Lecture # 17 – Stated Preference Approaches/ Estimating Benefits

I. Stated Preference Valuation

- Conjoint analysis example: (Welling *et al*., *Environmental and Resource Economics* 2022)
    - This paper estimates willingness to pay for climate adaptation investments
    - Used an on-line survey in Bremen, Germany
        - 1178 respondents
    - Socioeconomic information collected:
        - Age, gender, household size, # of children > 14, monthly income
    - Respondents were given different levels of background information
        - Goal: to test how information provided in surveys affects results
        - Randomly divided into two groups
            - *Script* sample given the following introduction:
                - *The Senate of Bremen adopted the climate change adaptation strategy for Bremen in April 2018. The strategy document explains the consequences of climate change for the city of Bremen. Strong rain, river and storm floods will become more likely. The strategy document predicts a rising risk of flooding with property damages, such as flooded basements and underground garages. According to the strategy document, heat waves will also become more likely. These can reduce your productivity and strain your cardiovascular system. The climate change adaptation strategy mentions several measures which the city of Bremen could apply. The first part of this survey focuses on some of these measures.*
        - *No Script* sample given no background on climate change:
            - *The first part of this survey focuses on possible urban green measures for the city of Bremen*
        - The survey was otherwise identical.

- Choice experiment
  - Each survey included 9 choice scenarios
  - Payment mechanism: city administration would implement the policy and collect payments depending on the survey outcome
  - Community characteristics considered:
    - Number of trees on streets
    - Green space
    - Green roofs
  - Provided background information on current level of characteristics in Bremen
  - Asked two questions to verify whether respondents believed whether the changes (policy consequentiality) and the payment (payment consequentiality) would have an impact
    - "To what degree do you believe that your responses will affect which measures will be implemented in the city of Bremen?"
    - "To what degree do you believe that your responses will affect whether you will have to pay the additional cost if the measures are implemented?"
- Results (Average WTP):
  - €29 for one additional tree per 100 meters
  - €23 per additional percentage point of green area
  - €2 per 1 additional "extensive" green roof/100 roofs
  - €12 per 1 additional "intensive" green roof/100 roofs
  - Information matters
    - WTP higher for those receiving background information on climate change
    - More likely to select an alternative
      - 28% chose status quo if given no information
      - 20% chose status quo if given background information
    - Information improved credibility of proposed changes, but did not affect consequentiality

II. Evaluating CV

- Validity assessment
    - Note that bias does not preclude validity
        - Need to consider how much bias is acceptable
        - Virtually all empirical studies have some bias, not just stated preference.
    - Economists consider four types of evaluating validity when assessing contingent valuation. I describe each below.
- Content validity: Do surveys lead to true values? Is best practice being followed?
    - Are the procedures used to design and implement a survey appropriate?
    - Can be evaluated using the guidelines from this article
        - Each study must be evaluated separately
    - Kling et al. point out that best practices are now well known. Three important ones are:
        - The environmental good needs to be defined with a high level of specificity
            - Both status quo and changed level need to be explained in a way accessible to laypeople
        - Constructed market should represent a realistic mechanism for bringing about the proposed change
            - Size of change needs to be physically plausible
            - Context matters: estimated values should not be independent of context
        - Understanding how to encourage and/or test for rationality and truthfulness has evolved
            - Not as simple as using a referendum format
            - Voluntary vs. coercive payment matters
            - Framing the survey to be consequential matters
            - Criteria to evaluate construct validity are case specific and depend on context

- Lessons from behavioral economics
  - Behavioral economics suggests that people are not always rational
  - Behavioral anomalies matter
    - Individual preferences may not be well-behaved in neoclassical sense
    - Individuals do not always optimize when making choices
  - If behavioral anomalies are observed in stated preference outcomes, is it a failure of stated preference methods or of the neoclassical paradigm?
    - Are anomalies because respondents aren't revealing true preferences or because they are simply inconsistent with the expectations of the researcher?
    - Note that true behavioral anomalies are not unique to stated preference. Can occur in revealed preference as well.
  - Types of anomalies
    - Individuals apply simplified decision rules to reduce cognitive burden (e.g. because the good is unfamiliar)
    - Scenario adjustment or rejection
      - E.g. value something other than intended by the researcher
      - Needs to be checked for during survey pre-testing
    - Anomalies due to survey design
      - Anchoring
        - Whether ask for value of a seal and then a whale, or a whale and then a seal, the first value given is around $100.
        - Therefore, WTP depends on order of questions.
      - Therefore, sequencing matters
        - If concerned about sequencing can:
          - Only elicit a single value per survey
          - Randomize order
    - Endowment effect
      - People require more compensation to part with something they already have than they would give up to acquire it
      - Explains divergence in WTP and WTA, which is often cited as evidence of CV's failings
    - Warm glow
      - Now understood as one of many reasons for pro-social behavior such as contributing to public goods
      - Social norms and other-regarding preferences such as altruism and reciprocity also lead individuals to value an environmental good more than its private benefits

- Reasons for departures from optimizing behavior:
    - People may make "mistakes" due to bounded rationality
        - E.g. mental accounting implies that it is not one's overall budget but a specific expense category that matters
        - Evidence: Li *et al* (2005) found that respondents had lower WTP for reduction of global warming when they received reminders about their discretionary income and its use for environmental causes, compared to only receiving reminders about their household budget
        - He and Zhang (2021) find WTP for air quality improvements depends on air quality when the survey is administered
    - Rationality may be the results of repeated participation in markets
        - Mistakes are costly, so people learn
        - Therefore, complex or unfamiliar decision environmental may lead to mistakes resulting from "rule of thumb" behaviors
            - CV studies may be particularly vulnerable to this
- These lessons from behavioral economics help explain anomalies, but do they also raise concerns?
    - If anomalies exist, comparing estimates of surveys and market data may not be valid
        - Must choose a paradigm first (e.g. behavioral or neo-classical) and then design a study to test accuracy of stated preference based on that paradigm
        - If stated preference fails the test, is it because stated preference is flawed or the paradigm is flawed?
    - But if so, how do we make use of the numbers from CV in policy?
- <u>Construct validity</u>: are stated preference estimates consistent with theoretical predictions?
    - Most common way of testing validity
    - Note that some constructs, such as adding-up test, are also failed by marketing studies
        - Thus, shouldn't invalidate the use of stated preference
    - Are the results similar and consistent with theory?
        - Values increase when income increases.
        - Environmentalists WTP more.
            - Hausman and Diamond argue that would happen despite flaws in survey.
            - For example, if measuring general taste for environmental quality.

- Kling et al. note four testable predictions:
  - Number of people willing to contribute to an environmental good in a survey should increase when requested payment falls
  - People should be willing to pay more to have a higher quantity of the good (scope effect)
    - While scope effects missing in some early studies, recent meta-analyses show scope effects present in well-designed studies
  - Income elasticity of WTP should be greater than one
    - Assumes environmental quality is a luxury good
  - WTP and WTA should not be substantially different
    - Observed divergence appears to be because of behavioral anomalies, not design of hypothetical studies
- The first generally holds, but the other three were often violated for stated preference data, particularly in early surveys
- Other problems noted by critics of CV
  - Embedding
    - Suppose 1 group asked about X
    - A second group is asked about Y
    - A third group is asked about X + Y
    - WTP of (1) + (2) should equal (3)
    - However, (1) + (2) often > 3
    - Critics see embedding as evidence of a <u>warm glow</u>.
      - It's not just that people value X or Y, but that people feel good about supporting the environment.
      - When you ask the values individually, the warm glow is added up twice.
      - Thus, adding up results of individual studies leads to double counting of benefits, making aggregation of results for policy difficult.
    - Defenses of embedding
      - Diminishing returns: protecting one lake is valuable. Protecting a second lake isn't as important.
      - Income effects could be a factor.
        - However, WTP is a small percentage of income.
      - Could be large substitution effects between X and Y

    - Kling *et al.* note that more recent work shows that these three predictions are sensitive to two common features of environmental goods:
        - Fixed quantities
            - Demand will be downward sloping
            - But relationship to income depends on:
                - Implied income elasticity of demand
                - Substitutability among all quantity-constrained goods
                - Share of augmented income allocated to market goods
                    - Implies income elasticity of WTP may be less than one even if income elasticity of demand for the fixed quantity is greater than one
        - Limited substitutability with other consumption goods
    - Kling *et al*. caution that as new stated preference techniques are developed, construct validity must be kept in mind
- <u>Convergent validity</u>
    - Are the results similar to results from revealed preference approaches?
    - Formal tests often accept revealed and stated preference equality.
        - Even when statistically different estimates occur, they appear to illustrate common economic phenomena.
    - Not possible for non-use values, but can be used in other cases (e.g. recreation)
        - In these cases, revealed and stated preference studies can be seen as complementing each other.

- Criterion validity: do stated preference estimates match real payments?
    - Compare prediction from stated preference to a standard thought to be a suitable proxy for true measurement
        - Examples include experimental studies and voting studies
    - Two types of lab experiments
        - Induced value experiment: Participants are assigned a value for the experimental good as prat of the research design
            - Advantage: can focus on value elicitation
            - Used to examine accuracy of hypothetical referendum vehicles versus binding real payments votes
            - Generally find that distribution of values from hypothetical votes matches induced-value
        - Homegrown value experiments: participants' actual values for a real commodity used as the criterion
            - E.g. consider a referendum where all participants must pay a given amount if the majority votes in favor
            - Results compared to real payment mechanism as a test of validity
            - Generally find that stated values are higher
                - This is known as hypothetical bias
                - Critique (Kling et al. p. 15): Not all of these studies satisfy the incentive compatibility and consequentiality requirements
                    - Recent work finds hypothetical bias goes away when consequentiality holds
    - Why might results differ?
        - Ongoing work attempting to answer this
            - E.g. asking people if they are certain and recoding uncertain "yes" as "no" improves results
        - Understanding why would make it possible to calibrate answers

III. Estimating Benefits: General Principles

- Having discussed techniques for valuing environmental benefits, we now discuss how to determine what those benefits are.
    - We begin with general principles for determining benefits, then look at two examples: ecosystem services and climate change
- In principle, comprehensive analyses of all environmental effects of a policy are desirable.
    - However, this is difficult.
    - Thus, with some exceptions (such as the social cost of carbon calculations), benefit calculations are done using an "effect-by-effect" approach.
    - Need to avoid double-counting in these cases.
        - For example, stated value techniques get at non-use values, but may double count use values already included in revealed preference measures.
        - Thus, caution is needed when combining results from different methods.
- Steps to "effect-by-effect" approach to benefits analysis
    1. Identify benefit categories potentially affected
        - Need to understand how policy will affect the environment, so as to know what the benefits will be.
        - Research physical effects of pollutants
            - Often relies on interdisciplinary literature
        - Consider potential changes in these effects from policy
        - Determine which benefits to include
            - E.g. health benefits, recreation, non-use benefits
        - Should co-benefits be included?
            - In the MATS case, including co-benefits made the net benefits of environmental regulations higher.
                - For example, an Obama Administration analysis found that most of the estimated benefits from reducing mercury pollution come from other related benefits.
                    - Reducing mercury pollution reduces the use of coal or required installation of abatement equipment. Both of which would also reduces fine particulate matter.
                - The Trump Administration said that these co-benefits should not be included.
                - In class we discussed arguments for and against, and discussed the importance of being transparent about what is and isn't included.

2. Quantify significant endpoints
    - Focus on changes to the environment resulting from the policy.
        - Consultation with natural scientists important (e.g. to understand transport of pollution through the environment)
    - These are done relative to a baseline (further discussed below)
    - Role of economics in this step:
        - Ensure information provided useful for subsequent valuation
            - Are endpoints appropriate?
        - Consider behavioral changes that may affect results of risk assessment.
            - Will people change behavior to reduce risk (e.g. stay indoors to avoid air pollution)
3. Estimate the value of these effects

- Because benefits analysis should capture *marginal* values, we need to properly define the baseline.
    - The baseline is "the best assessment of the world absent the proposed regulation or policy action." (EPA 2010, Chapter 5)
    - Note that the appropriate analysis of a policy is not *before* vs. *after*, but rather what happened compared to what would have happened without a policy intervention.
        - Because other factors in the economy will change whether or not the policy is enacted (e.g. changing demographics, economic activity, technology)
    - Usually use a single baseline, but multiple baselines used if policy is particularly complex (e.g. IEA energy projections)
- Examples of important variables for baseline consideration
    - Demographics
        - Population exposed to pollution matters.
        - Is the population changing?
            - Behavioral models needed if the policy may change demographic trends
    - Future economic activity
        - Is the regulated industry growing or in decline?
        - Is there increasing competition from abroad?
    - Changes in consumer behavior
        - Regulation may change prices
        - But would some behavior have changed anyway?
    - Technological change
        - Even without regulation, technology is changing
            - E.g. how will driverless vehicles affect energy demand?

- Consideration of baselines in statistical analyses
    - Observational studies may miss important factors influencing outcomes
    - These relate to defining a proper baseline
        - Example: Polluted areas may have higher population, different economic conditions, different crime rates
    - Quasi-experimental use regression analysis to compare similar treatment and control groups
        - Goal is to approximate random experiments in natural sciences
    - Differences-in-differences is one example
- Example: estimating the benefits of reducing mercury from coal-fired power plants.
    - Background
        - The Supreme Court invalidated the rule in 2015 for not appropriately considering costs.
        - It is also the rule discussed in the *Economist* reading, where co-benefits dominate the direct benefits discussed here.
            - Thus, in identifying potential benefit categories, the EPA considered a wide range of benefits
    - Here, I focus on the EPA's estimate of the direct benefits of reduced mercury pollution focus on the effects of mercury on the IQ of children exposed.
        - First, project the number of pregnant mothers exposed to mercury from recreationally caught freshwater fish.
            - Used 2000 Census data to find the number of females aged between 15 and 44 in each state
            - # of pregnant women = # of females aged 15-44 * state's fertility rate
        - Second, estimate number of prenatally exposed children
            - Estimate number of households with angler fisherman using results of a survey on recreational activities
            - # potentially exposed children = # of pregnant women * (# of anglers/adult population in state)
                - Assumes fertility rate the same for angler and non-angler households
        - Third, estimate daily exposure for 32 different sub-populations (urban/rural, high or low income, distance to water, etc.)
            - average daily mercury ingestion rate = average mercury concentration in fish in each region * average daily self-caught freshwater fish consumption
            - Based on scientific literature, divide by 0.08 to get maternal mercury concentration
        - Fourth, convert maternal mercury concentration to IQ reduction
            - Multiply maternal mercury concentration by 0.18
                - Again, this is based on a scientific study
                - Assumes a linear relationship between mercury & IQ
        - Fifth, estimate monetary loss for each IQ point lost

        - Two measures, both based on outside studies:
            - Lost earnings: 1.76-2.379%
            - Decreased schooling costs:
                - Lower IQ => less time in school => cost savings of $13,453
        - Combining these two effects, each IQ point results in a lifetime income loss of:
            - $8,013-$11,859 with 3% discount rate
            - $893-$1,958 with 7% discount rate
    - Sixth, multiply change in IQ times monetary loss per IQ point to get benefits from reducing mercury
        - Estimate ranges from $0.5 million to $6.1 million
        - Depends no discount rate and net monetary loss per IQ point
- Potential concerns:
    - Is IQ the best measure of effects of prenatal mercury exposure?
    - Is there a linear relationship between IQ loss and earnings
    - Uncertainties in calculation of exposure (e.g. # of fish consumed, demographics of angler population)
    - Note that the analysis uses scientific studies on different populations and applies them to this regulation
        - For example, the studies on mercury concentration come from New Zealand and the Seychelles Islands.
        - This is an example of *benefits transfer* described in the EPA reading

- Estimating co-benefits (Table 10)
  - Why do co-benefits occur?
    - Compliance options include:
      - Switching from coal to natural gas
      - Installing new abatement technologies
        - Expected many firms would install new abatement technology that would capture not only mercury, but also SO_2 and particulate matter.
  - Calculating benefits of reduced mortality
    - Estimate change in $PM_{2.5}$ and ozone due to MATS
    - Estimate change in population exposed
    - Use dose-response functions from epidemiological studies to predict changes in health and welfare due to changes in population exposed
      - Concerns:
        - WTP response might not be linear
        - Most wage studies done on working-age males
        - But elderly are more likely to be affected by air pollution.
    - Converted to dollar values using value of a statistical life
      - Used a value of $8.9 million
      - Sources: 21 studies using wage data & 5 CV studies
  - Other health effects
    - Reduction in chronic bronchitis, nonfatal heart attacks, lost school and work days, and lower medical costs
    - Used social cost of carbon to include benefits from reduced CO_2 emissions