Volume Title: Annals of Economic and Social Measurement, Volume 2, number 2

Volume Author/Editor: NBER

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm73-2

Publication Date: April 1973

Chapter Title: Development of a Microsimulation Model for Evaluating Economic Implications of Income Transfer and Tax Policies

Chapter Author: John F. Moeller

Chapter URL: http://www.nber.org/chapters/c9889

Chapter pages in book: (p. 183 - 187)

Annals of Economic and Social Measurement, 2/2, 1973

RESEARCH METHODOLOGY NOTES

Development of a Microsimulation Model for Evaluating Economic Implications of Income Transfer and Tax Policies

by John F. Moeller*

1. Introduction

Reform of our current system of welfare for needy families has been considered long overdue. Yet to date no scheme has emerged from Congress to replace the current system. During the past three years, every proposal for reform brought before Congress has required answers to the following questions:

1. How much will the new plan cost the Federal government and State and local governments in total, and net of what they are currently paying?
2. What types of persons and families receive coverage and what types pay additional taxes to finance the scheme?

A computerized procedure has been developed at the Urban Institute in response to the need for such information. In briefest form, the system simulates the eligibility conditions for grant proposals within each family record on a microdata file, computes grant payments or additional tax payments to finance the program, and aggregates the weighted results for nationwide analysis. Currently, the system runs off of the 1971 Current Population Survey (CPS) file for income year 1970, or the 1967 Survey of Economic Opportunity (SEO) file for income year 1966.

Early efforts were successful in producing reasonable results at relatively low dollar and time costs. Since the legislature became stalemated over the final shape of the reform legislation, we have had the opportunity to streamline the system several times. In mid-1971, a full-scale overhaul of the estimating procedure was undertaken by the Urban Institute with the computer assistance of the Hendrickson Corporation of Washington, D.C. This note reflects on past experience with the system, from infancy through maturity, and discusses the outlook for this research effort.[1]

2. Lessons from the Past

2.1. *Legislative Procedures*

For the uninitiated, one of the earliest lessons to be learned in the simulation work is that a grant/tax bill may wear many hats in the course of its journey from

* The research reported here was supported by funds made available to the Urban Institute through agency contracts with the Office of Economic Opportunity and the Department of Health Education and Welfare. The opinions expressed are those of the author and do not necessarily represent those of the Urban Institute or its sponsors. The author is indebted to Nelson McClung for assistance and support throughout the duration of this research work and to a host of research and programming specialists at the Urban Institute and the Hendrickson Corporation of Washington, D.C. who have contributed sizeable amounts of their resources to this modeling effort.

[1] It should be stressed that this note is written from the standpoint of an economic analyst who aided in the structural and logical system design. Programming and computer implementation of the simulation model is an altogether separate, yet equally important topic for consideration.

inception in the House Ways and Means Committee to the final act of approval —the President's pen. Since the simulator provides a crucial input for this process, he must be responsive to the waves of political sentiment which buffet the bill. The path of fiscal legislation must be closely followed, because as provisions of the bill are modified, the computer program must be adjusted to provide cost and coverage information.[2] In addition, the model must be used to service Congressional explorations of particular provisions of a bill (which may never be formally incorporated in the legislation as it winds its way through House and Senate). One becomes an expert with respect to current grant/tax legislation in simulation work.

As a corollary to this rule, the researcher also must learn how to interpret the language of a bill. This is no small feat since what does *not* appear in the legislation is often as important as what does appear. As a case in point, in defining chargeable income for a grant scheme,[3] the legislative language frequently takes the indirect route by specifying only nonchargeable income. In defining "employable" recipients of grant aid, a bill will instead specify "unemployables." A reform proposal sponsored by the then Senator Fred Harris (D.-Okla.) in early 1970 contains the following passage which illustrates this negative language:[4]

> Earned income of any individual or of any member of a family group during any month shall be disregarded to the extent of the first $75 of such income, plus one-half of the next $150 of such income for such month plus one-fourth of the remainder.

In translation the bill states that the first $900 of annual earned income per filing unit is exempted from the offset tax.[5] Amounts over $900 per year up to $2,700, are taxed at a 50 percent rate, and earned income in excess of $2,700 is taxed at a 75 percent rate. "Unearned income," by default, is presumed to be "regarded" at a 100 percent rate, although the bill makes no positive statement to this effect.[6] The analyst must be able to interpret what is explicit and what is implicit in the original legislation for input to the simulation model.

2.2. *System Economies in Data Manipulation*

Prior to making a computer run producing cost and coverage estimates (a production run), there often arises the need for certain one-time preprocessing runs. These routines may include (i) allocations of various sources of unearned income among persons, (ii) ageing the income and demographic information on the files to the year of the simulation request—if different from the base year of the microdata file, (iii) altering work experience on the file commensurate with that anticipated at a higher unemployment rate—if the simulation calls for a rate higher than that

[2] Unfortunately, the analyst receives little feedback from Capitol Hill regarding subjective aspects controlling a bill's fate. Political gossip would certainly brighten the job, but its absence from the effort enhances, rather than diminishes, the role of simulation models in policy analysis. Objectivity is an absolute necessity in the simulation work.

[3] Chargeable income is used interchangeably with the "offset tax" to indicate that income of the categorically eligible filing unit which is "offset" against the unit's gross allowance (or minimum income guarantee) for a grant plan in computing the new allowance.

[4] S.3433, 91st Congress, 2d. Session (1970), pp. 3–4.

[5] In model simulations, a monthly accounting period is replaced by an annualized one.

[6] Unearned income is also assumed to exclude current public assistance payments since these are to be replaced by the program simulated, though again this omission is not made explicit in the language of the bill.

prevalent in the base year of the file, (iv) data editing to impute essential information missing from the file or to assure consistency between family and person records, and (v) data condensation to eliminate nonessential information and shorten the length of the production run file.

In the initial version of the simulation model, the requestor would pay for nearly every one of these one-shot jobs with each production run made. The computer program had ageing and allocating preprocessing routines alongside routines determining eligibility for proposed reforms of current programs. Rather than requiring users to incur these multiple costs, the newer system is designed to offer the user a production run file with the results of all of these preprocessing one-time jobs written into the data file.

In addition, the revised system exhibits several other cost-saving features. Under the older system, costs soared for requests requiring large amounts of tabular information because output was stored in core while the program ran through each record on the data file. The standard tables printed at the end of production runs were quite extensive in detail and required much costly storage space. To avoid this situation, the user now may have the simulation results written onto a second file and utilize a report generator routine to produce standardized tabular information. If this route is not desired, under the new system, output stored in core may now be tailored more carefully to individual user needs. Finally, a request for several different simulations of the same basic plan may now be serviced with one pass of the microdata file; for example, one can vary some general grant or tax parameter such as exemption per filing unit. On the old system, separate passes of the file were required for each unique simulation. Reduction in computer and other costs is not unsubstantial, especially in terms of turnaround time.

2.3. *System Economies in Grant/Tax Eligibility Determination*

In simulating an imposing array of grant/tax proposals,[7] we recognized common threads among all redistributive fiscal programs. The sequence of eligibility determination stood out as a feature common to all simulations—(i) filing unit formulation, (ii) categorical "nonincome" filing unit eligibility determination, and (iii) means-tested "income" filing unit eligibility determination. A specific eligibility rule for one plan often had only slight variations from that of a previous simulation. For example, the definition of children in plans paying benefits to needy families with children usually set some maximum age beyond which all persons are considered adults; this maximum frequently varied from plan to plan. Below the maximum age, persons were only regarded as children within a certain age range if specified marital status and educational status conditions were met. The precise details of these requirements typically differed, but the general intent

[7] Over a 2½ year period, from September, 1969, to March, 1972, a total of 162 final production runs of different grant/tax simulations were produced on the original model. In addition, numerous special tabulations on program participants were requested during this period. Particular plans estimated varied from universal income-conditioned grant programs replacing current public assistance —see Nelson McClung, "Estimates of Income Transfer Program Direct Effects," *Technical Studies*, the President's Commission on Income Maintenance Programs (Washington, D.C.; 1970), pp. 135–142—to a package of 13 separate grant/tax program reforms simulated with one pass of the data file—see Robert S. Benson and Harold Wolman, ed., *Counterbudget* (New York: Praeger, 1971), pp. 47–67.

of the provisions remained invariant among the simulation requests. A more obvious commonality was the computational equivalence of income eligibility parameters (tax rates, credits, deductions, and exemptions) and grant gross-and-net-allowance computations.

In order to exploit these overlaps in cost and coverage estimation, the revised model was designed as a generalized grant/tax computational procedure. The flexibility inherent in this approach contrasts sharply with the rigid one-shot estimating procedures used in prior simulations. The two methods are distinguishable by the parameters and modules which characterize the general approach. Each separate model operation, such as constructing the filing unit or labeling persons as adults and children, is couched in a separate module of the system. Each module has its own generalized rule, with default settings for particular cases presumed to be most commonly requested. For nondefault settings, the appropriate parameters are simply juggled to yield specific rules for the simulation at hand. For example, one may wish to consider as unemployed only those persons reporting total duration of unemployment in excess of 5 weeks during the survey year. A parameter is provided in the generalized unemployment defined to allow specification of a minimum number of weeks unemployment during the year to qualify one as unemployed. Tripping the appropriate switch in the unemployment module yields the desired result.

The advantages of the generalized approach are many. Once the initial modularized system is linked and debugged, a particular job request may be serviced with far less programmer and analyst dollar and time cost than under the old system. The one-shot model required that a "blow-by-blow" account be submitted to the programmer in analytical language, and then be dutifully incorporated into the computer program. Each new request for estimates was grafted onto the program at its latest point of departure. Thus, there was no systematic retrieval system adopted for utilizing previously programmed eligibility rules reincarnated in part or in whole for a current simulation. In addition to providing such a retrieval system, the revised model also contributes an orderly, systematic framework for documenting the flood of decision rules that characterize most simulation requests. A record of mechanical parameter settings of a general rule is stored and later appears in the computer printout.

Furthermore, being familiar with the generalized definers greatly facilitates comparisons of simulations. A cost estimate for a family assistance plan made 2 years ago may in no way resemble an estimate made today; a wealth of detail may separate the two simulation runs and an accurate accounting of these divergencies is essential to avoid spurious comparisons. Haphazard, unsystematic documentation under the old model made such comparisons virtually impossible.

Of course, the new system is not perfect. Not every operation lends itself to parameterization; some tasks are by nature so program-specific that it is inefficient to generalize them.[8] As the complexity of a request mounts, so will mount the costs of implementation on the revised model. To the extent though that system economies may be exploited, the cost on the new system should be relatively

[8] The initial simulations on the new system include federal payroll taxes, federal income taxes, and public assistance. The system was thus able to subsume highly detailed tax and grant program specifications without sacrificing the generalized approach.

lower than cost of the same job on the old model. And offsetting the anticipated reduction in programmer/analyst cost per service request over time is the large initial sunk cost of creating and implementing the generalized structure. It remains to be seen how long it will be before the new system pays for itself. Also, in writing general rules, it is impossible to foresee the totality of policymaker fetishes in the area of income maintenance and tax programs. But the new system is elastic in the sense that it could incorporate a decision by policymakers to create a new screen, such as one on leisure time activities of married women, as a condition for program eligibility. Finally, the user should not be led to believe that simulations requiring information absent from the microdata file will stand any better chance of being estimated on the revised system.

3. Future Prospects

The market demand for microsimulation estimates of tax and grant programs should be steady for some time to come. Indeed, as potential users are made aware of the new technology, demand may increase substantially. With heightened awareness of the model's capabilities will unavoidably follow heightened awareness of the model's shortcomings. To date, the model contains no option to produce second-round cost and coverage estimates reflecting changes in labor supply; these changes will depend on characteristics of particular welfare reform proposals. In anticipation of this type of request, the research effort is currently trying to implement policy-relevant behavioral response coefficients into the model. Additional topics for research include lowering as well as raising the data file unemployment rate, and experimenting with part-year rather than annual accounting periods.

The revised system is designed to accommodate new CPS files as they become available, as well as other microdata sources. The new model runs off of a common-formatted, common-coded CPSEO file containing complete information from each contributing microdata set, the 1971 CPS and 1967 SEO respectively. Information common to the SEO file but not to the CPS is simply appended to the end of the common-format fields. As new data sources are added to the system, they will in turn be moulded into the common-format ordering and coding of the CPSEO. Thus, commonly coded information appearing in the same field on all microdata files has the same interpretation regardless of the data source. Once the initial cost of the data sorting is met for a newly acquired data set, the entire generalized grant/tax computational procedure is available for new simulations.[9] Although current data from CPS are of limited use for state and local decision-making, the common-format feature of the system makes analysis possible at these levels when adequate data become available.

If the micro-simulation system described here injects an element of objectivity into the grant/tax policy making domain and forces proposals into a rational framework, it will have been a successful effort.

Bates College, and
Urban Institute

[9] This statement is true to the extent that information required for particular modules is available on the newly acquired file.