Ex Ante Health Policy Evaluation: A Unified Approach

John A. Graves

Vanderbilt University

4/17/2019

Models projecting the impact of reforms to health insurance programs and markets play an important role in shaping U.S. health policy. In 2017, for example, Congressional attempts to repeal and replace the 2010 Affordable Care Act (ACA) were hampered by public outcry after the Congressional Budget Office (CBO) projected that upwards of 23 million people would become uninsured. The twists and turns of earlier debates over the ACA—and before it, the Clinton health plan—also were shaped by modelers' assessments of how reform would affect insurance coverage, premiums, health care spending, and government costs.[1]

Microsimulation models used by the CBO and by others to produce these estimates draw on economic theory and on a large and growing literature evaluating past state and federal reform efforts. Yet while models derive inputs from this shared evidence base, the evidence is uncertain and not in uniform agreement. Models also differ in their structure, underlying data sources and assumptions. It should come as no surprise, then, that models often produce varying projections of the same reform proposal.

This current state of affairs has subjected microsimulation models to criticism over their "black box" like qualities and their tendency to produce estimates with a limited accompanying sense of sensitivity to alternative parameters and assumptions. Moreover, modelers have shied away from producing comparative assessments of overall welfare impact. Existing models

[1]This history led one US Senator, Ron Wyden of Oregon, to remark that "The history of health reform is congressmen sending health legislation off to the Congressional Budget Office to die."

typically produce an array of intermediary point estimates on welfare-relevant outcomes (e.g., changes in coverage, premiums, spending and government costs) and leave it to policymakers to weigh those factors when comparing policy choices.

This approach to health policy modeling has a number of important shortcomings. First, despite modelers' attempts to caveat the high degree of uncertainty in their estimates, projections are often afforded a false sense of precision in policy debates. This results in key decisions being made without a full accounting of the uncertainty surrounding the budgetary and coverage impacts on millions of people. Second, despite recent efforts at greater transparency, the opacity of microsimulation models makes it difficult for researchers to know whether and how their work can inform modeling efforts. Finally, the development, execution, and maintenance costs of microsimulation models are considerable. Combined, these factors contribute to high barriers to conducting rigorous ex ante policy evaluation and a muddled sense of how the health economic research enterprise could be further refined to improve policy decision making.

This study outlines an approach to ex ante policy evaluation that addresses many of the above shortcomings. The first major contribution is a generalized discrete time and choice modeling framework for assessing the cost, coverage and welfare impact of health reform policies. This framework has roots in modeling methods commonly used for health technology assessment, and in the "sufficient statistics" approach to welfare evaluation developed in the public finance literature. I demonstrate that this modeling framework encompasses many existing approaches to health policy microsimulation, including elasticity-based and utility maximization-based models. Critically, however, the appraoch also facilitates simple yet powerful counterfactual policy assessments based on reduced form estimates. That is, the framework provides researchers with a tool to investigate the coverage and cost impacts of reform alternatives without the need for a detailed individual-level microsimulation model. As a proof of concept, I demonstrate how difference-in-differences evidence on the impact of Medicaid expansion on coverage take-up, combined with estimates on take-up of subsidized private health insurance derived from regression-discontinuity estimates (Finkelstein, Hendren and Shepard 2019) can be harnessed to model the coverage and cost impact of further expansion of coverage via public programs versus via increased subsidies for private coverage.

Second, within this framework I tie together diverse approaches to assessing uncertainty and the welfare impacts of policy. Specifically, I draw linkages between the marginal value of public funds (MVPFs), a summary measure of the costs and benefits of public policies (Hendren 2017), and value of information (VOI) methods. Intuitively, VOI quantifies the opportunity cost of policy decision making under uncertainty. At a given policy efficiency threshold (e.g., a MVPF value of 0.8, above which a policy might be desirable but below which it may not), modeling uncertaninty may or may not affect optional policy choices (i.e., choices that maximize relative comparisons of benefits to costs). If decisions based on comparative assessments of MVPF are insensitive to varying parameter values, then the value of uncertain information is low—i.e., it is not worth additional effort to reduce paramter uncertainty since the same decision would be made today as it would if we had better information. If decisions are sensitive to this uncertainty, however, then VOI methods quantify the opportunity cost of making policy decisions based on *current* information versus if we had perfect information on uncertain parameters. Variation in modeled outputs can be further decomposed to identify the relative degree to which specific parameters contribute to the overall value of perfect information. These assessments, in turn, can provide guideposts for refining and prioritizing future research to focus on domains where the value of information is high. I provide a concrete example of how VOI can enrich comparative welfare assessments by estimating the relative contribution of estimation precision and assumptions on the incidence of uncompensated care in contributing to uncertainty in MVPF estimates for policies that subsidize the purchase of private insurance coverage.

The remainder of this paper proceeds as follows. In the next section, I outline a discrete time modeling framework that provides a set of sufficient statistics to estimate the coverage and cost impact of health reform policies. I then demonstrate how existing approaches to microsimulation, including utility maximization and elasticity-based approaches, tie to this generalized modeling framework. Thereafter, I show the ability of the framework to accomodate modeling using parameters derived rom reduced form estimates. To do so, I draw on novel analyses of coverage changes estimtated in the Survey of Income and Program Participation, and on estimates of subsidized coverage take-up estimated in Finkelstein, Hendren and Shepard (2019). With this simple sufficient statistics model in hand, I show how the MVPF can provide a lens through which we can estimate the value of information on specific parameters in the model. I do this by estimating the VOI to decompose model output variation ...

Discrete Choice Model

Consider a model of insurance choice among J alternatives (including the choice not to insure). Define U_{itj} as the utility for choice unit i from selecting choice j at time t. For our purposes here we define the choice unit as the health insurance eligibilty unit (HIU), or the set of family members who could enroll under the same health insurance plan.

Suppose utility can be expressed in terms of a vector ($\mathbf{\tilde{x}_{itj}}$) of time-varying attributes of the choices and the HIU ($\mathbf{x_{itj}}$), including the history ($\mathcal{F}_{it}$) of the HIU's insurance selections, i.e., $\mathbf{\tilde{x}_{itj}} = \{\mathbf{x_{itj}}, \mathcal{F}_{it}\}$. Utility also depends on fixed attributes of the HIU ($\mathbf{z_i}$), and an unobservable component ϵ_{itj}, :

$$U_{itj} = V(\mathbf{\tilde{x}_{itj}}, \mathbf{z_i}) + \epsilon_{ij} \tag{1}$$

For HIU i, the choice of insurance y_{it} is based on maximizing utility across the J alternatives at time t:

$$y_{it} = \arg\max_j [U_{itj}, j = 1, \ldots, J]$$

We next assume a linear utility specification that depends on individual- and choice-specific attributes, the HIU's insurance selection at the previous occasion ($y_{i,t-1}$), and the idiosyncratic term (ϵ_{itj}) :

$$U_{itj} = \alpha_j + \beta_{\mathbf{j}}{}'\mathbf{x_{itj}} + \gamma'\mathbf{z_i} + \delta y_{i,t-1} + \epsilon_{itj} \tag{2}$$

Now define a function $B(\cdot)$ that maps the utility derived from choice j to the probability of selecting j. For example, if the error terms ϵ_{ij} are independent across units and are distributed Type I Extreme Value, the probability that unit i chooses insurance type k at time t is

$$\begin{aligned} C_{it}(k) &= P[U_{itk} > U_{itj} \forall k \neq j] \\ &= B(\mathbf{x_{itj}}, \mathbf{z_i}, y_{i,t-1}, \alpha_j \beta_{\mathbf{j}}, \gamma, \delta) \\ &= \frac{\exp(U(\mathbf{x_{itj}}, \mathbf{z_i}, y_{i,t-1}, \alpha_j \beta_{\mathbf{j}}, \gamma, \delta))}{\sum_{ij}[\exp(U(\mathbf{x_{itj}}, \mathbf{z_i}, y_{i,t-1}, \alpha_j \beta_{\mathbf{j}}, \gamma, \delta))]} \end{aligned}$$

This sets up a standard conditional logit model for insurance choice at time t. Alternative forms of $B(\cdot)$—such as based on a nested logit or multinomial logit model—could also be used.

Discrete Choice as a Markovian Process

In this theoretical setup, the HIU's selection depends on attributes of the HIU and insurance choice options, as well as on the HIU's previous choice ($y_{i,t-1}$). This latter assumption links the discrete choice framework to a discrete time Markovian process in which the future evolution of choices depends on the last choice made.

Cast in this light, the multinomial choice process at two periods can be specified in terms of a transition probability matrix $\mathbf{R_i} = [r_{irs}]$. Cells in this $J \times J$ matrix are defined by transition probabilities $r_{irs} = P(y_{it} = s | y_{i,t-1} = r)$, or the probability of unit i transitioning to choice s conditional on having previously selected choice r. Rows pertain to coverage at time t and columns to coverage at time $t+1$; thus, each row in this matrix sums to 1. At a population level (with size N) we define $\mathbf{R} = [r_{rs}]$ where $r_{rs} = \sum_{i=1}^{N} r_{irs}/N$.

Unifying Diverse Approaches to Ex Ante Policy Evaluation

A critical takeway is that along with estimates on the ex ante distribution of coverage, the set of population-level transition probabilities (r_{rs}) are sufficient statistics for modeling coverage changes under a proposed reform. Moreover, the individual-level transition probabilities (r_{irs}) also provide a critical crossing point between diverse approaches to ex ante policy modeling based on utility maximization and price elasticities.

To see this, consider a two period model at time t_0 (pre-reform) and t_0+1 (post-reform) and define a baseline occupancy vector $\mathbf{c}_{t_0} = c_{1t_0}, \ldots, c_{Jt_0}$ where

$$c_{jt_0} = \sum_{i}^{N} C_{it_0}(j)$$

For example, if there are four possible coverage categories (employer-sponsored insurance [esi], private non-group insurance [ng], public insurance [pub] and uninsured) then $\mathbf{c}_{t_0}$ summarizes the number of individuals in each category in the ex ante (i.e., pre-reform) period.

In principle, one could estimate $\mathbf{c}_{t_0}$ based on estimation of a choice model. Often, however, estimates of $\mathbf{c}_t$ are derived in part from a weighted sum or average of individual-level insurance data in a federal survey, such as the Curent Population Survey (CPS) or American Community Survey (ACS). It is also common for models to further calibrate survey-based estimates of $\mathbf{c}_{t_0}$ to match known and/or projected population totals.[2]

The distribution of coverage at time t_0+1 can be estimated by multiplying the baseline occupancy vector by the transition probability matrix:

$$\mathbf{c}_{t_0+1} = \mathbf{c}_{t_0}'\mathbf{R} \tag{3}$$

[2]This weighted sum is obtained by the Horvitz Thompson estimator for the sum of a pseudo-population based on survey data.

As we discuss in the section below, this discrete-time Markov trace is a key building block for assessing health insurance market changes using diverse appraoches to ex ante policiy modeling.

Modeling Policy Changes

Suppose our goal is to model an exogenous change to the choice set. This change could be brought about due to a reform that affects the price, availability, plan quality, and/or characteristics of insurance.

A standard assumption in behavioral microsimulation is that the exogenous change does not affect the unobserved disturbance term ϵ_{itj}. Rather, the policy change affects utility through its impact on prices, quality, offers of employment-based insurance, etc.

To model these changes via microsimulation, attributes of plans and individuals in the microdata are adjusted to reflect the reform scenario. In a utility maximization model, differences in predicted utility are used to derive new unit-level choice probabilities under the modeled reform. In an elasticity-based microsimulation model, price changes for each of the J insurance options are simulated for units in the microdata. Elasticities and further adjustments (e.g., income effects if subsidies or price increases are large relative to total income) are then applied to derive new choice probabilities.[3]

To formalize this process, we re-define key quantities in the framework above in terms of potential outcomes. $\mathbf{c}_{t_0+1}(0)$ is the distribution of coverage under the status quo, and $\mathbf{c}_{t_0+1}(1)$ is the distribution under the policy change. Similarly, $\mathbf{R}(0)$ is the transition probability matrix with no policy change, and $\mathbf{R}(1)$ is the transition probabilitiy matrix under reform.

While estimates of the effect of reform on transition probabilities—that is, $[\mathbf{R}(1) - \mathbf{R}(0)]$—

[3] In both a utility maximization model and an elasticity-based model, simulation might be carried out at several levels of a hierarchy. For example, an employed worker within an HIU faces a probability of having an offer of employer-sponsored insurance (ESI). This probability is often simulated based on a "synthetic firm" constructed around the worker's age and earnings profile. Specifically, the probability of an ESI offer faced by the worker is based on a synthetic firm-level simulation that accounts for observed ESI offer rates by industry and occupation, and that models potential changes to the structure, cost, and availability of ESI under the modeled reform scenario.

may be of interest, often the quantity of interest is the counterfactual change in coverage under a reform alternative:

$$\begin{aligned}\Delta \mathbf{c} &= \mathbf{c}_{t_0+1}(1) - \mathbf{c}_{t_0+1}(0) \\ &= \mathbf{c}_{\mathbf{t_0}}{}'\Big(\mathbf{R}(1) - \mathbf{R}(0)\Big)\end{aligned} \tag{4}$$

where $\mathbf{c}_{\mathbf{t_0+1}}(\mathbf{1}) = \mathbf{c}'_{t_0}\mathbf{R}(\mathbf{1})$ and $\mathbf{c}_{t_0+1}(0) = \mathbf{c}'_{t_0}\mathbf{R}(\mathbf{0})$.

Thus, to model a reform alternative three quantities are needed: the ex ante occupancy vector (i.e., $\mathbf{c}_{t_0}$) and estimates of the transition probability matrix under the status quo (i.e., $\hat{\mathbf{R}}(0)$) and under reform (i.e., $\hat{\mathbf{R}}(1)$).

The upshot of this framework is that estimates of $\hat{\mathbf{R}}(0)$ and $\hat{\mathbf{R}}(1)$ can be derived from any number of model types. For example, as noted above these estimates could derive from a microsimulaton model based on calibrated microdata and a specified utility model or price elasticity equations. Or, as we show later in this study, $\hat{\mathbf{R}}(0)$ and $\hat{\mathbf{R}}(1)$ can be based on transition probabilities estimated by or derived from a reduced form model.

Example: Congressional Budget Office Health Reform Model

A key difference between the CBO model and the discrete time framework above is that the CBO model is static: it does not simulate individual trajectories before and after a policy change.[4] Rather, the CBO model adopts a similar utility maximization approach and compares simulated baseline and counterfactual choices and costs for a given reform alternative. In other words, the CBO model produces estimates of $\mathbf{c}_{t_0+1}(0)$ and $\mathbf{c}_{t_0+1}(1)$.

Beyond this, however, the CBO modeling is similar in terms of its foundation in utility maximization theory and discrete choice. For a given choice unit (a health insurance eligiblity unit), the CBO model specifies a utility representation in which utility depends on a systematic

[4]Estimates in the CBO model are also adjusted to account for other policy changes simulated in different microsimulation models for the Medicaid program, federal taxes, etc. Multi-year projections are not based on a longitudinal simulation, but are rather constructed via updates to baseline coverage occupancy to reflect population changes, inflation, etc.

component (V_{in}) and an (unobservable) stochastic component ϵ_{in}:

$$U_{ij} = \beta_1 V_{ij} + \epsilon_{ij}$$

This utility specification is a close analogue to Equation 1, with the exception of an additional term (β) that adjusts the systematic component of utility. This scaling factor is designed to translate the utility value HIUs place on a given alternative into dollar terms.

The systematic component of utility is further modeled using microdata on individuals and simulated employer choices to offer insurance. For a single individual the modeled systematic component takes the following form:

$$V_{ij} = y_i - C_{ij} - E[H_{ij}] - \frac{1}{2}\rho_j \texttt{Var}(H_{ij}) + \delta_{1j}(y_i, a_i)$$

where y_i is the individual's income, a_i is the individual's age, C_{ij} is the out-of-pocket cost to the individual of coverage alternative j (e.g., premium, any mandate penalty, etc.), $E[H_{ij}]$ and $\texttt{Var}(H_{ij})$ are the expectation and variance of the individual's out-of-pocket health expenditure on coverage alternaive j, ρ_j is the coefficient of absolute risk aversion, and δ_{1j} is a utility shifter specific to each coverage type j.

The utility shifter in the above equation is designed to either increase or decrease the value of each coverage type–possibly varying by age or income–based on various factors. These factors could include the individul's awareness of their eligibility for the program, their ability to enroll in the program (e.g., through a website, or through a more or less cumbersome enrollment process), their preferences for or against certain types of coverage, etc. In addition, the individual's out-of-pocket spending (H_{in}) varies by insurance type, and is capped (based on income) to reflect the availability of uncompensated care and bankruptcy as implicit sources of insurance.

Each indiviual in the model faces a set of insurance options in their choice set (e.g., based on whether an offer of emloyment-based coverage is available to them, whether they are eligible for public insurance, etc.). These utilities are then fed through a nested logit framework to derive coverage take-up probabilities. Estimates of $\mathbf{c}_{t_0+1}(0)$ are based on aggregating these choice probabilities under the baseline scenario, and estimates of $\mathbf{c}_{t_0+1}(1)$ are based on aggregating choice probabilities under the modeled reform.

Example: Modeling Full Expansion of Medicaid Under the ACA

The set of parameters encompassed by $\mathbf{c}_{t_0}$, $\mathbf{R}(\mathbf{1})$, and $\mathbf{R}(\mathbf{1})$ are sufficient statistics for modeling the impact of potential reforms to US health insurance markets. Critically, these parameters could be generated based on a microsimulation model, or they could be derived simply from reduced form estimates of the relevant quantities of interest.

In this section, we provide a proof of concept of this reduced form approach by demonstrating how reduced-form evidence on coverage take-up under the ACA's state-level Medicaid expansion can be used to generate projections of the cost and coverage impacts of full expansion in all 50 states.

Scratch

Health Reform Modeling as a Discrete Time Markov Process

This section outlines a simple discrete time markov model for health insurance coverage in the U.S. population. We begin by defining an ex ante occupancy vector $\mathbf{p_{exa}}$ that summarizes the fraction of the population in each major health insurance type (employer-sponsored insurance, other private insurance, public insurance, and uninsured) in the pre-reform period.

$$\mathbf{p_{exa}} = \begin{pmatrix} p_{exa,esi} \\ p_{exa,pri} \\ p_{exa,pub} \\ p_{exa,unin} \end{pmatrix}$$

where $p_{exa,k}$ is the fraction of the population in each insurance category k in the ex ante period.

Now define the transition probability matrix:

$$\mathbf{R} = [r_{k,j}] = \begin{pmatrix} r_{esi,esi} & r_{esi,pri} & r_{esi,pub} & r_{esi,unin} \\ r_{pri,esi} & r_{pri,pri} & r_{pri,pub} & r_{pri,unin} \\ r_{pub,esi} & r_{pub,pri} & r_{pub,pub} & r_{pub,unin} \\ r_{unin,esi} & r_{unin,pri} & r_{unin,pub} & r_{unin,unin} \end{pmatrix}$$

where $r_{k,j}$ is the probability of transitioning from ex ante category k to ex post category j.

Finally, we can define an ex post occupancy vector:

$$\mathbf{p_{exp}} = \begin{pmatrix} p_{exp,esi} \\ p_{exp,pri} \\ p_{exp,pub} \\ p_{exp,unin} \end{pmatrix}$$

Basic matrix algebra links the two occupancy vectors as follows:

$$\begin{pmatrix} p_{exa,esi} \\ p_{exa,pri} \\ p_{exa,pub} \\ p_{exa,unin} \end{pmatrix}' \cdot \begin{pmatrix} r_{esi,esi} & r_{esi,pri} & r_{esi,pub} & r_{esi,unin} \\ r_{pri,esi} & r_{pri,pri} & r_{pri,pub} & r_{pri,unin} \\ r_{pub,esi} & r_{pub,pri} & r_{pub,pub} & r_{pub,unin} \\ r_{unin,esi} & r_{unin,pri} & r_{unin,pub} & r_{unin,unin} \end{pmatrix} = \begin{pmatrix} p_{exp,esi} \\ p_{exp,pri} \\ p_{exp,pub} \\ p_{exp,unin} \end{pmatrix}'$$

In the equation above, the set of transition probabilities $r_{k,j}$ can be considered sufficient statistics for evaluating the impact of a policy change on health insurance coverage in the population. That is, once we know these probabilities and how they change under a given reform option, we can simulate the impact on the overall coverage distribution in the population. By attaching costs to population movements among insurance types, we can simulate the cost impact to the government. And finally, as we show below, social welfare weights can also be attached to population movements. These weights can then be aggregated and compared across reform alternatives to make comparative evaluations of policy options.

Link to Existing Microsimulation Approaches

Utility Maximization Models

Elasticity