Microsimulations in the Presence of Heterogeneity

Constantijn W.A. Panis

"Microsimulations in the Presence of Heterogeneity"

Constantijn W.A. Panis
RAND

May 2003

Michigan Retirement Research Center
University of Michigan
P.O. Box 1248
Ann Arbor, MI 48104

Microsimulations in the Presence of Heterogeneity

Constantijn W.A. Panis

Abstract

This paper develops a method that improves researchers' ability to account for behavioral responses to policy change in microsimulation models. Current microsimulation models are relatively simple, in part because of the technical difficulty of accounting for unobserved heterogeneity. This is all the more problematic because data constraints typically force researchers to limit their forecasting models to relatively few, mostly time-invariant explanatory covariates, so that much of the variation across individuals is unobserved. Furthermore, failure to account for unobservables often leads to biased estimates of structural parameters, which are critically important for measuring behavioral responses. This paper develops a theoretical approach to incorporate (univariate and multivariate) unobserved heterogeneity into microsimulation models; illustrates computer algorithms to efficiently implement heterogeneity in continuous and limited dependent models; and evaluates the importance of unobserved heterogeneity by conducting Monte Carlo simulations.

1. THE ISSUE

Microsimulation has become an important component of the evaluation of the potential effects of changes in government policy. These methods are applied to a wide variety of projections, including the distribution of retirement income or pension availability. However, the usefulness of microsimulations has been limited by the almost-universal assumption that various dimensions of behavior are independent and that there is no unmeasured heterogeneity in either the pattern of behavior over time or correlations in related behaviors. In other words, the underlying models consist of one or more independently estimated equations, all without unobserved heterogeneity and without correlated residuals. Parameter estimates of potentially endogenous covariates may be biased, resulting in biased forecasts. This paper develops an straightforward approach to account for unobserved heterogeneity and correlated residuals in microsimulations, thus greatly enhancing the value of microsimulations for policy analysis.

For purposes of our discussion, it is useful to distinguish two types of microsimulation models. The first type simulates outcomes for an artificially generated population, the second type simulates outcomes for actual individuals. PENSIM is an example of the first type. PENSIM is a pension simulation model developed by the Policy Simulation Group for the Department of Labor (Holmer, Janney, and Cohen 1999). It artificially generates a population with known distributions of sex, school start dates, retirement dates, and death dates. Given these characteristics, it simulates educational achievement, employment spells, earnings, and pension entitlements. It is straightforward to incorporate unobserved heterogeneity in this type of model.

MINT is an example of the second type. MINT is a retirement income simulation model developed by RAND and the Urban Institute/Brookings Institution for the Social Security Administration (Panis and Lillard 1999; Toder et al. 1999, Toder et al., 2002). It is based on data from the 1990-1993 Surveys of Income and Program Participation (SIPP). MINT forecasts marriage, divorce, disability status, mortality, earnings, Social Security and pensions entitlements, and wealth accumulation. It starts at the last SIPP interview date and projects outcomes until death. Since it takes actual individuals with an observed history of outcomes, it is difficult to determine their distribution of unobserved heterogeneity. To illustrate this point, consider earnings projections. Suppose the underlying model of earnings is estimated with account for normally distributed unobserved heterogeneity. The distribution has mean zero for the population, but not for any particular individual. For example, someone whose past earnings have been very high (controlling for age, education, etc.) is likely to have a large positive draw

of the heterogeneity distribution in earnings, and thus likely to continue enjoying high earnings. Given his or her earnings history, it would be incorrect to assume that this person's future earnings are subject to the zero-mean population distribution of heterogeneity.

2. THREE APPROACHES

We propose three approaches to account for unobserved heterogeneity in microsimulations. Throughout this paper, unobserved heterogeneity may refer to univariate or multivariate heterogeneity. For example, the empirical application discussed below is a joint model of marriage and divorce, where an individual's unobserved (bivariate) propensities to marry and divorce may be correlated. The three approaches are:

1. *Sample Distribution Approach*. Simulate under the assumption that the individual's heterogeneity distribution is the same as the sample distribution. As indicated above, this is only appropriate for simulation models of individuals without a history of outcomes.

2. *Bayesian Approach*. Following elementary Bayesian statistics theory, derive the (posterior) distribution of heterogeneity, given observed past outcomes.

3. *Optimal Assignment Approach*. For each individual in the sample, find the value of his heterogeneity component that best matches his observed past outcomes. Treat this value as an observed characteristic and simulate using traditional algorithms that ignore unobserved heterogeneity.

We now discuss each approach.

2.1. Sample Distribution Approach

Under this approach, microsimulations are conducted using the heterogeneity distribution that was estimated on the estimation sample.[1] This is only appropriate for models in which individuals are simulated from birth, so that there are no past outcomes with information on one's heterogeneity draw. This applies to models with artificially generated populations, but may also be applied to models based on surveys. In those cases, the simulation must not be run from, say, the last interview date, but from birth. This approach thus becomes a tremendously

---

[1] The estimation and simulation samples need not be the same in microsimulation models. Throughout, we assume that the heterogeneity distribution of estimation and simulation samples are identical. This is only relevant for the Bayesian and optimal assignment approaches.

useful tool for assessing model fit. Simulations may be run from birth to the (exogenously determined) last interview date. The extent to which summary measures of simulated events correspond to summary measures of actual data indicates goodness-of-fit. The value of this exercise, even where future events are of primary interest, can hardly be overstated.[2] This approach was used, among others, in Heckman and Walker (1987) and Blundell and Reed (2000).

2.2.   Bayesian Approach

Bayes' central theorem relates a conditional distribution to a joint and a marginal distribution:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

From this, it follows that $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$, i.e.,

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)}.$$

Denote unobserved heterogeneity by ε (possibly a vector) and observed past outcomes by X.[3] Switching notation,

$$P(\varepsilon|X) = P(X|\varepsilon)\frac{P(\varepsilon)}{P(X)}.$$

For our purposes, observed past outcomes are given, so that $P(X)$ is an irrelevant constant. Thinking in terms of probability densities, it follows that:

$$f(\varepsilon|X) \propto f(X|\varepsilon)f(\varepsilon).$$

---

[2] This approach may also be used to find starting values in extraordinarily complex models, such as dynamic programming models. Researchers often begin with plausible values, simulate the implied distribution of outcomes, compare that to the observed distribution, and manually tweak parameter values until the implied distribution resembles the actual distribution. In dynamic programming, this method is known as calibration.

[3] While not correct, microsimulations typically assume that estimated model parameters β are not subject to stochasticity. We make the same assumption here, but the arguments generalize readily to account for the entire distribution of β.

The left-hand-side of this expression is the distribution of unobserved heterogeneity, conditional on observed past outcomes. The first term of the right-hand-side, $f(X|\varepsilon)$, is simply the conditional likelihood function, which is straightforward to compute. The term $f(\varepsilon)$, finally, is the density function, as assumed in the estimation. In our application, we assume that heterogeneity is distributed bivariate normally. In the Bayesian literature, $f(\varepsilon)$ is the "prior," $f(\varepsilon|X)$ the "posterior" distribution. We illustrate the empirical shape of the posterior distribution for a hazard model of duration until marriage, below.

2.3. Optimal Assignment Approach

The third approach determines, for each individual in the sample, the optimal value of his heterogeneity component. The optimality criterion we use here is that the values of all individuals in the sample combined yield the greatest likelihood.

This approach makes explicit the idea that each individual in the sample has a certain heterogeneity value (or "draw"). The researcher does not observe individuals' draws, but can estimate their distribution. The optimal assignment approach reverse-engineers the data to figure out individuals' (most likely) actual draws.

For linear models, the average (over replications within individual) predicted residual may well be a good approximation of an individual's heterogeneity value. For limited dependent models, the issue is much less straightforward. Consider a model of the timing of marriage. We observe multiple marriages for a subset of respondents and estimate a multilevel hazard model with unobserved heterogeneity assuming, say, normally distributed heterogeneity. What is the most likely heterogeneity draw for a sample member who, at age 20, has not yet married? The draw that maximizes the likelihood of observing the lack of a marriage is minus infinity. If we would use this value to predict future marriages, we would predict that this person will never marry. The same issue arises whenever all limited-dependent outcomes take on boundary values: the most likely draws would be plus or minus infinity. The resulting distribution of most likely draws is unlikely to resemble the assumed normal distribution.

In searching for individuals' optimal heterogeneity draws, we therefore impose the restriction that the draws must be distributed according to the distribution that was estimated. The procedure is as follows.[4] Consider a sample of N individuals and multiple outcomes per

---

[4] This approach was brought to our attention by Douglas Wolf.

individual for at least a subset of the sample. We estimate a model with unobserved heterogeneity. We then draw N values from the estimated heterogeneity distribution, and attempt to allocate those N draws to the N sample members, such that the joint likelihood is maximized. In other words, we compute the $N*N$ conditional likelihoods of N potential values for N individuals. From among these $N!$ permutations, we determine the N pairs of draw-individuals that maximize the joint conditional likelihood.

While theoretically possible, it is extremely CPU-intensive to evaluate all $N!$ permutations. A much more efficient approach was developed in the operations research literature. The problem is known as the assignment problem (Thorndike, 1950; Koopmans and Beckmann, 1957) and may be solved using linear programming techniques. Several algorithms have been developed, including the Hungarian, Out-of-Kilter, and Simplex algorithms (Carpaneto, Martello, and Toth, 1988).

3. SIMULATION RESULTS

3.1.   The Model

The ability to simulate models in the presence of heterogeneity is particularly valuable when several processes directly affect each other and are thought to be subject to correlated residuals. Correlated residuals may arise, for example, when unobservables affecting several outcomes pertain to a single decisionmaker. Ignoring correlated residuals will bias estimates of the direct effect of one outcome on another. For example, participation in a job training program and job performance may both be affected by unobserved ability or ambition of the worker. Ignoring that correlation will overstate the effect of training on job performance. A microsimulation based on a model with biased parameter estimates will thus also overstate the benefits of targeting additional workers for training.

We chose a more subtle application, namely one with correlated residuals but without a direct (and potentially biased) effect of one outcome on the other. To illustrate the techniques and their differences, we use a model of the timing of (re-)marriage and divorce. The model is based on 1990-1993 SIPP data. These data contain full retrospective marriage histories for all respondents, i.e., with multiple weddings and divorces for a subset of respondents. The model follows the specifications of marriage and divorce in MINT, but incorporates unobserved heterogeneity. We model the transitions into marriage using a continuous time hazard model, also known as a failure-time model (e.g., Kalbfleisch and Prentice, 1980). Its basic form is given

by piecewise-linear Gompertz. The multiplicative effects on covariates on the hazard are equivalent to additive effects on the log-hazard:

$$\ln h_{ij}^m(t) = \Gamma_m(t) + \theta_m' X_{ij} + \varepsilon_i,$$

where $\ln h_{ij}^m(t)$ is the log-hazard that individual i marries for the j-th time. The marriage baseline hazard, $\Gamma_m(t)$, captures duration dependencies on respondent age, calendar time, and (for second and higher-order marriages) the duration since the previous marriage dissolved. All duration dependencies are assumed to be piecewise-linear (piecewise-Gompertz, generalized Gompertz), so that the estimated patterns may adjust to whatever patterns are in the data. All covariates X_{ij} are constant within spells; some, such as the number of previous marriages, differ across marriages, but do not vary over time within a spell. We account for individual-specific unobserved heterogeneity, ε_i, to capture differences in unobserved characteristics that affect the propensity to enter (re-)marriage. We assume that ε_i is distributed normally with mean zero.

Once married, the individual becomes subject to the risk of divorce. Similar to the model for marriage formation, we model marriage dissolutions using a continuous time hazard model:

$$\ln h_{ij}^d(t) = \Gamma_d(t) + \theta_d' X_{ij} + \delta_i$$

where $\ln h_{ij}^d(t)$ is the log-hazard of divorcing (d) for the j-th time. The baseline hazard, $\Gamma_d(t)$, captures piecewise-linear duration dependencies on the duration since the wedding, respondent age, and calendar time. All covariates are constant over the duration of the divorce spell; some, such as the number of previous marriages, differ across marriages, but do not vary over time within a marriage. We account for individual-specific unobserved heterogeneity, δ_i, to capture differences in unobserved characteristics that affect the propensity to divorce. We assume that ε_i and δ_i are jointly normally distributed:

$$\begin{pmatrix} \varepsilon_i \\ \delta_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon\delta} \\ \sigma_{\varepsilon\delta} & \sigma_\delta^2 \end{pmatrix} \right),$$

where $\sigma_{\varepsilon\delta} = \rho_{\varepsilon\delta}\sigma_\varepsilon\sigma_\delta$ and $\rho_{\varepsilon\delta}$ is the correlation between ε_i and δ_i.

For details on the data and the model specifications (without heterogeneity) see Panis and Lillard (1999).

Marriage and divorce spells may be censored by the last survey wave or by mortality. Divorce spells may further be censored by spousal mortality. In the simulations, we ignore mortality and simulate everyone until their 65th birthday.

Table 1 and Table 2 show parameter estimates of the hazard of marriage formation and divorce, respectively, without and with heterogeneity, separately for men and women. Table 3 summarizes the heterogeneity structure. For both men and women, both marriage formation and divorce appear to be subject to highly significant unobserved heterogeneity. The magnitude of heterogeneity is roughly equal in both processes and for both sexes, except that heterogeneity in the propensity to divorce appears to play a smaller role for men. The correlation between heterogeneity components in marriage formation and divorce is strong (0.62 for men and 0.75 for women) and statistically significant. This implies that some "marriage-happy" individuals ($\varepsilon_i > 0$) will also have above-average inclinations to divorce ($\delta_i > 0$), leading to relatively many marriages over their lifetime. Conversely, individuals with negative draws of ε_i and δ_i tend to marry late and remain married, thus leading to mostly zero or one marriage over the lifetime. Ignoring correlated heterogeneity will compress the distribution of number of marriages.

Table 1. Hazard Estimates of Marriage Formation

	Male		Female	
	Without heterogeneity	**With heterogeneity**	**Without heterogeneity**	**With heterogeneity**
Constant	-24.8227 ***	-25.0884 ***	-22.0353 ***	-22.2538 ***
	(0.9536)	(0.9560)	(0.3948)	(0.3981)
Age slope 0-16 years	1.2473 ***	1.2508 ***	1.1841 ***	1.1818 ***
	(0.0604)	(0.0605)	(0.0251)	(0.0253)
Age slope 16-20 years	0.6483 ***	0.6617 ***	0.3749 ***	0.4424 ***
	(0.0089)	(0.0090)	(0.0050)	(0.0056)
Age slope 20-25 years	0.0769 ***	0.1626 ***	-0.0475 ***	0.0461 ***
	(0.0029)	(0.0047)	(0.0026)	(0.0039)
Age slope 25+ years	-0.0502 ***	-0.0341 ***	-0.0755 ***	-0.0676 ***
	(0.0007)	(0.0011)	(0.0008)	(0.0010)
Duration since last marriage	0.1246 ***	0.1702 ***	0.1104 ***	0.1519 ***
Slope 0-3 years	(0.0108)	(0.0112)	(0.0100)	(0.0103)
Duration slope 3-8 years	-0.1147 ***	-0.0800 ***	-0.0893 ***	-0.0639 ***
	(0.0071)	(0.0075)	(0.0064)	(0.0066)
Duration slope 8+ years	-0.0389 ***	-0.0390 ***	-0.0239 ***	-0.0236 ***
	(0.0054)	(0.0055)	(0.0040)	(0.0041)
Calendar time (relative to 1980)	-0.0082 ***	-0.0093 ***	-0.0045 ***	-0.0055 ***
	(0.0003)	(0.0004)	(0.0002)	(0.0003)
Second marriage	0.4550 ***	-0.3142 ***	0.3213 ***	-0.6436 ***
	(0.0232)	(0.0383)	(0.0212)	(0.0333)
Third marriage	0.6881 ***	-0.3870 ***	0.5595 ***	-0.8198 ***
	(0.0304)	(0.0550)	(0.0273)	(0.0469)
Fourth or higher marriage	1.2833 ***	-0.1754 **	1.1462 ***	-0.7337 ***
	(0.0416)	(0.0764)	(0.0361)	(0.0627)
Black	-0.3548 ***	-0.4587 ***	-0.5081 ***	-0.6518 ***
	(0.0146)	(0.0190)	(0.0126)	(0.0167)
Native American	-0.1013 *	-0.1011	0.0789 **	0.1364 **
	(0.0521)	(0.0675)	(0.0378)	(0.0552)
Asian	-0.2736 ***	-0.4168 ***	-0.2149 ***	-0.3506 ***
	(0.0334)	(0.0431)	(0.0296)	(0.0391)
Hispanic	-0.0770 ***	-0.1165 ***	-0.2647 ***	-0.3538 ***
	(0.0167)	(0.0217)	(0.0156)	(0.0205)
High school drop-out	-0.0612 ***	-0.0473 ***	0.1137 ***	0.2097 ***
	(0.0106)	(0.0147)	(0.0091)	(0.0131)
College graduate	-0.1707 ***	-0.2755 ***	-0.4163 ***	-0.6042 ***
	(0.0108)	(0.0149)	(0.0117)	(0.0163)
Widowed	0.2809 ***	0.1726 ***	-0.3789 ***	-0.3267 ***
	(0.0280)	(0.0398)	(0.0242)	(0.0312)
Permanent income	0.0157 ***	0.0078	-0.0223 ***	-0.0397 ***
	(0.0041)	(0.0056)	(0.0033)	(0.0048)
$\hat{\sigma}_{\varepsilon}$		0.6729 ***		0.7067 ***
		(0.0162)		(0.0134)

NOTE: Asymptotic standard errors in parentheses;
Significance: '*'=10%; '**'=5%; '***'=1%.

Table 2. Hazard Estimates of Divorce

	Male		Female	
	Without heterogeneity	**With heterogeneity**	**Without heterogeneity**	**With heterogeneity**
Constant	-1.0198 ***	-1.5013 ***	-1.7268 ***	-2.8312 ***
	(0.1136)	(0.1363)	(0.0950)	(0.1182)
Age slope 0-30 years	-0.1193 ***	-0.1027 ***	-0.1021 ***	-0.0622 ***
	(0.0039)	(0.0048)	(0.0032)	(0.0042)
Age slope 30+ years	-0.0400 ***	-0.0325 ***	-0.0523 ***	-0.0370 ***
	(0.0016)	(0.0020)	(0.0016)	(0.0019)
Duration since wedding Slope 0-1 years	0.4339 ***	0.4208 ***	0.7350 ***	0.7073 ***
	(0.0743)	(0.0744)	(0.0691)	(0.0691)
Duration slope 1-4 years	0.2395 ***	0.2265 ***	0.1526 ***	0.1248 ***
	(0.0121)	(0.0125)	(0.0107)	(0.0110)
Duration slope 4-15 years	-0.0228 ***	-0.0301 ***	-0.0156 ***	-0.0351 ***
	(0.0033)	(0.0036)	(0.0030)	(0.0034)
Duration slope 15-25 years	-0.0386 ***	-0.0447 ***	-0.0275 ***	-0.0403 ***
	(0.0049)	(0.0051)	(0.0044)	(0.0045)
Duration slope 25+ years	-0.0875 ***	-0.0951 ***	-0.0832 ***	-0.1009 ***
	(0.0062)	(0.0063)	(0.0052)	(0.0054)
Calendar time before 1980 (spline slope)	0.0401 ***	0.0399 ***	0.0429 ***	0.0434 ***
	(0.0010)	(0.0010)	(0.0008)	(0.0008)
Calendar time after 1980	-0.0025	-0.0021	0.0058 ***	0.0102 ***
	(0.0020)	(0.0021)	(0.0019)	(0.0020)
Second marriage	0.5737 ***	0.3580 ***	0.6368 ***	0.0625
	(0.0256)	(0.0455)	(0.0232)	(0.0440)
Third or higher marriages	1.2503 ***	0.8482 ***	1.3584 ***	0.2754 ***
	(0.0407)	(0.0887)	(0.0338)	(0.0786)
Black	-0.0274	-0.0309	-0.0085	0.0190
	(0.0213)	(0.0220)	(0.0185)	(0.0208)
Native American	-0.2117 ***	-0.2370 ***	-0.1068 ***	-0.2111 ***
	(0.0210)	(0.0220)	(0.0215)	(0.0243)
Asian	0.1198 ***	0.1009 ***	0.1786 ***	0.1124 ***
	(0.0282)	(0.0295)	(0.0238)	(0.0274)
Hispanic	0.3339 ***	0.3498 ***	0.3237 ***	0.3952 ***
	(0.0788)	(0.0822)	(0.0610)	(0.0756)
High school drop-out	-0.6198 ***	-0.6466 ***	-0.6378 ***	-0.7133 ***
	(0.0710)	(0.0721)	(0.0607)	(0.0646)
College graduate	-0.3015 ***	-0.3115 ***	-0.2076 ***	-0.2570 ***
	(0.0351)	(0.0359)	(0.0312)	(0.0342)
$\hat{\sigma}_\delta$		0.3296 ***		0.6276 ***
		(0.0544)		(0.0300)

NOTE: Asymptotic standard errors in parentheses;
Significance: '*'=10%; '**'=5%; '***'=1%.

Table 3. Residual Structure

	Male	Female
	With heterogeneity	**With heterogeneity**
$\hat{\sigma}_\varepsilon$	0.6729 *** (0.0162)	0.7067 *** (0.0134)
$\hat{\sigma}_\delta$	0.3296 *** (0.0544)	0.6276 *** (0.0300)
$\rho_{\varepsilon\delta}$	0.6202 *** (0.1176)	0.7532 *** (0.0371)

NOTE: Asymptotic standard errors in parentheses;
Significance: '*'=10%; '**'=5%; '***'=1%.

3.2.   Model Fit

To assess goodness-of-fit of our model, we simulate marriage histories in-sample from respondents' 12th birthday until their last SIPP interview date and compare simulated and actual marriage histories. We only simulate women's marriage and divorce experiences, in part because unobserved heterogeneity plays a larger role among women than among men (Table 3), and in part because marriage history determines eligibility for Social Security spousal benefits, which are more important for women than for men. There are 57,252 women in the sample.

Table 4 shows the actual and simulated numbers of times that women married between their 12th birthday and the last interview date.[5] The fraction that remained single throughout the observed period is 9.71 percent. Both simulations understate that fraction, but the simulation with heterogeneity (9.12 percent) performed better than the simulation without heterogeneity (8.77 percent). In the data, 3.85 percent married three or more times. Again, the simulation with heterogeneity reproduces that fraction more closely (3.12 percent) than the simulation without heterogeneity (2.93 percent).

---

[5] The maximum actual number is eight, the maximum simulated with heterogeneity nine. In both cases, only one respondent reached the maximum, so that the cellsize rounds to 0.00 percent.

Table 4. Actual and Simulated Number of Marriages

	Actual data	Simulations	
		Without heterogeneity	With normal heterogeneity
0	9.71	8.77	9.12
1	69.84	74.10	73.82
2	16.60	14.21	13.94
3	3.11	2.24	2.41
4	0.45	0.56	0.52
5	0.13	0.11	0.16
6	0.07	0.02	0.02
7	0.09	0.00	0.01
8	0.00	0.00	0.00
9	0.00	0.00	0.00
Total	100.00	100.00	100.00

Figure 1 shows the cumulative distribution of the timing of first marriage, as a function of age, for women in the actual data and for microsimulations without and with heterogeneity. Both simulations track the data closely, except in the fraction that remain never married, as noted above.
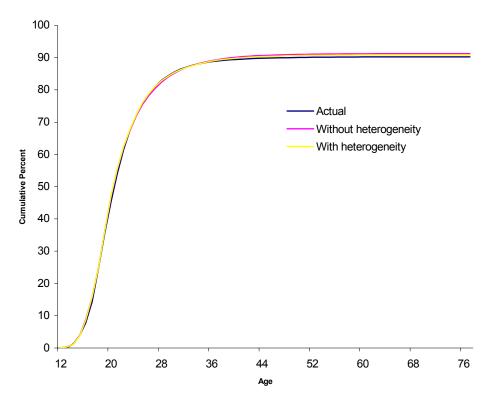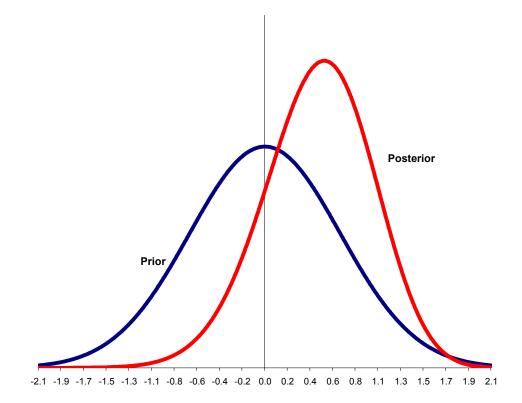

Figure 1. Cumulative Distribution of First Marriage as a Function of Age
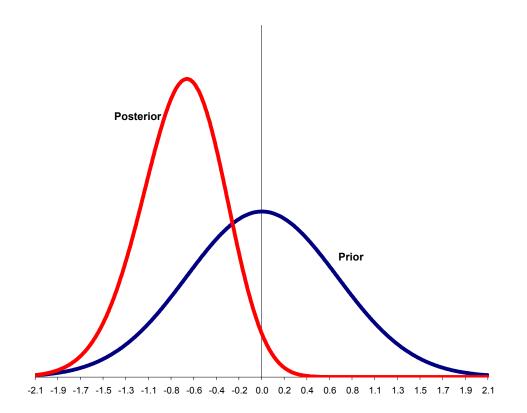
In summary, both in-sample simulations with and without heterogeneity track actual data fairly accurately. The simulation with heterogeneity appears to capture more dispersion in the outcomes.

3.3. Bayesian Posterior Distributions

As noted above, it is straightforward to derive the posterior density of unobserved heterogeneity as proportional to the product of the likelihood function and the prior density. In our application, the prior distribution is bivariate normal, resulting in a bivariate posterior distribution. To illustrate the shape of the posterior distribution, we restrict the model to marriage formation only, with a univariate normal heterogeneity distribution.

Figure 2. Prior and Posterior Distributions of Heterogeneity in Marriage: Female; married at ages 16, 30, and 44; divorced at last interview (age 48) Figure 2 graphs the prior and posterior distribution of the heterogeneity component in the hazard of marriage formation. The posterior distribution applies to a women who was born in 1943. She was married three times, from ages 16 to 25, ages 30 to 42, and ages 44 to 46. At the time of her last SIPP interview in 1992, she was 48 years old and divorced. The posterior distribution lies to the

right of the prior (sample) distribution, indicating that this woman has above-average propensity to marry.

**Figure 3. Prior and Posterior Distributions of Heterogeneity in Marriage:
Female; not yet married at last interview (age 55)**

Figure 3 illustrates the posterior distribution of a woman who had not yet married when she was last interviewed at age 55 in 1992. The posterior distribution lies to the left of the sample distribution, indicating that this woman has below-average propensity to marry.

In principle, microsimulations in the presence of unobserved heterogeneity may be conducted by drawing from individuals' (multivariate) posterior distributions. We leave this exercise for future work and simulate only using the optimal assignment approach.

3.4. Simulations Starting at the Last Interview Date

This subsection summarizes the results of simulations from the last interview date through age 65. Respondents who were age 65 or older at the time of the last interview are excluded from the

simulations. We simulated without and with heterogeneity; the simulation with heterogeneity used the optimal assignment approach.[6]

Table 5. Distribution of Simulated Number of Marriages at Age 65 (%; Women)

	Simulation without heterogeneity	Simulation with assigned heterogeneity
0	6.87	8.50
1	66.26	65.86
2	20.30	18.75
3	4.59	4.83
4	1.44	1.41
5	0.35	0.40
6	0.14	0.14
7	0.05	0.08
8	0.01	0.04
9	0.01	0.00
10	0.00	0.00
Total	100.00	100.00

Table 5 shows the distribution of simulated number of marriages as of individuals' 65th birthday. The fraction never married is higher in the simulation with heterogeneity (8.50 percent) than without heterogeneity (6.87 percent). The fraction individuals that will have been married three times or more is also higher: 6.90 percent with heterogeneity and 6.59 percent without heterogeneity. Both findings reflect greater dispersion of outcomes in the simulations that account for unobserved heterogeneity. Recall that the correlation between residual components in marriage formation and dissolution is 0.7532, i.e., large and positive. Individuals with positive draws of ε_i (quick-to-marry) are thus likely to also have positive draws of δ_i (quick-to-divorce), resulting in relatively many marriages over a lifetime. Conversely, individuals with negative draws of ε_i (slow-to-marry) are thus likely to also have negative draws of δ_i (slow-to-divorce), thus leading to zero or one marriage over a lifetime.

For purposes of Social Security policy, a particularly important outcome is whether an individual has been married for ten years or more, because this qualifies him or her for spousal benefits.[7] Table 6 shows the actual and simulated distributions of the fractions of individuals that were

---

[6] The simulation is based on over 43,000 women under age 65. The 43,000-by-43,000 assignment problem is very computer-intensive. In the current application, we therefore assign heterogeneity draws in blocks of 200 individuals. Larger blocks will improve the quality of assigned heterogeneity draws; we have not yet explored the practical implications of group size. Each draw is a 2-vector, i.e., a $(\varepsilon_i, \delta_i)$ pair.

[7] Similarly, a previous marriage of at least ten years qualifies the individual for ex-spousal benefits.

never married ten years or more, that were married to their current spouse for ten years or more when they turned 65 year old, and that were married to a previous spouse for ten years or more. The simulation with heterogeneity predicts a larger fraction women that were never married ten years or longer as of their 65th birthday. The difference, 2.34 percentage points or a relative difference of 17 percent, is substantively important. It implies that simulation models without heterogeneity may underestimate the number of women that are not eligible for (ex-)spousal benefits by one-in-six.

Table 6. Simulated Distributions of Ever Married 10+ Years (%; Women)

	Simulation without heterogeneity	Simulation with assigned heterogeneity
Never married 10+ years	13.80	16.14
Married 10+ years at 65th birthday	68.17	66.13
Previously married 10+ years	18.03	17.74
Total	100.00	100.00

4. CONCLUSION

This paper developed a method for simulating future outcomes of individuals with observed past outcomes in the presence of unobserved heterogeneity. It illustrated the feasibility in principle on a model of marriage and divorce. The results indicate that simulations with heterogeneity are better capable of capturing dispersion in the data, both in- and out-of-sample.

In future work, we intend to evaluate several computer algorithms that have been developed in the literature. The current applications uses the Hungarian algorithm (Carpaneto, Martello, and Toth, 1988), but initial explorations indicate that the Out-of-Kilter algorithm is more efficient for the problem at hand. We also intend to determine whether much may be gained from assigning all or very many observation-draws at once. In the current application, we drew 200 heterogeneity values (each a vector of two values) at a time and optimally matched them to 200 observations. Larger blocks will improve the quality of the matches, but it is not clear how much may be gained. Handling larger blocks dramatically increases the required computing time, at least with the Hungarian algorithm.

We also intend to further develop the Bayesian approach, that is, empirically derive the posterior distribution for every observation and draw from that distribution to obtain each observation's heterogeneity value for the microsimulation. How does this approach compare to the optimal assignment approach?

Finally, we intend to apply the methods to other substantive problems. In particular, we aim to apply them to problems in which independent treatment of correlated processes leads to a biased estimate of a structural parameter. While model fit in such cases may be quite good, out-of-sample predictions are likely very poor, and microsimulation models that incorporate correlated unobserved heterogeneity will greatly outperform simpler models.

REFERENCES

Blundell, Richard and H. Reed (2000), "The Employment Effects of the Working Families Tax Credit," *Fiscal Studies* 21: 75-103.

Carpaneto, G., S. Martello, and P. Toth (1988), "Algorithms and Codes for the Assignment Problem," *Annals of Operations Research*, 13: 193-223.

Heckman, James and James Walker (1987), "Using goodness of fit and other criteria to choose among competing duration models: A case study of Hutterite data." In C. Clogg, (ed.) *Sociological Methodology* 17: 247-307. American Sociological Association.

Holmer, Martin, Asa Janney, and Bob Cohen (1999), "PENSIM, a Pension Policy Simulation Model," Policy Simulation Group, July 1999.

Kalbfleisch, John D. and Ross L. Prentice (1980), *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, New York.

Koopmans, Tjalling, and Martin Beckmann (1957), "Assignment Problems and the Location of Economic Activities," *Econometrica* 25(1): 53-76.

Panis, Constantijn and Lee Lillard (1999), "Near Term Model Development, Part II," RAND, August 1999.

Thorndike, R.L. (1950), "The Problem of Classification of Personnel," *Psychometrica*, 15: 215-235.

Toder, Eric, Cori Uccello, John O'Hare, Melissa Favreault, Caroline Ratcliffe, Karen Smith, Gary Burtless, and Barry Bosworth (1999), "Modeling Income in the Near Term – Projections of Retirement Income Through 2020 for the 1931-60 Birth Cohorts," Urban Institute and Brookings Institution, September 1999.

Toder, Eric, Lawrence Thompson, Melissa Favreault, Richard Johnson, Kevin Perese, Caroline Ratcliffe, Karen Smith, Cori Uccello, Timothy Waidmann, Jillian Berk, Romina Woldemariam, Gary Burtless, Claudia Sahm, and Douglas Wolf. 2002. "Modeling Income in the Near Term: Revised Projections of Retirement Income Through 2020 for the 1931-1960 Birth Cohorts." Report prepared for the Social Security Administration, Urban Institute, June 2002.