A MATHEMATICAL MODEL OF PAY-FOR-PERFORMANCE FOR A HIGHER EDUCATION INSTITUTION

Matthew Hathorn, Weill Cornell Medical College
John Hathorn, Metropolitan State College of Denver
Lesley Hathorn, Metropolitan State College of Denver

ABSTRACT

This paper develops a mathematical model of the proposed pay-for-performance award system for an institution of higher education. Two constraints are imposed to ensure the fairness of the system. The model is general enough so that the payouts for the three performance levels (excellent, exceptional, and extraordinary) are clearly distinguished. Thus, the greater effort and performance that is required to achieve the highest level is rewarded with significantly higher monetary benefits. This outcome reinforces outstanding performance and should motivate faculty to perform at high levels in the future.

INTRODUCTION

The payoff for exceptional productivity must be substantial to make the increased effort of the performance, as well as the evaluation of this productivity, worthwhile (Baker, Jensen, & Murphy, 1988). Pay-for-performance (PFP) is a program that offers such an incentive in that it has been designed to improve the productivity of individuals by offering financial incentives for exemplary outcomes. That is, it is a one-off bonus associated with exceptional work.

The Board of Trustees of a college of two of the authors set aside an annual PFP budget line item equal to eight percent of the total faculty compensation, both salary and benefits, to reward those who perform at an exemplary level.

This paper develops a mathematical model of the PFP awards system and introduces two constraints to ensure the fairness of the proposed system. After the literature review the paper discusses a numerical example, develops the mathematical model and then discusses it before offering a conclusion in the final section.

BRIEF LITERATURE REVIEW

Despite the findings of a meta-analysis of 39 studies over 30 years that showed that there is a positive correlation between performance and financial incentives (Jenkins, Mitra, Gupta, & Shaw, 1998), not all studies have supported PFP. Some researchers have found little evidence of the effectiveness of PFP in, for example, health care settings (Rosenthal & Frank, 2006). The lack of evidence in this sector, however, has done nothing to stem the enthusiasm for the program as more than half of a sample of health management organizations (HMOs) use PFP (Rosenthal, Landon, Normand, Frank, & Epstein, 2006). It is important to note that one of the reasons given for the lack of effectiveness in health care PFP systems was attributed to a low bonus size (Rosenthal & Frank, 2006). In business, the demand for PFP continues to be very strong, despite a weak economy, according to Mercer's 2010 U.S. Executive Compensation and Performance Survey (2010).

For the purpose of this paper, PFP is distinguished from other types of incentives such as reinforcement on a ratio scale (e.g., piece work) and merit pay increase. It has been well-established that piece work increases productivity over reinforcement on an interval scale (e.g., fixed salary) in a number of domains (Skinner, 1974). For example, a ratio reinforcement strategy has been found to increase productivity in tree planters (Shearer, 2004) and logging (Haley, 2003). Ratio scale reinforcement is obviously out of place in all areas of higher education beyond the experimental laboratories. Merit pay increase is a system that is used by some colleges in which faculty receive a percentage increase of their current salaries when they meet or exceed minimum outcomes. This percentage is then added to their base salaries.

PFP is an alternative mechanism that has been proposed as a method of providing a one-off reward for exceptional work by faculty. However, evaluation and implementation can lead to disastrous outcomes (Terpstra & Honoree, 2008). Hence, it is necessary to present PFP in a tightly constrained mathematical model that gives structure to the implementation and consequent reinforcement process.

A mathematical model of a complex concept provides an objective abstract representation of that concept. Mathematical models allow for systematic adaptation to the assumptions by holding the variables in a constant ratio. By creating mathematical models, wordy verbal descriptions are coded into precise mathematical equations without unnecessary details. Issues identified by faculty as concerns in the criteria for earning the financial incentives can be systematically modeled and resolved without sacrificing the integrity of the model as a whole.

In a study conducted by Terpstra and Honoree (2008) in which almost 500 faculty members were surveyed on problems undermining the effectiveness of pay based on performance, the most salient problem identified by faculty was that “the merit pay increases that are given out are too small to motivate faculty” (p. 48). This echoes the findings in the health care sector settings (Rosenthal & Frank, 2006). The impact of this problem on the PFP model will be demonstrated. By using a realistic mathematical model to describe PFP, weights can be manipulated and the consequences can be evaluated before any costly mistakes are made in terms of money, time, and goodwill.

BACKGROUND

The Board of Trustees of a college of two of the authors has proposed a PFP model to reward faculty who perform at a high level. The process to be used to determine the awards will be the department evaluation guidelines. These guidelines identify the evaluation parameters for the four areas of (i) teaching, (ii) advising and student support, (iii) college and community service, and (iv) scholarship. These areas are assessed annually with reference to three categories: (i) "needs improvement," (ii) "meets standards," and (iii) "exceeds standards." Faculty who achieve the "exceed standards" category in one or more of the four evaluation areas and who "meet standards" in the remaining areas, will be considered for a PFP award in the category where they "exceed standards."

The performance categories for the PFP model are "excellent" (Level I), "exceptional" (Level II), and "extraordinary" (Level III), with increasingly higher standards and payouts for each level.

NUMERICAL EXAMPLE

For expository purposes, assume that the total number of PFP awards for a particular year is 330 and that the total funds available in that year are two million dollars. In this example the highest evaluation weight of six is given to teaching area since the college is a teaching institution. The next most important evaluation area is assumed to be scholarship with a weight of four. The remaining two evaluation areas are ranked as equally important with weights of one each. The weights assigned to the performance categories in this example are two for excellent, four for exceptional and eight for extraordinary.

One perceived problem with PFP is that the amount of the award is not viewed by the recipient as reflective of the effort and performance required to receive the honor (Terpstra & Honoree, 2008). Thus in this example the weights for the three performance categories are an attempt to overcome this drawback.

Table 1 details the information for the numerical example. The awards matrix gives the breakdown of the 330 honorees. The weights matrix is the product of the loads assigned to the evaluation area and the performance criteria for each element of the matrix. For example the weight of two-sevenths assigned to the extraordinary performance area for teaching is the product of the teaching weight of one-half (6/[6+4+1+1]) and the weight for extraordinary performance of four sevenths (8/[8+4+2]). The matrix for the fraction of funds to be paid out is the proportion of the total funds for each element. This is the product of the relevant fraction from the weights matrix and the ratio of the number of awards for that category as a function of the total number of awards. For the extraordinary teaching element the fraction of 0.051948 is the weight of two-sevenths multiplied by ratio of the number of extraordinary teaching awards to the total number of awards or .181818 (60/330). The payout matrix converts the fraction of funds matrix into the actual dollar amount of each award. Thus each recipient of an extraordinary teaching award in that year would receive a one-off payment of $14,828.54.

Table 1
Matrix Elements for the Numerical Example

Matrix	Performance criteria	Evaluation area			
		Teaching	Advising	Service	Scholarship
Awards	Excellent	80	30	5	12
	Exceptional	70	15	10	8
	Extraordinary	60	10	20	10
Weights	Excellent	1/14	1/84	1/84	1/21
	Exceptional	1/7	1/42	1/42	2/21
	Extraordinary	2/7	1/21	1/21	4/21
Fraction of funds	Excellent	0.017316	0.001082	0.00018	0.001732
	Exceptional	0.030303	0.001082	0.000722	0.002309
	Extraordinary	0.051948	0.001443	0.002886	0.005772
Payout	Excellent	$ 3,707.13	$ 617.85	$ 617.85	$ 2,471.42
	Exceptional	$ 7,414.27	$ 1,235.71	$ 1,235.71	$ 4,942.84
	Extraordinary	$ 14,828.54	$ 2,471.42	$ 2,471.42	$ 9,885.69

Note. The mathematical model denotes the awards matrix as N, the weights matrix as W, the fraction of funds matrix as P, and the payout matrix as O. Given this scenario, in the mathematical model parlance $N_T = 330$, $P_T = 0.116774892$, $f_T = \$ 2{,}000{,}000$, and $f_A = \$ 17{,}126{,}969.42$.

The model was developed with two constraints to ensure the fairness of the system. The first constraint is that the payments be the same if the weights are the same. In the example the weights for the advising and service evaluation areas were the same. In order to reflect their equal importance it is paramount that the payouts be the same for all three performance levels for these two evaluation areas. It can be seen from the payout matrix that this is the case thus the model does ensure compliance with the first constraint.

It can be seen from the payout matrix that across all three performance categories, regardless of the evaluation area, the dollar amount of the award is in the ratio of the weights assigned to the performance categories of two (excellent), four (exceptional) and eight (extraordinary). Likewise, for the four evaluation areas, regardless of the performance category, the payout is in the ratio of the weights assigned of: six for teaching, four for scholarship, and one each for advising as well as service. This meets the requirement of the second constraint that if the weight for one category is higher than another category then the amount of the award for the higher weighted category be no less than the award for the lower weighted category.

THE MATHEMATICAL MODEL

In this model, we seek a 3×4 matrix P such that the ij^{th} entry of P gives the fraction of total funds available allocated to the i^{th} performance criterion of the j^{th} evaluation area. In this case, i = 1, 2, 3, where 1 represents excellent, 2 represents exceptional, and 3 represents extraordinary. The evaluation areas are j = 1, 2, 3, 4 where 1 represents teaching, 2 represents advising, 3 represents service, and 4 represents scholarship. Each performance criterion and evaluation area is weighted by the elements from a matrix W of the normalized weights of the performance criteria and the evaluation areas. The independent variable is the number of awards, or population, in each performance criterion of each evaluation area, given by the elements of the matrix N.

Now, consider the matrix P, whose elements are:

$$p_{ij} = \frac{w_{ij}}{N_T}\, n_{ij} \qquad (1)$$

where n_{ij} = the population in each performance criterion, i, of each evaluation area, j,

$N_T = \sum_{i,j} n_{ij}$ is the total aggregate population over all categories, and

w_{ij} = the weight for the i^{th} performance criterion of the j^{th} evaluation area.

DERIVATION OF W

We can consider W to be the matrix product of two vectors, $\vec{w}_i$ and $\vec{w}_j$ where $\vec{w}_i$ is the column vector of normalized weights of performance criteria. Thus, if we have weights of two, three, and four:

$$\vec{w}_i = \begin{bmatrix} 2/9 \\ 3/9 \\ 4/9 \end{bmatrix} \qquad (2)$$

For example, $\vec{w}_{i_1}$ is the weight of the excellent performance criterion divided by the sum of the weights of all performance criteria, or 2/9 in this case.

Similarly, $\vec{w}_j$ is the row vector of normalized weights of the evaluation areas. Thus, if we have weights of 10, four, three, and three:

$$\vec{w}_j = \begin{bmatrix} \frac{10}{20} & \frac{4}{20} & \frac{3}{20} & \frac{3}{20} \end{bmatrix} \qquad (3)$$

For example, $\vec{w}_{j_1}$ is the weight of the teaching evaluation area divided by the sum of the weights of all the evaluation areas, or 10/20 in this case.

Thus, $W = \vec{w}_i\, \vec{w}_j$.

MODEL CONSTRAINTS

To ensure the fairness of the model, it is necessary to impose two constraints. First, if the evaluation weights are the same then the payouts must be the same. Second, if the evaluation weight in one area is greater than that for any other area then the payout for the higher weighted area cannot be less than that with the lower weight.

The model must satisfy the following constraints:

1. If the weights for any evaluation areas are the same then the payouts to each member of the population in those evaluation areas must be the same. That is, if $w_{ia} = w_{ib}$ for some $a \neq b$ and for all i, then $\frac{p_{ia}}{n_{ia}} = \frac{p_{ib}}{n_{ib}}$ for all i.

2. If the weight for one evaluation area is greater than that of another evaluation area, the payout to each member of the higher weighted evaluation area must not be lower than the payout to each member of the lower weighted evaluation area. That is, if $w_{ia} \geq w_{ib}$ for some $a \neq b$ and for all i, then $\frac{p_{ia}}{n_{ia}} \geq \frac{p_{ib}}{n_{ib}}$ for all i.

First, we check constraint 1. Take a and b such that $w_{ia} = w_{ib}$ for all i. From the definition of the model, this implies that

$$\frac{p_{ia}}{n_{ia}} = \frac{w_{ia}}{N_T} = \frac{w_{ib}}{N_T} = \frac{p_{ib}}{n_{ib}} \tag{4}$$

for all i. Thus, the first constraint is satisfied for all w_{ij}.

Next, we check constraint 2. Take a and b such that $w_{ia} > w_{ib}$ for all i. Then

$$\frac{p_{ia}}{n_{ia}} = \frac{w_{ia}}{N_T} > \frac{w_{ib}}{N_T} = \frac{p_{ib}}{n_{ib}} \tag{5}$$

Thus, the second constraint is satisfied for all w_{ij}. In fact, a stronger condition is also met. That is, having an evaluation area a weighted higher than another area b means that the payout to each member of the population of n_{ia} is *greater than* the payout to each member of the population of n_{ib}.

DETERMINATION OF PAYOUTS

Unfortunately, it is not possible to multiply matrix P by the total amount of funds available in order to determine the funds allocated to each category, because $\sum_{i,j} p_{ij} \leq 1$ with equality only in the case of $i = j = 1$. For example, take $n_{ij} = a$, for a a positive constant. That is, assume that the population of each category is the same. Further, assume that $\vec{w}_i$ has i_n elements and $\vec{w}_j$ has j_n elements. Then

$$\sum_{i=1}^{i_n} \sum_{j=1}^{j_n} p_{ij} = \sum_{i=1}^{i_n} \sum_{j=1}^{j_n} \frac{n_{ij} w_{ij}}{N_T} \tag{6}$$

Since each $n_{ij} = a$ and $N_T = i_n j_n a$, substituting both of these expressions into the above equation gives

$$\sum_{i=1}^{i_n} \sum_{j=1}^{j_n} p_{ij} = \sum_{i=1}^{i_n} \sum_{j=1}^{j_n} \frac{w_{ij}}{i_n j_n} \tag{7}$$

Because $i_n j_n$ is a constant, it can be taken out of the summation. This gives

$$\sum_{i=1}^{i_n} \sum_{j=1}^{j_n} p_{ij} = \frac{1}{i_n j_n} \sum_{i=1}^{i_n} \sum_{j=1}^{j_n} w_{ij} \tag{8}$$

By definition, each w_{ij} is equal to the product of the i^{th} component of $\vec{w}_i$ (denoted w_{i_i}) and the j^{th} component of $\vec{w}_j$ (similarly denoted w_{j_j}). Thus,

$$\frac{1}{i_n j_n} \sum_{i=1}^{i_n} \sum_{j=1}^{j_n} w_{ij} = \frac{1}{i_n j_n} \sum_{i=1}^{i_n} \sum_{j=1}^{j_n} w_{i_i} w_{j_j} \tag{9}$$

As w_{i_i} does not depend on j, it can be taken out of the inside summation. This gives

$$\frac{1}{i_n j_n}\sum_{i=1}^{i_n}\sum_{j=1}^{j_n} w_{i_i} w_{j_j} = \frac{1}{i_n j_n}\sum_{i=1}^{i_n} w_{i_i}\sum_{j=1}^{j_n} w_{j_j} \tag{10}$$

But $\vec{w}_i$ and $\vec{w}_j$ are normalized, so

$$\sum_{j=1}^{j_n} w_{j_j} = 1 = \sum_{i=1}^{i_n} w_{i_i} \tag{11}$$

Therefore, in this case, we get

$$\frac{1}{i_n j_n}\sum_{i=1}^{i_n} w_{i_i}\sum_{j=1}^{j_n} w_{j_j} = \frac{1}{i_n j_n}\sum_{i=1}^{i_n} w_{i_i} = \frac{1}{i_n j_n} \tag{12}$$

Since i_n and j_n are positive integers, this equation is less than or equal to 1. It can be shown that the sum of the entries of the matrix P is always less than or equal to 1 (see the appendix). Thus, there is an easy solution to determine the payout for each member.

Let P_T equal the sum of the entries of the matrix P. That is, let

$$P_T = \sum_{i,j} p_{ij} = \sum_{i,j}\frac{n_{ij}}{N_T} w_{ij} \tag{13}$$

Then, dividing each entry in P by P_T,

$$\sum_{i,j}\frac{p_{ij}}{P_T} = \sum_{i,j}\frac{n_{ij} w_{ij}}{N_T P_T} = \frac{1}{P_T}\left[\sum_{i,j}\frac{n_{ij} w_{ij}}{N_T}\right] \tag{14}$$

However, the term in the brackets is P_T by definition, so this reduces to

$$\sum_{i,j}\frac{p_{ij}}{P_T} = \frac{1}{P_T} P_T = 1 \tag{15}$$

Thus, multiplying the matrix P by the scalar $\frac{1}{P_T}$ forces the elements of P to sum to 1. This matrix can then be multiplied by the total amount of funds available, denoted by f_T. Then the quantity in each element of the matrix $\frac{f_T}{P_T} P$ is the total amount awarded to that category. More simply, since $\frac{1}{P_T}$ is a scalar, we can think of multiplying P by an adjusted amount of funds, denoted f_A, where $f_A = \frac{f_T}{P_T}$. Thus, the total amount awarded to each category would be given by $f_A P$. To determine how much each member of the population of element ij receives, divide the ij^{th} element of $f_A P$ by the ij^{th} element of N. Thus, each award recipient's payout for a particular evaluation area would be given by the elements of the matrix O

$$o_{ij} = \frac{f_A}{n_{ij}} p_{ij} \tag{16}$$

This model fulfills both constraints and uses all the funds allocated for the PFP awards.

DISCUSSION

It can be seen that the payouts to each faculty member (p_{ij}) are independent of the number of other faculty who have achieved the same award. This will discourage attempts by faculty to concentrate on achieving a high performance level in a sparsely populated evaluation area in an attempt to maximize their returns. This follows from the fact that the individual payouts depend only on the weight matrix W and in fact are an exact multiple of the smallest payout in O. The multiple is equal to the ratio of the element of w_{ij} to w_{min}, where w_{min} is the smallest value of W.

CONCLUSION

The proposed model develops a PFP system that satisfies the constraints imposed to ensure that the system is fair. The model is flexible enough to allow the performance weights to be adjusted to ensure that the amount of the award at each performance level is sufficiently differentiated to be perceived to reward the effort required to attain the level of performance necessary for the honor.

Although the weak economy has hindered expectations of any increases (merit or cost-of-living), and even resulted in hiring freezes or layoffs, it is in these lean times that colleges can prepare to implement financial incentives to reward the most productive faculty when the economy improves. By developing a plan and creating a model that will support it, future faculty development can be maximized.

REFERENCES

Baker, G. P., Jensen, M.C., & Murphy, K. J. (1988). Compensation and incentives: Practice vs. theory. *Journal of Finance*, *43*(3), 593-616.

Haley, M. R. (2003). The response of worker effort to piece rates. Evidence from the Midwest logging industry. *Journal of Human Resources*, *38*(4), 881-890.

Jenkins, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance? A meta-analytic review of empirical research. *Journal of Applied Psychology*, *83*(5), 777-787.

Mercer's 2010 U.S. Executive Compensation and Performance Survey. (2010). *Mercer*. Retrieved May 21, 2010, from http://www.mercer.com/summary.htm?idContent=1375180.

Rosenthal, M. B., & Frank, R. G. (2006). What is the empirical basis for paying for quality in health care? *Medical Care Research and Review*, *63*(2), 135-157. doi: 10.1177/1077558705285291.

Rosenthal, M. B., Landon, B. E., Normand, S.L., Frank, R. G., & Epstein, A. M. (2006). Pay for performance in commercial HMOs. *The New England Journal of Medicine, 355*(18), 1895-1902.

Shearer, B. (2004). Piece rates, fixed wages and incentives: Evidence from a field experiment. *Review of Economic Studies, 71*(2), 513-534.

Skinner, B. F. (1974). *About behaviorism*. New York, NY: Knopf.

Terpstra, D. E., & Honoree, A. L. (2008). Faculty perceptions of problems with merit pay plans in institutions of higher education. *Journal of Business and Management, 14*(1), 43-59.

APPENDIX

Proof

To prove that the sum of the entries of the matrix P is always less than or equal to 1 (that is, $\sum_{i,j} p_{ij} \leq 1$), first consider P to be the Hadamard (element-wise) product of W and a matrix $\boldsymbol{N'}$, where $\boldsymbol{N'} = \frac{1}{N_T}\boldsymbol{N}$. Thus,

$$p_{ij} = w_{ij}n'_{ij} \quad \text{(A1)}$$

Now recall that the Frobenius inner product on the space of all matrices is defined as

$$\langle A, B\rangle_F = \sum_i \sum_j a_{ij}\, b_{ij} \quad \text{(A2)}$$

Thus, we can write the sum of all the entries in P as the Frobenius inner product of $\boldsymbol{W}$ and $\boldsymbol{N'}$,

$$\sum_i \sum_j p_{ij} = \langle \boldsymbol{W}, \boldsymbol{N'}\rangle_F = \sum_i \sum_j w_{ij}n'_{ij} \quad \text{(A3)}$$

The Frobenius inner product admits the Frobenius norm, $\|\boldsymbol{A}\|_F$, defined for real numbers as

$$\|\boldsymbol{A}\|_F = \sqrt{\langle A, A\rangle_F} = \sqrt{\sum_i \sum_j a_{ij}^2} \quad \text{(A4)}$$

Recall the Cauchy-Schwarz inequality, which will be central to our proof, is

$$|\langle \boldsymbol{A}, \boldsymbol{B}\rangle| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\| \quad \text{(A5)}$$

for any inner product and the corresponding norm. Here, we use the Frobenius inner product and Frobenius norm. Notice that, in that case, the left-hand side of the Cauchy-Schwarz inequality can be written as the sum of all the entries in the matrix $\boldsymbol{P}$, $\langle \boldsymbol{W}, \boldsymbol{N'}\rangle_F$, as we have already shown above. This implies that

$$|\langle \boldsymbol{W}, \boldsymbol{N'}\rangle_F| \leq \|\boldsymbol{W}\|_F \|\boldsymbol{N'}\|_F \quad \text{(A6)}$$

In the case of our model, the absolute value is redundant because all entries of $\boldsymbol{W}$ and $\boldsymbol{N'}$ are positive, so the Frobenius inner product must be positive. Thus, we can discard the absolute value around the left-hand side. The result is

$$\langle \boldsymbol{W}, \boldsymbol{N'}\rangle_F \leq \|\boldsymbol{W}\|_F \|\boldsymbol{N'}\|_F \quad \text{(A7)}$$

Now consider the right-hand side. Since both $\boldsymbol{W}$ and $\boldsymbol{N'}$ are matrices in which every element is less than one, we can consider just one matrix and generalize our result to the other. Hence, without loss of generality, we consider $\boldsymbol{W}$. By definition of the Frobenius norm, $\|\boldsymbol{W}\|_F$ is

$$\|\boldsymbol{W}\|_F = \sqrt{\sum_i \sum_j w_{ij}^2} \quad \text{(A8)}$$

We can show that the values of $\boldsymbol{W}$ sum to 1. Notice that the ij^{th} element of $\boldsymbol{W}$ can be written $w_{i_i}w_{j_j}$. Thus, the sum of all the elements of $\boldsymbol{W}$ can be written

$$\sum_{i=1}^{m} \sum_{j=1}^{n} w_{i_i}w_{j_j} \quad \text{(A9)}$$

where *m* is the number of elements in $\vec{w}_i$ and *n* is the number of elements in $\vec{w}_j$. But w_{i_i} does not depend on *j*, so we can take it out of the first summation. The resulting equation is

$$\sum_{i=1}^{m} w_{i_i} \sum_{j=1}^{n} w_{j_j} = \sum_{i=1}^{m} w_{i_i} = 1 \tag{A10}$$

The equalities follow because $\vec{w}_i$ and $\vec{w}_j$ are normalized, and therefore their elements sum to 1. Therefore

$$\sum_{i=1}^{m} \sum_{j=1}^{n} w_{i_i} w_{j_j} = \sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij} = 1 \tag{A11}$$

Squaring both sides gives

$$\left(\sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij}\right)^2 = 1 \tag{A12}$$

Now let us order the elements of $\boldsymbol{W}$ so that w_k is the k^{th} element of $\boldsymbol{W}$, where $k = i + m(j-1)$, and where, as before, *m* is the number of elements in $\vec{w}_i$. Let w_{-k} denote the sum of all the elements of $\boldsymbol{W}$ except for the k^{th}. Then we can write the above expression as

$$w_1^2 + w_1 w_{-1} + w_2^2 + w_2 w_{-2} + \cdots + w_p^2 + w_p w_{-p} = 1 \tag{A13}$$

where *p* is the number of elements in $\boldsymbol{W}$.

Since all of the elements of $\boldsymbol{W}$ are positive, each term $w_k w_{-k}$ is positive. Therefore

$$w_1^2 + w_2^2 + \cdots + w_p^2 \leq w_1^2 + w_1 w_{-1} + w_2^2 + w_2 w_{-2} + \cdots + w_p^2 + w_p w_{-p} \tag{A14}$$

In fact, the two sides of this expression are only equal if all but one of the w_k are 0, in other words, $\boldsymbol{W}$ has only one element. The above equation implies that

$$w_1^2 + w_2^2 + \cdots + w_p^2 \leq 1 \tag{A15}$$

Recalling the definition of w_k, we can see that this is equivalent to

$$\sum_i \sum_j w_{ij}{}^2 \leq 1 \tag{A16}$$

Taking the (positive) square root of both sides gives

$$\sqrt{\sum_i \sum_j w_{ij}{}^2} \leq 1 \tag{A17}$$

Recall that the Frobenius norm of $\|\boldsymbol{W}\|_F$ is defined as the left-hand side of the above equation. Therefore,

$$\|\boldsymbol{W}\|_F \leq 1 \tag{A18}$$

A similar argument suffices to demonstrate that

$$\|\boldsymbol{N}'\|_F \leq 1 \tag{A19}$$

Therefore, by the Cauchy-Schwarz inequality

$$\sum_i \sum_j p_{ij} = \langle \boldsymbol{W}, \boldsymbol{N}' \rangle_F \leq \|\boldsymbol{W}\|_F \|\boldsymbol{N}'\|_F \leq 1 \tag{A20}$$

Thus, we prove that the sum of $\boldsymbol{P}$ is less than or equal to 1. In addition, if there is more than one element in $\boldsymbol{W}$ a strict inequality holds. Therefore the sum of the elements of $\boldsymbol{P}$ cannot possibly be 1 unless $i = j = 1$; that is, there is only one element in $\boldsymbol{W}$.