IRLE
Institute for Research on Labor and Employment

IRLE WORKING PAPER
#152a-13
August 2013

Estimating Net Costs of Adopting Family Friendly Policies at Universities: Description of a Simulation Program

Clair Brown and Eric Freeman

Cite as: Clair Brown and Eric Freeman. (2013). "Estimating Net Costs of Adopting Family Friendly Policies at Universities: Description of a Simulation Program." IRLE Working Paper No. 152a-13.
http://irle.berkeley.edu/workingpapers/152a-13.pdf

Berkeley
UNIVERSITY OF CALIFORNIA

Estimating Net Costs of Adopting Family Friendly Policies at Universities: Description of a Simulation Program

Clair Brown and Eric Freeman[1]
UC Berkeley

August 30, 2013
Working Paper No. 152a-13
Institute for Research on Labor and Employment
University of California at Berkeley

Abstract:

We describe some of the modeling behind a simulation program that estimates the costs and benefits to a university of adopting certain family friendly policies for faculty members. This working paper is part of the Tools for Change project, which is funded by the National Science Foundation. . The co-Principal Investigators of the Tools for Change project are Mary Ann Mason, Former Dean of the Graduate Division, University of California, Berkeley, and Distinguished Professor Joan C. Williams, 1066 Foundation Chair and Founding Director of the Center for WorkLife Law at the University of California, Hastings College of the Law. The tool described in this paper can be found (or will be hosted in the future) at the following website: http://www.toolsforchangeinstem.org/.

[1] This project required the survey data and insightful analysis previously conducted by Marc Goulden, Mary Ann Mason, Angelica Stacy, and Sheldon Zedeck. We deeply appreciate the time that Marc Goulden spent with us so that we could use the UCB surveys. We also are very grateful for detailed help from Jason Pontius and Sandra Gahn with their work from a study at Iowa State University. We also benefited from discussions and input from Marc Goulden, Joan Williams, Mary Ann Mason, Noelia Sanchez, Angelica Stacy, Susan Carlson, Sandra Gahn, Jason Pontius, Nick Wolfinger, Sabrina Gordon, and Kariz Matic. Funded by the National Science Foundation under grant number 1106411. Any findings are those of the authors and do not necessarily reflect the views of the National Science Foundation. Authors alone are responsible for any errors (with apologies for any errors).

Building on a study of two Family Responsive policies (FRP), Prof. Clair Brown and Dr. Eric Freeman created a statistical simulation (called the Simulator), based upon regression analysis, of the estimated benefits and costs at universities and colleges. The Simulator provides a user friendly online tool that Department Chairs or Deans (or anyone who is interested) can use to estimate the costs and benefits of these Family Responsive policies at their institutions. The benefits include improved family formation and job satisfaction when faculty have more control over their work and family decisions and can better balance work and family to fit their career and family goals; in turn, these benefits for faculty can lead to improved retention, thus saving the university funds for startup costs and interviewing, for example. The costs essentially consist of the direct costs of the programs.

Two family responsive policies are studied: Active Service-Modified Duties (ASMD), and Stopping the Tenure Clock.[2] ASMD enables a professor who has a child to have reduced duties, e.g. a reduced teaching load. Stopping the Tenure Clock allows a tenure-track professor to delay her/his tenure review for one year; they are then to be judged in the review as if they had the same amount of time as other professors who did not stop the clock. The analysis is focused on science, technology, engineering, and math (STEM) fields, although non-STEM fields are included when data are available. The Simulator and results presented below are based on a broad definition of STEM that includies the social sciences.[3]

Two central steps were used to estimate the benefits and costs of the FRP:

> (i) Estimating the effects of the Family Responsive policies on faculty relevant for retention and career progression, which include birth of children and job satisfaction of faculty, given age, gender, marital status, tenure status, race, and department.
>
> (ii) Estimating the costs of attrition among faculty in STEM, especially the costs of hiring and developing STEM faculty, which include start-up grants and costs of the application process, as well as the direct costs of the Family Responsive policies.

To estimate the child bearing and job satisfaction of faculty (described below), we used data from three surveys of UC Berkeley faculty: the 2003 and 2009 Faculty Climate Surveys and the 2002 Faculty Work and Family Survey. Our use of these three surveys has been granted by the principal investigators (Prof. Mary Ann Mason, Associate Vice Provost Angelica Stacy, and Prof. Sheldon Zedeck) and has been reviewed and allowed by the UC Berkeley Committee for Protection of Human Subjects. The approval process was completed in early March 2012.[4]

To estimate retention rates over a five year period, given faculty characteristics and department, we used the public use versions of the Survey of Doctorate Recipients (SDR) waves from 2003, 2006, and 2008 (pooled).[5] This data set does not include humanities, and so only non-humanities outcomes are

---

2 A detailed summary of UC Berkeley policies can be found at http://ucfamilyedge.berkeley.edu/initiatives.html

3 In preliminary analyses and throughout much of the analysis, STEM was defined both broadly (including the social sciences) and narrowly (excluding social sciences), and the results were not sensitive to the definition of STEM when department controls are included.

4 We appreciate the help and support that Marc Goulden has been providing in accessing, understanding, and using these data sets.

5 We thank Nick Wolfinger for drawing our attention to the public use version of the dataset.

estimated and reported.

An important aspect of this work is that our estimates of these retention rates are indirect. We first make estimates of the effects of the policies on the number of children of, and the job satisfaction of, faculty members, using the UC Berkeley survey data. Then we make estimates of the effects of the number of children and job satisfaction on the retention of faculty, using the SDR data. Combining the two sets of estimates gives estimates of the effects of the policies on retention of faculty members.

Although UC Berkeley had Family Responsive policies available, the 2002 and 2003 faculty surveys indicated that many faculty did not know about them, or thought that use of them caused a stigma. After 2003, the university instituted a major campaign to educate faculty and department chairs about how to use Family Responsive policies and to change the climate and practices on campus for using FRP. The 2009 survey, plus feedback from faculty, department chairs, and administrators, indicated that the campaign had been successful in implementing FRP (in the sense of increasing the usage of the policies). The survey also showed that men increased their usage of the policies to a large extent. UC Berkeley also enhanced the FRPs after 2002.[6]

The individuals in the 2002 and 2003 surveys are treated as a control group for an institution that has not successfully implemented FRPs, and the individuals in the 2009 survey as a treatment group for an institution that has successfully implemented FRPs. We think that this is a conservative yet realistic approach, because higher education institutions often have FRPs “on the books” yet these policies have not been successfully implemented.[7] For institutions that do not offer FRPs, we would expect an even larger faculty response from implementing FRPs than the responses reported here.

In step (ii), we used a study of costs of FRPs at Iowa State University[8] to estimate the costs of faculty attrition and associated recruitment costs. We combined steps (i) and (ii) into a simulation of faculty flow through the university under the two different policy regimes to estimate the benefits minus the costs of the outcome of successful implementation of FRPs. This provides an estimate of the net benefit, or the value-added, of the programs.

This simulator provides a higher education organization with outcomes in two possible situations: In one, the university has fully implemented a suite of best-practice family-friendly policies; in the other situation, the university does not have these policies implemented.[9] The predictions are provided at the present time (t=0) and outcomes in five years (t=5). A comparison of outcomes in the two situations provides an estimate of the value added (or net benefit) of implementing the FRPs.

---

[6] These enhancements included increased offerings for ASMD and part-time work options, as well as central funding of ASMD (i.e. funding that was no longer at the departmental level).

[7] We note that an interesting additional line of research, which we did not have the data to pursue, would be to attempt to compare UCB with another university that did not make any changes to family friendly policies (if it had them at all) between the years 2002 and 2009.

[8] We thank Susan Carlson, Sandra Gahn and Jason Pontius for the use of some of their estimates of hiring and other costs from their Iowa State University study. We are grateful to Gahn and Pontius for their generosity with their time in helping us to use their data. Many of the details of this work were provided to us in personal correspondence; however, there is some information about this study available at www.advance.iastate.edu/conference/conferencepdf/2008_10-11gahncarlson_hoc1.pdf and http://www.wiche.edu/info/walf/meetings/takingStock/carlson.pdf (Both accessed as of August 2013.)

[9] It is assumed that these policies have been in place for multiple years before the present time, so that they have had a chance to have an effect.

Note that, as in many social science models, our data allow us to establish only correlation, not causation. It is of course difficult to make such predictions of causation without a controlled randomized trial, which of course (and obviously rightly so) is impossible for this topic. However, our data sets allow potential estimates of the benefits and costs of FRPs. There are many assumptions on which the accuracy of these estimates can be judged, and we call attention to one key assumption, namely that the environment in 2009 (and the preceding few years) at UC Berkeley differed from that in 2002/2003 (and the preceding few years) mostly in terms of the change in implementation of the FRPs. We think that this is a reasonable assumption to make in order to explore how faculty respond to family friendly policies.

Summary of Creation of Simulator (see Flow Charts I, II, and III)

Input: Upload File with Characteristics of all Faculty Members: Age, Gender, Tenure, Under-Represented Minority, Married/Partnered, Department
Output: Predicted Costs and Benefits

Notes on the required input file format. The file you upload should be a csv (comma separated value) file. You can create it using Microsoft Excel or an alternative such as OpenOffice; if you like, you can first create it as a file with a .xls extension and then later use "Save As" and choose the file type as csv. (Data should be separated by commas, not tabs.) The first row must have the following entries in exactly the following form: "Age","Female","Tenured","UnderRepresentedMinority","MarriedOrPartnered","Department". (Note that, for example, capitalization matters). Each row after the first row represents the respective characteristics of a faculty member in a department or university. Age should be a positive number. Female, Tenured, UnderRepresentedMinority, and MarriedOrPartnered are all variables that must have values either 0 or 1. So, for example, 1 for Female means the faculty member is female, and 0 means the faculty member is male. The meaning of UnderRepresentedMinority, defined to conform to the definition used by the SDR data sets we use, is a person of any race other than White non-Hispanic or Asian non-Hispanic. For Department, the variable must be 1, 2, 3 or 4; 1 represents Math/CS/Physical Sciences; 2 represents Life Sciences; 3 represents Engineering; 4 represents Social Sciences.

Basic Summary of the Prediction Methods:

The Simulator first uses the UCB data sets to predict the presence of children (in certain age groups) in the household of, and the job satisfaction of, a faculty member, in the case of FRPs being implemented and in the case of their not being implemented. (This is done by letting the 2009 observations represent the former scenario and the 2002/2003 observations the latter.) Then, as we do not have access to UCB retention data, the SDR data sets are used to predict the likelihood that a faculty member will stay at a university given the earlier predictions about child indicators and job satisfaction (as well as other control variables such as age and race). As well, data from an ISU study are used to estimate the costs of hiring, and also the costs of implementing FRPs. (Users can also provide their own estimates for these costs.) Finally, the simulator compares the increased costs from implementing FRPs against the possible benefits of reduced hiring costs if retention rates are increased.

Key datasets

- UCB data
  - 2002 UCB portion of 2002-2003 UC Work Family Survey
  - 2003 UCB Faculty Climate Survey

    - 2009 UCB Faculty Climate Survey

- SDR data (Survey of Doctorate Recipients)
    - 2003/2006/2008 Survey of Doctorate Recipients
    - No data on humanities Ph.D.'s

Key Steps For Predictions for an Individual Faculty Member
(Note that group level predictions are aggregations of individual-level predictions, given age, gender, marital status, tenure status, minority status, department, so we only describe the individual-level predictions.)

- (I) Qualitative Predictions
    - (A) Predicting Child Indicator Variables at time t=0
        - Child < 6 in hh (household), Child 6-11 in hh, Child 12-18 in hh
        - Uses UCB data
    - (B) Predicting Job Satisfaction at time t=0
        - Uses UCB data
        - Important to note that the scales of Job Satisfaction differ between 2003 and 2009; see Data Appendix.[10]
    - (C) Predicting Child < 6 in hh at time t=5
        - Uses retrospective regressions to estimate number of children in previous 5 year period, and uses this as estimate.
        - Uses UCB data
    - (D) Predicting Probability of Retention at time t=5
        - Uses SDR data
        - Key inputs are job satisfaction and child indicators.

- (II) Take-up Predictions
    - (A) Estimate expected number of uses of ASMD
        - Uses UCB data
        - (i) Predicts total number of children that a faculty member will have in years t=0 through t=5
            - Uses retrospective regressions to estimate number of children in previous 5 year period, and uses this as estimate.
        - (ii) Estimate probability of ASMD usage given that one has a child
        - Multiply (i) by (ii) to get expected number of uses of ASMD in years t=0 through t=5
    - (B) Estimate Probability of Stopping Tenure Clock in years t=0 through t=5
        - Uses UCB data
        - Use probit regression over last 5 years in order to estimate this.

[10] The SDR scale for job satisfaction matches the 2003 scale. See the Data Appendix.

        - Assumes that tenure clock is stopped either zero or 1 times (i.e., assumes it is never stopped twice).

- (III) Cost/Benefit Predictions
    - (A) Estimate Average Costs for Recruiting and for Using Policies (if used)
        - Most of the estimates are provided from an Iowa State University (ISU) study; data provided to us by the authors: Susan Carlson, Sandra Gahn and Jason Pontius.
        - These estimates are provided as defaults, but can be replaced by the user's own estimates if preferred, e.g. if an administrator has more accurate estimates for their university or department.
        - Four estimates:
            - Hiring and Startup Costs (Total over first 5 Years, including Startup Costs). There are estimates of these for each of four departmental groups. The default estimates we provide from the ISU study depend on the department of the faculty member.
            - Administrative Cost of Using One Semester of ASMD or of Stopping the Tenure Clock Once. Default is from the ISU study.
            - Cost to Hire Replacement Lecturer for One Semester of ASMD Usage. The default value of $7000 is an estimate provided by a UCB administrator.
            - Cost of Stopping the Tenure Clock (for One Year). The default estimate is $0.
    - (B) Calculate Costs to Implement Policies for ASMD and Tenure Clock Stoppage
        - Assumes that a replacement faculty member uses the policies with the same likelihood as the original faculty member, so costs are the same whether the faculty member is retained or not.
        - Multiply expected number of uses of ASMD by: [(Cost to Hire Replacement Lecturer for One Semester of ASMD Usage) + (Administrative Cost of Using One Semester of ASMD or of Stopping the Tenure Clock Once)]
            - Assume there is at most one semester of ASMD usage for each child
        - Multiply expected probability of Tenure Clock Stoppage by: [(Cost to Stop the Tenure Clock) + (Administrative Cost of Using One Semester of ASMD or of Stopping the Tenure Clock)]
    - (C) Calculate Costs for Recruiting and Hiring
        - Multiply the user-provided (or default) estimate for hiring/startup costs by the probability that a faculty member is retained.
    - (D) Calculate Total Costs/Benefits
        - Sum the Policy Costs (B), and the Costs/Benefits for Recruiting and Hiring (C)

Discussion of the Method Used

In many of the estimated relationships using UCB data, the key variable is the 2009 indicator dummy variable. This is used in our design as a proxy variable for having stronger policies in place, as the university had undertaken important activities to educate faculty and department chairs about FRPs after 2003 in order to change the campus climate so that use of these policies became standard practice.

These policies had a higher take-up rate in 2009 than in 2002 or 2003, and also child-bearing rates increased, as we see below.

In order to estimate faculty responses by major department categories, we had to classify the three UCB surveys and the three SDR surveys by department categories, and match the categories across surveys as well as possible. The data appendix shows what departments or schools each major department category contains. UCB surveys included all departments and schools, and so we were able to create a non-STEM category (mostly humanities), which we use as our benchmark (i.e., excluded) category for the department dummies in the regressions.[11] The SDR does not include respondents in the humanities, and so only STEM categories are used in estimations using the SDR to estimate retention for each department category.

In many of the estimated relationships, we often restrict to particular age ranges to make the empirical work realistic (eg, to reflect realistic child-bearing ages for female faculty).

In general, we encountered two problems in estimating our simple regressions.

First, in some of the probit regressions, the estimation process did not converge or there was one coefficient that was not well specified. There is a technical reason for this called "complete separation" or "quasi-complete separation."
(See, e.g., http://www.ats.ucla.edu/stat/mult_pkg/faq/general/complete_separation_logit_models.htm)

An example of this kind of situation occurs when everyone (or almost everyone) who is male and not married/partnered has no children. Then to get the best possible statistical fit, the algorithm fitting the probit (i.e. maximizing a likelihood function) essentially "wants" to choose the "coefficient" on "not married" to be as negative as possible, and would go to "negative infinity" if it could, but the algorithm stops, because -5 is similar to "negative infinity" when they are inputs to the standard normal cumulative distribution function. So the approach we take is to predict that a male has no children if not married, and then restrict the regression to males who are married/partnered.

Second, some relationships that are assumed to be a function of age encountered estimation problems. (The regressions include an age and age-squared term.) For example, we might expect the predicted number of children under age 6 to increase until around age 35 or 40 and then decrease afterwards. In some cases in which the estimated model did not behave in this fashion, in order to have age function realistically, we instead chose to do such things as restricting to age ranges where the relationship made more sense (e.g. to females <= 50 years old).

Please see the data appendix for a detailed description of the variables and how they were created from the UCB faculty surveys and the SDR.

Estimated Relationships

Before discussing the relationships estimated using the pooled UCB faculty surveys, we look at the

[11] We ran some sensitivity tests for the models using the UCB surveys to investigate sensitivity of the estimateions to how departments are classified and to see if department dummy captured variation across departments, or if estimations needed to be done for each department category. These tests indicated that the estimations across departments displayed similar patterns and department dummies captured the variation.

number of cases available (informally, N). The following table gives the number of observations in each of the three UCB surveys, by gender and departmental group. (Note that in many of the regressions, we pool the 2009 survey together with one of the other two surveys.)

	Female			Male		
	2002	2003	2009	2002	2003	2009
(i) Math/Physics	9	15	12	83	96	76
(ii) Life Sciences	32	38	39	116	114	86
(iii) Social Sciences	47	50	65	91	98	79
(iv) Engineering	13	13	14	110	91	63
Non-STEM	80	84	88	114	129	90

Note in particular the low number for Females in Math/Physics and Engineering. Interestingly, there were actually multiple regressions for which estimates for the departmental indicator variable coefficients were not that different, and the estimates were seldom significant. But of course that could just be due to the small sample size.

Here are the number of observations in our SDR sample, by gender and departmental group (non-STEM not available): [12]

	Female	Male
(i) Math/CS/Physics	392	1185
(ii) Life Sciences	315	635
(iii) Social Sciences	663	990
(iv) Engineering	71	482

First we estimate child bearing behavior. The child bearing outcomes are measured by three child indicator variables:
Child < 6 in household,
Child 6-11 (inclusive) in household,
Child 12-18 (inclusive) in household

(I) (A) Predicting Child Indicator Variables at time t=0 in the two universes (with and without policies)

Use 6 probit regressions: 3 outcomes x 2 genders

Three outcomes: Child < 6, Child 6-11 (inclusive), Child 12-18 (inclusive)

Estimates a probit regression using data from UCB surveys. These probits pool the 2002 and 2009 data. Note that we do not use 2003 data here because the 2003 survey asks only the question "How many children under 18 have you taken care of?" for both children in the past and current children. It does not tell us any more about children's specific ages.

The six probit regressions have the following form (for females, and for males, separately) Prob(Child in age category in household) = Phi(constant, 2009 dummy, Age, Age Squared, department dummies, Under-Represented Minority dummy, Married or Partnered dummy, Tenured dummy) where the department dummies are for Math/CS/Physical Sciences; Life Sciences; Engineering; Social Sciences (non-STEM fields are the excluded category). (Phi here represents the cumulative distribution function for the standard normal distribution.)

(i) Child < 6 in hh, restricted to females (N = 223)

---

[12] Each of these observations appears in both the 2003 and 2008 waves (and each is counted only once in this chart).

Indicator for Having At Least One Child (Strictly) Younger than 6 Years Old (Inclusive), RESTRICTED TO FEMALES UNDER 51 YEARS OLD; Predict 0 probability if age > 50.

(ii) Child < 6 in hh, restricted to males (N= 628)
Indicator for Having At Least One Child (Strictly) Younger than 6 Years Old (Inclusive), RESTRICTED TO THOSE WHO ARE NOW MARRIED OR PARTNERED AND AGE AT MOST 65, MALE ONLY; predict 0 if not married/partnered, and predict 0 if >65.

(iii) Child 6 to 11 (inclusive) in hh, restricted to females (N= 274)
Indicator for Having At Least One Child 6 to 11 Years Old (Inclusive), RESTRICTED TO AGES 36-60, INCLUSIVE, FEMALE ONLY; Predict 0 probability if age <= 35 or > 60

(iv) Child 6 to 11 (inclusive) in hh, restricted to males (N = 651)
Indicator for Having At Least One Child 6 to 11 Years Old (Inclusive), RESTRICTED TO AGES 36-65, INCLUSIVE, MALE ONLY; Predict 0 probability if age <= 35 or > 65.

(v) Child 12 to 18 (inclusive) in hh, restricted to females (N=274)
Indicator for Having At Least One Child 12 to 18 Years Old (Inclusive), RESTRICTED TO AGES 36-60, INCLUSIVE, FEMALE ONLY; Predict 0 probability if age <= 35 or > 60.

(vi) Child 12 to 18 (inclusive) in hh, restricted to males (N=651)
Indicator for Having At Least One Child 12 to 18 Years Old (Inclusive), RESTRICTED TO AGES 36-65, INCLUSIVE, MALE ONLY; Predict 0 probability if age <= 35 or > 65.

Next we estimate job satisfaction. Unfortunately we cannot use child indicator variables as controls because of data limitations.

(I) (B) Predicting Job Satisfaction at time t=0

Use 2 linear regression models: 2 genders

Job satisfaction responses take the values 0,1,2, and 3. The model can predict non-integer values.[13] The numbers outside the range 0 to 3 are truncated to either 0 or 3.

Uses UCB data; these regressions pool the 2003 and 2009 data. Note that we do not use 2002 data here because the 2002 data does not have data on job satisfaction. (We cannot use child indicators as control variables because the 2003 data does not have data on child indicators.)

We noticed stronger associations for assistant professors, so we included a term that interacts the 2009 dummy term with the tenured variable.

The two regressions have the following form (for females and for males, separately)
Job satisfaction = (constant, 2009 dummy, Age, Age Squared, Department dummies, Under-Represented Minority dummy, Married or Partnered dummy, Tenured dummy, Tenured/Dummy Interaction Term)
Where the department dummies are for Math/CS/Physical Sciences; Life Sciences; Engineering; Social Sciences.

Female regression (N=354)

---

[13] One could of course imagine using an ordered probit model here; we chose to use a linear regression as we use the predicted values in a further probit regression that takes continuous values as inputs and thus fractional differences in the predictions are relevant to this further model.

Male regression (N=788)

Now we turn to estimating the costs of the FRPs. First we look at the faculty use of ASMD, and predict (i) the number of children a faculty member will have in the next five years in situation where FRPs are available, and (ii) the probability that the faculty member will use ASMD, using UCB data.

(I) (C) Predicting Child < 6 in hh at time t=5

Three cases:

(i) If female and EITHER (a) > 50 now or (b) had a child 12 to 18 in hh as of 5 years ago, then predict 0 for the probability of having a child under 6 now. NOTE THAT in practice, the input of "having a child 12 to 18 in household" is a predicted probability from another regression; so what the simulator does, is that if there is a child 12 to 18 in hh as of 5 years ago with 99% or greater probability, then it predicts 0 for the probability of having a child under 6 now. Otherwise it looks at case (ii).

(ii) If female and BOTH (a) <= 50 now AND (b) did NOT have a child 12 to 18 in hh as of 5 years ago (or, in practice in the simulator, had one with <1% probability), then use a probit, as described below.

(iii) If male, use the probit described below.

(ii) If female and BOTH (a) <= 50 now AND (b) do NOT have a child 12 to 18 in hh as of 5 years ago (or, in practice in the simulator, have one with <1% probability), then use this probit regression.

Uses pooled 2002 and 2009 UCB data, RESTRICTED TO THOSE AT MOST 50 YEARS OLD AT TIME OF SURVEY AND WITH NO CHILDREN 12 TO 18 YEARS OLD AS OF 5 YEARS AGO, AND RESTRICTED TO THOSE WHO WERE AT UCB AS OF 5 YEARS AGO, FEMALE ONLY Using Characteristics 5 Years Before the Survey.

PR(Has a Child < 6 Years Old) = Phi(constant, 2009 Dummy, Age 5 Years Before the Survey, Square of Age 5 Years Before the Survey, Indicator for Child Less Than 6 as of 5 Years Before the Survey, Indicator for Child 6 to 11 Years Old as of 5 Years Before the Survey, Tenured As of 5 Years Before the Survey, Under-Represented Minority, Department dummies)
Where the department dummies are for Math/CS/Physical Sciences; Life Sciences; Engineering; Social Sciences.
N=132

(iii) If male:
Characteristics 5 Years Before the Survey, using Pooled UCB 2002 and 2009 Survey, MALE ONLY

PR (Now Has a Child < 6 Years Old) = Phi(constant, 2009 Dummy, Age 5 Years Before the Survey, Square of Age 5 Years Before the Survey, Indicator for Child Less Than 6 as of 5 Years Before the Survey, Indicator for Child 6 to 11 Years Old as of 5 Years Before the Survey, Indicator for Child 12 to 18 Years Old as of 5 Years Before the Survey, Tenured As of 5 Years Before the Survey, Under-Represented Minority, Department dummies)
Where the department dummies are for Math/CS/Physical Sciences; Life Sciences; Engineering; Social Sciences.
N=817

(II) (A) Estimate expected number of uses of ASMD, using UCB data.

(i) Predict total number of children that a faculty member will have in years t=0 through t=5
We use linear regressions to estimate the number of children in previous 5 year period, and use this as an estimate of number of children in the next 5 years.
Note that in estimating the regression coefficients, we use characteristics from 5 years ago; in the prediction stage, we use characteristics as of t=0.

Use 2 linear regression models: 2 genders

Uses pooled 2002 and 2009 UCB data.

We truncate at 0 children (no negative values used).

The two regressions have the following form (for females and for males, separately)

Retrospective Regressions of Number of Children Born in Previous 5 Years (Exclusive of Current Year) on Characteristics 5 Years Before the Survey = (constant, 2009 dummy, Age 5 Years Before the Survey, Square of Age 5 Years Before the Survey, Indicator for Child Less Than 6 as of 5 Years Before the Survey, Indicator for Child 6 to 11 Years Old as of 5 Years Before the Survey, Indicator for Child 12 to 18 Years Old as of 5 Years Before the Survey, Tenured As of 5 Years Before the Survey, Under-Represented Minority, Department dummies),
Where the department dummies are for Math/CS/Physical Sciences; Life Sciences; Engineering; Social Sciences.

Female regression (N=364)

Male regression (N=817)

(ii) Estimate probability of ASMD usage given that one has a child
(Note that in estimating the probit coefficients, we use characteristics from the time of the most recent child.)
Use probability of ASMD usage for the most recent child; then assume it is the same conditional probability for each child.[14]

Use 2 probit regression models: 2 genders

Uses pooled 2002 and 2009 UCB data.

We truncate at 0 children (no negative values used).

The two regressions have the following form (for females and for males, separately)

Prob(Using ASMD for Most Recent Child) = Phi(constant, 2009 dummy, Age at Time of Most Recent Birth, Age at Time of Most Recent Birth Squared, Indicator for Child Age 6 to 11 Inclusive in Household at Time of Most Recent Birth, Indicator for Child Age 12 to 18 Inclusive, in Household at Time of Most Recent Birth, Under-Represented Minority, Department dummies, Tenured at Time of Most Recent Birth),
Where the department dummies are for Math/CS/Physical Sciences; Life Sciences; Engineering; Social Sciences.

Female regression (N=113)

Male regression (N=239)

(iii) Multiply predictions from (i) by predictions from (ii) to get expected number of uses of ASMD in years t=0 through t=5

Here we make the assumption that if a faculty member has multiple children in years t=0 through t=5, then their probability of using ASMD given that one has a child is the same for all of these children. This is a limitation of our data, as the 2009 UCB survey asked only about usage of ASMD for the most

---

[14] This is a limitation of the data- see below for more discussion of this.

recent child. (Of course, in practice, one might guess that if a faculty member uses ASMD for their most recent child, they are likely to also have used it with their second most recent child.) Having made this assumption, we can then argue as follows, in order to spell out why the simple multiplication of the predictions in (i) and (ii) is justified. (Note that we also ignore instances of twins in this heuristic argument.)

Expected Value Of: (Uses of ASMD from t=0 to t=5 for a particular faculty member)
= Expected Value Of: (Uses of ASMD | exactly 1 child from t=0 to t=5) * Prob(exactly 1 child in period t=0 to t=5)
+ Expected Value Of: (Uses of ASMD | exactly 2 children in period t=0 to t=5) * Prob(exactly 2 children in period t=0 to t=5)
+ Expected Value Of: (Uses of ASMD | exactly 3 children in period t=0 to t=5) * Prob(exactly 3 children in period t=0 to t=5)
+ . . .
= E[Uses of ASMD | exactly 1 child in period] * Prob(exactly 1 child in period t=0 to t=5)
+ { (Expected Value Of: (Uses of ASMD for first child in period| exactly 2 children in period t=0 to t=5)) + (Expected Value Of: (Uses of ASMD for second child in period | exactly 2 children in period t=0 to t=5)) } * Prob(exactly 2 children in period t=0 to t=5)
+ { (Expected Value Of: (Uses of ASMD for first child in period | exactly 3 children in period t=0 to t=5)) + (Expected Value Of: (Uses of ASMD for second child in period | exactly 3 children in period t=0 to t=5)) + (Expected Value Of: (Uses of ASMD for third child in period | exactly 3 children in period t=0 to t=5))} * Prob(exactly 3 children in period t=0 to t=5)
+ . . .
(by our assumption, defining EVU= Expected Value of the Number of Uses of ASMD for the most recent child, i.e. probability of usage of ASMD for the most recent child)
= EVU * Prob(exactly 1 child in period t=0 to t=5)
+ { EVU + EVU } * Prob(exactly 2 children in period t=0 to t=5)
+ { EVU + EVU + EVU } * Prob(exactly 3 children in period t=0 to t=5)
+ . . .
= EVU * { Prob(exactly 1 child in period t=0 to t=5)
+ 2 * Prob(exactly 2 children in period t=0 to t=5)
+ 3 * Prob(exactly 3 children in period t=0 to t=5) }
= EVU * Expected Value of Number of children in period t=0 to t=5

(II)(B) Estimate Probability of Stopping Tenure Clock in years t=0 through t=5

Assumes that tenure clock is stopped only once, which has only rare exceptions. To be eligible for Stopping the Tenure Clock, one must have a newborn (or a newly adopted or foster child under 5) and one must not have tenure. Therefore these estimates are restricted to those who do NOT have tenure at the time of the most recent birth, and we do not put an indicator for having a child <6 on the RHS. Note that in estimating the probits, we use age at time of most recent birth.

These probits only use three RHS variables as the sample size is fairly small. When we included indicators for children 6-11 in hh, and children 12-18 in hh, as control variables, the estimation process did not converge, probably because there were only a few cases where a faculty member had children this old and was now untenured. So we did not use these indicators as control variables.

Probit regression using pooled 2002 and 2009 data from UCB surveys.

For FEMALE or MALE, if older than 45, predict 0.

Probit regressions estimated for faculty not tenured at birth of most recent child, separate regressions for females and males $\leq$45

Prob(Indicator for Stopping Tenure Clock for Most Recent Child) = Phi(constant, 2009 Dummy, Age at Time of Most Recent Birth, Age at Time of Most Recent Birth Squared)

Female N=42

Male N=64

Next we predict the probability of retention of the faculty member at the end of 5 years. Here we use the Survey of Doctorate Recipients (SDR) data because we did not have this data for UCB faculty.

(I)(D) Predicting Probability of Retention at time t=5

These models use linked data from the 2003, 2006 and 2008 waves of the SDR. For an observation to be used, it must be present in all three waves. (This is necessary in order to detect retention, as the questionnaires only ask if they were at the same job as in the *previous* wave; i.e., the 2008 questionnaire does not ask if they were at the same job in 2003 as in 2008.)

We estimate 8 regressions (2 genders in 4 department categories).

The 4 departmental groups in the SDR are:

(i) Composite of both Postsecondary Teachers-Computer and Math Sciences and Postsecondary Teachers-Physical and Related Sciences [These had to be combined in order to line up with the UCB departmental classification.]

(ii) Postsecondary Teachers-Life Related Sciences.

(iii) Postsecondary Teachers-Social and Related Sciences

(iv) Postsecondary Teachers - Engineering

For each gender in a given department, we estimate the following regression:

Uses linked 2003 and 2008 SDR waves: Same Job in week of October 1 2008 as in October 1 2003 (as well as in week of April 1 2006), ALSO RESTRICTED TO Composite of both Postsecondary Teachers-Computer and Math Sciences and Postsecondary Teachers-Physical and Related Sciences.

Uses 2003 Characteristics with the exception of the 2008 variable indicating if there is a child less than 6 in the household in 2008. Only for sample of observations present in ALL three waves (so including 2006).

Uses survey weights from the SDR. The survey has a complex structure.[15] See http://www.nsf.gov/statistics/srvydoctoratework/ for details about the survey structure.

For each gender in each department (2x4 regressions):

PR (respondent in same job in 2008 as in 2003) = Phi(constant, Age, Age Squared, Child under 6 in household in 2003, Child 6 to 11 in household in 2003, Child 12 to 18 in household in 2003, Child under 6 in household in 2008, Job Satisfaction, Under-represented Minority.)

An Example of how to use the Simulator

These regression estimations provide the estimations of the benefits and costs of implementing FRP policies. Let us go through an example. It is important to note that the exact form of the simulator (and its inputs) may change after the date of this publication.

We look at the results of the simulator for a fictional department within the Life Sciences of 20 faculty members. We use a (fictitious) file of the characteristics of the 20 faculty members, which represents

[15] The public use dataset does not contain any information on stratification etc; thus standard errors are not correct.

the type of file that a department chair would input into the simulator to see the predicted results for her department.[16] (Note that a Life Sciences department, because of large startup costs, might be among those expected to respond most positively to FRP.) One can follow this example by copying the text precisely below into a Microsoft Excel file and saving it as a csv file and then using it as the input file for the simulator. There are other ways to create this file.[17]

<u>Input file for example:</u>

```
"Age","Female","Tenured","UnderRepresentedMinority","MarriedOrPartnered","Department"
31,1,0,0,1,2
33,0,0,1,0,2
34,1,0,0,1,2
34,1,1,1,1,2
36,0,1,0,0,2
39,1,1,0,1,2
44,0,0,0,1,2
46,1,0,0,1,2
51,1,1,0,1,2
52,0,0,1,1,2
53,1,1,0,1,2
54,0,1,0,0,2
55,1,1,0,1,2
56,0,1,0,1,2
58,0,0,0,1,2
59,0,1,0,1,2
60,1,1,0,1,2
65,0,1,1,0,2
65,1,1,0,1,2
72,0,1,0,1,2
```

For our example, we use the default values for the cost estimates.[18] The simulator makes the prediction that if this department of 20 people were at a university ***without*** best practice FRP policies in place, the expected costs for recruiting and hiring for the department cumulative over a five year period will be about $1.916 million, while these costs, if the department were at a university ***with*** best practice FRP policies in place, these costs would be estimated to be about $1.774 million. At a university ***without***

---

[16] This file of data is in part randomly generated and only partial attempts were made to have the data be representative of a typical Life Sciences department. So obviously no conclusions should be made based on this exercise. We note that we do have data on some such departments, but of course it is confidential data, so we obviously cannot use it for this example.

[17] Here is another way that should work although the wording may differ, as it has only been tested on an open source analogue of Microsoft Word:

Open a new (blank) file with a simple text editor such as Microsoft Word (or an analogue). Copy the text so that the first line of the file is exactly:
"Age","Female","Tenured","UnderRepresentedMinority","MarriedOrPartnered","Department"

The last line (72,0,1,0,1,2) should be the final line of data. Then paste this into the new file.

Then go to Save As, and save the file as a .txt file.

[18] For example, the estimate of hiring costs for Life Sciences is $275,581, the cost to hire a replacement lecturer is $7,000 and the administrative cost is $625.

such policies, of course there is no cost to implement the policies, while at a university ***with*** the policies, the costs are estimated at about $19,600.[19] Summing costs of both (i) hiring and (ii) policy implementation gives total costs of $1.916 million if the department were at a university ***without*** best practice policies and total costs of $1.794 million if the department were at a university ***with*** best practice policies. The estimated net savings from FRP policies for the department is approximately $122,000.[20]

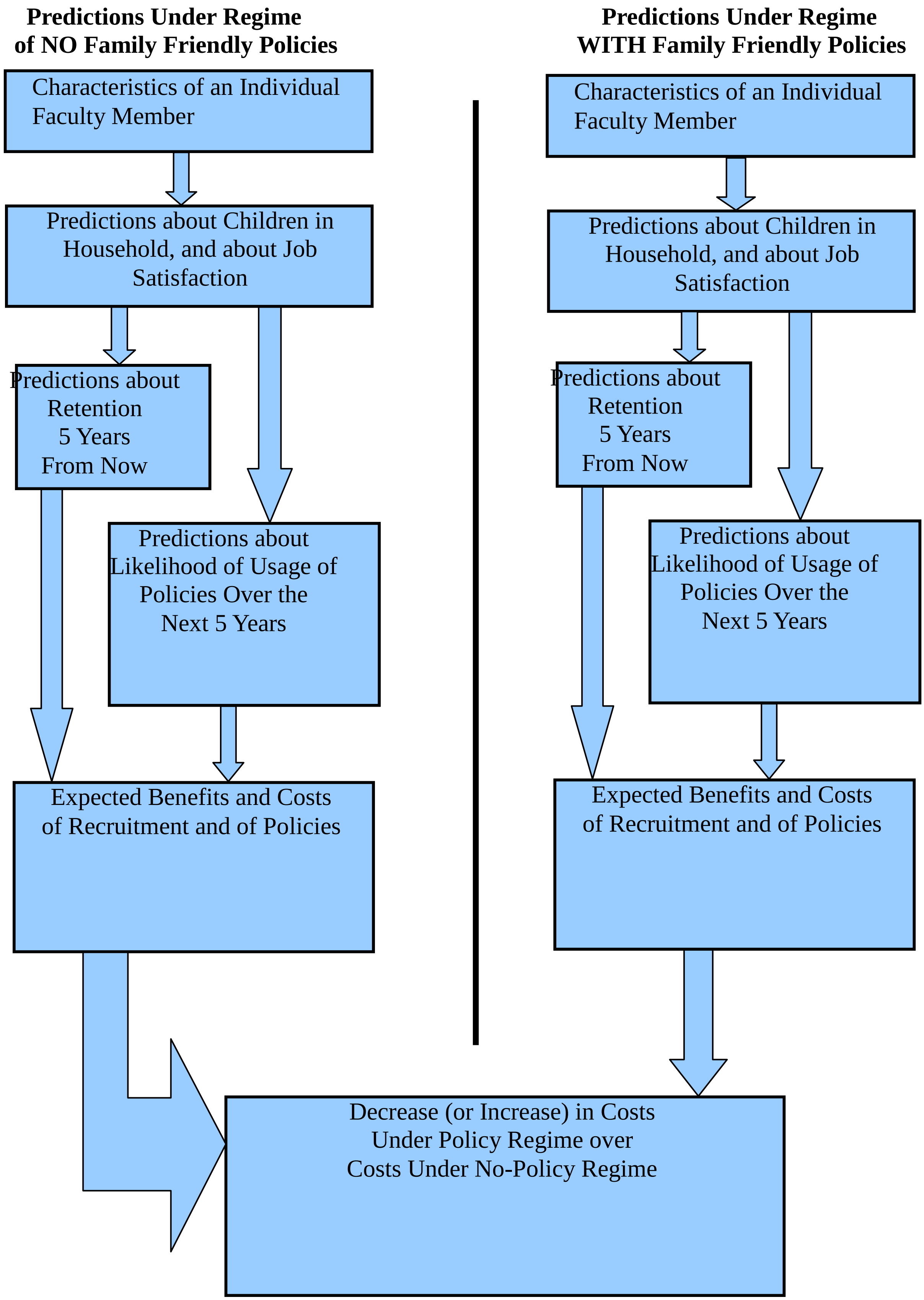
We hope that you find the simulator easy to use and of benefit to your organization. Please let us know!
Clair Brown cbrown@econ.berkeley.edu
Eric Freeman
UC Berkeley

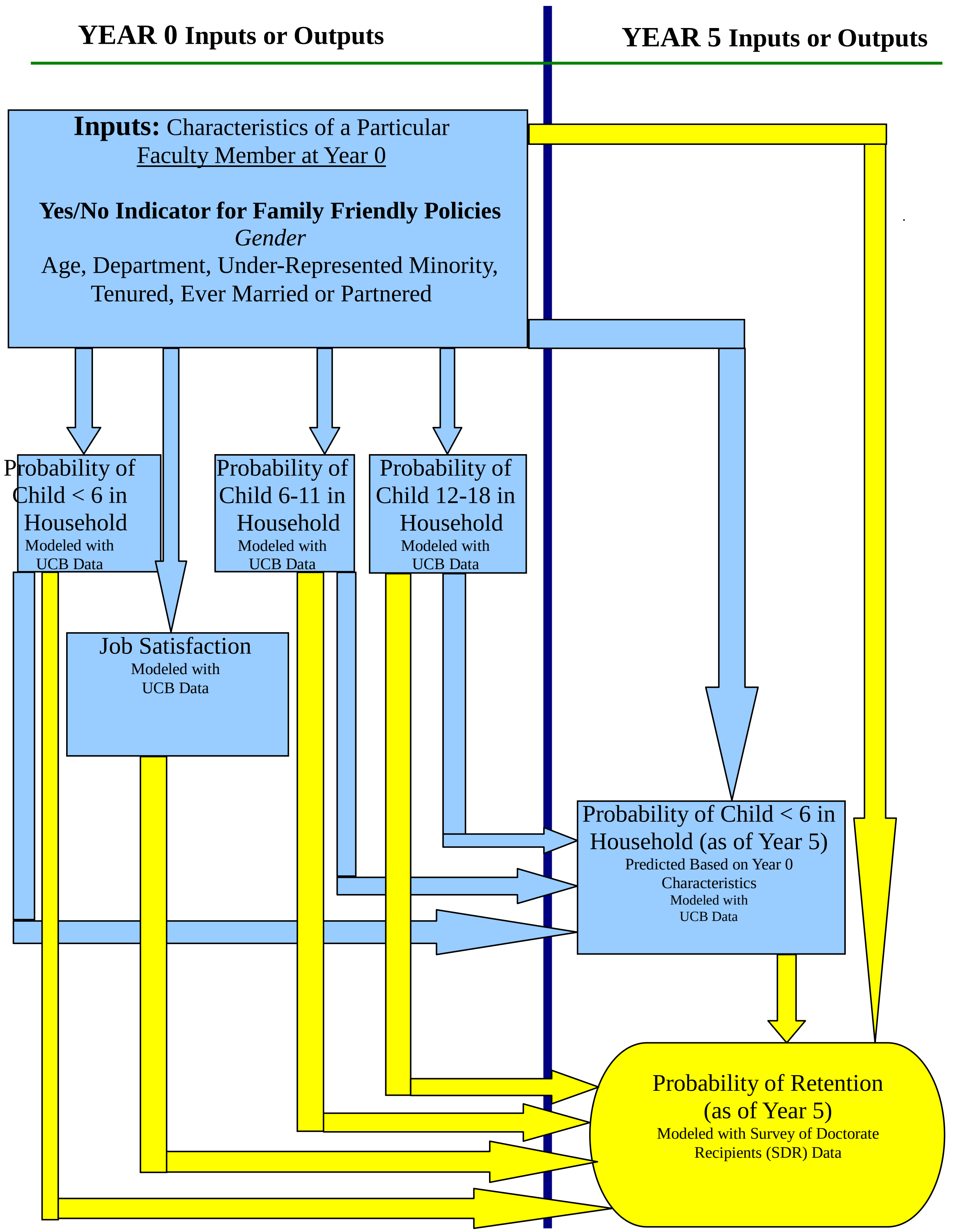
[19] Note that although this number might at first appear small, most of the faculty already have tenure and/or are older than the most common childbearing or adoption age, and thus the policies are likely to be used by only a fraction of the whole group. And of course this highlights that one of the key reasons that one might expect that these policies to have a potential financial benefit is that hiring costs can be very large relative to the costs of the policies.

[20] The estimates provided by the simulator assume that the observed faculty behavior is caused by the policies. As noted, we cannot establish causality but only associations. Also, we note that we report many significant digits here so that the result of the subtraction is easier to understand and not confused by rounding issues.

Overview of Simulation

(for an Individual Faculty Member)

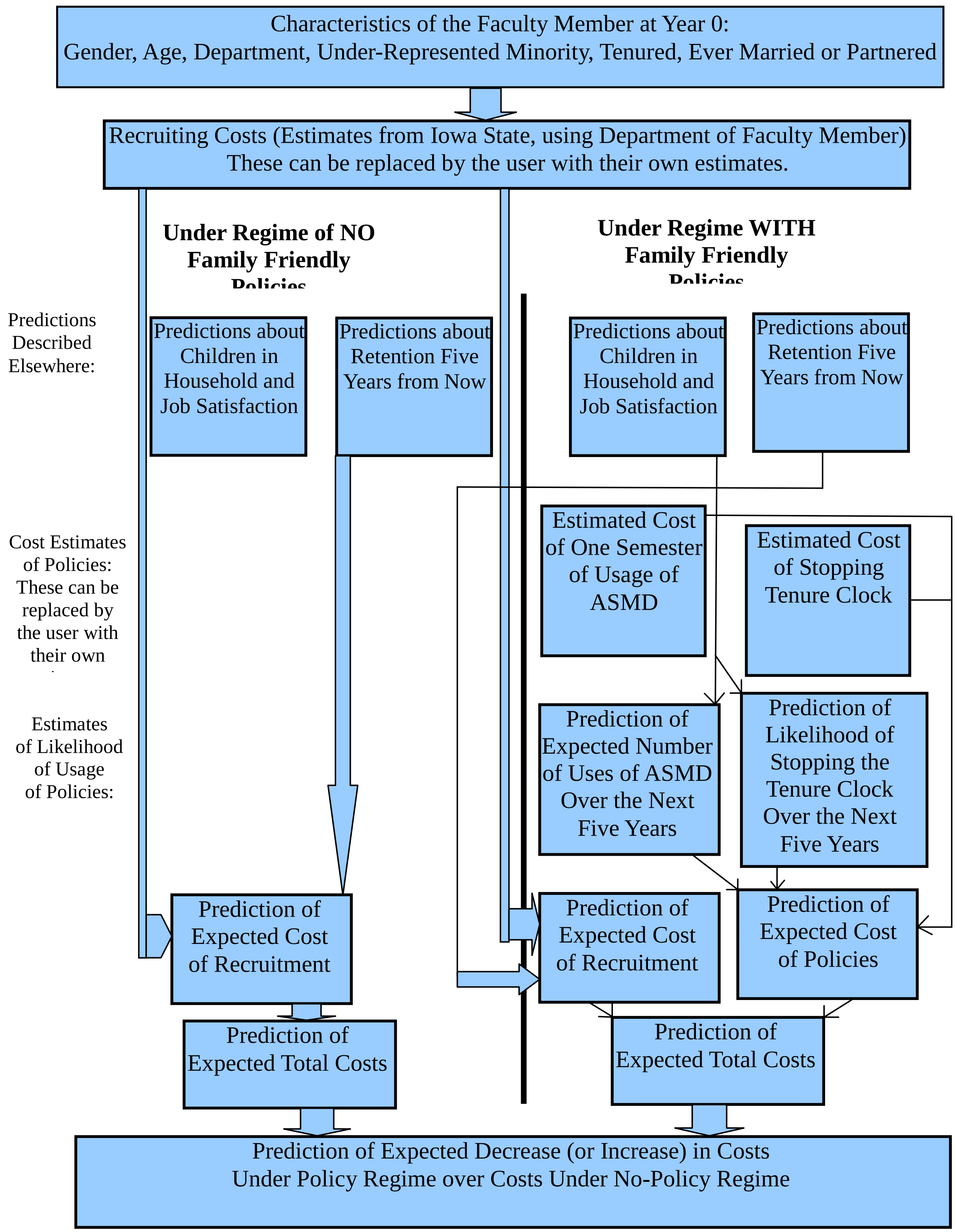




This entire process is done in parallel for two cases: (i) under an established regime of family-friendly policies, and (ii) under a regime without such policies. The difference in the final probability of retention can then be computed (as can differences in the intermediate outputs).

Flowchart for Cost Predictions in Simulation

(for an Individual Faculty Member)



Data Appendix

Variables used from the UCB survey.

Child Indicators (in UCB data): Child < 6 in household (hh), Child 6-11 (inclusive) in hh, Child 12-18 (inclusive) in hh (defined for 2002 and 2009 only): Indicator for children is imputed based on questions about youngest through fourth youngest child. Age is determined only by year of birth: 1997-2002 for 2002 sample; 2004-2009 for 2009 sample. This ignores the month of birth. Note that the 2002 survey was given in Fall 2002 while the 2009 survey was given in Spring 2009. 2002 Questions first ask 'Since you joined the faculty of University of California have you ever had substantial responsibility for raising one or more children under eighteen?' and if the answer is yes then asks 'For each child that you have raised or parented please enter their birth month and year (or month and year they entered your household if the child is adopted or is a stepchild)'. 2009 Questions first ask if respondent has any children and if so questions open that ask the following. 'For each child for whom you have had substantial responsibility for parenting please specify the month and year of the child's birth or adoption/foster placement with you. For stepchildren please specify the month and year when your parenting responsibilities began. (If you have parented more than four children please enter information for the four youngest children only.)'

Job Satisfaction (defined for 2003 and 2009 only): 2003 Question: 'All things considered, how satisfied are you with your current position at UC Berkeley?'; 2009 Question: 'All in all, how satisfied would you say you are with your job?'; 2003 coding: 0 Very dissatisfied; 1 Somewhat dissatisfied; 2 Somewhat Satisfied; 3 Very satisfied; 2009 coding: 0 Not at all satisfied; 1 Not too satisfied; 2 Somewhat satisfied; 3 Very satisfied

Indicator Variable for Math or School of Information or Physical Sciences. The category of Physical Sciences includes Chemistry, Physics, Astronomy, Geology and similar. In 2009 it also includes Statistics; in 2002 Statistics is categorized in the Social Sciences (in categories that cannot be disentangled). The reference group is all fields not covered by the indicator variables; these fields largely consist of the Humanities as well as some other professional schools. Excludes observations with field not known; there are fewer than 15 of these.

Indicator Variable for Life Sciences. Includes: Biological Sciences; the College of Natural Resources; and the Schools of Optometry and Public Health. The reference group is all fields not covered by the indicator variables; these fields largely consist of the Humanities as well as some other professional schools. Excludes observations with field not known; there are fewer than 15 of these.

Indicator Variable for Engineering. The reference group is all fields not covered by the indicator variables; these fields largely consist of the Humanities as well as some other professional schools. Excludes observations with field not known; there are fewer than 15 of these.

Indicator Variable for Social Sciences. Includes among other things: History; School of Public Policy; the Business School. In 2002 it includes Statistics- this is in Physical Sciences in 2009. The reference group is all fields not covered by the indicator variables; these fields largely consist of the Humanities as well as some other professional schools. Excludes observations with field not known; there are fewer than 15 of these.

Under-Represented Minority is defined here for 2002 and 2009 to be essentially all but White non-Hispanic and Asian non-Hispanic, i.e. everyone not in one of those two categories. In 2002 the race is imputed from the question 'How would you describe your race or ethnicity? Check all that apply.' In 2002 anyone giving at least one of the responses 'Black or African American','Mexican American','Hispanic or Latino','American Indian or Alaskan Native','Other' was coded as 1. The other responses for 2002 are 'White' and 'Asian American or Pacific Islander'. In 2009 the race is imputed from the question 'What is your ethnic category? Check all that apply.' and the question 'Of the ethnic categories below, which do you self-identify with most? Pick one.' The coding proceeded as follows: anyone responding to the first question with at least one of the responses 'White (not of Hispanic origin)','Other Asian','Chinese/Chinese American','Japanese/Japanese-American','Filipino/Pilipino', 'Pakistani/East Indian' was coded as 0. Then anyone responding to the first question with at least one of the responses 'Black/African American (not of Hispanic origin)','American Indian or Alaska Native','Mexican/Mexican American/Chicano','Latin American/Latino','Other Spanish/Spanish American' was coded as 1. The same process was then used for the second question, thus allowing the second question to trump the first question. We note that fewer than 5 respondents were coded to be 1 using the first question and then changed to 0 using the second question. We note also that each question also allowed the response 'Other, please specify:' with a spot for an open-ended response; but we did not have access to either whether this was chosen or the open-ended responses, so presumably these were coded as NA. For 2003, this variable is defined by a yes/no response to the question "Do you consider yourself a minority based on race or ethnicity?"

Tenured is 1 if Associate or Full; 0 if Assistant

Female is 1 if Female; 0 if Male.

Married or Partnered is 1 if the response is either Married or Partnered; it is 0 if Single, Divorced, Separated or Widowed or (a 2009 choice only) 'Never Married/Partnered'.

Indicator for Using ASMD for most recent child (defined for 2002 and 2009 only): imputed based on responses to multiple questions. In 2002, it is NA unless there is a response of yes to the question 'Since you joined the faculty of University of California, have you ever had substantial responsibility for raising one or more children under eighteen?' AND the date of birth for their youngest child is given as being later than 1988. In 2009, it is NA unless there was a birth/arrival of a child later than 1988 while the respondent was at UCB. The 2002 responses are for the question 'Received relief from teaching duties?' while the 2009 responses pertain to a question asking if they 'Received accommodation (if requested)' for 'Active service-modified duties (course relief)' and also to questions about the length. In 2002, questions were also asked about paid leave (other than sabbatical leave); a response of no to the ASMD question but yes to paid leave of '9-12 weeks' or '12+ weeks' (the two longest options) is also coded as a yes to ASMD. In 2009, questions were also asked about 'Paid pregnancy/birth Leave (disability leave)'; a response of no to the ASMD question but yes to paid leave of at '2 to 3 months', 'One semester', 'Two semesters', or 'More than a year' (the four longest options) is also coded as a yes to ASMD. In practice, these rules regarding coding for paid leave changed at most a handful of codings in both cases. Finally, for 2009, there are some who gave a response to the length question of >= 1 semester of asmd, but said NA to the simple yes/no question of getting the accommodation. This is interpreted as meaning that they are entitled to it, as in all these cases, they also said they 'Requested (or [were] entitled to) accommodation'. If someone said NO to getting the accommodation but still gave a length of >= 1 semester I interpret it to mean that they requested that length but did not get it. For both 2002 and 2009, respondents were asked about their four most recent children in order. If the years were given out of order, responses were corrected to reflect the most recent child, based on the birth years given; e.g., if a respondent reports their second youngest child to have a birth year after the youngest child, we use the birth years, rather than the reported order, to determine which is the youngest child.

Indicator for stopping tenure clock for most recent child (defined for 2002 and 2009 only): imputed based on responses to multiple questions. In 2002, it is NA unless there is a response of yes to the question 'Since you joined the faculty of University of California, have you ever had substantial responsibility for raising one or more children under eighteen?' AND the date of birth for their youngest child is given as being later than 1988. In 2009, it is NA unless there was a birth/arrival of a child later than 1988 while the respondent was at UCB. In 2002, there is simply a checkbox response to the question 'Stopped tenure clock?'. For 2009, there are some who gave a response to a related length question, stating that they used >= 1 semester of tenure clock stopping time, but said NA to the simple yes/no question of getting the accommodation. This is interpreted as meaning that they are entitled to it, as in all these cases, they also said they 'Requested (or [were] entitled to) accommodation'. So we code a 1 in these cases. If someone said NO to getting the accommodation but still gave a length of >= 1 semester we interpret it to mean that they requested that length but did not get it and thus we code a 0 in these cases. For both 2002 and 2009, respondents were asked about their four most recent children in order. If the years were given out of order, responses were corrected to reflect the most recent child, based on the birth years given; e.g., if a respondent reports their second youngest child to have a birth year after the youngest child, we use the birth years, rather than the reported order, to determine which is the youngest child.

Tenured at time of most recent birth (defined for 2002 and 2009 only): Imputed from questions about start years as assistant, associate and full professor. Tenured is 1 if Associate or Full; 0 if Assistant. If not at UCB at time of most recent birth (or if unknown whether or not at UCB at that time), this is coded as NA. If start year as assistant professor is given, but neither start year for associate nor full is given, then if the assistant professor start year precedes the most recent birth, this variable is coded as 0 if the start year is less than or equal to 7 years before the birth year, and 1 if at least 8 years before the birth year. (It is coded as NA if the start year is after the birth year.)

Variables used from the SDR

Retention in 2008, at the same university as in 2003: 1 if working for same employer during the week of October 1 2008 as during BOTH the week of April 1 2006 and the week of October 1 2003. 0 if working in first year and either (a) not working in second year or not working in third year; or (b) working at a different employer in second year or third year. NA if not working in first year. Note that the value is coded as 1 if the respondent says they are at the 'Same employer but in different type of job'. Note that this variable will be coded with the value 0 if they work at the same employer in 2003 and 2008 but somewhere else (or do not work) in 2006.

Age: Age recoded for public use. Topcoded at 75 or older.

Age squared: Age (recoded and topcoded) squared.

Child under 6 in household in 2003: Children living in household indicator: under age 6.

Child 6 to 11 in household in 2003: Children living in household indicator: age 6 to 11 inclusive.

Child 12 to 18 in household in 2003: Children living in household indicator: age 12 to 18 inclusive.

Child under 6 in household in 2008: NOTE THAT this is a variable from the 2008 data set. Children living in household indicator: under age 6.

Job Satisfaction: 0 Very dissatisfied; 1 Somewhat dissatisfied; 2 Somewhat satisfied; 3 Very satisfied. Only given for those working during the reference week. (October 1 2003 for 2003; April 1 2006 for 2006; October 1 2008 for 2008.) NA for those not working.

Under-represented Minority: Under-represented minority. 0 for 'Asian non-Hispanic ONLY' or 'White non-Hispanic ONLY'; 1 otherwise.

Departmental Variables: SDR observations that we use are restricted to those with one of five values for the variable "Job code for principal job [recoded for public use]" (coded NOCPRPB). The question asked is "Using the JOB CODES, choose the code that BEST describes the work you were doing on your principal job during the week of [survey reference week]". This question is asked of those working during the survey reference week, which depends on the wave of the survey, and is one of the week of October 1, 2003, the week of April 1, 2006, or the week of October 1, 2008. The five responses are:
Postsecondary Teachers-Computer and Math Sciences
Postsecondary Teachers-Physical and Related Sciences . . .
Postsecondary Teachers-Life Related Sciences
Postsecondary Teachers-Social and Related Sciences
Postsecondary Teachers - Engineering
The first two categories are combined in our classification (in order to match this classification to classifications used in the three UCB surveys as well as possible).

Summary Information for Variables Used in UCB Surveys

Number of observations for each survey:
2002: 743
2003: 861
2009: 633

Variable	Potential Values	Mean (by Survey Year)		
		2002	2003	2009
Female	0,1	.260	.277	.356
Age	Roughly 25 to 80*	50.9	51.0	50.7
Under-Represented Minority	0,1	.087	.164	.083
Tenured	0,1	.870	.854	.774
Married or Partnered	0,1	.8547	.860	.865
Job Satisfaction	0,1,2,3	NA	2.23	2.32
Use of ASMD for Most Recent Child	0,1	.311	NA	.535
Stoppage of Tenure Clock for Most Recent Child	0,1	.056	NA	.212
Child < 6 in Household	0,1	.161	NA	.203
Child 6-11 in Household	0,1	.144	NA	.190
Child 12-18 in Household	0,1	.209	NA	.158
Indicator for Math/ Information/Physics	0,1	.143	NA	.145
Indicator for Life Sciences	0,1	.209	NA	.202
Indicator for Social Sciences	0,1	.206	NA	.234
Indicator for Engineering	0,1	.174	NA	.127

* Exact ages not given as a precaution to protect privacy.

Summary Information for Variables Used in SDR Surveys

Number of Observations in Subsample Used: 4733

(This subsample restricts to those in all three of the 2003, 2006 and 2008 waves of the survey, and to those in the SDR survey who are postsecondary teachers in one of the four departmental groups. Here each observation represents one person linked across the waves; i.e., one person is counted only once, not two or three times.)

Variable	Potential Values	Mean
Female	0,1	0.256
Age	26 to 75	49.2
Under-Represented Minority	0,1	.0788
Job Satisfaction	0,1,2,3	2.45
Child under 6 in household in 2003	0,1	.166
Child 6 to 11 in household in 2003	0,1	.203
Child 12 to 18 in household in 2003	0,1	.204
Child under 6 in household in 2008	0,1	.116
Retention in 2008, at the same university as in 2003	0,1	.784
Indicator for Math/CS/Physics	0,1	.326
Indicator for Life Related Sciences	0,1	.200
Indicator for Social and Related Sciences	0,1	.354
Indicator for Engineering	0,1	.120