*Journal of Agricultural and Applied Economics*, 34,3(December 2002):531–546


# Student Numbers and Sustaining Courses and Fields in Ph.D. Programs

**George C. Davis and Ernesto Perusquia**

Many agricultural economics departments are concerned about the vitality of their Ph.D. programs. A particular problem is insufficient student numbers to justify teaching certain courses or fields. As a consequence, much faculty time can be spent debating alternative program structures without any real idea of the likelihood that a proposed program structure will succeed. This article presents a framework for deriving some analytical and empirical results for alternative Ph.D. program structures. A downloadable program is used to generate some representative results that will hopefully help others minimize speculations and time spent in committee or departmental meetings.

*Key Words:* Ph.D. programs, student numbers

**JEL Classifications:** A2, Q1

In the Department of Agricultural Economics at Lake Wobegon University, everything good about the Ph.D. program is above average: the number of students, the assistantships, the graduate faculty, the salaries, even the support staff. The students can choose any *JEL* code number as a specialty area, all classes have ample students, and the student to faculty ratio is high. In fact, the department head, Professor Twain, thinks the rumors about dying Ph.D. programs "are greatly exaggerated."

As a fantasy, the vibrant Ph.D. program at Lake Wobegon University is something many departments would like to experience. In reality, there are many agricultural economics departments around the country that are concerned about the vitality of their Ph.D. programs. To document this concern only requires a sympathetic ear at professional meetings or a cursory review of the literature (e.g., Huffman and Orazem; Norton et al.; Schrimper 1985, 1999). A particular problem for many departments is low student numbers.

Low student numbers are especially problematic for departments in the teaching area, where it would seem there are economies of class size, at least over a large range of student numbers. For the instructor, much of the total cost associated with teaching a class is fixed: for the same type of instruction, out-of-class preparation and in-class instruction time are basically the same for 5, 10, or 20 students. Although the variable cost is increasing, it is likely increasing at a decreasing rate. In general, it is more time efficient to teach a full class than an almost empty class. This argument is even more relevant when the opportunity cost of time is considered, given that every hour devoted to teaching takes away

---

George C. Davis and Ernesto Perusquia are associate professor and graduate research assistant, respectively, Department of Agricultural Economics, Texas A&M University, College Station, TX.

Appreciation is extended to Ron Schrimper for helpful discussions and for providing us with some of the data used in the present article and to James Richardson and an anonymous referee for helpful discussions. Recognition is also granted to Wade Griffin for assigning one of the authors to a subcommittee from which this research originated and challenging the author to turn the committee work into a journal article.

from time that could be spent doing research or extension activities. These cost economies are probably the reason that most universities have some type of policy on the minimum class size required for a course.

Many variables influence whether or not a course or a field will be successful: the number of students, which depends on assistantship levels and the number of assistantships available; the number of courses offered; the number of courses required; the number of field course credit hours required; the number of credits per course; etc. With so many variables, departmental debates about the small class problem can bog down as different factions argue for different instruments (variables). One faction claims, "the problem is not courses: we need more students, and to get more students we need more assistantship resources." Another faction says, "we are unlikely to get more resources, so we need to cut the number of offerings." Still another faction argues, "we don't have to cut course offerings—we can teach the same number of courses but just require the students to take more course hours." Each of these statements is true to some degree, but there is also a great deal of uncertainty as to the effectiveness of each of these alternatives. What is required is some evidence as to the efficacy of these alternatives.

There are two ways to obtain evidence as to the efficacy of alternative Ph.D. program structures: experimentally or analytically. Although an experimental approach of turning a Ph.D. program into a laboratory will provide observational evidence, this evidence comes at an extremely high price—excessive administrative duties for faculty members, program discontinuity, and varying program quality, to name a few. Alternatively, although an analytical approach will not provide observational evidence, it can provide likely outcome evidence. More important, it does not come with such a high price tag in terms of faculty time and program continuity. As a consequence, an analytical approach is attractive.

The purpose of this article is to present an analytical framework for determining the number of expected students in a field and in a course in some alternative Ph.D. program structures. We have found in our own department that these results helped minimize the amount of time spent debating and deciding the likely success of alternative program structures. We suspect others struggling with these issues may also find the results useful. Because there are many factors that can affect the number of students, and therefore the number of courses that are viable and vital, the next section gives a literature review of the main factors that have been identified as important on a national level. The following section then provides the analytical framework and the results. The article closes with a summary and some concluding remarks.

## Some Important Characteristics of Ph.D. Programs

To understand the concerns about Ph.D. programs, it is informative to first look at what has happened to the quantity of Ph.D. students over time. Schrimper (1999) provided perhaps the most recent data available over time on Ph.D. degrees granted.[1] Those data were an update of the data given in Schrimper (1985) and were compiled mainly from the May issues of the *American Journal of Agricultural Economics*. The data were compiled from 36 institutions, and the national and regional averages over time are shown in Figure 1. As can be seen, the average number of Ph.D.s per department for the entire United States per year has varied between four and six from 1985 to 1997. In general, the North Central and Western regions averaged granting more Ph.D.s than the national average, whereas the Northeast and Southern regions averaged granting fewer Ph.D.s than the national average.

Table 1 gives the summary statistics for each department and overall departments within each region between 1985 and 1997.

[1] Schrimper's data is for Ph.D. degrees granted. Although the emphasis here is on Ph.D. students, it seems reasonable to expect the number of Ph.D.s granted to be some relatively constant proportion of the number of students.

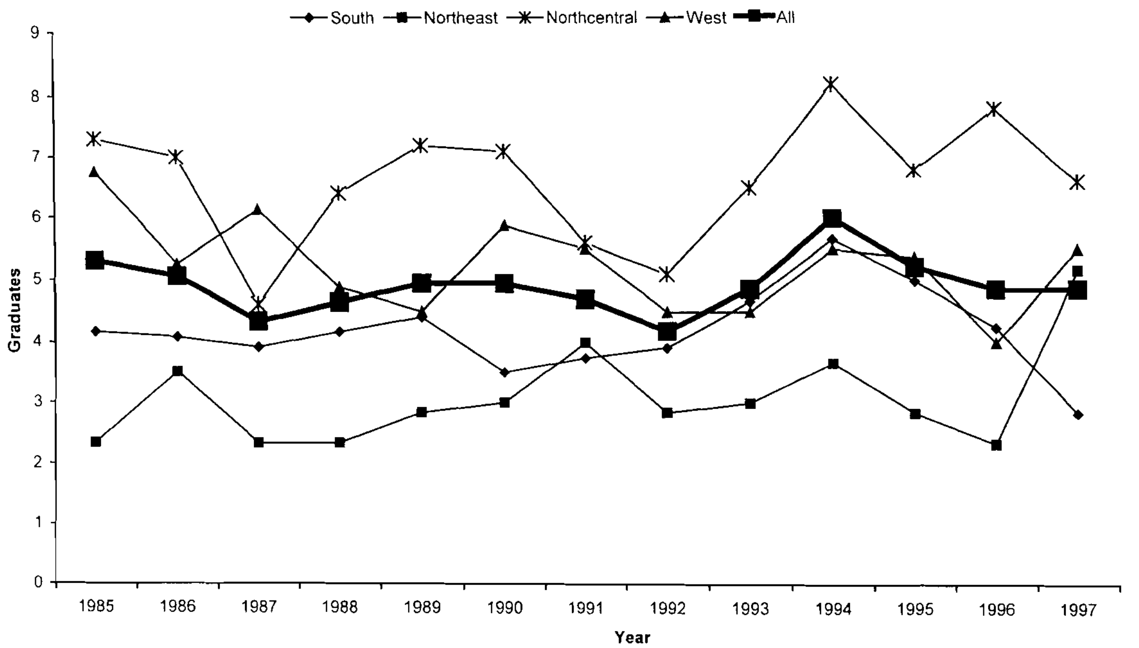


**Figure 1.** Average Number of Ph.D. Degrees Per Year 1985–1997

Of the 36 departments listed, 15 had averages above the national average, leaving 21 with averages below the national average. Of the 15 departments with averages above the national average, all had coefficients of variation less than .5. In fact, the correlation between the means and the coefficients of variation is −.79, which suggests that size and stability are positively correlated. This distribution of students is likely due to many factors, which can be partitioned into supply and demand factors.

In the only published empirical analysis of the market for Ph.D. graduate students that we are aware of, Huffman and Orazem developed a theoretical model of the demand and supply for graduate students. Overall, the empirical results were in agreement with their theory and intuition. On the demand side, they found that the wage rate for graduate assistantships, total state farm income, and experiment station expenditures were negatively related with quantity demanded. They also found that total state personal income, the average wage rate for assistant professors, total state agricultural extension expenditures, and the number of undergraduates were all positively associated with quantity demanded. The only coefficient having a sign in conflict with theory was that on the agricultural extension expenditures. On the supply side, they found that available graduate student assistantships, the wage rate for graduate assistantships, the wage rate for assistant professors, and the size of the faculty were all positively associated with the quantity supplied. They found that both measures of opportunity cost were negatively associated with the quantity supplied. Of particular interest, they found that the supply elasticity, with respect to the assistant's wage rate net of tuition, was .57. Thus, for every 10% increase in the assistantship wage rate net of tuition, the number of students is expected to increase by 5.7%. Although one could quibble over some of the specific variables in the model, the major determinants seem to be captured, mainly the price (available assistantships, assistantship wage rate, and tuition), opportunity cost, expected return, and institutional effects.

In looking at these major determinants, the individual departments have the greatest control over the two price determinants: available

**Table 1.** Summary Statistics on Number of Ph.D.s Granted Per Year by Department, 1985–1997

| University | Average Number of Ph.D.s Granted | Standard Deviation of Ph.D.s Granted | Coefficient of Variation of Ph.D.s Granted |
|---|---|---|---|
| **South** | **4.2** | **3.2** | **0.8** |
| Texas A&M | 9 | 3.3 | 0.4 |
| Oklahoma State | 7.3 | 2.7 | 0.4 |
| NCSU | 7.2 | 2.5 | 0.3 |
| Florida | 4.9 | 2.5 | 0.5 |
| VPI | 4 | 1.7 | 0.4 |
| Kentucky | 3.7 | 2.2 | 0.6 |
| Mississippi State | 3.4 | 2.8 | 0.8 |
| Georgia | 2.6 | 2.0 | 0.8 |
| Clemson | 2.3 | 1.6 | 0.7 |
| Tennessee | 2.2 | 1.5 | 0.7 |
| Auburn | 2.1 | 1.6 | 0.8 |
| Texas Tech | 1.4 | 1.0 | 0.7 |
| **Northeast** | **3.1** | **2.9** | **0.9** |
| Cornell | 8.1 | 2.8 | 0.3 |
| Rhode Island | 2.8 | 1.5 | 0.5 |
| Maryland | 2.5 | 1.8 | 0.7 |
| Penn State | 2.3 | 2.4 | 1.0 |
| Connecticut | 1.6 | 1.4 | 0.9 |
| Massachusetts | 1.1 | 0.8 | 0.7 |
| **North Central** | **6.6** | **4.2** | **0.6** |
| Iowa State | 12.1 | 3.3 | 0.3 |
| Minnesota | 10.1 | 3.1 | 0.3 |
| Michigan State | 8.4 | 3.4 | 0.4 |
| Illinois | 7.8 | 1.6 | 0.2 |
| Purdue | 7.5 | 3.5 | 0.5 |
| Ohio State | 7.1 | 3.9 | 0.5 |
| Wisconsin | 5.5 | 1.5 | 0.3 |
| Missouri | 3.4 | 2.3 | 0.7 |
| Kansas State | 2.3 | 2.2 | 0.9 |
| Nebraska | 2.1 | 1.8 | 0.8 |
| **West** | **5.3** | **3.2** | **0.6** |
| Berkeley | 9.8 | 3.6 | 0.4 |
| Davis | 6.4 | 3.1 | 0.5 |
| Stanford | 5.5 | 1.7 | 0.3 |
| Washington State | 4.8 | 1.9 | 0.4 |
| Oregon State | 4.2 | 2.7 | 0.6 |
| Hawaii | 3.7 | 2.4 | 0.6 |
| Colorado State | 3.7 | 1.9 | 0.5 |
| Utah State | 3.9 | 2.9 | 0.7 |
| **National** | **4.9** | **3.6** | **0.7** |

assistantships and the assistantship wage rates. Yet the control over these determinants has likely declined over the past 20 years as the source of funding has changed (see Alston and Pardey; Huffman and Just 1994, 1999; Just and Huffman; Norton et al.; Perry; Rubenstein et al.).

Much of the funding discussion in the literature has focused on the difference between formula funds and competitive funds. There is no competition for formula funds between universities; formula funds are allocated based on a specific formula.[2] Huffman and Just (1994, 1999) have documented empirically the benefits of formula funds over competitive funds for agricultural productivity, but there would seem to also be some advantages associated with formula funds in controlling assistantships.

As Huffman and Just (1994, 1999) pointed out, formula funds have less risk and uncertainty than competitive funds. In addition, the research projects associated with formula funds are continuing with no real deadline or deliverable product. Competitive funds usually have a short timeline with a specific deliverable product. For these reasons, it is much easier to make budgeting plans in recruiting graduate students and offering assistantships with formula funds. Furthermore, first- and second-year students are more difficult to fund with competitive funds than with formula funds, because these students usually do not possess the necessary skills to be very productive on a short-term competitive fund project. Alternatively, because of the longer timeline and no specific deliverable product associated with formula funds, first- and second-year students can be "subsidized" with formula funds until they are at a more productive stage.

That said, creative administrators can make

[2] Alston and Pardey (chapter 2) provided a nice historical account of the changing formula. As Huffman and Just (1999) stated in their footnote two, "In the Amended Hatch Act (1995), . . . , [a] new formula was established: 20% is divided equally among states, 26% is allocated according to a state's share of farm population, 26% is allocated on a state's share of national rural population, 25% is allocated to regional research, and 3% is allocated to administrative cost."

formula funds and competitive funds highly substitutable, but only if the competitive funds have been captured. The relevant follow-up questions then are: what has happened to the amount of formula funds versus competitive funds over time? what has happened to the distribution of these funds across departments over time?

Perry showed that *total* research and extension expenditures actually have been on the rise since 1982. Norton et al. found similar results between 1974 and 1993. Norton et al. and Perry both showed that, in real terms, the level of formula funds has been decreasing since the early 1980s, whereas the level of competitive funds have been increasing. What is perhaps most interesting about this fact is the distribution of the competitive funds across states.

Norton et al. found that competitive funds are highly skewed toward a few states. Using the data reported in Norton et al. (their Table 3) and from Schrimper, our Table 2 shows the rank of the 36 departments by states in terms of funding for 1986 and 1992, along with the number of Ph.D. degrees granted 4 years later. The states are sorted in descending order by the average number of Ph.D.s granted, and the underline indicates the cutoff for being above the national average of 4.91 degrees. It should be kept in mind that the Perry and Norton et al. data are with respect to *total* agricultural research and extension monies, not just those devoted to agricultural economics.[3] With this caveat in mind, in 1990 there were 16 departments (associated with 13 states) whose average number of Ph.D.s granted was above the national average. Four years earlier, these 13 states accounted for about 54% of the competitive funds, implying that the other 39 states accounted for 46%.[4] In 1996, there were once again 16 departments (associated with 14 states) whose average number of Ph.D.s granted was above the national average. Four years earlier, these 14 states accounted for about 56% of the competitive funds, implying that the other 38 states accounted for 44%.

This casual empiricism is only suggestive of the possible relationship between the size of Ph.D. programs and the changing distribution of funds between formula and competitive funds. The data are too imprecise and the statistics too crude to draw any strong conclusions. However, the results do seem to be consistent with intuition and anecdotal evidence. Although the number of Ph.D.s granted has remained relatively stable, the real decline in formula funds and the increasing reliance on competitive funds has likely placed more stress on the majority of departments, given that a majority of the competitive funds go to a minority of the departments.

## Analytical Approach

Certainly, there are many factors that will affect the number of students in a class, and these will vary over time and by department. Estimating an appropriate multivariate model, such as a count data system, could be done if sufficient data existed and there was enough program structure variation. However, for most departments, these data are either not readily available or there is not enough variation in the program structure to draw reliable econometric inferences about alternative structures. For these reasons, an alternative approach is pursued, to shed some light on the question at hand.

### *The Analytical Approach*

Many Ph.D. programs are structured such that a core set of courses are required to be taken by all students, followed by a set of elective courses from which the students can choose.

[3] Norton et al. indicated (p. 1344) that economics projects were not eligible for competitive grants during the first decade of funding. In addition, most of the rate of change between 1974 and 1993 seems to be due to changes between 1986 and 1990. The relationship therefore between agricultural economics funding and total agricultural research funding is by no means clear, but it would be expected to be positive.

[4] As Norton et al. indicated, there are 52 "states" because the data include the District of Columbia and Puerto Rico.

**Table 2.** Average Ph.D.s Grated Per State and Competitive Fund Rank 1986 and 1992[a]

| State | 1990 Student Average | 1986 Competitive Funds % | 1986 Competitive Funds Rank | State | 1996 Student Average | 1992 Competitive Funds % | 1992 Competitive Funds Rank |
|---|---|---|---|---|---|---|---|
| Minnesota | 12.3 | 2.75 | 15 | Iowa | 10.5 | 2.66 | 13 |
| Iowa | 11.7 | 2.06 | 17 | Minnesota | 10.5 | 2.19 | 19 |
| New York | 11.3 | 7.90 | 2 | Illinois | 8.5 | 3.96 | 5 |
| California | 10.3 | 12.35 | 1 | Indiana | 8.5 | 2.35 | 16 |
| Indiana | 8.7 | 2.80 | 14 | North Carolina | 8 | 3.95 | 6 |
| Ohio | 8.3 | 3.55 | 9 | Ohio | 8 | 2.49 | 15 |
| Michigan | 8.0 | 4.69 | 5 | California | 7.5 | 11.52 | 1 |
| Illinois | 7.0 | 5.08 | 4 | Michigan | 7.5 | 3.32 | 7 |
| Oklahoma | 6.3 | 0.64 | 33 | New York | 7 | 8.05 | 2 |
| Texas | 5.8 | 3.69 | 7 | Florida | 7 | 3.11 | 9 |
| North Carolina | 5.7 | 3.09 | 11 | Wisconsin | 6.5 | 5.88 | 4 |
| Washington | 5.3 | 4.30 | 6 | Colorado | 5.5 | 1.79 | 21 |
| Utah | 5.3 | 0.95 | 29 | Pennsylvania | 5 | 2.81 | 10 |
| Wisconsin | 4.0 | 6.73 | 3 | Maryland | 5 | 2.21 | 18 |
| Virginia | 4.0 | 1.82 | 20 | Texas | 4.8 | 5.93 | 3 |
| Florida | 3.3 | 3.66 | 8 | Missouri | 4.5 | 2.68 | 12 |
| Oregon | 3.3 | 3.37 | 10 | Kansas | 4.5 | 1.42 | 23 |
| Missouri | 3.3 | 2.90 | 12 | Tennessee | 3.5 | 1.43 | 22 |
| Hawaii | 3.3 | 0.25 | 45 | Virginia | 3.5 | 1.08 | 28 |
| Georgia | 3.0 | 1.96 | 19 | Rhode Island | 3.5 | 0.35 | 49 |
| Kentucky | 3.0 | 1.52 | 23 | Washington | 3 | 3.17 | 8 |
| Colorado | 3.0 | 0.99 | 28 | Nebraska | 3 | 1.15 | 25 |
| Rhode Island | 3.0 | 0.37 | 42 | Oklahoma | 3 | 0.98 | 30 |
| Tennessee | 2.7 | 0.58 | 34 | Utah | 3 | 0.83 | 34 |
| Mississippi | 2.7 | 0.54 | 35 | Hawaii | 3 | 0.48 | 44 |
| Pennsylvania | 2.3 | 2.32 | 16 | Alabama | 2.5 | 1.11 | 26 |
| Nebraska | 2.3 | 1.21 | 25 | South Carolina | 2 | 0.84 | 32 |
| South Carolina | 2.3 | 0.51 | 36 | Kentucky | 1.5 | 1.11 | 27 |
| Alabama | 2.0 | 0.79 | 31 | Connecticut | 1.5 | 0.94 | 31 |
| Massachusetts | 1.0 | 2.90 | 13 | Georgia | 1 | 2.71 | 11 |
| Maryland | 1.0 | 2.02 | 18 | Oregon | 1 | 2.05 | 20 |
| Connecticut | 1.0 | 1.14 | 26 | Mississippi | 1 | 0.70 | 38 |
| Kansas | 0.7 | 1.58 | 22 | Massachusetts | 0.5 | 2.22 | 17 |

[a] It was assumed that it takes 4 years to complete a Ph.D. The competitive grant numbers are from Norton et al. and are 3-year centered moving averages. The 1990 student numbers are the averages for 1989, 1990, and 1991, a centered moving average. The 1996 student numbers are the 2-year average between 1996 and 1997, because the Schrimper data stop in 1997.

Often a field or specialty area is defined as a set of designated courses, or, as an alternative interpretation, a student may define his own field with the only requirement being that a set number of courses within a group of courses must be taken. Because of the cost economies alluded to above, most universities have a minimum student number requirement for a course to be taught. In an informal survey of about 10 universities, we found that the most common minimum requirement is about five students, although some universities leave that decision to the department. Regardless of whether or not there is a formal required class size minimum, at some point the economics of the class size becomes a pertinent departmental issue.

Core courses are often taken in other departments (e.g., economics), with additional students from those departments helping to easily surpass the required student minimum. The required student minimum is more of a problem in field courses taken only by agricultural economics students. As a consequence, the intended focus of the following analysis is on field courses.

The ultimate question of interest is how many students will take a given course under different program structures, different student numbers, and different probabilities of taking a field? Before presenting the general approach to the problem, consider a simple example. Suppose a department has three courses from which fields can be defined: $c_1$, an advanced econometrics course; $c_2$, a demand theory course; and $c_3$, an industrial organization theory course. Let the set of courses be defined as $\mathfrak{C} = \{c_1, c_2, c_3\}$. Assume the structure of the program is such that two courses are required for a field, so there are three possible fields: $\mathcal{F} = \{\{c_1, c_2\}, \{c_1, c_3\}\{c_2, c_3\}\} = \{f_1, f_2, f_3\}$, where each field represents a unique set or $f_1 = \{c_1, c_2\}$, $f_2 = \{c_1, c_3\}$, and $f_3 = \{c_2, c_3\}$. The first field, $f_1$, could be defined as the empirical demand analysis field; the second field, $f_2$, as the empirical industrial organization field; and the third field, $f_3$, as the theoretical microeconomics field. Now suppose that the department has a strong reputation in the area of econometrics and industrial organization and the graduate coordinator believes that probabilities associated with each field being taken are $P(f_1) = .40$, $P(f_2) = .50$, and $P(f_3) = .10$. With 10 students each taking a field, the expected numbers of students in each field are $f_1$, $E(N_1) = 10 \times .40 = 4$; $f_2$, $E(N_2) = 10 \times .50 = 5$; and $f_3$, $E(N_3) = 10 \times .10 = 1$. Because the courses appear in more than one field, the expected numbers of students in a course are $c_1$, $E(n_1) = 4 + 5 = 9$; $c_2$, $E(n_2) = 4 + 1 = 5$; and $c_3$, $E(n_3) = 5 + 1 = 6$.

Of course, the above results depend on several key parameters and the results will vary as these parameters change.[5] As a consequence, it is important to understand the more general structure. Suppose there are $N$ students in a cohort and a department has $C$ courses from which fields can be defined. Now a field will be defined as a set of courses ($k$) to be taken out of the $C$ available courses. Placing no restrictions on the number of fields then the number of possible fields is the combination $\boldsymbol{F} = \binom{C}{k}$, or from $C$ courses choose $k$.[6] The total number of fields defines the event space. For each field in the event space, a subjective probability of the field being taken by a student $P_i$ is assigned. This then defines a multinomial distribution

$$(1) \qquad P(N_1, N_2, \ldots, N_F) = \frac{N!}{N_1!N_2! \cdots N_F!} \times P_1^{N_1}P_2^{N_2} \cdots P_F^{N_F},$$

which gives the probability that out of $N$ students, $N_1$ choose field 1, $N_2$ choose field 2, etc., and the subscript $\boldsymbol{F}$ indicates the number of fields.[7] Let $I = \{1, 2, 3, \ldots, C\}$ be the indexing set for the courses and $\mathbf{J} = \{\mathbf{1}, \mathbf{2} \ldots, \boldsymbol{F}\}$ be the indexing set for the fields. Note that these indexing sets imply that subsets of the $I$ indexing set define an element in the $\mathbf{J}$ indexing set (e.g., $f_1 = \{c_1, c_2\}$, so $\mathbf{1} = \{1, 2\}$). For a multinomial distribution, the expected number of students taking a field $\mathbf{j} \in \mathbf{J}$ is then

$$(2) \qquad E(N_j) = NP_j.$$

[5] These parameters are taken as being exogenously determined. For example, as was discussed in the earlier section, the number of students will depend on many factors, such as recruiting efforts, assistantship levels, and departmental reputation. However, by taking the number of students as exogenous, the analysis herein does not consider those factors that may affect student numbers but states only what is expected to happen *with a given number* of students.

[6] The conclusions discuss how the approach can be easily generalized to allow the student to choose more than one field from all possible fields and also how to restrict the number of fields to less than all possible fields.

[7] The multinomial distribution and its properties can be found in just about any mathematical statistics book. See, for example, Mendenhall, Scheaffer, and Wackerly, page 214.

The expected number of students in a course is then just the sum of the expected number of students in each field requiring that course or

$$(3) \qquad E(n_{i\in I}) = \sum_{\forall i\in \mathrm{j,j}=1}^{F} E(N_{\mathrm{j}}).$$

But simple substitution of Equation (2) into Equation (3) gives

$$(4) \qquad E(n_{i\in I}) = \sum_{\forall i\in \mathrm{j,j}=1}^{F} NP_{\mathrm{j}} = N \sum_{\forall i\in \mathrm{j,j}=1}^{F} P_{\mathrm{j}} = Np_i,$$

where $p_i = \Sigma_{\forall i\in \mathrm{j,j}=1}^{F} P_{\mathrm{j}}$ is the probability that a student takes a specific course. Thus, the probability of a specific course being taken can be recovered from the information about the structure of the program and the probability of a specific field being taken. However, note that although the rules of probability require that $\Sigma_{\mathrm{j}=1}^{F} P_{\mathrm{j}} = 1$, there is *no* such requirement that $\Sigma_{i=1}^{C} p_i = 1$. This result is due to the layering or overlapping structure of the sets involved in that a course will appear in more than field. In addition, although Equations (3) and (4) can be used to estimate the expected number of students for a specific field or course, there may also be interest in the average number of expected students in a field $\bar{N}$ and the average number of expected students in a course $\bar{n}$. Given the formulas above, these averages have rather simple forms

$$(5) \qquad \bar{N} = \frac{1}{F}\sum_{\mathrm{j}=1}^{F} E(N_{\mathrm{j}}) = \frac{1}{F}\sum_{\mathrm{j}=1}^{F} NP_{\mathrm{j}} = \frac{N}{F}$$

$$(6) \qquad \bar{n} = \frac{1}{C}\sum_{i=1}^{C} Np_i = \frac{N}{C}\sum_{i=1}^{C} p_i.$$

### *Comparative Statistics*

What can be said analytically about the general procedure summarized in Equations (2)–(6)? Some limited analytical insights can be obtained by considering the partial derivative of these equations with respect to some of the parameters. Equation (2) indicates that for each additional student, the expected number of students in a field will increase by $P_{\mathrm{j}}$. Equation (2) also indicates that for each additional unit increase in $P_{\mathrm{j}}$, the expected number of students in a field increases by $N$. Because $P_{\mathrm{j}} \in [0, 1]$ and $N \geq 1$, then a one-unit change in the probability of a field being taken has a larger impact on student numbers in a field than increasing the total number of students by one student. Similar results apply for Equation (4). Equations (5) and (6) indicate that increasing the number of students ($N$) by one will increase the average number of students in the fields and courses by $\boldsymbol{F}^{-1}$ and $1/C\ \Sigma_{i=1}^{C} p_i$, respectively. Equation (5) indicates that increasing the number of fields ($\boldsymbol{F}$) by one will decrease the average number of students in a field by $-N/\boldsymbol{F}^2$. However, a similar result *does not* necessarily hold for Equation (6) because as the number of classes ($C$) increases, the denominator and the summation term in the numerator in (6) will increase.

With respect to courses, stating that the results will change as $p_i$ changes is not very enlightening and just begs the question, what causes $p_i$ to increase within the present framework? By definition, $p_i = \Sigma_{\forall i\in \mathrm{j,j}=1}^{F} P_{\mathrm{j}}$ and anything that causes this sum to increase will cause $p_i$ to increase. It is true that for a fixed number of fields (**F**) with fixed probabilities ($P_{\mathrm{j}}$), increasing the number of required courses in a field ($k$) will increase the number of terms in the summation—$i$ is an element of more **j**—therefore, $p_i$ will increase. But beyond this, there are no clear signable analytical results. This is mainly because several of these parameters are jointly determined and also affect the summation term in a nonlinear manner. For example, if there are at least four courses, the number of course combinations (i.e., number of possible fields $\boldsymbol{F}$) will increase, reach a maximum, and then decline as the number of required courses in a field ($k$) increases. In addition, as the number of fields ($\boldsymbol{F}$) change, this will change the probability of a particular field ($P_{\mathrm{j}}$), but all do not have to decrease—only some more than others. Because of results such as these, we turn to a simulation analysis of some possible scenarios.

### *Parameter Settings*

The analytical structure given above depends on several key parameters: the number of students ($N$), the number of courses from which fields will be defined ($C$), the number of courses required for a field ($k$), and the probability that a field will be taken by a student ($P_j$). Obviously, different departments will have different values for these parameters, and as these parameters change, so too will the expected number of students in a course. Here some examples will be presented that are representative of some reasonable structures, but there is no claim that these results are exhaustive. Other parameter settings will lead to different results, and for those interested in other parameter settings, the program used to calculate the results is available at http://agecon.tamu.edu/faculty/gdavis/gdavis.htm.[8] The program is a simple spreadsheet program that is very user friendly and allows all the parameters of the general framework to be altered to any specification that is desirable.

For the representative or demonstration cases presented herein, the following parameter settings are considered. The number of students are allowed to range from 3 to 15 ($N = 3, 4, \ldots, 15$). The number of courses from which fields can be constructed are 3, 4, and 5 ($C = 3, 4, 5$). The number of required courses are 2, 3, and 4 ($k = 2, 3, 4$). By the combinatorics, the parameters $C$ and $k$ then define the number of possible fields ($\boldsymbol{F}$): ($C = 3$, $k = 2$, $\boldsymbol{F} = 3$), ($C = 4$, $k = 3$, $\boldsymbol{F} = 4$), ($C = 4$, $k = 2$, $\boldsymbol{F} = 6$), ($C = 5$, $k = 4$, $\boldsymbol{F} = 5$), ($C = 5$, $k = 3$, $\boldsymbol{F} = 10$), and ($C = 5$, $k = 2$, $\boldsymbol{F} = 10$).[9] Subjective probabilities are then assigned to each field ($P_j$), such that the probabilities of the fields within a specific program structure are somewhat normally distributed. Again, someone interested in other parameter settings is encouraged to download the spreadsheet and tailor the parameter settings to their preferences.

## Results

Tables 3–8 give the results. The courses defining the fields are reported in the top part of each table, along with the assigned probabilities for each field and then the implied probabilities for each course. The lower part of the tables show the number of students, the corresponding number of expected students in each field, the average number of students in a field, the expected number of students in a course, and the average number of students in a course.[10]

### *Some Representative Results and "What if" Questions*

Table 3 shows the results for a program in which there are three courses, with two courses required per field, or equivalently interpreted, from three courses the student must take two. Because of the higher probability associated with field two ($f_2$), naturally field two will have more students than field one ($f_1$) or field three ($f_3$). Also, note that for each additional student, the average number of students in a field increases by $\boldsymbol{F}^{-1} = 1/3 = .33$, as implied by Equation (5). In terms of courses, course one ($c_1$) and course three ($c_3$) constitute field two, and given that field two has a higher probability than the other fields, then courses one and three have more expected students than course two ($c_2$). Also, and as implied by Equation (6), for each additional student the average number of expected students in a course is $1/C \sum_{i=1}^{C} p_i = 2/3 = .67$. The other tables can be interpreted in a similar manner.

As has been argued, redesigning Ph.D. programs is administratively very costly, especially in terms of removing or adding new fields or courses, and one would like an idea of how successful a new program structure will be before it is implemented. The frame-

---

[8] Once at this web page, be sure to read the "readme" file first.

[9] Having 6 or 10 fields may seem high, but an equivalent way to interpret this structure is that the student defines his own field and then the only restriction is that the student must choose $k$ courses out of $C$.

[10] Obviously, the mathematics can lead to noninteger values for the number of students. Given that we are measuring physical presence and not mental presence, noninteger values technically are inappropriate, so to be conservative, one may want to round the numbers down to the closest integer.

**Table 3.** Three Courses–Choose Two Model

| Field | Content | Field Probability | Course | Course Probability |
|---|---|---|---|---|
| $f_1$ | $\{c_1, c_2\}$ | 0.25 | $c_1$ | .75 |
| $f_2$ | $\{c_1, c_3\}$ | 0.50 | $c_2$ | .50 |
| $f_3$ | $\{c_2, c_3\}$ | 0.25 | $c_3$ | .75 |
| Sum | | 1.00 | | 2.00 |

| Student Numbers | Expected Numbers in a Field | | | | Expected Numbers in a Course | | | |
|---|---|---|---|---|---|---|---|---|
| $N$ | $f_1$ | $f_2$ | $f_3$ | Average | $c_1$ | $c_2$ | $c_3$ | Average |
| 3 | 0.75 | 1.50 | 0.75 | 1.00 | 2.25 | 1.50 | 2.25 | 2.00 |
| 4 | 1.00 | 2.00 | 1.00 | 1.33 | 3.00 | 2.00 | 3.00 | 2.67 |
| 5 | 1.25 | 2.50 | 1.25 | 1.67 | 3.75 | 2.50 | 3.75 | 3.33 |
| 6 | 1.50 | 3.00 | 1.50 | 2.00 | 4.50 | 3.00 | 4.50 | 4.00 |
| 7 | 1.75 | 3.50 | 1.75 | 2.33 | 5.25 | 3.50 | 5.25 | 4.67 |
| 8 | 2.00 | 4.00 | 2.00 | 2.67 | 6.00 | 4.00 | 6.00 | 5.33 |
| 9 | 2.25 | 4.50 | 2.25 | 3.00 | 6.75 | 4.50 | 6.75 | 6.00 |
| 10 | 2.50 | 5.00 | 2.50 | 3.33 | 7.50 | 5.00 | 7.50 | 6.67 |
| 11 | 2.75 | 5.50 | 2.75 | 3.67 | 8.25 | 5.50 | 8.25 | 7.33 |
| 12 | 3.00 | 6.00 | 3.00 | 4.00 | 9.00 | 6.00 | 9.00 | 8.00 |
| 13 | 3.25 | 6.50 | 3.25 | 4.33 | 9.75 | 6.50 | 9.75 | 8.67 |
| 14 | 3.50 | 7.00 | 3.50 | 4.67 | 10.50 | 7.00 | 10.50 | 9.33 |
| 15 | 3.75 | 7.50 | 3.75 | 5.00 | 11.25 | 7.50 | 11.25 | 10.00 |

Note: Numbers in table may differ from those implied by formulas due to rounding.

work presented herein can be used to shed some informative light on several "what if"–type questions. For example, suppose a department has five courses, students must take two courses out of the five, and the present parameter settings apply—remember that you can select your own parameter settings in the downloadable program. Without any more restrictions on the program structure, this implies 10 possible fields. Suppose that the department wants to keep the existing program structure but wants to know how many students are needed for the average expected number of students in a course to be greater than five? Looking at the five–choose two program, Table 6 indicates that it would take at least 12 students on average in this rather flexible program to reach that average. In this program structure, course two ($c_2$) meets the minimum class size with about seven or eight students, but the other courses are in much worse shape.

Now suppose, because of resources, or other constraints, that the department realizes that it is simply not viable to attract more than eight Ph.D. students in a cohort. The question is now what program structure is best suited for eight students in order for the class size on average to meet the minimum of five? Rather than working with a fixed program structure and a variable number of students, this question just fixes the number of students and allows the program structure to vary. With the present parameter settings, Tables 3–8 indicate that there are three structures that may support this criterion: 5.33 students with a three courses–choose two program (Table 3), 6.00 students with a four courses–choose three program (Table 5), and 6.40 students with a five courses–choose four program (Table 8). If the administrator is willing to go down to four students on average per course, then the four courses–choose two option also becomes viable (Table 4). This provides a sample of the types of questions that can be addressed within this framework.

## Conclusions and Extensions

Although funding for Ph.D. programs has likely increased over the past two decades,

**Table 4.** Four Courses–Choose Two Model

| Field | Content | Field Probability | Course | Course Probability |
|---|---|---|---|---|
| $f_1$ | $\{c_1, c_2\}$ | 0.05 | $c_1$ | .50 |
| $f_2$ | $\{c_1, c_3\}$ | 0.10 | $c_2$ | .50 |
| $f_3$ | $\{c_1, c_4\}$ | 0.35 | $c_3$ | .50 |
| $f_4$ | $\{c_2, c_3\}$ | 0.35 | $c_4$ | .50 |
| $f_5$ | $\{c_2, c_4\}$ | 0.10 | | |
| $f_6$ | $\{c_3, c_4\}$ | 0.05 | | |
| Sum | | 1.00 | | 2.00 |

| Student Numbers | Expected Numbers in a Field | | | | | | | Expected Numbers in a Course | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ | Average | $c_1$ | $c_2$ | $c_3$ | $c_4$ | Average |
| 3 | 0.15 | 0.30 | 1.05 | 1.05 | 0.30 | 0.15 | **0.50** | 1.50 | 1.50 | 1.50 | 1.50 | **1.50** |
| 4 | 0.20 | 0.40 | 1.40 | 1.40 | 0.40 | 0.20 | **0.67** | 2.00 | 2.00 | 2.00 | 2.00 | **2.00** |
| 5 | 0.25 | 0.50 | 1.75 | 1.75 | 0.50 | 0.25 | **0.83** | 2.50 | 2.50 | 2.50 | 2.50 | **2.50** |
| 6 | 0.30 | 0.60 | 2.10 | 2.10 | 0.60 | 0.30 | **1.00** | 3.00 | 3.00 | 3.00 | 3.00 | **3.00** |
| 7 | 0.35 | 0.70 | 2.45 | 2.45 | 0.70 | 0.35 | **1.17** | 3.50 | 3.50 | 3.50 | 3.50 | **3.50** |
| 8 | 0.40 | 0.80 | 2.80 | 2.80 | 0.80 | 0.40 | **1.33** | 4.00 | 4.00 | 4.00 | 4.00 | **4.00** |
| 9 | 0.45 | 0.90 | 3.15 | 3.15 | 0.90 | 0.45 | **1.50** | 4.50 | 4.50 | 4.50 | 4.50 | **4.50** |
| 10 | 0.50 | 1.00 | 3.50 | 3.50 | 1.00 | 0.50 | **1.67** | 5.00 | 5.00 | 5.00 | 5.00 | **5.00** |
| 11 | 0.55 | 1.10 | 3.85 | 3.85 | 1.10 | 0.55 | **1.83** | 5.50 | 5.50 | 5.50 | 5.50 | **5.50** |
| 12 | 0.60 | 1.20 | 4.20 | 4.20 | 1.20 | 0.60 | **2.00** | 6.00 | 6.00 | 6.00 | 6.00 | **6.00** |
| 13 | 0.65 | 1.30 | 4.55 | 4.55 | 1.30 | 0.65 | **2.17** | 6.50 | 6.50 | 6.50 | 6.50 | **6.50** |
| 14 | 0.70 | 1.40 | 4.90 | 4.90 | 1.40 | 0.70 | **2.33** | 7.00 | 7.00 | 7.00 | 7.00 | **7.00** |
| 15 | 0.75 | 1.50 | 5.25 | 5.25 | 1.50 | 0.75 | **2.50** | 7.50 | 7.50 | 7.50 | 7.50 | **7.50** |

Note: Numbers in table may differ from those implied by formulas due to rounding.

the source of the funds has shifted from noncompetitive to more competitive funds, which has apparently affected the distribution of funds across states and therefore departments. The shifting distribution of funds has likely increased the amount of stress many departments face concerning the number of Ph.D. students and number of field offerings and requirements. The purpose of this article was to provide an analytical approach and program that may be useful for departmental debates and decisions about Ph.D. programs.

The problem was cast in a simple but flexible combinatorial framework consisting of a few key parameters: the number of students in the cohort ($N$), the number of courses ($C$), the number of required courses to make a field ($k$), and the subjective probabilities associated with a field being taken by students ($P_j$). From this information, the expected number of students within each field and course can be calculated, along with the average number of expected students in a field and course. A few analytical generalizations do emerge from the analysis, some of which are obvious, some of which are not.

- Increasing the number of required courses will increase the average number of students in a course, *ceteris paribus.*
- Increasing the number of fields by one decreases the average number of students in a field by the number of students in the cohort divided by the number of fields squared, *regardless of the probabilities assigned to the fields, ceteris paribus.*
- Increasing the number of students in the cohort by one increases the average number of students in a field by the fraction one over the number of fields, *regardless*

**Table 5.** Four Courses–Choose Three Model

| Field | Content | Field Probability | Course | Course Probability |
|---|---|---|---|---|
| $f_1$ | $\{c_1, c_2, c_3\}$ | 0.10 | $c_1$ | .90 |
| $f_2$ | $\{c_1, c_3, c_4\}$ | 0.40 | $c_2$ | .60 |
| $f_3$ | $\{c_1, c_2, c_4\}$ | 0.45 | $c_3$ | .60 |
| $f_4$ | $\{c_2, c_3, c_4\}$ | 0.10 | $c_4$ | .90 |
| Sum | | 1.00 | | 3.00 |

| Student Numbers | Expected Numbers in a Field | | | | | Expected Numbers in a Course | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | $f_1$ | $f_2$ | $f_3$ | $f_4$ | Average | $c_1$ | $c_2$ | $c_3$ | $c_4$ | Average |
| 3 | 0.30 | 1.20 | 1.20 | 0.30 | **0.75** | 2.70 | 1.80 | 1.80 | 2.70 | **2.25** |
| 4 | 0.40 | 1.60 | 1.60 | 0.40 | **1.00** | 3.60 | 2.40 | 2.40 | 3.60 | **3.00** |
| 5 | 0.50 | 2.00 | 2.00 | 0.50 | **1.25** | 4.50 | 3.00 | 3.00 | 4.50 | **3.75** |
| 6 | 0.60 | 2.40 | 2.40 | 0.60 | **1.50** | 5.40 | 3.60 | 3.60 | 5.40 | **4.50** |
| 7 | 0.70 | 2.80 | 2.80 | 0.70 | **1.75** | 6.30 | 4.20 | 4.20 | 6.30 | **5.25** |
| 8 | 0.80 | 3.20 | 3.20 | 0.80 | **2.00** | 7.20 | 4.80 | 4.80 | 7.20 | **6.00** |
| 9 | 0.90 | 3.60 | 3.60 | 0.90 | **2.25** | 8.10 | 5.40 | 5.40 | 8.10 | **6.75** |
| 10 | 1.00 | 4.00 | 4.00 | 1.00 | **2.50** | 9.00 | 6.00 | 6.00 | 9.00 | **7.50** |
| 11 | 1.10 | 4.40 | 4.40 | 1.10 | **2.75** | 9.90 | 6.60 | 6.60 | 9.90 | **8.25** |
| 12 | 1.20 | 4.80 | 4.80 | 1.20 | **3.00** | 10.80 | 7.20 | 7.20 | 10.80 | **9.00** |
| 13 | 1.30 | 5.20 | 5.20 | 1.30 | **3.25** | 11.70 | 7.80 | 7.80 | 11.70 | **9.75** |
| 14 | 1.40 | 5.60 | 5.60 | 1.40 | **3.50** | 12.60 | 8.40 | 8.40 | 12.60 | **10.50** |
| 15 | 1.50 | 6.00 | 6.00 | 1.50 | **3.75** | 13.50 | 9.00 | 9.00 | 13.50 | **11.25** |

Note: Numbers in table may differ from those implied by formulas due to rounding.

*of the probabilities assigned to the fields, ceteris paribus.*

- Increasing the number of students in the cohort by one increases the average number of students in a course by the sum of the probabilities over all classes—which is not required to be one—divided by the number of classes, *ceteris paribus.*
- Increasing the probability that a field or course is taken by one unit has a larger impact on increasing the number of students in a field or course than increasing the number of students in the cohort by one unit, *ceteris paribus.*

Other comparative static results are ambiguous and will depend on the specific parameter settings.

The analysis presented herein is much more flexible than it may first appear and can be easily extended if so desired. For example, there may be concern that the procedure is limited in that only one field is chosen or the number of possible fields is unrestricted. Adding the requirement of more than one field would just add another layer to the problem, but the procedure would be the same. More specifically, one would start as before and determine the number of possible fields from the number of classes and the specific structure of the combinatorial problem. Once the number of fields is determined, then a second-level combinatorial problem would be defined wherein the student would choose a specific number of fields from the total number of fields available, and this second combinatorial problem would define the new event space. Subjective probabilities would then be assigned to the field combinations, and one would then work backward to determine the probabilities of specific fields and classes making and the number of students in specific fields and classes. With respect to the number of pos-

**Table 6.** Five Courses–Choose Two Model

| Field | Content | Field Probability | Course | Course Probability |
|---|---|---|---|---|
| $f_1$ | $\{c_1, c_2\}$ | 0.00125 | $c_1$ | .23 |
| $f_2$ | $\{c_1, c_3\}$ | 0.025 | $c_2$ | .70 |
| $f_3$ | $\{c_1, c_4\}$ | 0.05 | $c_3$ | .37 |
| $f_4$ | $\{c_1, c_5\}$ | 0.15 | $c_4$ | .38 |
| $f_5$ | $\{c_2, c_3\}$ | 0.27375 | $c_5$ | .33 |
| $f_6$ | $\{c_2, c_4\}$ | 0.27375 | | |
| $f_7$ | $\{c_2, c_5\}$ | 0.15 | | |
| $f_8$ | $\{c_3, c_4\}$ | 0.05 | | |
| $f_9$ | $\{c_3, c_5\}$ | 0.025 | | |
| $f_{10}$ | $\{c_4, c_5\}$ | 0.00125 | | |
| Sum | | 1.00 | | 2.00 |

| Student Numbers | Expected Numbers in a Field | | | | | | | | | | | Expected Numbers in a Course | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ | $f_7$ | $f_8$ | $f_9$ | $f_{10}$ | Average | $c_1$ | $c_2$ | $c_3$ | $c_4$ | $c_5$ | Average |
| 3 | 0.00 | 0.08 | 0.15 | 0.45 | 0.82 | 0.82 | 0.45 | 0.15 | 0.08 | 0.00 | **0.30** | 0.68 | 2.10 | 1.12 | 1.13 | 0.98 | **1.20** |
| 4 | 0.01 | 0.10 | 0.20 | 0.60 | 1.10 | 1.10 | 0.60 | 0.20 | 0.10 | 0.01 | **0.40** | 0.91 | 2.80 | 1.50 | 1.50 | 1.31 | **1.60** |
| 5 | 0.01 | 0.13 | 0.25 | 0.75 | 1.37 | 1.37 | 0.75 | 0.25 | 0.13 | 0.01 | **0.50** | 1.13 | 3.49 | 1.87 | 1.88 | 1.63 | **2.00** |
| 6 | 0.01 | 0.15 | 0.30 | 0.90 | 1.64 | 1.64 | 0.90 | 0.30 | 0.15 | 0.01 | **0.60** | 1.36 | 4.19 | 2.24 | 2.25 | 1.96 | **2.40** |
| 7 | 0.01 | 0.18 | 0.35 | 1.05 | 1.92 | 1.92 | 1.05 | 0.35 | 0.18 | 0.01 | **0.70** | 1.58 | 4.89 | 2.62 | 2.63 | 2.28 | **2.80** |
| 8 | 0.01 | 0.20 | 0.40 | 1.20 | 2.19 | 2.19 | 1.20 | 0.40 | 0.20 | 0.01 | **0.80** | 1.81 | 5.59 | 2.99 | 3.00 | 2.61 | **3.20** |
| 9 | 0.01 | 0.23 | 0.45 | 1.35 | 2.46 | 2.46 | 1.35 | 0.45 | 0.23 | 0.01 | **0.90** | 2.04 | 6.29 | 3.36 | 3.38 | 2.94 | **3.60** |
| 10 | 0.01 | 0.25 | 0.50 | 1.50 | 2.74 | 2.74 | 1.50 | 0.50 | 0.25 | 0.01 | **1.00** | 2.26 | 6.99 | 3.74 | 3.75 | 3.26 | **4.00** |
| 11 | 0.01 | 0.28 | 0.55 | 1.65 | 3.01 | 3.01 | 1.65 | 0.55 | 0.28 | 0.01 | **1.10** | 2.49 | 7.69 | 4.11 | 4.13 | 3.59 | **4.40** |
| 12 | 0.02 | 0.30 | 0.60 | 1.80 | 3.29 | 3.29 | 1.80 | 0.60 | 0.30 | 0.02 | **1.20** | 2.72 | 8.39 | 4.49 | 4.50 | 3.92 | **4.80** |
| 13 | 0.02 | 0.33 | 0.65 | 1.95 | 3.56 | 3.56 | 1.95 | 0.65 | 0.33 | 0.02 | **1.30** | 2.94 | 9.08 | 4.86 | 4.88 | 4.24 | **5.20** |
| 14 | 0.02 | 0.35 | 0.70 | 2.10 | 3.83 | 3.83 | 2.10 | 0.70 | 0.35 | 0.02 | **1.40** | 3.17 | 9.78 | 5.23 | 5.25 | 4.57 | **5.60** |
| 15 | 0.02 | 0.38 | 0.75 | 2.25 | 4.11 | 4.11 | 2.25 | 0.75 | 0.38 | 0.02 | **1.50** | 3.39 | 10.48 | 5.61 | 5.63 | 4.89 | **6.00** |

Note: Numbers in table may differ from those implied by formulas due to rounding.

**Table 7.** Five Courses–Choose Three Model

| Field | Content | Field Probability | Course | Course Probability |
|---|---|---|---|---|
| $f_1$ | $\{c_1, c_2, c_3\}$ | 0.00125 | $c_1$ | 0.77 |
| $f_2$ | $\{c_1, c_2, c_4\}$ | 0.025 | $c_2$ | 0.30 |
| $f_3$ | $\{c_1, c_2, c_5\}$ | 0.05 | $c_3$ | 0.63 |
| $f_4$ | $\{c_1, c_3, c_4\}$ | 0.15 | $c_4$ | 0.63 |
| $f_5$ | $\{c_1, c_3, c_5\}$ | 0.27375 | $c_5$ | 0.67 |
| $f_6$ | $\{c_1, c_4, c_5\}$ | 0.27375 | | |
| $f_7$ | $\{c_2, c_3, c_4\}$ | 0.15 | | |
| $f_8$ | $\{c_2, c_3, c_5\}$ | 0.05 | | |
| $f_9$ | $\{c_2, c_4, c_5\}$ | 0.025 | | |
| $f_{10}$ | $\{c_3, c_4, c_5\}$ | 0.00125 | | |
| Sum | | 1.00 | | 3.00 |

| Student Numbers | Expected Numbers in a Field | | | | | | | | | | | Expected Numbers in a Course | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ | $f_7$ | $f_8$ | $f_9$ | $f_{10}$ | Average | $c_1$ | $c_2$ | $c_3$ | $c_4$ | $c_5$ | Average |
| 3 | 0.00 | 0.08 | 0.15 | 0.45 | 0.82 | 0.82 | 0.45 | 0.15 | 0.08 | 0.00 | **0.30** | 2.32 | 0.90 | 1.88 | 1.88 | 2.02 | **1.80** |
| 4 | 0.01 | 0.10 | 0.20 | 0.60 | 1.10 | 1.10 | 0.60 | 0.20 | 0.10 | 0.01 | **0.40** | 3.10 | 1.21 | 2.51 | 2.50 | 2.70 | **2.40** |
| 5 | 0.01 | 0.13 | 0.25 | 0.75 | 1.37 | 1.37 | 0.75 | 0.25 | 0.13 | 0.01 | **0.50** | 3.87 | 1.51 | 3.13 | 3.13 | 3.37 | **3.00** |
| 6 | 0.01 | 0.15 | 0.30 | 0.90 | 1.64 | 1.64 | 0.90 | 0.30 | 0.15 | 0.01 | **0.60** | 4.64 | 1.81 | 3.76 | 3.75 | 4.04 | **3.60** |
| 7 | 0.01 | 0.18 | 0.35 | 1.05 | 1.92 | 1.92 | 1.05 | 0.35 | 0.18 | 0.01 | **0.70** | 5.42 | 2.11 | 4.38 | 4.38 | 4.72 | **4.20** |
| 8 | 0.01 | 0.20 | 0.40 | 1.20 | 2.19 | 2.19 | 1.20 | 0.40 | 0.20 | 0.01 | **0.80** | 6.19 | 2.41 | 5.01 | 5.00 | 5.39 | **4.80** |
| 9 | 0.01 | 0.23 | 0.45 | 1.35 | 2.46 | 2.46 | 1.35 | 0.45 | 0.23 | 0.01 | **0.90** | 6.96 | 2.71 | 5.64 | 5.63 | 6.06 | **5.40** |
| 10 | 0.01 | 0.25 | 0.50 | 1.50 | 2.74 | 2.74 | 1.50 | 0.50 | 0.25 | 0.01 | **1.00** | 7.74 | 3.01 | 6.26 | 6.25 | 6.74 | **6.00** |
| 11 | 0.01 | 0.28 | 0.55 | 1.65 | 3.01 | 3.01 | 1.65 | 0.55 | 0.28 | 0.01 | **1.10** | 8.51 | 3.31 | 6.89 | 6.88 | 7.41 | **6.66** |
| 12 | 0.02 | 0.30 | 0.60 | 1.80 | 3.29 | 3.29 | 1.80 | 0.60 | 0.30 | 0.02 | **1.20** | 9.29 | 3.62 | 7.52 | 7.50 | 8.09 | **7.20** |
| 13 | 0.02 | 0.33 | 0.65 | 1.95 | 3.56 | 3.56 | 1.95 | 0.65 | 0.33 | 0.02 | **1.30** | 10.06 | 3.92 | 8.14 | 8.13 | 8.76 | **7.80** |
| 14 | 0.02 | 0.35 | 0.70 | 2.10 | 3.83 | 3.83 | 2.10 | 0.70 | 0.35 | 0.02 | **1.40** | 10.83 | 4.22 | 8.77 | 8.75 | 9.43 | **8.40** |
| 15 | 0.02 | 0.38 | 0.75 | 2.25 | 4.11 | 4.11 | 2.25 | 0.75 | 0.38 | 0.02 | **1.50** | 11.61 | 4.52 | 9.39 | 9.38 | 10.11 | **9.00** |

Note: Numbers in table may differ from those implied by formulas due to rounding.

**Table 8.** Five Courses–Choose Four Model

| Field | Content | Field Probability | Course | Course Probability |
|---|---|---|---|---|
| $f_1$ | $\{c_1, c_2, c_3, c_4\}$ | 0.025 | $c_1$ | 0.35 |
| $f_2$ | $\{c_1, c_2, c_3, c_5\}$ | 0.15 | $c_2$ | 0.85 |
| $f_3$ | $\{c_2, c_3, c_4, c_5\}$ | 0.65 | $c_3$ | 0.98 |
| $f_4$ | $\{c_1, c_3, c_4, c_5\}$ | 0.15 | $c_4$ | 0.85 |
| $f_5$ | $\{c_1, c_2, c_4, c_5\}$ | 0.025 | $c_5$ | 0.98 |
| Sum | | 1.00 | | 4.01 |

| Student Numbers | Expected Numbers in a Field | | | | | | Expected Numbers in a Course | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | Average | $c_1$ | $c_2$ | $c_3$ | $c_4$ | $c_5$ | Average |
| 3 | 0.08 | 0.45 | 1.95 | 0.45 | 0.08 | **0.60** | 1.05 | 2.55 | 2.93 | 2.55 | 2.93 | **2.40** |
| 4 | 0.10 | 0.60 | 2.60 | 0.60 | 0.10 | **0.80** | 1.40 | 3.40 | 3.90 | 3.40 | 3.90 | **3.20** |
| 5 | 0.13 | 0.75 | 3.25 | 0.75 | 0.13 | **1.00** | 1.75 | 4.25 | 4.88 | 4.25 | 4.88 | **4.00** |
| 6 | 0.15 | 0.90 | 3.90 | 0.90 | 0.15 | **1.20** | 2.10 | 5.10 | 5.85 | 5.10 | 5.85 | **4.80** |
| 7 | 0.18 | 1.05 | 4.55 | 1.05 | 0.18 | **1.40** | 2.45 | 5.95 | 6.83 | 5.95 | 6.83 | **5.60** |
| 8 | 0.20 | 1.20 | 5.20 | 1.20 | 0.20 | **1.60** | 2.80 | 6.80 | 7.80 | 6.80 | 7.80 | **6.40** |
| 9 | 0.23 | 1.35 | 5.85 | 1.35 | 0.23 | **1.80** | 3.15 | 7.65 | 8.78 | 7.65 | 8.78 | **7.20** |
| 10 | 0.25 | 1.50 | 6.50 | 1.50 | 0.25 | **2.00** | 3.50 | 8.50 | 9.75 | 8.50 | 9.75 | **8.00** |
| 11 | 0.28 | 1.65 | 7.15 | 1.65 | 0.28 | **2.20** | 3.85 | 9.35 | 10.73 | 9.35 | 10.73 | **8.80** |
| 12 | 0.30 | 1.80 | 7.80 | 1.80 | 0.30 | **2.40** | 4.20 | 10.20 | 11.70 | 10.20 | 11.70 | **9.60** |
| 13 | 0.33 | 1.95 | 8.45 | 1.95 | 0.33 | **2.60** | 4.55 | 11.05 | 12.68 | 11.05 | 12.68 | **10.40** |
| 14 | 0.35 | 2.10 | 9.10 | 2.10 | 0.35 | **2.80** | 4.90 | 11.90 | 13.65 | 11.90 | 13.65 | **11.20** |
| 15 | 0.38 | 2.25 | 9.75 | 2.25 | 0.38 | **3.00** | 5.25 | 12.75 | 14.63 | 12.75 | 14.63 | **12.00** |

Note: Numbers in table may differ from those implied by formulas due to rounding.

sible fields being unrestricted and determined by the combinatorial solution, this is also easily handled. For example, the five courses–choose two structure generates 10 possible fields, but one may feel that this is too many fields for the number of courses, given that many fields only differ by one course. This is easily handled by just pruning out of the field set the fields one thinks are illegitimate and then assign subjective probabilities to the remaining fields. This simply reduces the number of available fields in the event space, but the event space would still have a multinomial distribution and one could proceed as described above. The possible program structures are really only limited by the imagination, and the above procedures will hopefully prove useful in exploring the likely outcomes of some of the alternative program structures imagined.

*[Received October 2001; Accepted May 2002.]*

## References

Alston, J.M., and P.G. Pardey. *Making Science Pay: The Economics of Agricultural R&D Policy.* Washington, DC: AEI Press, 1996.

———. "Agricultural Research: Benefits and Beneficiaries of Alternative Funding Mechanisms." *Review of Agricultural Economics* 21,1(Spring/Summer 1999):2–19.

Huffman, W.E., and R.E. Just. "Funding, Structure, and Management of Public Agricultural Research in the United States." *American Journal of Agricultural Economics* 76(November 1994): 744–59.

Huffman, W.E., and P. Orazem. "An Econometric Model of the Market for New Ph.D.s in Agricultural Economics in the United States." *American Journal of Agricultural Economics* 67(December 1985):1207–14.

Just, R.E., and W.E. Huffman. "Economic Principles and Incentives: Structure, Management, and Funding of Agricultural Research in the United States." *American Journal of Agricultural Economics* 74,5(December 1992):1101–8.

Mendenhall, W., R.L. Scheaffer, and D.D. Wackerly. *Mathematical Statistics with Applications*, 2nd ed. Boston: Duxbury Press, 1981.

Norton, V., D. Colyer, N.A. Norton, and L. Davis-Swing. "Issues and Trends in Agricultural and Agricultural Economics Research Funding." *American Journal of Agricultural Economics* 77(December 1995):1337–46.

Perry, G.M. "Research and Extension Expenditures Rising." *Choices*, 2nd quarter 2000, pp. 24–25.

Rubenstein, K.D., P.W. Heisey, C. Klotz-Ingram, and G.B. Frisvold. "Competitive Grants and the Funding of Agricultural Research in the U.S." Paper presented at the annual meetings of the American Agricultural Economics Association, Tampa, Florida, July 30–August 2, 2000.

Schrimper, R.A. "Trends and Characteristics of Ph.D. Degrees in Agricultural Economics in the United States." *American Journal of Agricultural Economics* 67,5(November 1985): 1200–6.

______. "Output and Employment Characteristics of Recent Ph.D.'s in Agricultural Economics in the South." Paper presented at the annual meetings of the Southern Agricultural Economics Association, Memphis, Tennessee, February 1, 1999.