

Preferences for new and existing contraceptive products

Denzil G. Fiebig^a, Stephanie Knox^b, Rosalie Viney^b, Marion Haas^b
Deborah J. Street^c

^a*School of Economics, University of New South Wales*

^b*Centre for Health Economics Research and Evaluation (CHERE),
University of Technology, Sydney*

^c*Department of Mathematical Sciences, University of Technology,
Sydney*

Abstract

New contraceptive methods provide greater choice in terms of effectiveness, management of side-effects, convenience and frequency of administration and flexibility, but make the decisions about contraception more complex. There are limited data on the factors that determine women's choices among these alternatives, to inform providers about the factors which are most important to women, or to predict uptake of new products. This paper reports on a choice experiment designed to elicit women's preferences in relation to prescribed contraception and to forecast the impact of the introduction of two new products into the Australian market. A generalized multinomial logit model is estimated and used in the simulation exercise. The model forecasts that the hormonal patch would be well received among women, achieving a greater market share than current non-pill products, but the vaginal ring would have limited appeal.

Keywords - contraceptive choice; new product demand; discrete choice experiment; generalized multinomial logit.

ACKNOWLEDGEMENTS

This research is supported by an Australian Research Council Linkage grant. The industry partners are Family Planning NSW, Jannssen-Cilag Pty Ltd, Organon Australia Pty Ltd and Bayer-Schering Pty Ltd. We would like to thank Edith Weisberg and Deborah Bateson for their continuing support and invaluable input into this particular paper and the overall project. The thoughtful comments of two referees and Peter Sivey and other participants at the 1st Australasian Workshop on Econometrics and Health Economics are gratefully acknowledged.

Corresponding Author:

Professor Denzil G Fiebig

School of Economics

UNSW

Sydney, 2052, Australia

Ph +61-2-9385 3958; Fax +61-2-9313 7691; Email d.fiebig@unsw.edu.au

1. Introduction

The range and complexity of contraceptive choices available in developed countries has increased dramatically, with implications for consumers and health care providers. While the combined contraceptive pill has been the dominant method of contraception for many decades, in the last five years there has been rapid development of new formulations of the pill and new long-acting methods such as hormonal IUDs, implants, intra-vaginal methods and hormonal patches. These new methods provide women with much greater contraceptive choice in terms of effectiveness, management of side-effects, convenience and frequency of administration and flexibility, but make the decisions about contraception more complex. The availability of new products also requires providers to determine what information is of most interest and relevance to women to help them make decisions in relation to contraception.

This study investigates factors which influence a woman's choice of method of contraception. The study uses discrete choice experiment methods to quantify the trade-offs that women make in assessing different contraceptive alternatives, to provide information about how they choose under different circumstances, and to predict uptake of contraceptive alternatives new to the market. Previous research has used choice experiment methods to investigate contraceptive choices. Seston et al (2007) use a choice experiment to investigate trade-offs in relation to emergency contraception, focussing on attributes such as wait time, staff attitude and opening hours. However the trade-offs in relation to planned contraception are quite different. Terris-Prestholt et al (2009) have investigated trade-offs in relation to prevention of sexually transmitted diseases, prevention of pregnancy and cost in different barrier methods of contraception for women in South Africa. Delavande (2008) estimates willingness to pay for a hypothetical method of contraception that is 100% effective. A key focus of the current study is the prediction of uptake of new products in the planned contraceptive setting, as well as providing information that will be useful to providers in giving relevant advice and information to women about the characteristics of new and existing products. The focus on new products is a key motivation for using a discrete choice experiment method. Such methods are of particular use when there are limited data from actual market transactions as will be the case when the introduction of new products is being considered.

The complexity of prescribed contraceptive methods available and the existence of a complex range of “outside” options present particular design challenges not faced in typical discrete choice experiments. The choice experiment described in this paper incorporates eight alternative products, using seven attributes to describe these products. Some attributes, such as frequency of administration, have levels that are specific to only some of the products, whereas the levels of other attributes are common across all products. This presents specific challenges in developing a near-optimal design, in terms of statistical efficiency, while still satisfying requirements of respondent efficiency. The DCE was constructed using a modification of a popular construction strategy, variously called simultaneous option generation or the L^{MA} approach, and was tested in simulation studies to confirm that it could successfully recover a range of assumed prior parameter values.

In terms of estimation, the need to predict new product demand requires a model that allows for flexible substitution patterns across very different alternatives. Random parameter specifications, such as mixed logit (MXL), are very popular in choice modelling applications partly because they do provide a means of capturing more flexible variance-covariance structures for the unobservables in choice models and hence avoid problems associated with independence of irrelevant alternatives (IIA). However, they neglect another potentially important aspect of heterogeneity associated with variations in scale. This study uses the generalized multinomial logit (G-MNL) framework recently developed in Fiebig, Keane, Louviere and Wasi (2010). Their work highlights the empirical importance of accommodating this extra dimension of heterogeneity over and above the flexibility provided by standard MXL model specifications.

In broad methodological terms, this paper illustrates how the methods for designing and analysing choice experiments can be adapted to more complex choice problems. This has important implications for the application of choice experiments in health economics, where simplification of choice contexts and neglect of important sources of heterogeneity in the resultant responses can lead to incorrect inference.

2. What is known about contraceptive choices?

Approximately 56 per cent of all Australian women aged 18 to 49 years use some form of contraception (National Health Survey, 2001 collected by the Australian Bureau of Statistics (ABS)). Since its introduction in 1961, the oral contraceptive pill has been the most commonly used method of contraception. New formulations of the pill including the progestogen only pill (mini pill) and lower dose formulations of the combined pill that reduce unwanted side-effects have increased its popularity. Use of the contraceptive pill ranges from almost 70 per cent for 18 to 24 year olds to about 41 per cent amongst 45 to 49 years olds (National Health Survey, 2001). Over 75 per cent of women aged 18 to 44 years were reported to have ever used the contraceptive pill, with half of those being under 20 years of age (Yusuf and Siedlecky, 2007).

The proportion of women using other types of contraception such as the IUD, the diaphragm and the injection is generally lower. One reason may be lower awareness and availability of information about these alternative contraceptive methods. There are also differences in patterns of contraceptive use among some socio-demographic groups. The language spoken at home and the woman's level of schooling are significant factors influencing contraceptive choices (Yusuf and Siedlecky, 2007). Most studies on contraceptive choice model the relationship between contraceptive choice and age. However, there is evidence that life stage, rather than age, drives contraceptive decisions (Gray and McDonald, 2007).

In the past decade there have been considerable developments in the contraceptive market, leading to the introduction of a wider array of oral contraception formulations, as well as newer methods of contraception. In particular, there have been developments of new long-acting methods such as hormonal Intra-Uterine Devices and implants, injections, dermal patches and vaginal rings.

Acceptability of new methods depends on familiarity with them, as demonstrated by results from trials of the vaginal ring in Europe and North America (Novák et al., 2002). At the start of the trial in which women used the vaginal ring, almost 66 per cent of women indicated the pill was the best perceived method of contraception, but after 13 months of using the ring, 81 per cent reported that the ring was the best

perceived method. Gilliam et al (2007) found that a much larger percentage of North American college students expressed willingness to use the transdermal patch and vaginal ring than were actually using it, despite their availability. There was a negative correlation between willingness to use new methods and current use of the oral pill. This reluctance to switch from the pill may explain the slower uptake of newer contraceptive options.

The availability of so many different contraceptive choices poses challenges for GPs in providing the balanced and comprehensive information necessary for women to be able to make a fully informed choice. More information about the factors that influence women's choices can assist GPs in providing this information. While market or revealed preference data are available that characterize the types of contraceptive choices women are currently making (Yusuf and Siedlecky (2007) and Gray and McDonald (2007)), these data provide little or no detailed information about what is driving these choices. In particular, there would only be limited variability in key product characteristics across different contraceptives and it is this type of information that is required in order to predict how women will respond to new products that constitute different bundles of characteristics.

3. Development of the choice experiment

3.1 Overview of the experiment

Stated preference data from a discrete choice experiment (DCE) are used to evaluate women's preferences for attributes of alternative contraceptive products. The choice experiment was designed to reflect realistically the exchange of information and the choices that are likely to be made in a typical consultation between a woman and her GP in relation to contraception. The complexity of prescribed contraceptive methods available and the existence of a complex range of "outside" options present particular design challenges not faced in typical discrete choice experiments. The aim was to reflect the context of the consultation in the attributes of the experiment and present the range of prescribed contraceptive products that might be discussed without leading to cognitive overload, and allow for women to choose a realistic outside option (stay with her existing method). The experiment includes contextual factors and a range of contraceptive products, and varies the attributes of the products within a realistic range.

3.2 Development of attributes

Attributes were selected from a review of the literature and in discussion with clinical experts in the field of family planning and a series of focus groups with women generally, young women specifically and GPs. The final set of attributes consisted of a labelled product attribute (8 levels) and 7 other attributes (effect on acne, effect on weight, frequency of administration, contraceptive effectiveness, doctor's recommendation, effect on periods and cost). Attribute levels were selected to include a range of realistic values, either actual or hypothetical. The levels of the attributes "frequency of administration" and "contraceptive effectiveness" were nested within subsets of products to allow each product to be described in terms of plausible levels of these two attributes. This ensured that all of the options that were eventually shown to respondents were realistic, even if some were hypothetical. See Table 1 for a full list of attributes and levels.

3.3 Design of the Choice Experiment

Presenting all eight products in each choice set was considered to be both cognitively demanding, and unlikely to reflect the real choice sets women would be offered at a

consultation. However, given that the product category was one with which respondents were familiar, choice sets of size three were considered to be feasible, and realistic. In another component of our research, we have asked GPs how many products they would discuss with different patients and find that the median and modal number of products discussed is 3. Each choice task consisted of choosing both the best and the worst from a set of three options and each respondent completed 32 such tasks.

Usually if there are products described by attributes with different levels then each choice set contains one option of each product. The particular choice sets to use are determined by constructing the three options to be presented in each choice set at the same time. This was achieved by constructing a resolution 3 fraction of the appropriate factorial design. As well as having the attributes associated with each product, we have extended this idea further and included an additional attribute to indicate the product that should be included for each of the three options in each choice set. Each product is described by eight attributes. As the first has 8 levels, the second has 3 levels and so on we choose the first option as a combination from a $(8 \times 3 \times 5 \times 4 \times 4 \times 8 \times 3 \times 3)$ factorial. This is true for each of the other options in each choice set. Thus we have found the choice sets by constructing a resolution 3 fraction of an $(8 \times 3 \times 5 \times 4 \times 4 \times 8 \times 3 \times 3) \times (8 \times 3 \times 5 \times 4 \times 4 \times 8 \times 3 \times 3) \times (8 \times 3 \times 5 \times 4 \times 4 \times 8 \times 3 \times 3)$ factorial, where the brackets indicate the three options and of course play no substantive role in determining the fraction. Thus we are using a modification of the L^{MA} construction technique, where we have $M=3$ options in each choice set and $A=8$ attributes for each option, but the attributes do not all have L levels.

Indeed from the mathematical perspective the order of the attribute levels is immaterial and thus we want to find a resolution 3 fraction of the $8^6 \times 5^3 \times 4^6 \times 3^9$ complete factorial. If we work directly with these values and insist on equal occurrence of pairs of attribute levels for all pairs of attributes (to ensure a diagonal covariance matrix), then the smallest design would have $N = 9 \times 25 \times 64 = 14400$ choice sets which was not a tractable number. Instead we decided that we would find a resolution 3 fraction of a $8^9 \times 4^{15}$ complete factorial and then collapse three 8-level attributes to obtain the 5-level attributes and collapse nine 4-level attributes to obtain the 3-level attributes. The initial design that we used was the $2^6 \times 4^{12} \times 8^{10} \times 16$ with 128 treatment

combinations given in Kuhfeld (2006). We then replaced the 16 level attribute by three 4-level attributes to give a $2^6 \times 4^{15} \times 8^{10}$. We then removed one of the 8-level attributes and all of the 2-level attributes. The most common levels were chosen to have higher replication to increase the plausibility of options. These 128 choice sets were blocked into 4 versions each with 32 choice sets. Respondents were randomly allocated to one of these versions. Although the theoretical covariance matrix from such a collapsing is known to be non-diagonal (see Street and Burgess 2007 for a comprehensive discussion), we performed simulation studies that confirmed that we were able to accurately recover a range of assumed prior values of the parameters using this design.

One consequence of this method of construction is that in some choice sets a product will appear twice. This not only reflects the reality of a consultation (in which two different brands of a product might be presented), but from a design perspective is perfectly acceptable, as a fraction was chosen in which the levels of the attributes for the two occurrences were not the same and so two different items were available in the choice set. In addition, in the DCE that we constructed the generic attributes were not in fact nested within product (although as described above the construction method might have made it appear that they were) and so the main effects of the generic attributes could be, and indeed were, fitted across the products. Simulation studies confirmed that the two-factor interaction effects of all of the attributes with products could be estimated if desired and a range of assumed prior values could be recovered.

3.4 Choice task

As noted above, each choice set presented three product options of which respondents were asked to choose the best and the worst based on the described attributes of each option. Respondents were then asked whether they would choose the selected best option over their current method of contraception. Figure 1 depicts a representative example of a choice set. We restrict our attention to the initial decision to choose the

best option amongst the three alternatives. Analysis of the other choice data will be pursued in later work¹.

The experiment had labelled options with the 8 levels of product as the option labels. Since there were only 3 options per choice set, not all products were available in each choice task. Therefore a woman's preference for a particular product could only be expressed in those choice sets where the preferred product was one of the three options.

3.5 Recruitment and data collection

The DCE was completed on-line. Respondents were recruited from an on-line panel by a market research company. Women aged 18-49 years, currently using or planning to use contraception to prevent pregnancy in the next 5 years were eligible to participate.

¹ The experimental design was repeated over three information conditions and a woman was randomised to complete only one of these conditions. The conditions were (a) basic information condition: women in this arm of the DCE were given information about the products and attributes to assist them with the experiment without any particular emphasis on any product. (b) A "scare" condition: as well as the basic information on each product, women were shown a short article on the risks associated with the use of the contraceptive pill before completing the DCE. (c) A "promotion" condition: as well as the basic information on each product, women were shown a short promotional article on the features of the patch and vaginal ring before completing the DCE. In this paper we report results only for the basic information condition.

4. Econometric analysis

4.1 The Generalized Multinomial Logit Model

The analysis of the choice data relies on a random utility model where the utility that woman i derives from choosing alternative j in choice scenario s is given by

$$(1) \quad U_{isj} = X'_{isj}\beta + \varepsilon_{isj}; i = 1, \dots, n; j = 1, \dots, J; s = 1, \dots, S$$

where X_{isj} is a $K \times 1$ vector of explanatory variables and β is a conformable vector of coefficients. Assuming the disturbance terms ε_{isj} to be identically and independently distributed as extreme value, the standard multinomial logit (MNL) specification results. Simplicity of estimation and interpretation are among the main advantages of this model but these come at the cost of some restrictive assumptions that are clearly unrealistic in our choice context.

Random parameter specifications such as mixed logit (MXL) have become very popular in choice modelling applications; see for example Revelt and Train (1998), Brownstone and Train (1999), Hall et al (2006) and Hole (2008). They provide an appealing and tractable way to extend basic binary and multinomial choice models, and the development of widely available software has made MXL an attractive estimation approach for empirical research. Random parameter choice models have two compelling justifications. First, preference heterogeneity which is pervasive in modelling individual choice behaviour is quite naturally formulated in terms of random parameter models. Second, these specifications provide a potential means of capturing more flexible variance-covariance structures for the unobservables in choice models and hence avoid problems associated with independence of irrelevant alternatives (IIA). This second justification is especially relevant in the current analysis where predicting new product demand is a key component of our analysis. Such an analysis requires a model that allows for flexible substitution patterns across very different alternatives.

Here we adopt the generalized multinomial logit (G-MNL) framework recently developed in Fiebig, Keane, Louviere and Wasi (2010). This framework includes standard MXL specifications as special cases but additionally allows for individual

scale heterogeneity. The work of Fiebig, Keane, Louviere and Wasi (2010) highlight the empirical importance of accommodating this extra dimension of heterogeneity.

In order to motivate the G-MNL model, first extend (1) by allowing for scale heterogeneity, that is rather than assuming ε_{isj} has a fixed scale (or variance), we allow it to vary across respondents to obtain:

$$(2) \quad U_{isj} = X'_{isj} \beta + \varepsilon_{isj} / \sigma_i; i = 1, \dots, n; j = 1, \dots, J; s = 1, \dots, S$$

where the scale has been denoted by the scalar σ_i . Now rewrite (2) to obtain an equivalent formulation:

$$(3) \quad U_{isj} = X'_{isj} (\beta \sigma_i) + \varepsilon_{isj}; i = 1, \dots, n; j = 1, \dots, J; s = 1, \dots, S.$$

The scale heterogeneity model is observationally equivalent to a particular type of heterogeneity in the utility weights, an observation that has led Louviere et al (2008) to be critical of the standard MXL model. At the very least one should be aware that apparent preference heterogeneity discovered through a random parameter MXL specification might be better modelled as the result of scale heterogeneity.

The G-MNL model accommodates both sources of heterogeneity. For example, taste heterogeneity or “residual” taste heterogeneity can be introduced by extending (3) to obtain:

$$(4) \quad U_{isj} = X'_{isj} (\beta \sigma_i + \eta_i) + \varepsilon_{isj}; i = 1, \dots, n; j = 1, \dots, J; s = 1, \dots, S$$

whereby the so-called G-MNL-I model results. The full G-MNL model of Fiebig, Keane, Louviere and Wasi (2010) provides flexibility in how scale and taste heterogeneity are combined. Implementation within this general framework still requires answers to important specification questions related to how scale heterogeneity is captured, what other parameters are assumed random, what distributions to choose to represent taste heterogeneity and whether to allow correlations across random coefficients. These important issues are discussed below.

4.2 Model specification

One attraction of a model of scale heterogeneity, as is specified below, is a considerable amount of flexibility with the addition of only one parameter. Because σ_i is required to be positive we specify an exponential transformation given by:

$$(5) \quad \sigma_i = \exp(\bar{\sigma} + \tau \varepsilon_{0i});$$

where $\varepsilon_{0i} \sim N(0,1)$ and to ensure identification of β we set $\bar{\sigma} = -\tau^2 / 2$.

The additional parameter τ provides a measure of scale heterogeneity. A τ of zero would imply that the G-MNL model reduces to a standard MXL specification.

In its most general form, the G-MNL model can involve a large number of parameters. For example, with the number of attributes in our study, allowing all their coefficients to be random and correlated would require the estimation of over 600 parameters. A more parsimonious specification is clearly required. As our other primary concern is capturing flexible substitution patterns, alternative specific constants (ASCs), that here represent the different contraceptive products, are assumed to be random and correlated. Fiebig, Keane, Louviere and Wasi (2010) argue that ASCs are fundamentally different from most observed attributes and as such recommend that they not be subject to scaling. This specification, together with scale heterogeneity as defined in (5) applied to other attributes, are combined to form our base model:

$$(6) \quad U_{isj} = (\beta_{0j} + z_i' \gamma_j + \eta_{0ij}) + X'_{isj} (\beta \sigma_i) + \varepsilon_{isj}; i = 1, \dots, n; j = 1, \dots, J; s = 1, \dots, S$$

where X is now interpreted as including observed attributes exclusive of ASCs and we have allowed for interactions between the ASCs and individual specific characteristics. In what follows we simply refer to (6) as the G-MNL model.

In this framework there is flexibility in the choice of the distribution of the vector of ASC random components η_{0i} . Here a multivariate normal distribution with zero mean and covariance matrix Σ is assumed. Assuming normality seems sensible because

there are no strong priors on the signs of these coefficients and importantly allowing correlation introduces error dependence across the alternatives in each choice situation. Also, assuming random effects to be individual specific induces correlation across choice situations, accounting for the dependence structure in unobserved utility among the repeated choices of a woman which comes from the panel structure of the data. This would be expected, since the same unobserved factors affect a specific respondent, to a certain degree, over the repeated choices. This correlation is not perfect because of the presence of the independent extreme value terms ε_{isj} .

Estimation by maximum simulated likelihood (MSL) is undertaken using a program kindly provided by Dr Nada Wasi. All estimation results reported below were generated using 1000 Halton draws to simulate the likelihood functions to be maximized (Train, 2003).

5. Estimation

5.1 Data

In total, 700 eligible women entered the study and 528 completed the survey, giving a completion rate of 75%. Of those that completed the survey 173 were allocated to the basic information condition but two of these respondents were subsequently dropped because of missing covariate information. Current contraceptive use and socio-demographic variables for the entire sample and for the 171 respondents assigned to the basic information sample are summarized in Table 2.

The mean age of the respondents in the basic information condition was 30 years. The median household income was over \$67,000 pa and 44% were in full-time employment. Sixty-seven percent were currently using some form of contraception, 58% were currently using the pill, 54% were using condoms (50% were using pill only and 24% condoms only). Twelve percent of women used a non-oral prescribed method of contraception, 4% were currently using DepoProvera injection and 4.7% were using an implant. Seventeen percent of women had been advised against the use of a particular method of contraception for health reasons and 9.8% had been advised against using the pill.

When compared with the sub-population of Australian women aged 18-49 the survey sample was somewhat younger than the sub-population, with 83% of women in the sample in the age group 18 to 39 years, compared with 67% in the sub-population (ABS, 2006). The study only included women who were either using or were likely to use contraception in the near future, thus excluding a larger proportion of women in the older age group. The women's current use of contraceptive methods was typical of other samples of Australian women; compare Tables 2 and 3.

5.2 Estimation results

In the models estimated, the base model given by (6) includes the women's income and their current contraceptive method as covariates interacted with each of the product effects. In the case of current method, women were classified into three mutually exclusive and exhaustive groups: a base group who do not use any prescribed method (i.e. condoms only, emergency, other non-prescribed or no method), those who only use the pill and those who use a prescribed method other than the pill. As is common in studies similar to this, the addition of random individual-specific effects renders insignificant most effects associated with socio-demographic characteristics. So while much more detailed information is available for these women, these effects cannot be precisely estimated if added to our model and we retain only this small set of covariates. The base model also includes necessary normalizations whereby the ASC for the IUD and associated interactions with socio-demographic characteristics are set to zero.

Table 4 provides a comparison of fit across several different models. Starting with a basic MNL model, the separate addition of either correlated random ASCs (MXL) or scale heterogeneity (G-MNL-A) leads to a dramatic improvement in fit. Recall that the latter comes at the expense of only a single extra parameter but even after making degrees of freedom adjustments using either AIC or BIC, MXL fits better than G-MNL-A.

G-MNL-B and G-MNL-C incorporate both sources of heterogeneity but the former does not allow for correlated product effects. Our expectation was that these effects would need to be correlated in order to capture differential substitution patterns and

here we find confirmation, at least in terms of fit. Thus G-MNL-C is our preferred model providing the best fit amongst the models considered as measured by AIC, BIC or Pseudo- R^2 .

In these data it seems that both scale and preference heterogeneity are present. To explore further the relative contributions of these two sources of heterogeneity a comparison of model fit is provided in terms of each woman's log-likelihood contribution. In Figure 2, these contributions for MXL, G-MNL-A and G-MNL-C are overlaid, ordered by their contribution according to G-MNL-C. As in a similar figure provided in Fiebig, Keane, Louviere and Wasi (2010), there appear to be three distinct segments. In the left most segment of the figure, where G-MNL-C fits the data best, G-MNL-C dominates both MXL and G-MNL-A and often by a substantial margin. In the middle section MXL and G-MNL-C are somewhat comparable in terms of fit. But the performance of G-MNL-A is more variable in this middle section; sometimes it does well compared to G-MNL-C but often it does very poorly. In the right most segment, where the relative fit of G-MNL-C is poorest, G-MNL-C is again seen to dominate MXL. However, in this last segment G-MNL-A does well, dominating MXL and often being superior to G-MNL-C.

These results indicate that scale heterogeneity better enables us to model extremes of behaviour. This is most pronounced in the last segment where choice behaviour is highly random and women's choices are little affected by attribute values. In the first segment adding scale heterogeneity can often dramatically improve the performance of MXL. However, scale heterogeneity is clearly not the entire story and for the majority of the sample scale heterogeneity alone does not provide a good fit of individual choice behaviour. As might be expected for this particular choice problem, scale heterogeneity is important but preference heterogeneity is relatively more important for explaining behaviour over the entire sample of women.

A full set of estimated parameters for the preferred G-MNL-C model are provided in Table 5. The estimated value of τ in this model is 0.781 with an estimated standard error of 0.085. Thus a Wald test indicates that τ is significantly different from zero

and further confirms the need to account for scale heterogeneity even after allowing for correlated random product effects.

There are two key inferences to be drawn from the estimated means and standard deviations associated with the random product effects. First, all of the estimated mean product coefficients are positive except for the ring but only the effect for the injection is significantly different from zero at the 5% level. These coefficients need to be interpreted relative to the base product choice that was the IUD. This implies that after controlling for attribute differences, on average, there is little in the way of additional perceived product characteristics that lead women to prefer these products when compared to an IUD.

Because current method and income has been interacted with the product ASCs, these inferences about mean effects relate to the base case of no prescribed method and no income. But while there are some distinct patterns in how the estimated product effects vary with the covariates, none of these individual effects are significant at the 5% level.

These are comparisons of average effects. The second key inference comes from considering the estimated standard deviations associated with the random product effects where our results suggest considerable variation around these mean tendencies. All estimated product effect standard deviations are large relative to their respective estimated means coefficients and all except that associated with the hormonal IUD are precisely estimated.

In order to facilitate interpretation of the estimates associated with the remaining attributes, Figure 3 provides estimates of marginal willingness to pay (MWTP) calculated by dividing the attribute coefficient estimates by the estimated price coefficient, with the resultant ratios converted back into (Australian) dollars. Comparisons are easily made across different attributes as they represent, in dollar terms, the amount women are willing to trade off different attribute levels relative to the base level. To enhance comparability across attributes, the bases for the frequency attributes for injection, patch and ring and IUD implant were changed from the lowest to the highest period; so that the most preferred level is now the base for all attributes.

Note that none of the small negative estimates that appear in Figure 3 are associated with estimated coefficients that were precisely estimated.

Recall from Table 1 that attributes have varying numbers of levels. Also note that there is nesting within the attributes “effect on bleeding” and “effect on weight” and hence for these two attributes the estimates of MWTP are broken up into like blocks. For the bleeding attribute, irregular bleeding and heavy and light periods with different pain levels are each compared to the common base of no period. Similarly different levels of weight gain and loss are compared to the common base of no effect on weight.

It is clear that women are very concerned about bleeding and period pain. Compared to the base level of no period, the estimates indicate women are prepared to pay a substantial amount to avoid irregular bleeding and heavy periods associated with any level of pain. This is also the situation when a light period is combined with increased pain but women are not concerned about a light period with either less or no change in pain. In terms of other side effects, women want to avoid worsening acne and any weight gain but they place little value on improving acne or losing weight.

Doctor recommendations also have a substantial impact on women’s choices. The difference between a positive recommendation (the base) and a negative recommendation has the second highest MWTP of \$41.

As noted earlier, effectiveness and frequency were defined differently for different products. Women were willing to pay substantial amounts to avoid the declining effectiveness of the pill, patch and ring defined as chance of pregnancy over 12 months going from the base of 1 in 100 to 5 in 100 to 10 in 100. They were willing to pay similar amounts for avoiding declining effectiveness of the IUD, implant and injection going from 1 in 1000 to 1 in 500 and then to 1 in 100.

Relative to effectiveness, there was more variation across product groups in terms of frequency of administration. For the IUD and implant, there was almost no difference in terms of MWTP between administration every one, three or years. Women were willing to pay to avoid more frequent administration of the injection, patch and ring

but were indifferent between one and three months compared to the base of six months. For the pill women were willing to pay to avoid having to take the pill at particular times but were indifferent between taking it at the same time or within a particular interval.

6. Simulations

These estimation results can now be used to predict the market share potentially captured by the new products, the hormonal skin patch and hormonal vaginal ring, if they were introduced into the current Australian market. Note that the current market is interpreted to be at the time of the survey which was conducted in 2007. Since this time, the ring has actually been introduced into the Australian market.

6.1 Calibration

The method described by Train (2003) was used to implement the simulations. First an out of sample data set was created that included the sub-set of products on the Australian market at the time of the survey. This subset contained the six products: combined pill, the minipill, the implant, the injection, the hormonal IUD and the IUD. The attribute levels for each of these products were set at the levels that most realistically describe that product and the covariates were set at their sample means. Estimates from the G-MNL-C model were used to simulate the predicted probabilities of the current set of products. This provides a predicted market share distribution that will not necessarily reflect the actual distribution of market shares for these products in Australia. Therefore using the method described by Train (2003) we recalibrated the product coefficients for the current product set so that the predicted probabilities estimated from the model more closely reflected the actual market share for each product currently available in Australia. The coefficients for the new products (vaginal ring and hormonal patch) were not constrained in the calibrated model. The distribution of contraceptive products that were currently being used by women in 2005 according to the HILDA study (Gray and McDonald 2007) was selected as the benchmark of market share in Australia.

The recalibration was an iterative process. The ratio of the predicted market share from the model to the actual market share from HILDA was used to recalibrate the product coefficients in the G-MNL-C according to:

$$\alpha_j^1 = \alpha_j^0 + \ln(S_j / \hat{S}_j^0)$$

where:

S_j = actual market share in the forecast population

\hat{S}_j^0 is the predicted market share from the model

α_j^0 is the model coefficient for product j used in prediction of \hat{S}_j^0

α_j^1 is the new constant for product j to be fitted as a constraint in the next model.

The model with the new product effects produced a new predicted share which was compared with the actual market share and new estimates were calculated for the products and the predictions were run again. This process was repeated until a set of estimated product effects was fitted that resulted in a predicted market share that approximated the actual market share. The outcomes of this process are provided in Table 6. Calibration raised the market share of the oral products from 46% to 74%, making it much closer to the population estimate of 82%.

Once the model had been calibrated so that the predicted probabilities of the current product set reflected actual market share, the model was used to predict the probabilities of the full product set including the patch and vaginal ring to simulate the changes in market share with the introduction of new products. Again the attributes for the new products were set at realistic values. The out of sample probability estimates were simulated from 200 Halton draws.

6.2 Simulation results

The graph in Figure 4 depicts the market shares for each contraceptive product under four alternative scenarios. The current situation is that predicted by the calibrated model for the six existing products. The other three scenarios represent the predicted

market shares if the hormonal patch and vaginal ring were introduced together or separately into the current Australian market.

The model predicts that the hormonal patch would achieve around 14% share of the market, whether introduced alone or with the vaginal ring. Relatively more of the share for the patch would come from the oral pill products, relatively less from the injection and the implant and the market for the Hormonal IUD and IUD would be relatively unaffected. The model predicts that the vaginal ring would capture approximately 3% of the current market share, in this case drawing more equally from other product shares. This market share for the ring is not sensitive to whether it is introduced alone or in conjunction with the patch. Together these results indicate that the patch and the ring appeal to somewhat different sectors of the market since introducing them together did not appreciably affect their separate market shares.

The relative lack of appeal of the ring suggests some further simulations in order to determine the impact of making the ring a more attractive option. Two possible situations are considered: (i) the introduction of the ring together with a promotional program which is implemented in the simulation by specifying a recommendation for the ring by the GP; and (ii) allowing a price discount for the ring consistent with the product being subsidized under Australia's Pharmaceutical Benefits Scheme. All other attribute levels remain the same as the initial simulation and the new results are compared to a base case from Figure 4 where both the ring and patch are introduced. The results are presented in Figure 5 for three alternative scenarios: promotion and price reduction separately and then in combination.

While the predicted increases in market share for the vaginal ring are large in relative terms, the actual shares remain small. When the ring is promoted and has a substantial price discount its market share is still only 6.7% compared to the predicted share for the patch of 12.6%. Because the induced increase in the demand for the ring is small in magnitude, the impact on the other market shares are also small in magnitude but proportionately the largest declines, across each scenario, are seen for the injection and the hormonal IUD.

7. Conclusion

Our results have provided a clear picture of what women value when making choices between alternative contraceptive products. In our econometric modeling we also stressed the need to specify a model allowing flexible substitution patterns between products. This was important because there are likely to be very different patterns of substitutability across existing products but also because a key objective of our work was to simulate the introduction of the hormonal patch and vaginal ring into the current Australian market. As expected these substitution patterns were very different from the proportional substitution patterns that would arise using say conditional logit models. One distinctive finding from these simulations was the lack of sensitivity of the market shares of the new products across scenarios whether they were introduced separately or together. This is clear evidence that the ring and patch appeal to different segments of the market. The other important finding was the clear preference for the patch over the ring even when the latter was promoted and made cheaper.

The current work has concentrated on the preferences of women but as with many health-related decisions, doctors can exert considerable influence on the decisions of patients through their recommendations. Doctor recommendations have been included amongst our set of attributes and hence have been exogenously given. It is of interest to investigate whether the preferences of doctors align with those of women and how doctors decide on what types of contraceptive products to discuss with women. These questions are left for future research.

Table 1: Attributes and levels of the contraceptive product options

Attribute	Level
product	0 Combined Pill
	1 Mini pill
	2 Depot Injection
	3 Implant
	4 Intra-uterine hormonal device
	5 Patch
	6 Ring
7 IUD	
Generic attributes	
effect on acne	0 This product has no effect on acne symptoms
	1 In some women, this product improves acne symptoms
	2 In some women this product worsens acne symptoms
effect on weight	0 This method has no effect on weight
	1 Some women using this method may lose up to 1 kg in weight
	2 Some women using this method may gain up to 1 kg in weight
	3 Some women using this method may lose up to 3 kg in weight
Doctor's recommendation	4 Some women may gain up to 3 kg in weight
	0 Your doctor recommends this product for you
	1 Your doctor says this may not be a good method for you
	2 Your doctor says this is a suitable method for you
cost	3 Your doctor makes no recommendation
	0 \$1 per month
	1 \$7 per month
	2 \$20 per month
effect on bleeding	3 \$60 per month
	0 Most women using this method experience no periods
	1 Most women using this method experience irregular bleeding
	2 Most women using this method experience light periods with less pain
	3 Most women using this method experience light periods with no change in pain
	4 Most women using this method experience light periods with increased pain
	5 Most women using this method experience heavy periods with less pain
6 Most women using this method experience heavy periods with no change in pain	
7 Most women using this method experience heavy periods with increased pain	
Alternative specific attributes (nested)	
frequency of administration pill	0 1 per day
	1 1 per day within interval
	2 1 per day at same time
frequency of administration injection, patch and ring	0 Once a month
	1 Once every 3 months
	2 Once every six months
frequency of administration IUD implant	0 Once a year
	1 Once every 3 years
	2 Once every 5 years
contraceptive effectiveness implant, injection and IUD	0 1/1000 women using this product get pregnant in a 12 month period
	1 1/500 women using this product get pregnant in a 12 month period
	2 1/100 women using this product get pregnant in a 12 month period
contraceptive effectiveness pill, patch and ring	0 1/100 women using this product get pregnant in a 12 month period
	1 5/100 women using this product get pregnant in a 12 month period
	2 10/100 women using this product get pregnant in a 12 month period

**Table 2: Sample characteristics:
Total sample and basic information treatment***

	Total sample	Condition 1: Basic information
Sample size	528	171
	%	%
Age (years)		
18-19	5.13	5.29
20-24	22.05	25.29
25-29	19.39	21.76
30-34	19.20	17.06
35-39	17.68	13.53
40-44	11.60	11.76
45-49	4.94	5.29
Household Income		
No income	2.29	2.92
\$1-\$149	0.19	0.58
\$150-\$249	2.48	1.75
\$250-\$399	4.00	4.09
\$400 - \$599	6.86	9.94
\$600 - \$799	8.95	8.19
\$800 - \$999	9.90	8.19
\$1000 - \$1299	15.43	12.87
\$1300 - \$1599	17.52	20.47
\$1600 - \$1999	14.48	13.45
> \$2000	17.90	17.54
Employment		
Full time	43.81	44.97
Part time	29.33	31.95
Unemployed	26.86	23.08
Smoking status		
Regular	21.48	20.7
Occasional	5.32	7.10
Non-smoker	73.19	71.35
Current Contraception		
Oral pill	50.57	57.89
Condoms	52.08	53.80
Emergency	6.06	5.85
Implant	5.49	4.68
Vaginal Ring	0.95	0.00
IUD	2.65	1.75
Injection	4.17	4.09
No effective	14.39	14.62

*There were a small number of missing observations in the total sample. Sample statistics were calculated over complete cases.

Table 3: Patterns of contraceptive use in Australia*

	All	Using contraception	Using prescribed contraceptive
Sample size	3044	1847	929
	%	%	%
Oral	25.3	41.3	82.1
Condom	19.3	31.5	
Tubal/hysterectomy	8.4	13.8	
Vasectomy	8.8	14.5	
Intra-uterine device	1.4	2.3	4.6
Injectable	1.8	2.9	5.8
Implant	2.3	3.8	7.6
Withdrawal	2.1	3.5	
Safe period method	1.4	2.2	
Other	1.0	1.6	

* Source: HILDA 2005

Table 4: Comparison of model fit*

	MNL	MXL	G-MNL-A	G-MNL-B	G-MNL-C
Random scale	No	No	Yes	Yes	Yes
Random ASCs	No	correlated	No	uncorrelated	correlated
Log-likelihood	-5393.3	-4932.5	-5260.9	-5030.6	-4836.3
Parameters	55	83	56	63	84
AIC	10896	10031	10633	10187	9840
BIC	11259	10579	11003	10603	10395
Pseudo- R^2 *	0.093	0.170	0.115	0.154	0.186

*Null model for comparison is the product intercepts only MNL model with log likelihood = -5942.94.
All models estimated with attributes and selected socio-demographic variables.

Table 5: G-MNL-C estimation results

		Coefficient	Std. Err.	p-value
Random Product Intercepts				
combined pill	Mean	0.645	0.536	0.23
	Std Deviation	2.130	0.221	0.00
minipill	Mean	0.347	0.528	0.51
	Std Deviation	2.112	0.223	0.00
injection	Mean	0.910	0.451	0.04
	Std Deviation	0.964	0.225	0.00
implant	Mean	0.363	0.350	0.30
	Std Deviation	0.854	0.228	0.00
hormonal IUD	Mean	0.175	0.290	0.55
	Std Deviation	0.355	0.242	0.14
Patch	Mean	0.463	0.520	0.37
	Std Deviation	1.685	0.222	0.00
ring	Mean	-0.543	0.541	0.32
	Std Deviation	1.278	0.316	0.00
Fixed parameter attributes				
effect on acne				
	improves acne symptoms	0.030	0.070	0.67
	worsens acne symptoms	-0.367	0.059	0.00
effect on weight				
	lose up to 1 kg in weight	-0.064	0.076	0.40
	gain up to 1 kg in weight	-0.304	0.079	0.00
	lose up to 3 kg in weight	0.062	0.086	0.47
	gain up to 3 kg in weight	-0.509	0.094	0.00
frequency of administration				
	1 per day within interval	-0.207	0.154	0.18
	1 per day at same time	-0.135	0.163	0.41
	Once every 3 months	0.003	0.125	0.98
	Once every six months	0.217	0.132	0.10
	Once every 3 years	0.062	0.129	0.63
	Once every 5 years	0.061	0.121	0.62
contraceptive effectiveness				
	1/500 women get pregnant in a 12 month period	-0.205	0.091	0.02
	1/100 women get pregnant in a 12 month period	-0.534	0.115	0.00
	5/100 women get pregnant in a 12 month period	-0.336	0.098	0.00
	10/100 women get pregnant in a 12 month period	-0.597	0.109	0.00
Doctor's recommendation				
	may not be a good method for you	-0.805	0.102	0.00
	suitable method for you	-0.197	0.095	0.04
	no recommendation	-0.428	0.087	0.00
	linear cost (\$100's)	-1.978	0.201	0.00
effect on bleeding				
	irregular bleeding	-0.743	0.117	0.00
	light periods with less pain	0.023	0.102	0.82
	light periods with no change in pain	-0.007	0.101	0.95
	light periods with increased pain	-0.635	0.108	0.00
	heavy periods with less pain	-0.540	0.115	0.00
	heavy periods with no change in pain	-0.631	0.095	0.00
	heavy periods with increased pain	-1.378	0.160	0.00
Covariates				
Household weekly income midpoints (\$1000's)				
	Effect on combined pill	-0.474	0.339	0.16
	Effect on minipill	-0.393	0.327	0.23

Effect on injection	-0.460	0.237	0.05
Effect on implant	-0.191	0.187	0.31
Effect on hormonal IUD	-0.166	0.179	0.35
Effect on patch	-0.126	0.321	0.69
Effect on ring	-0.111	0.300	0.71
Current method = other than pill			
Effect on combined pill	-1.219	1.056	0.25
Effect on minipill	-1.254	1.027	0.22
Effect on injection	0.092	0.695	0.90
Effect on implant	0.224	0.601	0.71
Effect on hormonal IUD	-0.130	0.664	0.85
Effect on patch	-1.466	0.874	0.09
Effect on ring	-0.950	0.902	0.29
Current method = pill only			
Effect on combined pill	0.314	0.446	0.48
Effect on minipill	0.443	0.428	0.30
Effect on injection	0.000	0.358	1.00
Effect on implant	0.181	0.317	0.57
Effect on hormonal IUD	0.166	0.275	0.55
Effect on patch	-0.423	0.420	0.31
Effect on ring	-0.091	0.432	0.83
Scale heterogeneity (τ)	0.781	0.085	0.00
Log-likelihood	-4836.3		
Number of observations	5472		

Table 6: Calibration results

	Unconstrained		Calibrated		Actual
	Product effects	Predicted market shares	Product effects	Predicted market shares	Market share*
Combined Pill	0.64	0.18	5.03	0.37	0.41
Mini pill	0.35	0.16	4.84	0.37	0.41
Injection	0.91	0.23	1.49	0.07	0.06
Implant	0.36	0.12	2.53	0.09	0.08
Hormonal IUD	0.18	0.21	-2.07	0.07	0.02
IUD	0.00	0.12	0.00	0.03	0.02

* Based on HILDA where rates for the oral pill have been allocated equally between the combined pill and the minipill. Similarly the HILDA rates for the intra-uterine device have been equally allocated to the hormonal and non-hormonal IUD.

Figure 1: Example of a choice set

Compare the characteristics of options A B & C then answer the questions below			
	A	B	C
<u>Product</u>	Mini Pill	Intra-uterine Hormonal Device	Mini Pill
<u>Effect on Acne</u>	In some women, this product improves acne symptoms	In some women this product worsens acne symptoms	In some women this product worsens acne symptoms
<u>Effect on Weight</u>	This product has no effect on weight	Some women using this product may gain up to 1 kg in weight	Some women using this product may lose up to 1 kg in weight
<u>Frequency of Administration</u>	1 per day	Once every 5 years	1 per day within interval
<u>Contraceptive Effectiveness</u>	10/100 women using this product get pregnant in a 12 mth period	1/500 women using this product get pregnant in a 12 mth period	1/100 women using this product get pregnant in a 12 mth period
<u>Doctor's Recommendation</u>	Your doctor says this may not be a good method for you	Your doctor makes no recommendation	Your doctor says this is a suitable method for you
<u>Effect on Periods</u>	Most women using this method experience heavy periods with less pain	Most women using this method experience heavy periods with less pain	Most women using this method experience heavy periods with less pain
<u>Cost</u>	\$7 per month	\$20 per month	\$20 per month
Of the options presented above, which do you like the most?			
Of the options presented above, which do you like the least?			
If you had to choose from the option above that you like the most, or your current method of contraception which would you choose?			

Figure 2: Individuals' Log-likelihood comparison

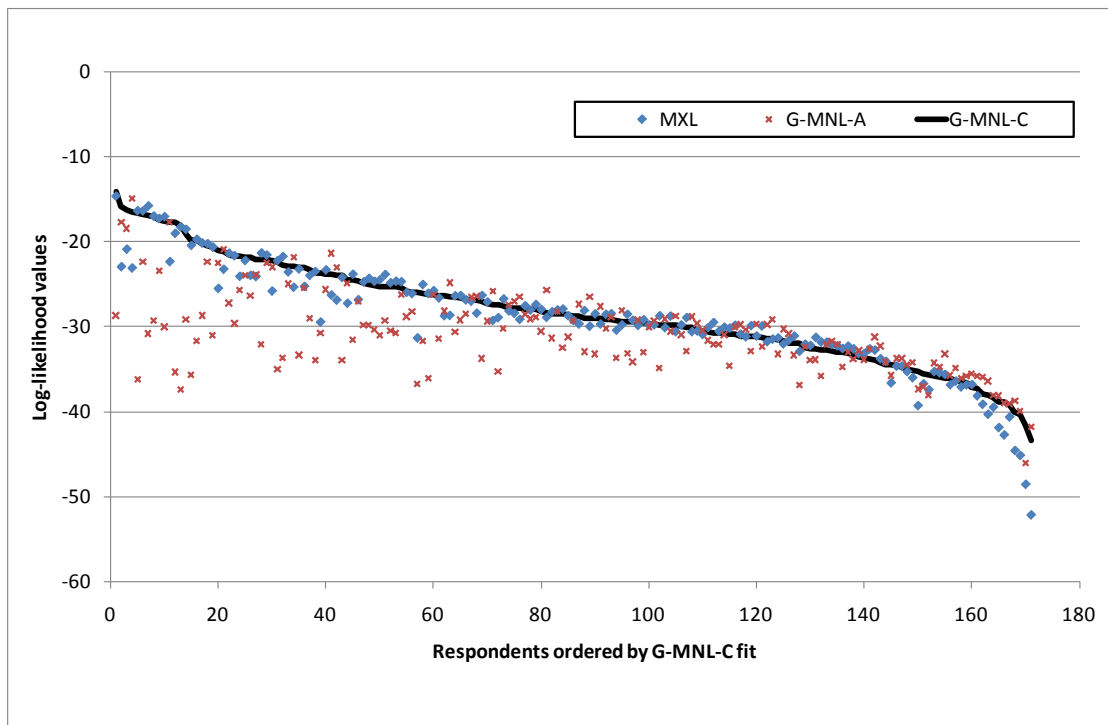


Figure 3: Relative impact of contraceptive attributes on women's choices: Marginal WTP relative to base case (\$A)

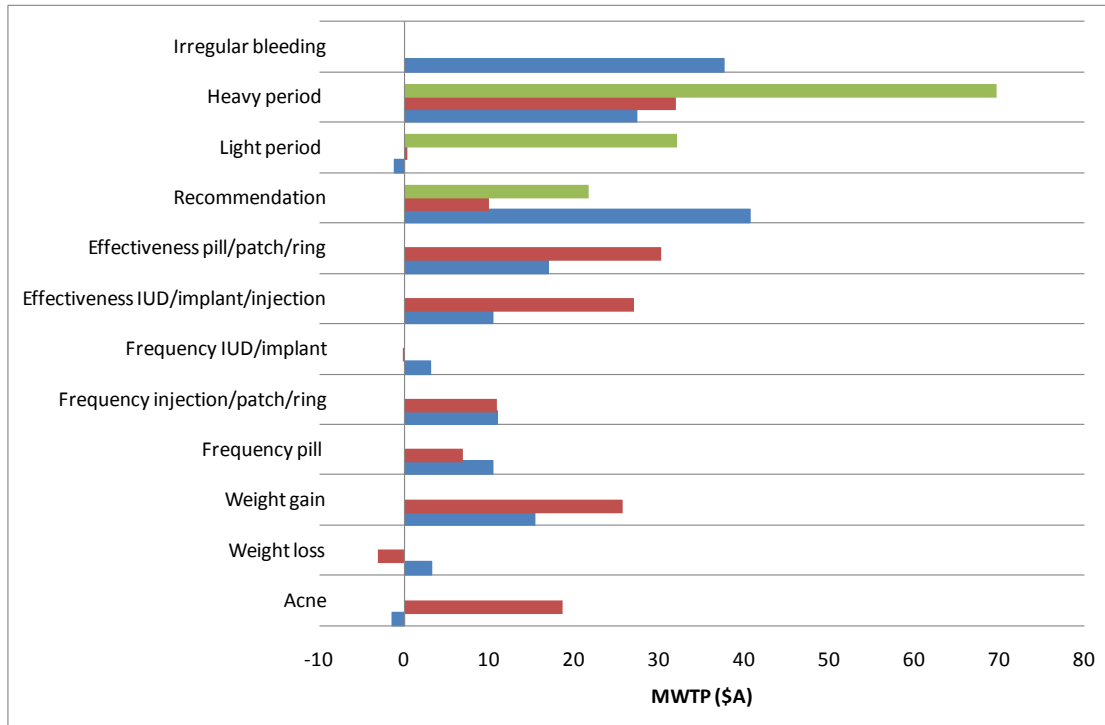
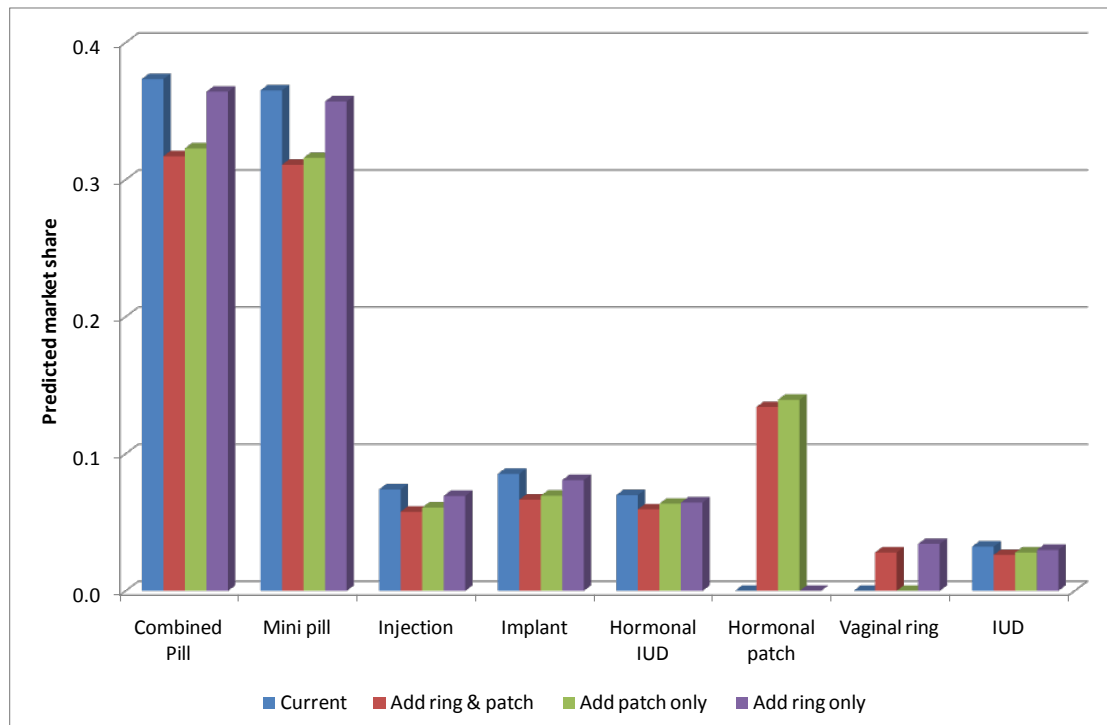
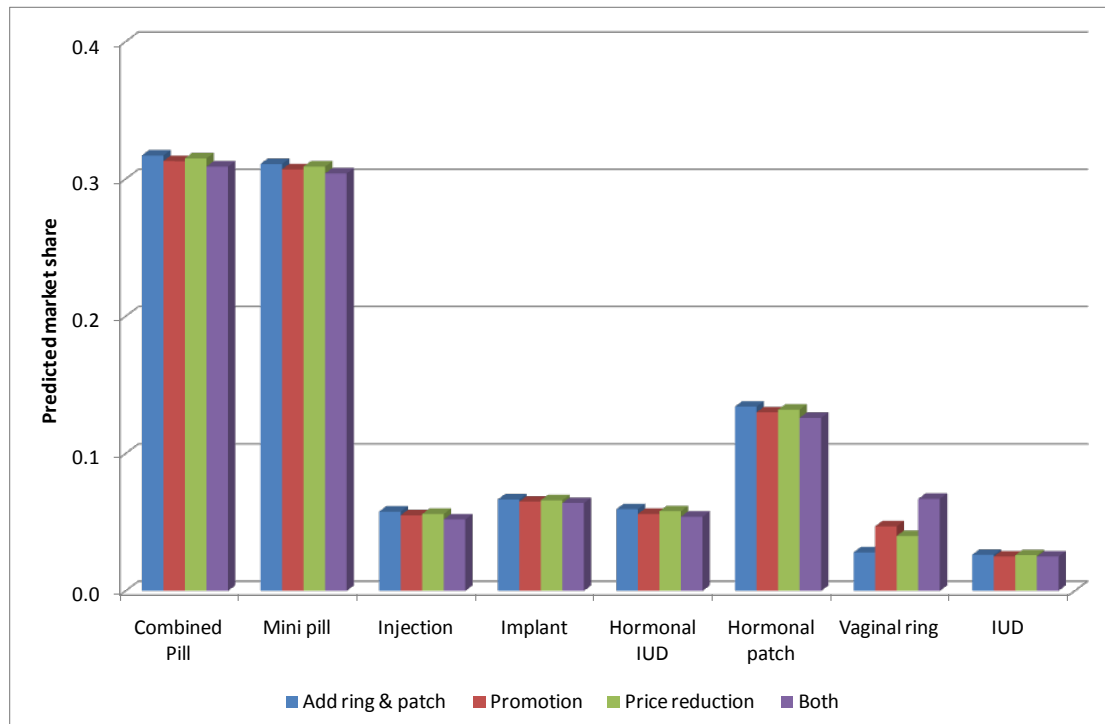


Figure 4: Forecasts of market shares after introduction of new products*



* Current situation represents market shares calibrated to represent market existing in 2007.

Figure 5: Forecasts of market shares after introduction of new products with promotion and price incentives for the vaginal ring*



* The promotion involves a positive GP recommendation for the ring and the price reduction is consistent with a subsidized price under the Australian PBS.

References

- ABS, (2006), *Australian Historical Population Statistics*. Australian Bureau of Statistics: Canberra. cat. no. 3105.0.65.001.
- Brownstone, D. and Train, K. (1999), "Forecasting new product penetration with flexible substitution patterns", *Journal of Econometrics*, 89, 109-129.
- Delavande, A. (2008), "Pill, Patch or Shot? Subjective expectations and birth control choice", *International Economic Review*, 49, 999-1042.
- Fiebig, D.G., Keane, M.P., Louviere, J.J. and Wasi, N. (2010), "The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity", *Marketing Science*, 29, 393-421.
- Gilliam, M., Holmquist, S. and Berlin A.(2007), "Factors associated with willingness to use the contraceptive vaginal ring", *Contraception*, 76, 30-34.
- Gray, E. and McDonald, P. (2007), "Contraceptive practice and the reproductive life course", *HILDA Survey Research Conference*. Melbourne, 2007.
- Greene, W.H. (2008), *Econometric Analysis*, 6th edition, Prentice Hall.
- Hall, J.P., King, M.T., Fiebig, D.G., Hossain, I., and Louviere, J.J. (2006), "What influences participation in genetic carrier testing? Results from a discrete choice experiment", *Journal of Health Economics*, 25, 520-537.
- Hole, A.R. (2008). "Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment", *Journal of Health Economics*, 27, 1078-1094.
- Kuhfeld, W.F. (2006), *Orthogonal Arrays*, SAS Institute, <http://support.sas.com/techsup/technote/ts723.html>.
- Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000), *Stated choice methods: Analysis and application*, Cambridge University Press.
- Louviere, J.J., Street, D.J., Burgess, L., Wasi, N., Islam, T. and Marley, A.A.J. (2008), "Modelling the choices of individual decision makers by combining efficient choice experiment designs with extra preference information, *Journal of Choice Modelling* , 1, 128-163.
- Novák, A., de la Loge, C., Abetz, L., and van der Meulen, E.A. (2002), "The combined contraceptive vaginal ring, NuvaRingS: An international study of user acceptability", *Contraception*, 67, 187-194.
- Revelt, D. and Train, K. (1998), "Mixed logit with repeated choices: Household choices of appliance efficiency level", *Review of Economics and Statistics*, 80, 647-657.
- Seston, E.M., Elliott, R.A., Noyce, P.R. and Payne, K. (2007), "Women's preferences for the provision of emergency hormonal contraception services", *Pharmacy World and Science*, 29(3), 183-9.
- Street, D.J. and Burgess, L. (2007), *The construction of optimal stated choice experiments: theory and methods*. New Jersey: Wiley.
- Terris-Prestholt, F. (2009), "Will microbicide introduction affect condom use? Analysis of a discrete choice experiment among urban South African women", *iHEA 7th World Congress*. Beijing, China.
- Train, K.E. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press.
- Yusuf, F. and Siedlecky, S. (2007), "Patterns of Contraceptive Use in Australia: Analysis of the 2001 National Health Survey", *Journal of Biological Science*, 39, 734-744.