	Cost: A+$40	-0.418	0.000	-0.449	0.000
Completed uni	0.160	0.658	0.139	0.755	Cost: A+$10	0.041		0.146	
Inc $50- $80,000	0.342	0.343	-0.594	0.183	FP: 1/250, 1/500	0.021	0.652	0.113	0.015
Inc >$80,000	0.243	0.511	0.085	0.850	FP: 1/150, 1/150	-0.029	0.524	-0.059	0.208
Inc missing	-0.972	0.036	0.671	0.586	FP: 1/100, 1/100	-0.205	0.000	-0.321	0.000
Not Australian-born	0.075	0.853	0.284	0.425	FP: 1/1000, 1/2000	0.214		0.268	
Current smoker	-0.452	0.182	0.598	0.164	FN: 1/15, 1/33	0.024	0.590	0.041	0.379
Ex-smoker	-0.114	0.762	1.094	0.017	FN: 1/10, 1/20	0.028	0.533	0.055	0.239
Common					FN: 1/5, 1/10	-0.239	0.000	-0.296	0.000
Interval: 1 year	0.715	0.000	0.811	0.000	FN: 1/20, 1/100	0.186		0.201	
Interval: 3 years	-0.127	0.062	-0.137	0.061	**Alt-spec: HPV test**				
Interval: 5 years	-0.858	0.000	-1.021	0.000	Rec: HPV test	0.571	0.000	0.517	0.000
Interval: 2 years	0.270		0.347		Rec: no HPV test	-0.571		-0.517	
Last screen: 2 years	-0.061	0.374	-0.205	0.004	HPV cost: $100	-0.028	0.719	0.005	0.955
Last screen: 3 years	0.295	0.000	0.406	0.000	HPV cost: $150	-0.411	0.000	-0.540	0.000
Last screen: 5 years	0.951	0.000	1.102	0.000	HPV cost: $200	-0.784	0.000	-0.694	0.000
Last screen: 1 year	-1.185		-1.303		HPV cost: $50	1.223		1.229	
GP: new	-0.522	0.000	-0.605	0.000	**Intercepts (ASCs)**				
GP: seen before	0.522		0.605		*P*	-0.445	0.414	-1.500	0.007
GP: male	-0.469	0.000	-0.705	0.000	Std. dev	2.671	0.000	2.489	0.000
GP: female	0.469		0.705		*L*	-1.331	0.014	-2.307	0.000
Rec: standard	0.540	0.000	0.492	0.000	Std. dev	2.481	0.000	2.236	0.000
Rec: liquid	0.067	0.318	0.182	0.010	*PH*	-2.299	0.000	-2.898	0.000
Rec: any Pap	0.365	0.000	0.337	0.000	Std. dev	2.943	0.000	2.786	0.000
Rec: no test	-0.972		-1.013		*LH*	-2.541	0.000	-3.127	0.000
GP: get finc incentive	-0.049	0.215	0.059	0.153	Std. dev	3.975	0.000	2.940	0.000
GP: no finc incentive	0.049		-0.059		**Correlation***				
					P, L	0.498	0.000	0.657	0.000
					P, PH	0.507	0.000	0.504	0.000
					P, LH	0.057	0.115	-0.012	0.844
					L, PH	0.558	0.000	0.387	0.000
					L, LH	0.765	0.000	0.449	0.000
					PH, LH	0.675	0.000	0.715	0.000
					N	26,720		24,355	
					Log L	-5,158		-4,681	

Note: Reported under Coeff column are MXL coefficients, and under p column is the probability value that the respective coefficient is equal to zero. Coefficients without probability values are coefficients of the reference group. 'Rec' stands for GP's recommendation. For standard Pap, the cost levels are $0, $10, $20 and $30, and the costs of liquid-based test add to these cost. For false positive (FP) and false negative (FN) rates, the first figure is for standard Pap and the second is for liquid-based test. * p-values of the covariance terms. The number of draws for the simulated probabilities is R = 2,000.

Table 5: MXL results: a sample of women over 30 years old in DCE2

	Restricted DCE1		Restricted DCE2			Restricted DCE1		Restricted DCE2	
	Coeff.	p	Coeff.	p		Coeff.	p	Coeff.	p
Socio-demographics					**Alt-spec: Pap test**				
Age	0.034	0.035	-0.040	0.036	Cost: A+$20	0.375	0.000	0.308	0.000
Trade certificates	0.214	0.682	0.249	0.606	Cost: A+$30	0.074	0.166	-0.053	0.428
Some university	0.503	0.249	0.279	0.696	Cost: A+$40	-0.478	0.000	-0.436	0.000
Completed university	0.599	0.214	0.425	0.410	Cost: A+$10	0.029		0.180	
Inc $50 - $80,000	0.095	0.819	-0.280	0.554	FP: 1/250, 1/500	-0.012	0.826	0.183	0.005
Inc >$80,000	0.473	0.277	-0.783	0.157	FP: 1/150, 1/150	-0.002	0.971	-0.071	0.288
Inc missing	-1.651	0.001	0.179	0.837	FP: 1/100, 1/100	-0.225	0.000	-0.364	0.000
Not born in Australia	-0.415	0.246	0.487	0.257	FP: 1/1000, 1/2000	0.238		0.251	
Current smoker	-0.049	0.927	0.110	0.840	FN: 1/15, 1/33	0.037	0.487	0.059	0.369
Ex-smoker	-0.367	0.339	1.743	0.000	FN: 1/10, 1/20	0.033	0.532	0.038	0.565
Common					FN: 1/5, 1/10	-0.247	0.000	-0.293	0.000
Interval: 1 year	0.754	0.000	0.851	0.000	FN: 1/20, 1/100	0.177		0.196	
Interval: 3 years	-0.095	0.234	-0.085	0.417	**Alt-spec: HPV test**				
Interval: 5 years	-0.969	0.000	-1.092	0.000	Rec: HPV test	0.527	0.000	0.537	0.000
Interval: 2 years	0.309		0.326		Rec: no HPV test	-0.527		-0.537	
Last screen: 2 years	-0.060	0.449	-0.285	0.005	HPV cost: $100	-0.051	0.581	-0.053	0.664
Last screen: 3 years	0.308	0.000	0.445	0.000	HPV cost: $150	-0.473	0.000	-0.558	0.000
Last screen: 5 years	1.038	0.000	1.340	0.000	HPV cost: $200	-0.838	0.000	-0.497	0.000
Last screen: 1 year	-1.286		-1.501		HPV cost: $50	1.362		1.108	
GP: new	-0.568	0.000	-0.702	0.000	**Intercepts (ASCs)**				
GP: seen before	0.568		0.702		*P*	-1.936	0.079	0.717	0.436
GP: male	-0.414	0.000	-0.734	0.000	Std. dev	2.880	0.000	3.083	0.000
GP: female	0.414		0.734		*L*	-2.204	0.044	0.314	0.735
Rec: standard	0.505	0.000	0.436	0.000	Std. dev	2.865	0.000	2.408	0.000
Rec: liquid-based	0.038	0.633	0.139	0.173	*PH*	-3.674	0.001	-1.025	0.297
Rec: any Pap	0.354	0.000	0.264	0.009	Std. dev	3.030	0.000	3.389	0.000
Rec: no test	-0.897		-0.838		*LH*	-3.143	0.003	-0.859	0.359
GP: get finc incentive	-0.070	0.127	0.046	0.441	Std. dev	5.118	0.000	3.647	0.000
GP: no finc incentive	0.070		-0.046		**Correlation***				
					P, L	0.432	0.000	0.526	0.000
					P, PH	0.473	0.000	0.421	0.000
					P, LH	-0.074	0.018	-0.179	0.004
					L, PH	0.459	0.000	0.570	0.000
					L, LH	0.728	0.000	0.521	0.000
					PH, LH	0.598	0.000	0.732	0.000
					N	20,160		12,940	
					Log L	-3,691		-2,301	

Note: The number of replication for simulated probabilities are R=1000. * p-values based on covariance terms. Only 25 percent of DCE1 sample are aged less than 30. The number of women in the restricted sample is 81. Results for unrestricted case DCE2 are reported in Table 4.

Table 6: Simulation results on the probability of no test

Sample	N	$\overline{E_1}$ [%]	Std. Dev	$\overline{E_2}$ [%]	Std. Dev	$\overline{E_3}$ [%]	Std. Dev
DCE1	167	-0.051	0.006	-0.176	0.062	-0.079	0.011
		[17.41%]		[43.92%]		[32.60%]	
DCE2	153	-0.076	0.007	-0.196	0.014	-0.118	0.016
		[19.63%]		[40.24%]		[38.09%]	
95% CI of DCE1 means		(-0.067, -0.037)		(-0.222, -0.135)		(-0.112, -0.054)	
DCE2 Sub-samples							
No Prior	77	-0.097	0.014	-0.197	0.023	-0.127	0.030
		[23.03%]		[38.80%]		[39.36%]	
Prior	76	-0.068	0.009	-0.223	0.025	-0.126	0.026
		[17.29%]		[43.15%]		[39.06%]	
Uninformed	44	-0.074	0.019	-0.170	0.039	-0.106	0.033
		[16.65%]		[31.69%]		[28.62%]	
Informed	33	-0.132	0.034	-0.242	0.054	-0.157	0.062
		[31.87%]		[47.95%]		[55.42%]	
Prior, Informed	43	-0.074	0.017	-0.181	0.038	-0.100	0.033
		[18.09%]		[36.92%]		[29.86%]	
95% CI of No prior means		(-0.012, -0.055)		(-0.234, -0.122)		(-0.173, -0.055)	
95% CI of Prior means		(-0.085, -0.035)		(-0.276, -0.129)		(-0.178, -0.053)	

Note: 95% CI denotes 95% confidence interval: $\overline{x} \pm 2\sigma_{\overline{x}}$. $\overline{E_1}$, $\overline{E_2}$ and $\overline{E_3}$ reports the average price effect, provider effect and R&D effect, respectively, in terms of change in probability of no test. In square brackets are the effects as a proportion change from the reference probability.

Figure 1A-D: MXL estimates comparisons