Chapter 0 – Simulation Model Description

January 14, 2018

Table of Contents

1 Notation and Definition of Variables

2 Data, Rates of Change and Model Assumptions

3 Gender Diversity Change

3.1 Attrition process

3.2 Promotion process

3.3 Hiring process

4 Three Models Using Different Replacement Strategies

4.1 Hire-Promote

4.2 Promote-Hire

4.3 Multinomial hiring (including changes in department size)

4.4 Example of mobility process

5 Department Size

5.1 Department size variation, year by gender by level

5.2 Department size variation with upper and lower boundaries

5.3 Growth models

5.3.1 No change

5.3.2 Linear change

5.3.3 Chunks of N-year change

6 Simulation Output

6.1 Generate plots

6.2 Create .csv files

6.3 Print output files

6.4 Using the Hoffman2 cluster

**1 Notation and Definition of Variables**

This following is the mathematical notation used in the discussion below.

Variable Description Range

|  |  |
| --- | --- |
| *R* the set of professorlevels:Assistant, Associate, Full denoted numerically  *r* level of a particular group of professors  *N* total number of professors inthe department  *nr* number of professors at a given level | ${1,2,3} $  *rER  NEZ+  nr E Z+* |
|  |  |

*G* the set of professor genders *{m*, *f}*

*g* the gender of a particular *gEG*

professor

*qr* Target percentage of *qr E [*0,1*]*professors at level r

Variable Description Range

*TN* Target number for the size of *TNEZ+*

the department

*vr* number of vacancies at level r *vr E Z+*

*vN* total number of vacancies in *vN E Z+*the department

*hrg* number of professors hired at *hrg E Z+*level *r* and gender

**2 Data, Rates of Change and Model Assumptions**

This simulation uses a Markov-based model to mimic the internal labor market of university professors, including assistant, associate and full professors. The results project how various personnel decisions may affect the future gender diversity of faculty. The data around which this simulation were built come from the faculty personnel office of a university in the Western United States. Twelve years of faculty data were obtained, 2004-05 through 2015-16. The data follow the attrition, hiring and promotion of each full-time ladder faculty member. Attrition includes all faculty who leave for any reason. Hiring includes faculty hired externally for assistant, associate and full professor positions. Promotions include faculty who move from assistant to associate professor or from associate to full professor.

Rates of Change

Attrition, hiring and promotion rates are averages of yearly changes in the proportion of women and men by level. For each year, this represents one transition, the time between the end of one year and the beginning of the next. If there are four years of data, for instance, there are three transitions: one between years 1 and 2, one between years 2 and 3 and one between years 3 and 4. The average of these three transitions is the rate. Thus, if 10 of 50 male Full Professors leave after year one (.20), 8 of 40 male Full Professors leave after year two (.16) and 0 out of 45 male Full Professors leave at the end of year three (.00), the average attrition rate for male Full Professors over the three years is: the total number who left (10+8+0=18) divided by the total number of male full professors (50+40+45=135), which produces an attrition rate of 18/135=0.133. Table 1 shows the department’s historical attrition, hiring and promotion rates: the average of eleven transition rates by gender by level calculated at the end of each year.

Table 1

Historical Average Transition Rates: Professorial Level By Gender

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Professorial Level | | | | | | | | | | |
| Baseline Transition Rates | Assistant | |  | | Associate | | |  | Full | | |
| Women | Men | |  | | Women | Men |  | | Women | Men |
|  |  |  | |  | |  |  |  | |  |  |
| Attrition | 0.0556 | 0.0687 | |  | | 0.0000 | 0.0574 |  | | 0.0741 | 0.0399 |
|  |  |  | |  | |  |  |  | |  |  |
| Hiring | 0.1724 | 0.8276 | |  | | 0.4000 | 0.6000 |  | | 0.1667 | 0.8333 |
|  |  |  | |  | |  |  |  | |  |  |
| Promotion | 0.0555 | 0.0611 | |  | | 0.1905 | 0.1149 |  | | NA | NA |
|  |  |  | |  | |  |  |  | |  |  |

Example of rate calculation using women assistant professors:

1. Attrition - Of all women assistant professors during the years data were collected, the percent who left: 2/36=5.56%.

2. Hiring - Of all assistant professors who were hired during the years data were collected, the percent who were women: 5/29=17.24%.

3. Promotions - Of all women assistant professors during the years data were collected, the percent who were promoted to associate professor: 2/36=5.56%.

Model Assumptions

The simulation provides several stochastic models. The main difference between them is the order in which hiring decisions are made and promotions awarded. In Model 1 hiring occurs before promotions, in Model 2 promotions occur before hiring and in Model 3 hiring and promotions are independent. Otherwise, the models are the same. Models 1 and 2 were designed to test how the simulation behaved under different cascading rules. However, unless otherwise noted, all further documentation refers to Model 3 because it best represents how this department makes faculty personnel decisions. The models operate using several assumptions:

1. The number of faculty is always an integer rather than a fraction.

2. Department size, defined as the total number of faculty, changes based on the process users think best emulates the future of their department.

3. All promotions, hires and attrition occur at the end of each time step. The last thing that occurs before each transition is the adjustment for department size. Assistant professor adjustments are always hires. Associate and full professor adjustments involve hires and promotions. The lower boundary for the number of faculty in any career level is constrained >= 0.

4. After all changes and adjustments are applied to the number of faculty during the year, the result becomes the distribution of faculty at the beginning of the next year.

5. Mobility in all models uses a cascade beginning with full professor and ending with assistant professor.

6. Attrition occurs before hiring and promotions.

7. Change for women occurs before change for men.

8. Parameters can be modified (see section 7).

**3 Gender Diversity Change**

All faculty departments undergo a yearly process of expected and unexpected changes in FTE (full time equivalents). In this simulation, expected changes occur through a mobility cascade in which actual average attrition, hiring and promotion rates are applied first to full professors, then to associate professors and finally to assistant professors. Rates are applied to women first and then to men. The options for filling vacancies are selected by the user and described in Section 5.0.

The following extended example is for Model 3 in which faculty hiring and promotion decisions are independent of one another.

**3.1 Attrition process**

Each year a number of faculty leave, perhaps due to retirement, alternative job opportunities or tenure rejection. Attrition rates are fixed for each level and vary by gender and level. Obviously, the attrition rate is greatest for full professors, with lower attrition rates for associate and assistant professors

The attrition rate by gender and level is denoted *ar g E [*0, 1*]*. Note that the attrition rate is a fixed constant and does not vary with time, meaning *arg(t) = arg*.

Within a given year, the number of attritions of a particular level and gender group at each level is:

Number of professor attritions at gender g, level r, in time t *= arg nrg(t)* (7)

for example, the number of attritions of female assistant professors in a given year would be:

Number of female assistant professor attritions at time t *= a*1*f n*1*f (t)* (8)

Thus, the total number who leave a given level is equal to the sum of female and male attrition at that level.

Total number of who leave at level r in year t *= arf nrf (t) + arm nrm(t)* (9)

**3.2 Promotion process**

Each year a portion of the faculty at the assistant and associate levels are promoted. The promotion rate is fixed by the department’s average rate and invariant over time. The promotion rate is denoted by *prg E [*0, 1*]*, where *r* represents the current level of the professor and *g* the gender. Thus *p*1*f* represents the promotion rate for a female Assistant professor to an Associate professor.

Number of promotions from level r to level r+1, for gender g, at time t *= prg nrg(t)* (10)

Promotion rates are independent of department churn, meaning that hiring and attrition do not influence the number of promotions. Promotion is based on department standards such as publication rates, teaching and participation in department committees.

The promotion rates for male and female professors are based on department level data. However, setting a common promotion rate does not seem to change the underlying dynamics.[???]

**3.3 Hiring process**

Each year the department makes decisions to hire additional faculty. Some of these hires fill vacancies created by attrition while others may arise from unique chances to add highly qualified individuals to the department.

The question addressed here involving hiring is how many faculty are hired by gender by level. The number of FTE vacancies in a department by gender by level [DO ANY OF THE MODELS CONSTRAIN THE NUMBER OF FACULTY IN A CAREER LEVEL? ISN’T THE DISTRIBUTION OF HIRES BY VACANCY A PARAMETER SOMEWHERE?] in a given year is based on the user’s selected growth model (see Section 5). If the department is under its target size, then the department will add additional faculty until the current number equals the target number. There is no guarantee that all vacancies will be filled within a single year, as availability of suitable candidates and other factors may limit the total number of hires. Thus, department size changes. Formally, the number of vacancies in the department is:

*vN* = max(*TN − N*, 0) (11)

Including the *max*() function ensures that the number of vacancies always is non-negative. Even if the department size is above its target size, this would not generate negative vacancies-meaning layoffs. Instead, the department would simply stop hiring until the department size fell below the target again.

4.4? Multinomial hiring process? Or fold this into 4.3?

We model hiring as a multinomial probabilistic process where all of the vacancies are filled by a mix of assistant, associate, full professors, or are left unfilled based upon a probability for each outcome. The multinomial distribution generalizes the binomial distribution. Formally the definition of the multinomial distribution is that "it models the probability of counts for rolling a k-sided dice n times."[Wikipedia: [https://en.wikipedia.org/wiki/Multinomial\_distribution]](https://en.wikipedia.org/wiki/Multinomial_distribution%5D). For example, suppose we have a die with six sides and a 16 probability of getting any value between 1 and 6. If we rolled this die 1 time, we might get a vector such as [0,0,1,0,0,0] meaning that the die roll yielded a 3. Each roll will yield a different result, and the next roll might yield a value of 1 with corresponding vector [1,0,0,0,0,0]. When we sum the number of times we obtained each value after 10 rolls, we could obtain a vector like [1,1,3,2,2,1]. This final vector of counts after *k*-rolls is what the multinomial distribution yields, namely the probability distribution over this final set of category counts.

In our model there is a probability over the hiring of professors at each level at each time step. The probability of hiring new professor increases as the difference between the target and actual department size increases. Given that a professor hire is approved, the probability that this professor is of a particular level and gender is fixed. The probability of hiring a professor at level *r*, gender *g* at time *t* is denoted *prg*(*t*). Using these hiring probabilities, we can generate a random draw from the multinomial distribution. The rather ugly looking multinomial density function for professor hiring is:

|  |  |
| --- | --- |
| *vN*(*t*)!  *P*(*hm*, *f*1,2,3(*t*), *hnone*(*t*); *vN*, *pm*, *f*1,2,3(*t*), *pnone*(*t*)) =  *h*1,2,3  *m*,*f* (*t*)! *· hnone*( *t*)! | *× p*1,2,3  *m*,*f* (*t*) *× · pnone* (12) |

The *hnone*, *pnone* represent the probability of not hiring anyone for a given vacancy. While expression (??) might look ugly, it is simply the standard multinomial distribution with the relevant model parameters.

Note also, that in the case where the department is over the upper boundary, *vN*(*t*) is equal to zero and no hiring would likely occur that year (unless some opportunistic hiring opportunity presents itself).

**4 Three Models Using Different Replacement Strategies**

The simulation provides three models with different replacement strategies. After attrition, Model 1 (Hire-Promote) hires faculty first and fills remaining vacancies with promotions. After attrition,

Model 2 (Promote-Hire) promotes faculty first and fills remaining vacancies with hires. After attrition, in Model 3 (Independent) hiring and promotions are calculated independent of one another.

[STRUCTURE HERE TBD. I THINK WE SHOULD DESCRIBE MODEL 3 AND PUT THE DESCRIPTIONS FOR MODELS 1 AND 2 IN AN APPENDIX.]

A. Full Professors - Vacancies and Replacements (Level 3).

*Attrition*. At the end of each time step, full professors leave at the department's full professor attrition rate. Assume there are 20 full professors and the attrition rate for both women and men is 0.25. At the end of the year, five full professors leave creating five vacancies. If the attrition rate produces a fractional FTE, the number is rounded up to the next integer.

*Hiring*. The full professor hiring rate for women is used to decide how many of the vacancies will be filled by women. Assume the hiring rate for women full professors is 0.20. For each vacancy, draw randomly from a Bernoulli distribution with a 20% chance of getting a 1, in which case a woman is hired, and an 80% chance of getting a 0, in which case a woman is not hired. After drawing five times, once for each vacancy, the sum of the five 0/1 draws yields 0 to 5 women full professor hires, leaving N-(0 to 5) positions that must be considered for men.

Assume the full professor hiring rate for men is 0.50. For each vacancy where a woman was not hired, the hiring rate for men is applied to another random draw from a Bernoulli distribution with a 50% chance of getting a 1, in which case a man is hired, and a 50% chance of getting a 0, in which case a man is not hired. Thus, the probability of making a hire for each full professor vacancy, if the number of vacancies is N, is the joint probability of the hiring rate for women times the hiring rate for men, in this case N(0.20+0.50) so the total number of vacancies is N – [N(0.20+0.50)].

*Promotions*. In this example, there were five full professor vacancies after attrition. If 1 woman and 1 man were hired, there are three remaining full professor vacancies to be filled by promotions. Assume the promotion rate for women associate professors is 0.40. For each of the three remaining full professor vacancies, draw randomly from a Bernoulli distribution with a 40% chance of drawing a 1 in which case a woman associate is promoted to full professor and a 60% chance of drawing a 0 in which case no women associates are promoted. This leaves somewhere between 0 and 3 (0+0+0, 0+0+1, -1+1+1) vacancies to be filled by promotion from men associates.

After this attrition, hiring and promotion process, it is possible vacancies still exist. If there are five full professor vacancies and only four are hired, the fifth position remains unfilled and the vacancy disappears. This causes department size to shrink. In order to keep department size within boundaries set by the user, a separate set of decision rules is applied after the entire cascade. Afterwards the added or subtracted fte are distributed randomly across the levels.

B. Associate Professors - Vacancies and Replacements (Level 2). Associate professor mobility is calculated after that for full professors. The calculation proceeds in the same way with attrition rates applied to women and then men. Vacancies created by attrition are filled by hiring. Women are hired first by applying the associate professor hiring rates for women to the Bernoulli draw for each vacancy. If any vacancies remain, the same process is repeated for men using their associate professor hiring rates. After attrition and hiring, if there are remaining vacancies, the process is applied using promotion rates of assistant professors to fill these vacancies first with women who are currently assistant professors and then with men. Unfilled vacancies disappear at the end of the calculation.

C. Assistant Professors – Vacancies and Replacements (Level l).  
Assistant professor mobility is calculated after that for associate professors. The calculation is the same as for full professor and associate professor. The only difference is that the cascade finishes after attrition and hiring. There are no “promotions” because the only way to become an assistant professor is to be hired. Similar to before, unfilled vacancies disappear at the end of the calculation.

This cascade may result in a faculty of the same size for the following year or one that decreases in size because some vacancies remain unfilled. It never results in a faculty that increases in size beyond the department size boundaries. Vacancies are considered overall department vacancies, that is, they are not tied to the assistant, associate or full professor level from which they originated. Department size is maintained using a set of decision rules (see Appendix B, 9.4).

*[Note that this will change in the simultaneous choice model where we can rule out unfilled vacancies if we wish ]*

4.4 Example of mobility process for Model 3

The following extended example is for Model 3 in which faculty hires and promotions are independent and the cascade begins with full professors and ends with assistant professors.

*Initially professors leave according to the attrition rate at their level. These vacancies are filled either through promotion or hiring. In the original model there was a serial process where either hires were completed first and then remaining vacancies were filled by promotion, or vice versa. In the simultaneous choice model, both hiring, promotion, or no fill are completed at the same time. We achieve this by assigning a probability to all five potential choices:*

* *Hire a woman*
* *Hire a man*
* *Promote a woman*
* *Promote a man*
* *No fill*

*The sum of the probabilities for the choice must sum to 1. For example, if there are 5 total vacancies at level 3, then there may be a 20% chance for a male hire, a 20% chance for female hire, a 20% for male promotion, a 20% chance for female promotion, and a 20% chance for leaving the position unfilled. We take a random draw from a uniform distribution and if the value of that draw falls between 0-0.20, then a male is hired, if the draw falls between (0.20 to 0.40] then a woman is hired, and so forth. Once a choice is made for all vacancies, the model proceeds to the next level, just as in the original simulation. [IS THIS STILL ACCURATE?]*

**5 Department Size**

Department size is included as a parameter because changes is department size influence the speed at which gender diversity can change. It is generally easier to increase the proportion of women in a department that is increasing in size than it is if the department is decreasing. This results because increasing size generally means more vacancies to fill.

To formalize these basic definitions, we define *nr* as the number of professors at a given level. That number is equal to the sum of male *nrm* and female *nrf* professors at that level, in the department.

*nr = nrf + nrm* (1)

The total size of the department *N* is equal to the number of professors at each level.

N = n1+n2+n3 = ni

i=1

(I had to rewrite the equation above because I couldn’t get rid of the Section Break otherwise)

Note that because department size and number of professors by gender by level changes from one transition to the next, time is a parameter. Although used implicitly in every calculation, we omit time unless its inclusion facilitates clarity. For example, the department size equation above is more accurately described by:

3

*N(t) = n*1*(t) + n*2*(t) + n*3*(t) =*  *ni(t)* (3)

*i=*1

Department size is a parameter that may vary over time. The simulation can be run either with a fixed (See 4.2.1) or dynamic department size (See 4.2.2 through 4.2.4). Users can also specify upper and lower boundaries for change.

In general, department size *TN* differs from actual department size *N*. This difference varies stochastically from year to due to differences between actual and projected faculty hiring and attrition. In some years attrition exceeds the number of hires and department size falls. In other years, hiring is greater than attrition, increasing department size. This might happen, for instance, when the department makes an unexpectedly large number of opportunistic hires, which exceed the number who leave. When department size exceeds an upper boundary or falls below a lower boundary, countermeasures such as increased hiring or injunctions on new hires bring the department back within an acceptable range.

Because both department size and number of professors by gender by level may fluctuate from year to year, it is convenient to express the number of professors as proportions. These are represented by *qr E* [0, 1]. For example, a department may be composed of 20% assistant professors, 15% associate professors, and 65% full professors in a given year. These proportions are calculated at the end of the last transition when all attrition, hiring and promotion changes and the new department size have been accommodated. Due to professor churn, the proportions change from year to year. The professor proportions are constrained by the requirement:

3

 *qr*(*t*) = 1 (4)

*r*=1

To calculate the number of professors at a particular level at a given time, simply multiple the current professor share times the target department share: *qr \* TN*. To prevent this equation from providing fractional numbers of professors, we use a floor function (*[ ]*) to round the values down to the nearest integer.

*nr [qr ' TN]* (5)

As an example consider a department with a target department size *TN* of 100. A possible professor share for full professors *q*3 may be 70%. Thus the target number of full professors in the department would be 0.70 *\** 100 = 70.

We track the percentage deviation of the actual department size from target in a given year with the following formula.

*N*(*t*)

Deviation from Target department size at time t = 1 *−* (6)

*TN*

The deviation is positive if the department size is less than the target and will be negative if the deviation is in excess of the target. The sign of the deviation indicates whether the deviation inflates or deflates the subsequent hiring probability at the beginning of the transition for the following year.

5.1 Department size variation, year by gender by level

The final step in each transition is change in department size. Such changes may result from enrollment, financial resources or university policy decisions. We provide three models of department growth. The user can select the model that most closely represents change in their department.

1. No change

2. Linear change

3. Chunks of N-year change

4. Projected change

5.2 Department size variation with upper and lower boundaries

The model allows for some random department size fluctuation around the target size. Depart­ments rarely remain the same size from year to year. Occasionally opportunities arise to hire promising new faculty, while other times a number of senior faculty decide to retire in the same year. To improve the flexibility of the model, we incorporate the change for random variation into the model. The simulation allows the user to specify a maximum and minimum allowable excess yearly variation in faculty FTEs. For example, the user might decide that as many as two or three faculty can be added or subtracted at the end of each time step.

For each addition or subtraction, the model randomly selects between three possible outcomes: add one FTE, subtract one FTE or make no changes, coded as (+1, -1, or 0). For instance, if the user specifies the maximum allowable change as 3 faculty, the program randomly chooses from the set (+1, -1, or 0) three times, once for each FTE, producing a maximum variation of +3 (+1+1+1) or -3 (-1-1-1). If the maximum allowable change is 2, the program randomly chooses from the set (+1, -1, or 0) two times, again, once for each FTE for a maximum variation of +2 (+1+1) or -2 (-1-1).

Before these additional FTEs are incorporated into hiring activity, the model checks whether the new department size--including these additional FTEs remains within the department upper and lower size bounds. If the new department size is within bounds, then the model will add the additional FTEs to the number of vacancies for hire and execute the hiring process as specified above. If the new department size falls outside of the bounds of the department size, then the candidate number of additional FTEs is discarded and a new random draw of additional FTEs is made and rechecked--until a value is found that keeps the department size within bounds.

Note that if the number of excess FTEs is negative, this does not imply any sort of layoff or forced retirement. Instead these negative FTES are deducted from the number of potential hires for that same year. Thus if there were to be 4 new hires but there were -2 excess FTEs, then only 2 faculty would be hired instead of 4. Based on the definition of the hiring model above, the number of vacancies or hires can never go below 0.

5.3 Growth Models

The simulation provides several options for depicting department size over time. The user can select which option best matches the department’s past changes and future projections.

5.3.1 No Change

As a baseline we include a version of the model that explicitly ignores changes in department target size over time. This model still constrains department variation within an upper and lower boundary, however there is (a) no explicit mechanism for changes in department target size nor (b) an explicit mechanism for changes in the department upper and lower boundary.

The following is text from “notebook text 5.21.17.”

START HERE. Every year, departments either lose and gain faculty or stay the same size. In order to keep department size within a reasonable lower and upper boundary, adjustments are made after the mobility cascade and just before the end of the time step. This works as follows.

Step 1. The user sets an upper and lower boundary for department size.

Step 2. After the mobility cascade, the program assesses whether department size remains within the range defined by the user’s boundaries. If it does, the program stops and moves to the next time step. If it doesn’t, the program uses a stochastic process to make adjustments before moving to the next time step. If department size exceeds the upper boundary, all hiring ends until the numbers come down. [I’VE FORGOTTEN. IS IT POSSIBLE FOR DEPARTMENT SIZE TO EXCEED THE UPPER BOUNDARY?] If it decreases below the lower boundary, hiring and promotions are increased until the numbers return within range. Thus, the closer department size gets to its extreme values, the greater the pressure to revert towards the middle of the size boundaries.

Step 3. When department size requires adjustment, the user specifies a maximum and minimum allowable yearly variation in faculty FTE. For example, the user might decide that as many as two or three faculty can be added or subtracted at the end of each time step. When expected attrition, hires and promotions leave department size outside its boundaries, these faculty FTE are used to bring it back within.

For each addition or subtraction, the program randomly selects between three possible outcomes: add one FTE, subtract one FTE or make no changes, coded as (+1, -1, or 0). For instance, if the user specifies the maximum allowable change as 3 faculty, the program randomly chooses from the set (+1, -1, or 0) three times, once for each FTE, producing a maximum variation of +3 (+1+1+1) or -3 (-1-1-1). If the maximum allowable change is 2, the program randomly chooses from the set (+1, -1, or 0) two times, again, once for each FTE for a maximum variation of +2 (+1+1) or -2 (-1-1). Following these random choices, the sum of these values represents the total change in FTE that will be accommodated at the end of the time step.

Step 4. If department size plus this sum is within the department size range, the change is accepted. If it remains outside its range, the change is rejected. The random choice of (+1, -1, or 0) for each yearly allowable change is repeated until the sum of these changes keeps the department within range.

Note that if the department falls below its minimum size, the program automatically sets all allowable yearly changes to +1 to ensure that in the next time step, the department size will be above the lower boundary. [I THOUGHT THAT IF THE DEPARTMENT DECLINES BELOW THE LOWER BOUNDARY, THE CODED OPTIONS WERE JUST REPEATED UNTIL THEY RETURNED WITHIN RANGE? AT LEAST FOR THE NO CHANGE OPTION] Similarly, if the department exceeds the its maximum size at the end of a simulation time step, the program automatically sets the selected number of unexpected changes to 0. This ensures that the department will not continue to grow beyond its upper boundary in subsequent timesteps. [THIS DOESN’T OCCUR UNLESS THE TOTAL NUMBER OF VACANCIES EXCEEDS THE YEARLY ALLOWABLE CHANGE?]

[**KRISHNA DO THIS**]

Step 5. After the yearly allowable changes are computed, summed and accepted, the FTE are randomly distributed across the three faculty levels—assistant, associate and full professor—each with a probability .33. If the FTE is added to assistant professors, it is designated as a hire. If it’s added to associate or full professors, it may be a promotion or a hire. [IS THE PROBABILITY AN ASSOCIATE OR FULL PROFESSOR ADDITION IS A PROMOTION OR A HIRE DECIDED BY A COIN FLIP?]

5.3.2 Linear Change

The simplest model of growth is a fixed linear percentage change in the department size every year. Thus, the model allows department size to grow by a fixed rate--say 1%--each year of the simulation. The upper and lower department size boundary also increase by an equivalent amount each year.

The simulation provides two options for the fixed rate of linear growth, either: 1) the average change in department size across all years of data; or 2) the average change in department size using a moving average of two or three years across all years of data.

While linear growth is a useful model, the department can only ever increase or decrease in size over the course of the simulation. Periods of growth cannot be mixed in with periods of decline.

5.3.3 Chunks of N-year change

5.3.4 Change using projections

Finally, we include a model for projected change. The user enters a vector of department size target sizes. The number of years in the simulation duration is divided by the number of forecasts, which gives a sense of how many years each forecast spans. The model then annualizes a rate of change in the department size target to bring the actual department size in line with the forecast.

An example is useful in this instance. The user will specific three items: - The duration of the simulation: e.g. 20 years. - The initial department size: e.g. 100 professors   The vector of forecasts: [100, 105, 115, 120]

Since there are four forecasts, each forecast covers a span of 5 years or 20/4.

There is an open question about how quickly a new forecast is implemented by a department. Does hiring activity start immediately to fill a large target or is there a more gradual implementation of the new targets. It is difficult to *ex ante* assign a rule for how quickly forecasts are implemented as different departments may face different constraints. Hence models for both processes are included in the package.

In the model version with immediate implementation of the forecast, the department target size will change immediately at the end of the current forecast period and at the start of the next one. Hence if the department size changes abruptly from one period to the next, there will be an immediate jump in the hiring at the start of the subsequent period.

In the gradual version, the forecast adjusts hiring to reach the new department size by the middle of the forecast period.

Note that the two versions of the forecast model only differ in the case of department growth between forecast periods. The model does not change attrition rates, hence a decline in department size is controlled by the fixed attrition rate.

**6 Simulation Output**

The simulation first executes attritions and promotions overall or by level, and then generate a set of hires at each level. These hires will be added to the department size numbers in the subsequent transition.

This number of years the simulation runs is a user-defined parameter. The default is twenty years.

6.1 Instructions for plotting functions

6.2 Creating .csv files of the results

6.3 Producing a printed version

6.4 Returning to default parameters

6.1 Instructions for plotting functions

This section lays out instructions for customizing the plots displayed throughout the notebook. The different types of charts are denoted by special tags at the top of the plot cell. The four plot types are denoted as:

1. `#<overall>`: An overall chart displays the result of a department-wide summary statistic for a single model.

2. `#<bylevel>`: A bylevel chart displays the results of a single summary statistic for a each professor-level of a single model.

3. `#<comparison-overall>`: A comparison-overall chart compares two or more models on one department-wide statistic.

4, `#<comparison-bylevel>`: A comparison-bylevel compares two or more models one a single professor-level statistic.

The chart tag is also augmented with an indication of which statistic is displayed in the chart. The common summary statistics are:

5. gender\_proportion: the proportion of women in the department or in the department by level.

6. probability\_proportion: the probability that the department may achieve a target gender proportion in a given year

7. unfilled\_vacancies: The number of unfilled vacancies remaining at the end of year, after all hiring and promotion activity is complete.

9. male\_female\_numbers: The number of males and females in a department for a given year.

plot\_settings = {'plottype': 'gender proportion',

'number\_of\_runs': 100, # number simulations to average over

'target': 0.25, # target percentage of women in the department

'caption': 'The red line shows a goal of .25 women',

# Main plot settings

'xlabel':'Years',

'ylabel': 'Proportion of Women Overall',

'title': 'Figure 4.1.1a: Model 1 - Change in \*Proportion\* Women Overall',

'line\_width': 2,

'xmin': 0,

'ymin': 0,

#'xmax': 40,

#'ymax': 1.0,

'transparency': 0.25,

'marker\_shape': None,

'linecolor': 'green',

'model\_legend\_label': 'Simulated Probability',

'legend\_location': 'upper right',

#Optional Parameters

#Target value plot settings

'target\_plot': True,

'color\_target': 'red',

'color\_percent\_line': 'red',

'target\_plot\_line\_style': '--',

'target\_plot\_linewidth': 2,

'target\_plot\_legend\_label': 'Target',

#Percent plot settings

'percent\_line\_plot': False,

'percent\_line\_value': 0.5,

'percent\_line\_style': '-.',

'percent\_linewidth': 2,

'percent\_legend\_label': 'Reference Line',

#Male Female numbers plot settings

'male\_female\_numbers\_plot': False,

'mf\_male\_color' : 'black',

'mf\_target\_color' : 'red',

'mf\_male\_label' : 'Male',

'mf\_target\_label': 'Target',

'mf\_male\_linestyle' : '-',

'mf\_target\_linestyle': '-',

'mf\_male\_linewidth' :2,

'mf\_target\_linewidth' : 2

}

t.plot\_overall\_chart(\*\*plot\_settings)

6.2 Creating .csv files from the results

The cell below generates a `csv` file with a requested number of model runs. To generate the csv, the user must specify three pieces of information, and then run the cell. The first piece of information is the `model choice` variable, which identifies the model and scenario requested. The set of possible choices is listed in the cell below.

The second piece of information is the `model label` which is the first part of the title for the exported file. This allows the user to indicate the particular model selected as well as any additional information about the run.

The final piece of information is the `number\_of\_runs` variable which specifies the number of model iterations included in the exported file.

To generate the file, please execute both of the code cells in this section in order. If the cell produces an error, the most likely reason is a misprint in the exact spelling of the `model\_choice` model selection.

6.3 Print output files

As of January 14, 2018, we have not found software that will take the plotted output and format it for .pdf. This means plots often fall across two pages, which makes them difficult to read. There is no work around. However, users can download and print the .html file.

6.4 Using the Hoffman2 cluster to run simulation

At present, the simulation lives on UCLA’s Hoffman2 cluster computer. This means users need a Hoffman2 account which can be obtained by ….

The procedure for running the simulation will no doubt change. However, as of January 14, 2018, it works as follows:

1. Open terminal window on computer

2. Login to Hoffman2

3. Check and make sure all relevant software has been uploaded—Ask Krishna

4. Call Python environment: python2.7 h2jupynb -u blawrenc -v 3.4

[Note: substitute your account id for blawrenc]

5. This opens a tab in your browser titled: notebooks/

6. Select: [mgmt\_notebook05092016.ipynb](http://localhost:9898/notebooks/notebooks/mgmt_notebook05092016.ipynb)

7. This opens a second tab in your browser: [mgmt\_notebook05092016.ipynb](http://localhost:9898/notebooks/notebooks/mgmt_notebook05092016.ipynb)

8. At the beginning, there is a Table of Contents. There should also be a “Contents” sidebar that you can move and resize in your browser.

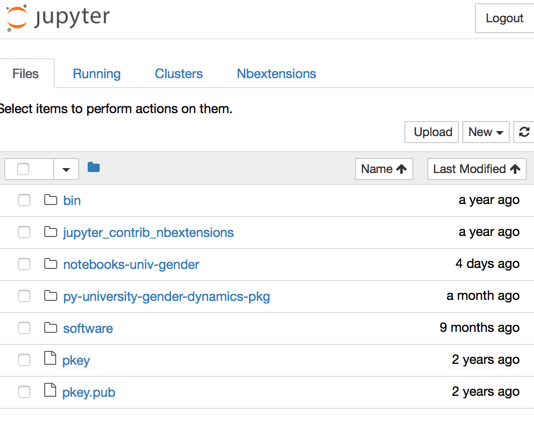
9. Select: 4.1.1 Model 1: Hire-Promote. Run the first cell. This loads the plotting software and initialization cells.

10. At this point you should be able to select and run any model in the notebook. More detailed instructions follow below.

Step 1

1. After logging in to the Hoffman2 cluster and calling the python environment, you will be taken to a browser tab that looks something like this.

2. Select “notebooks-univ-gender. This will take you to the next tab.



Step 2

The simulation is divided into five chapters.

1. Chapter\_0\_Gender\_Diversity\_model\_description

Describes how the simulation works (You are in Chapter 0 now)

2. Chapter\_1\_Gender\_Diversity\_overall\_simlation

Provides options for exploring the results at the department level.

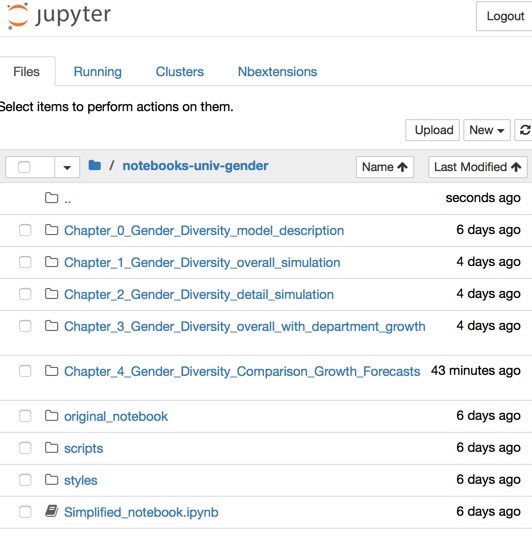
3. Chapter\_2\_Gender\_Diversity\_detail\_simulation

Provides options for exploring the results by career level: assistant, associate and full professor

4. Chapter\_3\_Gender\_Diversity\_overall\_with\_department\_growth

5. Chapter\_4\_Gender\_Diversity\_Comparison\_Growth\_Forecasts

Provides results on how the department’s growth or decline in size influences the overall proportion of women faculty.



Step 3

If you select a chapter, you will be taken to a browser tab with options. This example is for Chapter 4. The others are similarly constructed.

1. Chapter\_4\_Gender\_Diversity\_Comparison\_of\_Forecast\_Models.ipynb.

Select this option if you want to run the growth simulations.

2. Template\_ Chapter\_4\_Gender\_Diversity\_Comparison\_of\_Forecast\_Models.ipynb.

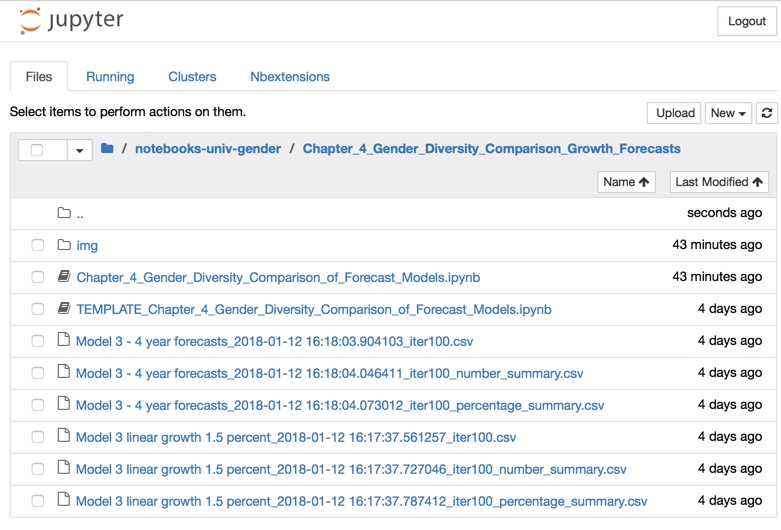
This is a fixed copy of the model with all the original rates and numbers used to run the simulation. It is provided so that no matter what changes you explore, you can always return to the original numbers.

3. The remaining six options are output files (.csv) from simulations that have been previously run. Each simulation run produces three files distinguished by their suffixes, in this case:

a. iter100.csv

b. iter100\_number\_summary.csv and

c. iter100\_percentage\_summary.csv.



Step 4

In order to download a .csv file, select the file in the left hand column, then select “Download” in the menu at the top. You can only download one file at a time.

