eAppendix

Identification of Levels

The attributes and levels were chosen based on the available literature and ongoing policy discussion. Travel time was based on average travel times to primary care doctors in an urban area.\(^1\) Wait times to see a primary care doctor were based on average wait times reported in a recent Commonwealth Fund study.\(^2\) Premium differences were based on a Kaiser Health News analysis of U.S. Health and Human Services Department data from 2016 which compares average premium difference between open and closed network plans in 13 different cities across the United States (Atlanta, Charlotte, Chicago, Cincinnati, Las Vegas, Manchester, N.H., Milwaukee, Minot, N.D., New Orleans, Pittsburgh, Portland, Ore., St. Louis, Topeka, Kansas).\(^3\) This data suggested that, in 2016, the premium difference for provider choice was, on average, 53 dollars per month, or 636 per year. We took the mean and the standard deviation of those average premiums to set the levels for the premium in our DCE.

Efficient Design

A DCE is an attribute-based experimental design to collect stated preferences. When a good or service has a multitude of identified attributes and levels, it is impossible to elicit patient’s preferences for all possible combinations of the levels of the attributes. The attribute levels in a full factorial design of our study would give rise to 162 possible scenarios (3*3*3*3*2) that can be combined into 1,291,040 potential choice sets. In practice, it is impossible to have patients rate all possible combinations, so the number of choices must be reduced.

A number of alternatives have been proposed to reduce the dimensions of the design, but the preferred measure within the literature is a D-efficient design which minimizes the determinant of the AVC matrix under the assumption of a vector of prior coefficients \(\beta_n\).\(^4\) Since in most cases one cannot guarantee to have found the optimal design, since this would require evaluating all possible designs, these designs are often referred to as D-efficient instead of D-optimal. Alternative approaches are A-efficient designs which minimizes the trace of the AVC matrix or S-efficient designs which minimize the maximum sample size required for statistically significant parameter estimates. This was not necessary in our case because we had a sample size of nearly one thousand respondents.
We used priors for the parameter estimates based on the literature and from a small pilot study we conducted. For wait time, we used -2, where wage rates represented the opportunity cost of time. The premium prior (-0.005) was based on prior work as well as the prior for breadth of network (0.5) and continuity of care (0.5). The travel time prior (-.04) was based on car travel time and accessibility by bus to general practitioner services. Upon gathering data from a small pilot study, we estimated the parameters and created efficient design. We used Ngene 1.1.2 to create the efficient design as this software allows the inclusion of specific priors in the utility functions. Our algorithm eliminated dominant alternatives (meaning better than the other two choices in that choice situation on all dimensions), repeated choice sets, and choice sets for which the answer can be inferred from the previous one (assuming transitivity and monotonicity).

One limitation in any discrete choice experiment is that the levels of the attributes are arbitrary cut-off points. The tradeoffs presented must be realistic. Therefore, it is important to establish a feasible set of attribute combinations and exclude the potentially implausible combinations. Once feasible sets of plausible combinations are chosen, the analysis can be focused on trade-offs. Green and colleagues (1988) described “unacceptable levels” in detail referring to levels which are sufficiently high or low that the respondent will ignore other attributes and discard that option. The findings of their study indicated that the form of the instructions also matters, but that respondents often ignore attributes altogether when responding to options containing unacceptable attribute levels. Because our premium values were based on 13 large metropolitan areas, predictions in other settings may not be valid. The results of our study should be generalized outside of the attribute range.

Willingness to Pay and Willingness to Wait
The measure of WTP is based on random utility theory. In WTP models in general, the measure of WTP is based on random utility theory estimated by a conditional logit model. WTP is given by the negative of the ratio of the coefficients for any particular attribute and the cost attribute, in this case monthly premium. Estimating the WTP of a particular plan attribute is thus simply estimating the ratio of the coefficients (or: marginal utilities) of the premium cost and the coefficient of the particular attribute we are interested in:

\[
WTP^x_k = \left( \frac{\beta_{attribute_k}}{\beta_{premium}} \right)
\]
In the case of the WTP estimates, we are operating in preference space so we keep the premium coefficient fixed. However, if the scale parameter (the standard deviation of unobserved utility) varies randomly (which is true for wait time), then the utility coefficients are correlated and WTW should be estimated in “WTW space” meaning that the distributions for WTW are specified directly.\textsuperscript{11} To estimate the WTW, we replace the denominator in the above equation with the estimate for the wait time attribute.

\begin{equation}
\text{WTW}_{x}^{k} = \frac{\hat{\beta}_{\text{attribute}_{xk}}}{\hat{\beta}_{\text{waittime}}}
\end{equation}

WTW estimates the relative importance of each attribute in terms of waiting time (days) to see a primary care doctor. All coefficients besides (fixed) premium were random.

The expected variation in WTW (with n dimensional space) is distinguished from the variation in the wait time coefficient, which incorporates scale.

REFERENCES


