MODEL-BASED MICROSIMULATION ESTIMATES FOR STATES WHEN STATE PROGRAMS VARY

Allen L. Schirm, Mathematica Policy Research, Inc. and Alan M. Zaslavsky, Harvard University
Allen Schirm, Mathematica Policy Research, 600 Maryland Ave., S.W., Suite 550, Washington, DC 20024-2512

Key Words: Small Area, Welfare Reform, Food Stamps

1. INTRODUCTION

The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA) made sweeping changes to our welfare system, including giving states great flexibility to design their own programs. States have exercised this new authority and developed programs funded by the Temporary Assistance to Needy Families (TANF) block grant established by the act. Yet, little is known about the effects of these changes to our welfare system. We are just beginning to learn, for example, about what happens to families who leave welfare. Even so, what we are learning pertains to a prolonged and vigorous economic expansion. Which households are affected by welfare reform and how they are affected may be substantially different if the economy enters a recession. Thus, welfare reform will likely continue. As state policymakers consider further program reforms, they will need reliable estimates of the cost and distributional effects of alternative policies.

Microsimulation has been used extensively to obtain national estimates of the effects of proposed reforms to federal programs. However, the current microsimulation models that use data from national surveys or samples of administrative records can often produce only imprecise estimates for states because sample sizes are small.

Schirm and Zaslavsky (1997) described a method for reweighting a microsimulation database to "borrow strength" and obtain more precise state estimates. A Poisson regression model is fitted to obtain an estimated prevalence in each state of every household in the database. This model is specified to control important aggregates at the state level, and the prevalences are expressed as a matrix of weights, with each household having a weight for every state. Estimates for a state are obtained by passing through the microsimulation model all households in the database, not just the households actually in that state. By applying the appropriate weight for each household, the database is weighted to look like the state, rather than the whole country.

In this paper, we review how microsimulation methods are used for policy analysis, and discuss the general problem of imprecision in microsimulations for states. Then, we describe the basic ideas of our reweighting method, and present the formal reweighting model. Finally, we discuss two applications of our reweighting method, and give a simple example to illustrate our method.

2. WHAT IS MICROSIMULATION?

A microsimulation model simulates how proposed changes to a government program affect the program and its participants. The model has two elements: (1) a micro database and (2) a computer program. The database is constructed from administrative or survey data with information on households in the population targeted by the government program. The model's computer program codes the rules of the government program under both the "baseline" policy, which is typically the current policy, and a "reform" policy, which is an alternative under consideration. The computer program also simulates what a caseworker does--that is, it determines whether a household is eligible for the government program and the benefits for which the household would qualify. In addition, the computer program simulates a household's behavioral response, determining whether the household will participate in the program. Processing all the households in the database, the model counts participants to estimate the caseload of the government program and adds up their benefits to estimate costs. By performing these operations under both baseline and reform policies and comparing the results, the model estimates the cost and caseload effects of the proposed reform. The model can also estimate the distributional effects of the reform, identifying the population subgroups that gain and lose benefits.

We have focused our work on the Micro Analysis of Transfers to Households (MATH®) family of models. These models, developed by Mathematica Policy Research beginning in 1974, have been used extensively to simulate reforms to the Food Stamp Program (FSP), the Aid to Families with Dependent Children (AFDC) and TANF programs, and the Supplemental Security Income (SSI) program. The database for the current MATH SIPP model was constructed by combining data for January 1994 from Waves 7 and 4 of the 1992 and 1993 Panels of the Survey of Income and Program Participation (SIPP). The database for the current QC Minimodel--another member of the MATH family--was constructed from the fiscal year 1996 Integrated Quality Control System (IQCS) sample, an administrative records database.

3. MICROSIMULATION FOR STATES

Although the MATH models have been used in recent years to estimate the national effects of hundreds of potential reforms to national programs (mainly the FSP), welfare devolution and the resulting need for state analyses and estimates create a new challenge. Because a database developed from a national survey like the SIPP or even a national administrative database like the IQCS sample has state samples that are small for general purposes or, at least, for some important applications, direct microsimulation estimates are imprecise. Thus, there is substantial uncertainty about the likely impacts of proposed reforms and little guidance for policymakers.

Borrowing strength with an indirect estimator is a common solution to the problem of imprecise direct estimates, and has been used successfully in many applications. For example, an indirect estimator has been used for several years to obtain state estimates for allocating federal funds under the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). Estimates of the numbers of infants (under age 1) and children (ages 1 to 4) below 185 percent of poverty are derived in three steps: (1) calculate direct sample estimates from the Current Population Survey (CPS); (2) calculate regression estimates, that is, predictions based on census and administrative data (measuring, for example, food stamp program participation); and (3) calculate "shrinkage" estimates by averaging the sample and regression estimates using empirical Bayes methods. The shrinkage estimates are substantially more accurate than the direct sample estimates.

In the microsimulation context, the problem with using an empirical Bayes or similar estimator is that the modeling is purely specific to the estimates being produced--estimates of WIC eligibles, for example. Such methods do not provide an "off-the-shelf" approach for estimating a microsimulation model's many outputs, which include baseline and reform caseload and cost figures as well as numerous measures of distributional effects. Instead, these conventional indirect estimators are best for deriving a single estimate or a few closely related estimates for each state. A different approach to indirect estimation is needed for microsimulation because it would not be practically feasible to develop a model for every estimate being produced.

4. REWEIGHTING TO BORROW STRENGTH AND IMPROVE PRECISION

In Schirm and Zaslavsky (1997), we proposed to borrow strength by reweighting sample observations in the microsimulation database. The basic idea of our reweighting approach is to use households from many states to borrow strength and improve precision when deriving estimates for any one state. How reweighting can be used to borrow strength is illustrated by comparing (1) the direct estimator that uses the original sample weights and does not borrow strength with (2) the indirect estimator that uses reweighted data and does borrow strength. To calculate an estimate for Virginia, for example, the direct estimator uses only the sample households for Virginia and their original sample weights. Observations for other states are ignored. This is equivalent to using all the observations in the database weighted by "Virginia weights" that equal the original sample weights for households in Virginia but are zero for households in all other states. In contrast, for indirect estimation, nonzero Virginia weights would be assigned to households in not only Virginia but also other states. Although this introduces some bias, a simulation using many more observations that are similar except for state of residence should have substantially improved precision, especially when we are estimating--as we often are--the effects of a narrowly targeted policy or the effects of a broader policy on a small subgroup of the population. The objective of reweighting and, more generally, indirect estimation is to enhance accuracy as measured by a standard like mean squared error (MSE) that reflects the tradeoff between bias and variance.

Under our proposed approach, we derive a matrix of state weights. Every household in the database gets as many new weights as there are states (51 counting the District of Columbia as a state). For every state, there is a weight for each household in the database, although some weights may be small or (by design) zero. To derive estimates for any one state, we pass all households in the database--regardless of actual state of residence--through the microsimulation model, and apply the appropriate set of weights. Virginia weights are used to derive estimates for Virginia, Maryland weights are used to derive estimates for Maryland, and so forth. Thus, Virginia borrows strength from other states that have households with nonzero Virginia weights.

Using a Poisson regression model, our reweighting method assigns a Virginia weight to a household according to how prevalent that "type" of household is in Virginia. The more prevalent it is, the more Virginia weight it gets. A household's type is defined by all the characteristics in the database, some of which are measured directly while others are calculated or simulated. A household's prevalence is determined, under the model, by a set of household characteristics that (1) are policy-relevant (e.g., characteristics determining program eligibility and benefits), (2) capture the key dimensions along which households in different states are different, and (3) have about the same meanings across states. This third property has important implications when a characteristic--such as an indicator of cash welfare receipt--is highly relevant to a simulation, but means something different in different states because of differences in state policies. In particular, we will discuss later how our reweighting approach needs to be modified when state welfare programs differ substantially.

The variables included in the reweighting model serve as control variables, and households are reweighted so that weighted sums (indirect estimates) equal specified control totals. These totals can be direct sample estimates, indirect estimates smoothed using empirical Bayes methods, or administrative totals. For example, if the number of people in the household is a control variable, the total state population is a control total. If household income is a control variable, total personal income in the state is a control total.

With the original (national) weights, the database looks like the entire United States. With Virginia weights, the database looks like Virginia in terms of some key aggregates (the control totals). We then conjecture that the reweighted database resembles Virginia in terms of many other relevant aggregates for which we cannot control, including, especially, the main estimands of the microsimulation model. Our ongoing evaluation will address the extent to which this is accomplished.

5. THE REWEIGHTING MODEL

Our reweighting model is:

$$w_{hs} = \gamma_{hs}\, e^{\beta_s' x_h + \delta_h},$$

where w_{hs} is the expected number of households of type h in (the population of) state s. A type is, practically speaking, unique on the database because no two households are exactly alike. Therefore, each household in our database represents its own type, and w_{hs} is the weight that will be given to household h when deriving estimates for state s. γ_{hs} is an indicator set by the modeler to one if state s is allowed to borrow from the state in which household h actually resides, and zero otherwise. x_h is a column vector of I control variables, that is, household characteristics for household h. β_s is a vector of I unknown parameters to be estimated for each state. δ_h is an unknown parameter to be estimated for each household. The first term in the exponent on the right side of the model ($\beta_s' x_h$) reflects the "general" prevalence in state s of households like household h, that is, households with similar characteristics. The second term (δ_h) reflects the "specific" prevalence of household h.

The β_s and δ_h parameters are estimated by maximum likelihood and satisfy two constraints:

$$\text{Constraint 1: } \sum_s w_{hs} = W_h \quad \text{for each } h,$$

where W_h is the control weight, that is, the original sample weight or national weight of household h, and

$$\text{Constraint 2: } \sum_h w_{hs} x_{hi} = X_{si} \quad \text{for each } s \text{ and } i,$$

where X_{si} is the control total for control variable i in state s. According to the first constraint, reweighting does not change the total weight given to a household across all states, that is, at the national level, ensuring that the household contributes the same to a national estimate after reweighting as before. According to the second constraint, all control totals are satisfied for every state. Schirm and Zaslavsky (1997) describe an iterative two-step procedure for estimating the parameters of the reweighting model.

6. REWEIGHTING THE MATH SIPP DATABASE

We have completed a preliminary reweighting of the MATH SIPP database. For the reweighting, we fit for each of three income groups a model with the following control variables: an intercept, number of persons, number of Hispanics, number of blacks, number of Asians, presence of members unrelated to the head (0/1), presence of kids under age 5 (0/1), presence of an elderly member (0/1), shelter expenses, utility expenses, home ownership (0/1), receipt of interest income (0/1), receipt of earned income (0/1), receipt of SSI income (0/1), number of working age adults (ages 18 to 59), number of working age adults with less than a high school education, and number of working age adults with no job. These variables measure household composition, income, and expenses, and each means the same thing in different states. They were selected based on their policy relevance and analysis of variance results showing that their distributions varied across states. All control totals used for the reweighting were direct sample estimates.

Schirm and Zaslavsky (1997) report the findings from a preliminary evaluation of this reweighting of the MATH SIPP database. Three findings are most striking. First, the variability of the direct sample estimates is unacceptably large. Second, the variability of the model-based estimates obtained from the reweighted data is substantially less and at a level at which the estimates can provide useful guidance to policymakers. For the estimands considered so far, the variances of the model-based estimates are typically at least 90 percent smaller than the variances of the direct sample estimates. This suggests that unless the biases of the model-based estimates are enormous, those estimates will have smaller MSEs than the direct sample estimates. Finally, reweighting the data smooths the estimates in a sensible way, shrinking toward the middle the most extreme direct sample estimates. Thus, although there are important limitations to the first steps that have been completed in our evaluation of the reweighting approach, the preliminary findings are encouraging.

7. WELFARE REFORM AND THE FSP

In addition to reforming substantially our welfare system, PRWORA included several important provisions pertaining to the FSP. One of those provisions permits states to establish Simplified Food Stamp Programs (SFSPs). An SFSP allows a state to use the same rules to calculate benefits under its food stamp and cash welfare (TANF) programs, rather than using two different sets of rules. Under the law and current practice (written regulations have not been completed), an SFSP cannot increase total food stamp benefit costs in a state or reduce benefits by too much for too many of the households in which only some members receive TANF benefits.

Microsimulation has been used to assess whether SFSPs designed by states comply with these conditions. When feasible, the microsimulations have been conducted using databases that were constructed from states' automated case records systems and contained the entire population of active food stamp cases. When this approach has not been feasible, microsimulations have been conducted using the QC Minimodel and its database.

8. REWEIGHTING THE QC MINIMODEL DATABASE

As mentioned earlier, the current QC Minimodel database is constructed from the fiscal year 1996 IQCS sample. The first purpose of the IQCS is to review the accuracy of food stamp eligibility and benefit determinations. To support reviews, each state draws every month a random sample of its food stamp caseload according to IQCS specifications. The state samples are included in a national database, which consists of about 50,000 households for a full fiscal year.

In addition to supporting quality control reviews, the national QC database provides valuable data for microsimulation and other analyses of the FSP. Among its advantages relative to a national survey database like the MATH SIPP database are that all--rather than a small fraction--of the households in the QC database receive food stamps. Furthermore, the database contains all of the information needed to determine a household's eligibility and benefits, and this information was gathered through the actual administration of the program, rather than a survey.

Despite these strengths, the QC database also has some limitations for analytic purposes. Because it contains only food stamp households, it cannot be used to simulate program reforms that expand eligibility. It also has some limitations for state-level analyses. One is that state samples may not be big enough for reliably assessing the effects of narrowly targeted program reforms or the effects of broader reforms on small population subgroups. Reweighting the QC database could address this limitation. However, another limitation is that key data items for both reweighting and simulating program reforms can mean something different from state to state because of differences in state policies.

Both of these latter two limitations arise in simulating a state SFSP. State samples may be too small for gauging the effect of a proposed SFSP on small but politically important subgroups. Potential sample size inadequacies are made worse by eligibility for an SFSP being restricted to cash welfare recipients, which are about 40 percent of the national food stamp caseload and a much lower fraction in some states.

That SFSP eligibility is restricted to cash welfare recipients raises the second of the two limitations and complicates the application of our reweighting approach. Clearly, to simulate accurately a state's proposed SFSP, we must identify accurately the households eligible for the SFSP. That implies that we must identify accurately the households receiving cash welfare. Because cash welfare receipt is recorded in the QC database, that may seem straightforward, and it is for households that actually reside in the state whose SFSP we are simulating. If we want to use households from other states to borrow strength, however, the substantial differences in states' cash welfare programs and, specifically, in the generosity of those programs make it hard to determine whether a household receiving welfare in one state would receive it in another state.

AFDC, the precursor to TANF, was still operating in fiscal year 1996, the year to which the current QC database pertains. Although waivers of federal laws and regulations were widespread, many differences in state AFDC programs were minor. However, there were large differences in the maximum benefit level, a program parameter that is a key determinant of a household's cash benefit (given its net income) and that substantially captures differences in generosity. Ranking states by their AFDC maximum benefit levels (for three-person families) and displaying roughly every tenth state shows the substantial variation in generosity:

State	AFDC Maximum Benefit ($)	Percentage of Food Stamp Households Receiving AFDC
Mississippi	120	24
Georgia	280	35
Virginia	354	24
New Jersey	424	44
Minnesota	532	43
Vermont	650	41

Not surprisingly, higher fractions of food stamp households participated in AFDC in more generous states, although the relationship is not monotonic.

In applying our reweighting approach, the problem created by these program differences is that a variable indicating whether a household receives cash welfare and is, therefore, categorically eligible for an SFSP is critical to the simulation of the SFSP, but does not meet one of the other requirements of a good control variable for reweighting. That indicator variable can mean something very different in two states because households in those states face different programs and choices. A household that participates in a generous state's program may not participate in a stingy state's program.

There are at least two strategies for handling this problem. The first strategy entails calculating for each household the benefit that it would receive in each state. Then, we would develop and apply a participation model to determine whether the household participates in each state. Besides being cumbersome, with 51 benefit and 51 participation variables, this strategy is unattractive because efforts to develop strong predictive models of program participation have not been terribly successful.

An alternative approach that seems preferable to us is to borrow strength from just some states, not all states. For a given state, we would borrow only from other states with similar welfare programs so that a household faces about the same participation decision in its home state as in the state for which we are borrowing. With this restriction, we give up some precision to avoid some bias.

We restricted borrowing when reweighting the QC Minimodel database. First, we ranked states by their AFDC maximum benefit levels. Then, in calculating state weights for each state, we borrow strength from a "band" of neighboring states in that ranking. We include enough states in a band to obtain about 3,000 borrowed households in which all members received AFDC. In the middle of the maximum benefit distribution, each state's band is roughly symmetrical, with about 1,500 households from lower states and 1,500 from higher states. Typically, this means borrowing from 5 to 6 states in each direction, with maximum benefit levels no more than about \$30 to \$50 higher or lower than in the state for which we are borrowing. To reach our (somewhat arbitrary) target of 3,000 borrowed households for states in the tails of the distribution, we have to extend the necessarily asymmetrical bands farther away, including states that may have maximum benefit levels that differ by \$100 or more from the level in the state for which we are borrowing. We will evaluate in future work whether a higher or lower target number of households gives the best tradeoff between bias and variance.

9. A NUMERICAL EXAMPLE

We next illustrate our reweighting approach with restricted borrowing using a simple artificial example. As shown in Table 1, our example has 7 states and 35 households (5 per state), each with a weight of 100. Besides an intercept, the reweighting model has just one control variable, x. In each state, x is uniformly distributed over a set of consecutive values. The state means rise steadily from State 1 to State 7, so the distributions of x overlap for "nearby" but not "far apart" states.

We specified the following pattern of restricted borrowing:

$$
\begin{array}{ll}
 & \text{lender} \\
 & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\text{borrower} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} &
\left[\begin{array}{ccccccc}
\mathbf{1} & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & \mathbf{1} & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & \mathbf{1} & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & \mathbf{1} & 1 & 1 & 0 \\
0 & 0 & 1 & 1 & \mathbf{1} & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & \mathbf{1} & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & \mathbf{1}
\end{array}\right]
\end{array}
$$

which allows for borrowing (indicated by a "1") from states up to two places away, with the exception of State 7. Although State 1 cannot borrow from lower states because there are none, it can borrow from States 2 and 3. State 2 can borrow from just State 1 on the lower end, but can borrow from two higher states--States 3 and 4. State 3 can borrow from the maximum of two states in each direction. The exception to our general rule for borrowing is that State 7 can borrow from State 6, but not State 5. This introduces some (intentional) asymmetry into the pattern of borrowing.

We estimated a reweighting model that has an intercept and x as control variables, and conforms to these restrictions on borrowing. The state weights are displayed in Table 1. It is easy to verify that the weights satisfy the first constraint discussed earlier: they sum to the original national weight (100) for each household (up to roundoff error). Likewise, the state weights satisfy the second constraint: weighted sums equal control totals within a specified tolerance (0.1 percent). Using the five households from State 3 and their original national weights, we estimate that the state has 500 households and an x total of 3,000 in its population. Using the model-based State 3 weights, which are nonzero for 25 households (20 from outside State 3), we estimate that State 3 has 502 households and 3,015 of x--values close to the control totals. (The rounded sums calculated from unrounded weights are 500 and 3,002, satisfying the tolerance criterion.) Finally, inspection of Table 1 confirms that the state weights conform to the restricted pattern of borrowing specified earlier. For example, nonzero State 7 weights are given only to households from States 6 and 7.

In addition to satisfying the borrowing restrictions and the two constraints, the weights in Table 1 exhibit sensible patterns. Considering the State 3 weights, for example, we see that households that are outside of State 3 but near its mean of 6 for x generally get larger weights than households farther from that mean. The State 3 weights for households from State 1 rise monotonically from 5 to 39 as x rises from 2 to 6. A second sensible pattern is that households that have the same x but are from different states tend to get similar weights for a given state. For instance, the households with x equal to 4 from States 1, 2, and 3 get State 3 weights of 17, 16, and 15.

That the estimated weights seem sensible to us does not demonstrate that they support more accurate microsimulation estimates for states. For that, we will return to the evaluation that we began after the reweighting of the MATH SIPP database. Having now also reweighted the QC Minimodel database, we can add to the list of issues to be explored in our evaluation. One such issue is whether the bands of states from which we borrow are too wide, too narrow, or about right. If future results are as encouraging as the preliminary results, microsimulation analyses based on reweighted data should provide much helpful guidance to policymakers as welfare and related policies evolve in the coming years.

REFERENCE

Schirm, Allen L. and Alan M. Zaslavsky (1997), "Reweighting Households to Develop Microsimulation Estimates for States," *1997 Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association, 306-311.

Table 1. An Example of Reweighting with Restricted Borrowing

			National Weight	State Weights						
Household	State	x	W	w_1	w_2	w_3	w_4	w_5	w_6	w_7
1	1	2	100	72	23	5	0	0	0	0
2	1	3	100	61	29	10	0	0	0	0
3	1	4	100	49	34	17	0	0	0	0
4	1	5	100	36	38	27	0	0	0	0
5	1	6	100	24	38	39	0	0	0	0
6	2	3	100	59	28	9	4	0	0	0
7	2	4	100	45	32	16	8	0	0	0
8	2	5	100	30	32	23	15	0	0	0
9	2	6	100	18	28	29	26	0	0	0
10	2	7	100	9	21	32	38	0	0	0
11	3	4	100	44	31	15	8	3	0	0
12	3	5	100	28	30	21	14	7	0	0
13	3	6	100	15	24	24	22	15	0	0
14	3	7	100	7	16	24	28	25	0	0
15	3	8	100	3	9	20	31	37	0	0
16	4	5	100	0	40	28	19	9	3	0
17	4	6	100	0	26	27	24	16	7	0
18	4	7	100	0	15	22	26	23	15	0
19	4	8	100	0	7	15	24	28	27	0
20	4	9	100	0	3	9	18	29	41	0
21	5	6	100	0	0	36	33	22	9	0
22	5	7	100	0	0	25	30	27	17	0
23	5	8	100	0	0	16	25	30	29	0
24	5	9	100	0	0	9	19	30	42	0
25	5	10	100	0	0	4	12	27	56	0
26	6	7	100	0	0	0	33	29	19	18
27	6	8	100	0	0	0	21	25	24	30
28	6	9	100	0	0	0	12	19	26	43
29	6	10	100	0	0	0	6	13	26	55
30	6	11	100	0	0	0	3	8	24	66
31	7	8	100	0	0	0	0	32	30	38
32	7	9	100	0	0	0	0	21	30	49
33	7	10	100	0	0	0	0	13	28	59
34	7	11	100	0	0	0	0	8	25	67
35	7	12	100	0	0	0	0	5	21	74