J Prod Anal (2012) 38:95–107
DOI 10.1007/s11123-011-0259-6

Grade inflation or productivity growth? An analysis of changing grade distributions at a regional university

Rebecca Summary · William L. Weber

Published online: 23 December 2011

Abstract We examine departmental grade distributions for the school years 1985–1986, 1995–1996, and 2004–2005 at Southeast Missouri State University. Mean undergraduate grade point averages (GPAs) increased from 2.6 in 1985–1986 to 3.1 in 2004–2005. Although higher student GPAs might be evidence of grade inflation, university departments might have experienced productivity improvements that enhanced student learning, given inputs. We represent the technology by the directional distance function. Departments produce two outputs-grade points earned by students and the information content of those grades-using faculty and student inputs. The entropy index is used to proxy the information content of grade distributions. The estimates indicate; no systematic changes in inefficiency over time; a movement along the production frontier toward a mix of outputs with relative more grade points and less entropy; a shift toward non-tenure track faculty that increases the shadow price of entropy relative to grade points.

Keywords Efficiency · Productivity · Grade inflation

JEL Classification D24 · I23

Rising average grades assigned by university professors have been attributed to grade inflation by some (Mansfield 2001; Johnson 2003), but have been defended by others (Kohn 2002) as an outcome of good teaching, a more accurate teacher assessment of student abilities, and a reduced higher educational focus on sorting students. Of course, the classic paper by Spence (1973) showed that students invest in human capital to sort, screen, and reveal their abilities to employers who face imperfect information on potential employees' marginal products. In an expansion of this idea, Stiglitz (1975) argues that the function of educational institutions is to impart knowledge and sort students into their areas of comparative advantage. However, Stiglitz also writes that "(t)he school system can decide on the fineness or coarseness of screening." (p. 294) Although a coarser screen need not entail grade inflation, rising average grades might be part of an implicit decision on the part of university faculties to have a coarser screen.

In this paper we examine the changing grade distributions of university departments at a regional public university, Southeast Missouri State University, from 1985 to 2005. With a four point grading scale the highest grade of A earns 4 points, a grade of B earns 3 points, a grade of C earns 2 points, a grade of D earns 1 point and a failing grade earns 0 points. The number of grade points a student can earn equals the sum over all classes of grade points multiplied by credit hours. A student's GPA (grade point average) equals total grade points divided by total credit hours. At Southeast Missouri State University GPAs increased from 2.6 in 1985–1986 to 3.1 in 2004–2005. We build a model where each department combines student and faculty inputs to produce a grade distribution. We proxy the knowledge component of the education process by the number of grade points students receive in the department and we use Shannon's (1948) entropy index to proxy the fineness of the screen provided by education. If students have varying interests and abilities a higher value

R. Summary · W. L. Weber (✉)
Department of Economics and Finance, Southeast MO State University, One University Plaza, Cape Girardeau, MO 63701, USA
e-mail: wlweber@semo.edu

R. Summary
e-mail: rsummary@semo.edu

of the entropy index provides information to students on their comparative advantage, information to potential employers who might find it less costly to identify high ability students, and information to graduate schools in determining admission and scholarships. We estimate a stochastic directional output distance function to measure department inefficiency for the school years of 1985–1986, 1995–1996, and 2004–2005.

After lackluster labor productivity growth in the late 1970s and 1980s, labor productivity in the US business sector grew at a relatively rapid rate throughout the 1990s until the present. (2005 Economic Report of the President Table B-50) If universities have experienced similar rates of productivity growth in educating students, that productivity growth might be reflected in rising student GPAs. However, as Baumol (1967) eloquently argues, there are some services, teaching for one, "in which labor is an end in itself, in which quality is judged directly in terms of amount of labor." (p. 416) Baumol's theory of unbalanced productivity growth between two sectors predicts that if the demand for the service produced by the slow growth sector is price inelastic and income elastic, a cost disease occurs, where the slow growth sector absorbs an increasing share of income. According to the College Board and reported by Kleiner (2004), tuition at 4 years public schools grew by 308% between 1984 and 2004 while real per capita income grew by only 138%. (2005 Economic Report of the President Table B-31) Are these figures for higher education evidence of Baumol's cost disease, or, are the higher expenditures consistent with enhanced outcomes? Universities might have been able to raise productivity with departments producing more output with less input due to technological progress or greater efficiency.

The directional output distance function is an effective way of modeling a multi-output, multi-input production technology and allows us to address several important questions: What is the level of departmental inefficiency? Are higher observed grades in 2004–2005 the result of productivity growth or evidence of grade inflation? Has the frontier tradeoff between grades points earned and the information content of those grades changed during the 20 years period? What effect does achieving program accreditation have on department efficiency? The last two decades have seen many universities attempt to control costs by substituting non-tenure track faculty for tenure track faculty. Is the changing mix of faculty inputs responsible for rising grades? In the next section we review the literature on the causes and consequences of grade inflation. We present our model of the educational production process in Sect. 2. In Sect. 3 we discuss the data and present our estimates of efficiency and productivity growth for university departments. The final section offers a summary of our work and its policy implications.

1 Grade inflation

A number of factors have been cited as causing grade inflation. Declining state appropriations for higher education has meant higher tuition. As consumers of higher education, students and parents might expect higher grades as compensation for paying higher tuition. To alleviate higher costs, many universities have increased their use of part-time faculty in the classroom. These faculty members tend to be evaluated primarily on their teaching and thus might have an incentive to try and boost student ratings via higher grades (Capozza 1973; Zangenehzadeh 1988). Johnson (2003) finds that grading differences across disciplines divert students and resources away from hard grading disciplines toward easy grading ones. Based on 17,000 matched comparisons of the same student in two different courses at Duke University, Johnson finds that chemistry, physics, mathematics, biology, and economics were the departments most penalized as students were twice as likely to take a course from a professor who assigned a mean grade of A- versus B. Sabot and Wakeman-Linn (1991) identify hard grading and easy grading disciplines at Williams College and seven other colleges and find that male students were 18.2% less likely to take a second course in economics if they received a B rather than an A in the introductory course. Bar et al. (2009) provide further evidence of students selecting easy grading courses. In 1998, Cornell University began an online median grade report for each class to provide students with a more accurate idea of their performance. In an analysis of this policy change these researchers find that while the share of courses offering a median grade in the A range was relatively stable, the share of students enrolled in those courses increased after 1998.

At most universities the cost per credit hour is the same for each course, with certain add-ons for laboratory and technology fees. Freeman (1999) argues that "given equal money prices per credit hour across disciplines, departments manage their enrollments by 'pricing' their courses with grading standards commensurate with the market benefits of their courses, as measured by expected incomes." (p. 344) Using data from the National Center for Education Statistics on 10,800 students from 648 institutions and 59 fields of study, Freeman finds evidence supporting his hypothesis: disciplines where students expect to earn higher incomes have lower GPAs.

When inflationary expectations are fully realized and relative prices remain constant, there are no costs of price inflation. However, most price inflations are not fully anticipated and are accompanied by changing relative prices between goods which redistribute income and reallocate resources. The same might be true for grade inflation. A grading scale that assigns five ranks of A through F to student grades is no better or worse than a five rank scale of A, A−, B+, B, and B− as long as all users of the grade distribution understand the equivalence of the two ranking schemes. However, while prices can rise without limit, a given grading scale has an upper bound so that grade inflation compresses grades and lowers the information content of those grades. Achen and Courant (2009) find evidence of grade inflation for disciplines in the University of Michigan's College of Literature, Science, and the Arts from 1992 to 2008, but very little change in the relative rankings of hard grading and easy grading disciplines.

There is growing evidence that average grades at universities have been increasing by about 0.15 grade points per decade for the last 35 years (GradeInflation.com 2005). Moreover, as Johnson (2003) finds for Duke University, grade inflation has also been accompanied by changes in the exchange rate or relative difficulty of courses across disciplines. In a signaling model, Chan et al. (2007) find that grade inflation can arise when employers are unable to distinguish between schools that have good students and schools that give easy grades. Since elite schools are more likely to have good students, employers treat grade inflation by these schools with less skepticism. However, the authors also show that when one school can exaggerate grades, other schools find it easier to inflate grades in order to fool the market. Thus, grade inflation becomes contagious and the information content of grades is reduced in a tragedy of the commons. Babcock (2010) finds that students study 50% less when taking a class where the expected grade is A versus a class where the expected grade is C. To the extent that university classes add to a student's human capital and are not just a method of signaling, a consequence of reduced study time will be reduced productivity. Given that the university we study is not an elite institution; the effects of grade inflation may be borne by high ability students as grades are compressed. Furthermore, the findings of Babcock suggest that other lower ability students will acquire less human capital.

To counter grade inflation Johnson proposes that additional information be provided on college transcripts about the average GPA per course so that students who take hard courses and earn less than an A grade can have their grade compared to the average grade received by the class. However, Ostrovsky and Schwarz (2003) argue that if universities can choose how informative to make transcripts, some universities might find it optimal to lump high-performing and low-performing students together with a coarser grading system. In their model, if universities disclose greater information an "unraveling" occurs where students and employers engage in early contracting. In addition, to the extent that some jobs entail high rents, informative transcripts that signal high ability students might reduce welfare.

2 Method

Because of the importance of scientific knowledge to productivity growth found by Adams (1990), various researchers have examined university efficiency in the production of knowledge outputs. In these studies the typical teaching outputs include full-time undergraduate and graduate students and research outputs include research grant dollars, number of publications, and ranks of research quality. Cost functions have been used to examine scale and scope economies by Cohn et al. (1989) for 1,887 US universities operating in 1981–1982, by deGroot et al. (1991) for 147 doctoral granting universities in 1982–1983, and by Glass et al. (1995) for 61 UK universities in 1989. In general, these studies find scale economies in the production of teaching and research and some scope economies for US, but not UK universities. Glass et al. (1998) use data envelopment analysis to estimate cost indirect distance functions and productivity indexes for UK universities during 1989–1992. During the study period, government policy aimed to improve productivity in teaching and research, but productivity declined by 4%, primarily due to biased technological change with the production frontier shifting out for teaching outputs, but inward for research outputs.

Although the traditional teaching outputs are full-time students, we model university departments as producers of a grade distribution to examine the knowledge output and the information screening output. The knowledge component of education is reflected in the number of grade points generated by the department. We measure the screening output or information content of those grade points by Shannon's entropy index. The number of departmental grade points equals the product of average departmental GPA, students, and class hours. For instance, a department that has a GPA of 3.0 on a 4 point scale with 500 students each completing three hours of credit would generate 4,500 grade points. Each department uses student and faculty inputs to produce the two outputs. Let the N inputs of department k be represented by $x^k = (x_1^k, \ldots, x_N^k)$ and let the M outputs of department k be represented by $y^k = (y_1^k, \ldots, y_M^k)$. The $k = 1, \ldots, K$ departments face a common technology, represented by the production possibility set, $P(x)$, where $P(x) = \{y : y \text{ can be produced by } x\}$. For our purpose we have $M = 2$ where $y_1 =$ departmental grade points and $y_2 =$ Shannon's entropy index. Student grades are assigned into one of five categories, A, B, C, D, and F and the proportion of grades in each category is w_i. Shannon's entropy index takes the form: $E = -\sum_{i=1}^{5} w_i \ln(w_i)$. When all students receive the same grade, the information content of the grade distribution is low and $E = 0$. The entropy index takes its maximum value when equal proportions of students are in each of the five categories: $E = -5 \cdot (0.2)(\ln(0.2)) = 1.61$. To our knowledge the entropy index has not been used as an output in studies examining the efficiency of educational institutions. We assume that the entropy index is a proxy indicator for group heterogeneity and the screening function of universities. While screening by educational institutions may increase income inequality, Stiglitz (1975) writes that "attempts to curtail educational screening may simply shift the focus of screening (for example, to on-the-job screening), with the possibility of lowering net national output without any commensurate gain in equality." (p. 299)

The properties of the production possibility set (Färe and Primont (1995) are:

$$
\begin{aligned}
&i. \quad 0_M \in P(x) \quad \text{for all } x \in R^N_+ \\
&ii. \quad P(0) = 0_M \\
&iii. \quad \text{if } x' \geq x, \quad \text{then } P(x') \supseteq P(x) \\
&iv. \quad \text{if } y \in P(x) \quad \text{then } \theta y \in P(x) \text{ for } 0 \leq \theta \leq 1 \\
&v. \quad \text{for all } x \in R^N_+, \quad P(x) \text{ is a closed and bounded set.}
\end{aligned} \tag{1}
$$

In order, these properties mean that inaction is possible; no output can be produced if no input is available; inputs are strongly disposable; output is weakly disposable; and finite amounts of input can only produce finite amounts of output. An additional assumption that is sometimes imposed is strong disposability of outputs which means that if $y \in P(x)$ and $y' \leq y$, then $y' \in P(x)$. Although $P(x)$ set might satisfy strong disposability of outputs, we do not require this assumption for our analysis.

We use the directional output distance function as a functional representation of the technology and as a measure of departmental inefficiency. For the directional vector $g = (g_1, \ldots, g_M)$ the directional output distance function seeks the maximum simultaneous expansion of all outputs for the g-directional vector. This distance function takes the form:

$$\vec{D}_o(x, y; g) = \max\{\beta : (y + \beta g) \in P(x)\}. \tag{2}$$

When departments produce on the frontier of $P(x)$ they are efficient with $\vec{D}_o(x, y; g) = 0$. Inefficiency is indicated by $\vec{D}_o(x, y; g) > 0$. Each of the two outputs are scaled to the frontier along the directional vector $g = (g_1, g_2)$. Different directional vectors can be chosen. For the unit directional vector, $g = (1, 1)$, the directional output distance function gives the maximum simultaneous expansion in grade points and the entropy index that is feasible given inputs. For a directional vector $g = (1, 0)$ the distance function gives the maximum feasible expansion in grade points holding entropy constant. For a directional vector $g = (0, 1)$ the distance function gives the maximum feasible expansion in entropy holding grade points constant.

For a directional vector $g = (y_1, y_2)$ the distance function gives the simultaneous percentage expansion in the two outputs. In such a case, the directional output distance function can be recovered from the Shephard output distance function as

$$\vec{D}_o(x, y; y) = \frac{1}{D_o(x, y)} - 1, \tag{3}$$

where $D_o(x, y) = \min\{\lambda : \frac{y}{\lambda} \in P(x)\}$ is the Shephard output distance function. The Shephard output distance function can be estimated as a translog form.

The directional output distance function has been adapted for production theory by Chambers et al. (1996) who derived it from Luenberger's (1992) benefit function. The properties of the directional distance function are inherited from the output sets, $P(x)$, $x \in R^N_+$. One property is that the directional output distance function provides a complete functional characterization of the technology in that

$$y \in P(x) \text{ if and only if } \vec{D}_o(x, y; g) \geq 0. \tag{4}$$

The directional output distance function also has the translation property:

$$\vec{D}_o(x, y + \phi g; g) = \vec{D}_o(x, y; g) - \phi, \tag{5}$$

which implies that if outputs are scaled by ϕg, the directional output distance function (inefficiency) declines by ϕ. Strong disposability of inputs is represented by

$$\vec{D}_o(x', y; g) \geq \vec{D}_o(x, y; g) \quad \text{for } x' \geq x \tag{6}$$

and weak disposability of outputs is represented by

$$\text{if } \vec{D}_o(x, y; g) \geq 0 \text{ then } \vec{D}_o(x, \theta y; g) \geq 0 \quad \text{for } 0 \leq \theta \leq 1. \tag{7}$$

We use the directional output distance function to estimate the frontier tradeoff between the two departmental outputs of grade points and the entropy index. Assuming differentiability of $\vec{D}_o(x, y;\ g)$, let $\frac{dy_m}{dy_{m'}}$ indicate the frontier tradeoff between outputs m and m'. Taking the total differential of Eq. 2 and evaluating it on the frontier (that is, for $\vec{D}_o(x, y;\ g) = 0$) we obtain

$$d\vec{D}_o(x, y; g) = \sum_{n=1}^{N} \frac{\partial \vec{D}_o(x, y; g)}{\partial x_n} dx_n + \sum_{m=1}^{M} \frac{\partial \vec{D}_o(x, y; g)}{\partial y_m} dy_m = 0 \tag{8}$$

Holding $dx_n = 0$, $n = 1, \ldots, N$ so that $P(x)$ is constant, a rearrangement of (8) yields

$$\frac{dy_m}{dy_{m'}} = -\frac{\partial \vec{D}_o(x, y; g)/\partial y_{m'}}{\partial \vec{D}_o(x, y; g)/\partial y_m}, \tag{9}$$

which is the marginal rate of transformation or shadow price ratio between the two outputs. The sign of the tradeoff in (9) depends on the signs of the two partial derivatives. A common assumption in production theory is strong disposability of outputs or output monotonicity, which implies that if $y' \leq y \in P(x)$, then $\vec{D}_o(x, y'; g) \geq \vec{D}_o(x, y; g)$. That is, decreases in output do not increase efficiency. If output monotonicity holds, then the two partial derivatives are non-positive and the production possibility frontier has a negative slope. As mentioned before, we only require that outputs are weakly disposable which allows for the possibility that the production possibility frontier might be positively sloped over some range.

The translog functional form has been widely used to estimate stochastic Shephard input distance functions by Atkinson et al. (2003a, b), and Atkinson and Dorfman (2005a, b), and to estimate stochastic output distance functions by Grosskopf et al. (1997), Paul et al. (2000), and Balcombe et al. (2007). Chambers (1998) suggests a quadratic form as a flexible approximation for directional distance functions. The quadratic form provides a second-order approximation to the true but unknown distance function and can be restricted in accordance with the translation property. In contrast, Shephard distance functions have a homogeneity property that can be restricted within the translog form, but the translog form cannot be restricted to satisfy the translation property. We choose a common directional vector, $g = (1,1)$, for all departments so that the directional vector does not have to be part of the specification. In addition, a common directional vector allows for the aggregation of departmental technical inefficiency to a university technical inefficiency indicator (Färe and Grosskopf 2004).

Efficient departments produce on the frontier with $0 = \vec{D}_o(x^{kt}, y^{kt}; 1, 1)$. Because we use a quadratic approximation to the true but unknown directional distance function and to allow for the existence of technical inefficiency we introduce a two component error term, $\varepsilon^{kt} = v^{kt} - \mu^{kt}$, where v^{kt} represents the noise component which is assumed to be *iid* with zero mean and μ^{kt} is a one-sided *iid* error term representing technical inefficiency. Let $Q\vec{D}_o(x^{kt}, y^{kt}; 1, 1)$ represent the quadratic approximation to the true directional distance function. Accounting for inefficiency and noise yields $0 = Q\vec{D}_o(x^{kt}, y^{kt}; 1, 1) + v^{kt} - \mu^{kt}$. In the period we examine, the number of accredited programs increased from five to thirteen. We control for this quality attribute using a binary indicator variable, AC^{kt}, which takes a value of one if an accredited program resided in department k in period t and zero otherwise. Data on departmental research output are available for only some departments for the most recent period. We argue that the accreditation variable partially controls for departmental research output since most accreditation certifications require some research. Substituting the quadratic form for the true but unknown directional output distance function yields:

$$0 = \alpha_o + \sum_{n=1}^{N} \alpha_n x_n^{kt} + \frac{1}{2} \sum_{n=1}^{N} \sum_{n'=1}^{N} \alpha_{nn'} x_n^{kt} x_{n'}^{kt} + \sum_{m=1}^{M} \beta_m y_m^{kt} + \frac{1}{2} \sum_{m=1}^{M} \sum_{m'=1}^{M} \beta_{mm'} y_m^{kt} y_{m'}^{kt} + \sum_{n=1}^{N} \sum_{m=1}^{M} \delta_{nm} x_n^{kt} y_m^{kt} + \theta \cdot AC^{kt} + \varepsilon^{kt}. \quad (10)$$

Symmetry of the cross input and output effects implies $\alpha_{nn'} = \alpha_{n'n}$, $n \neq n'$ and $\beta_{mm'} = \beta_{m'm}$, $m \neq m'$. The translation property implies $Q\vec{D}_o(x^{kt}, y^{kt}; 1, 1) = Q\vec{D}_o(x^{kt}, y^{kt} + \phi^{kt} \cdot 1) + \phi^{kt}$. For the quadratic form the translation property requires that $\sum_{m=1}^{M} \beta_m = -1$, $\sum_{m}^{M} \beta_{mm'} = 0$, $m = 1, \ldots, M$, and $\sum_{m=1}^{M} \delta_{nm} = 0$, $n = 1, \ldots, N$. Using the translation property and rearranging (10) yields

$$-\phi^{kt} = \alpha_o + \sum_{n=1}^{N} \alpha_n x_n^{kt} + \frac{1}{2} \sum_{n=1}^{N} \sum_{n'=1}^{N} \alpha_{nn'} x_n^{kt} x_{n'}^{kt} + \sum_{m=1}^{M} \beta_m (y_m^{kt} + \phi^{kt} \cdot 1) + \frac{1}{2} \sum_{m=1}^{M} \sum_{m'=1}^{M} \beta_{mm'} (y_m^{kt} + \phi^{kt} \cdot 1)(y_{m'}^{kt} + \phi^{kt} \cdot 1) + \sum_{n=1}^{N} \sum_{m=1}^{M} \delta_{nm} x_n^{kt} (y_m^{kt} + \phi^{kt} \cdot 1) + \theta \cdot AC^{kt} + \varepsilon^{kt}. \quad (11)$$

We estimate the inefficiency component in two ways. First, we follow Atkinson et al. (2003a) who estimated an input distance function using the approach of Cornwell et al. (1990). Since the outputs and inputs might be endogenous we estimate (11) using the generalized method of moments. We assume that the μ^{kt} can be modeled using a fixed–effects approach for time varying inefficiency:

$$\mu^{kt} = \sum_{k=1}^{K} \Omega_{0k} \cdot \mathrm{DEPT}^k + \Omega_1 \cdot t + \Omega_2 \cdot t^2 - v^{kt}, \quad (12)$$

where DEPT^k is an indicator variable for departments, t is a time trend, and v^{kt} is the noise component of ε^{kt}. The coefficients Ω_{0k} capture time invariant, department specific differences in the technology and the coefficients Ω_1 and Ω_2 capture the time varying differences in the technology that are common to all departments. We estimate Eq. 12 in a second stage regression. Second, we estimate (11) and (12) simultaneously using maximum likelihood methods following Battese and Coelli (1995) and Coelli et al. (1998).

In a recent innovative paper O'Donnell (2007) estimates the parameters of a directional output distance function using Bayesian methods. However, his method does not impose the translation property which is necessary for using the directional output distance function to estimate shadow price ratios. In the next section, we explain our choice of ϕ^{kt} and the inputs and outputs for our model.

3 Empirical results

Our data constitute an unbalanced panel of departments at Southeast Missouri State University for three periods extending over two decades. The university was established by statute in 1873 as the "Third District Normal School" with a mission of educating teachers for public schools in the region. The university service region includes 24 Missouri counties extending from St. Louis County in the north to the Boot-heel of Missouri in the south. Today, the mission of the university is to provide student centered education in the liberal arts and sciences and provide experiential learning. Some discipline specific research occurs, but most grant monies received are for providing service activities to the region such as older adult education and nutrition services, small-business planning and advisement, economic development, and early childhood health awareness programs.

Courses in each of the seven colleges (Business, Liberal Arts, Science and Mathematics, Education, Health and Human Services, University Studies, and Polytechnic Studies) not affiliated with their respective college departments are treated as a departmental unit. For the school year 1985–1986 there were complete data on inputs and outputs for 33 departments. During 1995–1996, 39 departments had complete data and 33 departments had complete data for 2004–2005. We assume that departments produce grade points (y_1) and screening or information content on the grade distribution (y_2) measured by the entropy index. Data on departmental research output at this predominantly teaching institution are only available for departments in some of the seven colleges for the most recent year. We partially control for research output by including an indicator variable for departments that have an accredited program or reside in a college that is accredited. A department that teaches twenty-five classes of three credits with an average class size of thirty with students receiving an average GPA of 3.0 would have $y_1 = 6{,}750$. Departments use inputs of undergraduate classes taught by non-tenure track faculty (x_1), undergraduate classes taught by tenure track faculty (x_2), graduate student classes (x_3), and student credit hours measured at the end of the fourth week of the semester (x_4). Substitution between the various inputs occurs as some departments, like Music, offer numerous 1 h classes but use relatively few student credit hours. Other departments offer fewer classes that have large numbers of students.[1] The number of student credit hours at the end of week four is almost always greater than the number of student credit hours earned at the end of the semester since students drop classes between week four and the end of the semester.

Descriptive statistics are presented in Table 1. From 1985–1986 to 1995–1996 the number of undergraduate classes taught by tenure track faculty relative to non-tenure track faculty increased. The mix changed toward non-tenure track faculty from 1995–1996 to 2004–2005. The number of graduate classes offered per department increased by about three from 1985–1986 to 1995–1996 and then increased by thirteen from 1995–1996 to 2004–2005. During the 20 years period there is clear evidence of higher grades and less information content in the grade distribution. While the number of class hours attempted by students at the end of the fourth week of the semester declined by more than 1,000 per department, the number of grade points awarded per department increased by 1,300 and the entropy index declined from 1.37 in 1985–1986 to 1.15 in 2004–2005. The university experienced a decline in enrollment from 1985–1986 to 1995–1996 that is evident in the number of class hours attempted by students and in the total number of grade points. The overall university grade point average increased from 2.6 in 1985–1986, to 2.9 in 1995–1996, to 3.1 in 2004–2005. The university wide entropy index was equal to 1.43 in 1985–1986, 1.36 in 1995–1995, and 1.26 in 2004–2005. No data were available to control for student quality at the department level. However, the average ACT score of first-time, full-time students at the university in 1995 was 22.6 ($s = 3.76$, $N = 1{,}331$ students) and was 22.4 ($s = 3.63$, $N = 1{,}346$ students) in 2004–2005, so there is no evidence of a change in student quality.

We normalize departmental inputs and outputs by dividing by the mean value of the respective inputs and outputs. Thus, an average department uses inputs $x = (1,1,1,1)$ to produce outputs $y = (1,1)$. To impose the translation property we let $\phi^{kt} = -y_2^{kt}$ for each observation. We estimate the directional distance function in two ways. First, we control for potential endogeneity between the outputs and inputs and estimate (11) using the generalized method of moments (GMM). Second, we estimate the

[1] Using department grade points as an output might introduce aggregation bias within a department as lower level and upper level courses are treated equally. Aggregation bias might also arise when comparing departments that primarily teach lower level general education courses and departments that primarily teach upper level courses offered to majors, or between departments where test questions have right and wrong answers versus departments where test questions are ambiguous and the cost to the professor of defending assigned grades is high (Achen and Courant 2009).

Table 1 Descriptive statistics

Variable	Symbol	Mean	Std. Dev.	Minimum	Maximum
1985–1986 ($K = 33$)[a]					
Undergrad classes taught by non-tenure track faculty	x_1	38.8	40.9	0	139
Undergrad classes taught by tenure track faculty	x_2	83.9	97.3	0	555
Grad classes offered	x_3	18.2	22.5	0	87
Weekly 4 h	x_4	7,012.5	4,556.0	445	19,534
Grade points	y_1	15,759.2	10,087.8	589	40,115
Entropy	y_2	1.37	0.14	1.02	1.54
1995–1996 ($K = 39$)					
Undergrad classes taught by non-tenure track faculty	x_1	26.0	47.2	0	252
Undergrad classes taught by tenure track faculty	x_2	82.6	100.4	3	630
Grad classes offered	x_3	21.7	46.6	0	268
Weekly 4 h	x_4	4,689.2	3,709.4	79	17,029
Grade points	y_1	11,563.6	9,371.2	161	46,890
Entropy	y_2	1.23	0.28	0.53	1.53
2004–2005 ($K = 33$)					
Undergrad classes taught by non-tenure track faculty	x_1	50.4	62.4	0	235
Undergrad classes taught by tenure	x_2	86.9	83.2	6	478
Grad classes offered	x_3	34.8	59.5	0	253
Weekly 4 h	x_4	6,334.2	4,667.9	100	20,125
Grade points	y_1	17,059.2	12,711.7	180	59,559
Entropy	y_2	1.15	0.33	0.22	1.52
All years ($K = 105$)					
Undergrad classes taught by non-tenure track faculty	x_1	37.7	51.3	0	252
Undergrad classes taught by tenure track faculty	x_2	84.4	93.5	0	630
Grad classes offered	x_3	24.7	45.7	0	268
Weekly 4 h	x_4	5,936.4	4,372.8	79	20,125
Grade points	y_1	14,609.4	10,905.1	161	59,559
Entropy	y_2	1.25	0.28	0.22	1.54

[a] K = number of departments

directional output distance function using the stochastic frontier approach described in Coelli et al. (1998).[2]

Roodman (2009) finds evidence that when a large number of instruments are used in the GMM, equations for endogenous variables may be over-fit and the Hansen test for over-identifying restrictions is weakened. Four different instrument vectors are used and described in Table 2 along with the Hansen J statistic. We follow Atkinson et al. (2003a) and report the estimates for the set of instruments that gives the lowest test statistic for the null hypothesis that the over-identifying restrictions are valid. The smallest J statistic (highest p value) is for the instrument vector that consists of dummy variables for each department, time dummy variables, and the accreditation dummy variable. Given forty-six departments, three periods (one time indicator is dropped), and the accreditation variable we have 49 instrumental variables to estimate the 22 coefficients of the directional output distance function. The Hansen J statistic of the null hypothesis that the over-identifying restrictions are satisfied has a X^2 distribution with degrees of freedom equal to the number of over-identifying restrictions.

Table 2 Hansen test for overidentifying restrictions

Instruments[a]	Number of instruments	J statistic	p value
Dept., t_1, t_2, AC,	49	13.92	0.98
Dept. t_1, t_2, AC, x_n, x_n^2, $x_n x_j$	63	25.83	0.97
Dept., t_1, t_2, AC, y_1, y_1^2	51	16.44	0.97
Dept. t_1, t_2, AC, x_n, x_n^2, $x_n x_j$, y_1, y_1^2, $x_n y_1$	69	31.02	0.96

[a] Dept. = department indicator variables, AC = indicator variable if department is accredited, t_1 = indicator for 1985–1986, t_2 = indicator for 1995–1996

[2] Estimates using the stochastic frontier approach are from the program Frontier 4.1 of Tim Coelli.

Table 3 Parameter estimates of the directional output distance function

Parameter	Variable	GMM[a] Estimate (std. error)	Stochastic frontier[b] Estimate (std. error)
α_0	Constant	0.367 (0.080)**	0.924 (.050)**
α_1	x_1	0.040 (0.066)	−0.019 (.029)
α_2	x_2	−0.344 (0.126)**	−0.179 (.052)**
α_3	x_3	−0.036 (0.051)	−0.037 (.040)
α_4	x_4	1.190 (0.080)**	0.369 (.064)**
α_{11}	x_1^2	−0.034 (0.019)*	0.034 (.010)**
$\alpha_{12} = \alpha_{21}$	x_1x_2	−0.107 (0.053)**	0.043 (.010)**
$\alpha_{13} = \alpha_{31}$	x_1x_3	0.084 (0.030)**	−0.036 (.009)**
$\alpha_{14} = \alpha_{41}$	x_1x_4	0.079 (0.056)	−0.019 (.028)
α_{22}	x_2^2	0.145 (0.094)	−0.068 (.016)**
$\alpha_{23} = \alpha_{32}$	x_2x_3	−0.061 (0.036)*	0.028 (.011)**
$\alpha_{24} = \alpha_{42}$	x_2x_4	0.407 (0.118)**	0.146 (.051)**
α_{33}	x_3^2	0.018 (0.016)	−0.003 (.007)
$\alpha_{34} = \alpha_{43}$	x_3x_4	−0.080 (0.051)	0.027 (039)
α_{44}	x_4^2	−1.059 (0.118)**	−0.245 (.055)**
$\beta_1 = -1-\beta_2$	y_1	−0.707 (0.133)**	−0.047 (.077)
$\beta_{11} = -\beta_{12}$	y_1^2	−0.295 (0.125)**	0.193 (.100)*
$\delta_{11} = -\delta_{12}$	x_1y_1	−0.005 (0.047)	−0.033 (.031)
$\delta_{21} = -\delta_{22}$	x_2y_1	−0.206 (0.119)*	−0.248 (.041)**
$\delta_{31} = -\delta_{32}$	x_3y_1	−0.020 (0.056)	−0.012 (.026)
$\delta_{41} = -\delta_{42}$	x_4y_1	0.460 (0.067)**	0.094 (.061)
θ	AC^t	−0.027 (0.009)**	−0.034 (.015)**
σ^2			0.006 (.001)**
γ			0.968 (.015)**

* $p \leq .10$

** $p \leq .05$

[a] Generalized method of moments estimates

[b] Estimated with Frontier 4.1 of Tim Coelli

The smallest estimated J statistic is 13.9 with 27 degrees of freedom. Given a critical $X^2(\alpha = .05) = 41.1$ we cannot reject the null hypothesis. In Tables 3 and 4 we report coefficient estimates for the quadratic form and the inefficiency component of the error term.

To compute technical inefficiency using GMM we calculate the negative of the residuals, $-\hat{\varepsilon}^{kt}$, from Eq. 11 and regress $-\hat{\varepsilon}^{kt}$ on the right-hand side of (12) to obtain estimates of $\hat{\mu}^{kt}$. Define $\hat{\mu} = \min_{k,t}\{\hat{\mu}^{kt}\}$. Adding and subtracting $\hat{u}$ to the fitted model (11) yields

$$-\hat{\phi}^{kt} = Q\hat{\vec{D}}_o(x^{kt}, y^{kt} + \phi_{kt} \times 1; 1, 1) + \hat{v}^{kt} - \hat{\mu}^{kt} + \hat{\mu} - \hat{\mu}$$

$$-\hat{\phi}^{kt} = Q\hat{\vec{D}}^*_o(x^{kt}, y^{kt} + \phi_{kt} \times 1; 1, 1) + \hat{v}^{kt} - \hat{\mu}^{kt*} \qquad (13)$$

where $Q\hat{\vec{D}}^*_o(x^{kt}, y^{kt} + \phi_{kt} \times 1; 1, 1) = Q\hat{\vec{D}}_o(x^{kt}, y^{kt} + \phi_{kt} \times 1; 1, 1) - \hat{\mu}$ is the estimated frontier distance function and $\hat{\mu}^{kt*} = \hat{\mu}^{kt} - \hat{\mu}$ is the non-negative estimate of technical

Table 4 Inefficiency equation estimates

	GMM		Stochastic	
	Coef.	Std. Err.	Coef.	Std. Err.
Time trends				
T	−0.032	0.196	0.049	0.069
T[2]	0.011	0.049	0.010	0.018
Department indicators				
College of business	0.011	0.244	−0.061	0.102
Accounting, finance, and business law	0.084	0.238	−0.658*	0.358
Accounting and management information systems	0.007	0.284	−0.433	0.305
Administrative services	0.168	0.238	−0.293	0.269
Economics	0.048	0.238	−0.708**	0.073
Economics and finance	0.004	0.284	−0.889	0.570
Management	0.003	0.238	−0.121	0.135
Management and marketing	−0.003	0.284	−0.184	0.135
Marketing	−0.011	0.238	0.019	0.096
Elementary and special education	0.030	0.215	0.302**	0.086
Educ. administration and counseling	0.017	0.215	0.208**	0.098
Special education	0.020	0.215	−0.488**	0.168
Physical education	0.022	0.238	0.083	0.094
Secondary education	−0.060	0.244	0.227**	0.092
Aerostudies	−0.077	0.215	0.212**	0.087
Speech pathology and audiology	0.009	0.244	−0.041	0.106
Criminal justice	0.028	0.238	−0.368	0.484
Criminal justice and sociology	0.002	0.284	−0.396	0.300
Recreation and tourism	−0.041	0.305	−0.786	0.681
Health and recreation	0.002	0.284	−0.131	0.129
Human environmental studies	0.026	0.215	−0.184	0.113
Military science	0.023	0.305	0.123	0.122
Nursing	−0.020	0.215	0.068	0.083
Social work	0.019	0.215	−0.084	0.095
Art	0.037	0.215	−0.217	0.132
English	0.000	0.215	−0.698**	0.091
Foreign languages	0.029	0.238	−0.203	0.163
Foreign language and geography	−0.002	0.284	−0.245	0.153
History	−0.006	0.215	−0.316	0.718
Communications	−0.017	0.215	−0.316	0.718
Music	−0.051	0.215	−0.552**	0.226
Philosophy and religion	0.086	0.238	−0.703**	0.077
Political science	0.043	0.238	−0.610	0.479
Political science, philosophy, and religion	−0.009	0.284	−0.683	0.468
Psychology	0.047	0.215	−0.350*	0.207

Table 4 continued

	GMM		Stochastic	
	Coef.	Std. Err.	Coef.	Std. Err.
Sociology and anthropology	0.002	0.238	−0.705**	0.076
Speech and theatre	0.058	0.215	−0.153	0.109
Biology	−0.014	0.215	−0.250	0.155
Chemistry	0.007	0.215	−0.223	0.145
Computer science	0.065	0.215	−0.411**	0.182
Geosciences	0.048	0.215	−0.382**	0.173
Mathematics	0.033	0.215	−0.165	0.152
Physics	−0.038	0.215	−0.169	0.102
Agriculture	0.006	0.215	−0.091	0.098
Industrial technology and engineering	0.024	0.215	−0.280*	0.157
University studies	−0.019	0.244	−0.451**	0.134

* $p \leq .10$

** $p \leq .05$

inefficiency (Atkinson et al. 2003b; Atkinson and Dorfman 2005a, b). To estimate the standard deviation of technical inefficiency of each department from (12) we use the method of Krinsky and Robb (1986) to draw 5,000 coefficient vectors from a multi-variate normal distribution with mean equal to the estimated coefficients of Eq. 12 and the associated estimated variance–covariance vector. After each draw we calculate the simulated value of the left-hand side of (12) given the draw of the coefficients and the departmental indicators and time trend. We define these simulated values as $\tilde{\mu}^{kt}$. For each draw we calculate $\tilde{\mu} = \min_{k,t}\{\tilde{\mu}^{kt}\}$ and the simulated value of technical inefficiency as $\tilde{\mathrm{TI}}^{kt} = \tilde{\mu}^{kt} - \tilde{\mu}$. The mean and sample standard deviation of $\tilde{\mathrm{TI}}^{kt}$ are then calculated from the 5,000 draws and are reported in the Table 5.[3]

Table 6 summarizes technical inefficiency for the two estimation approaches. Mean technical inefficiency (GMM approach) over all departments is 0.099 in 1985–1986, 0.093 in 1995–1996, and 0.102 in 2004–2005. Given the normalization of each output by its pooled mean, the average hypothetical department in 1985–1986 could increase grade points by 0.099 × 14,609 = 1,446 and increase the information content of the grade distribution by 0.099 × 1.25 = 0.12 if it were to reduce inefficiency and produce on the frontier of $P(x)$. The stochastic frontier yields mean estimates of inefficiency of 0.044 in 1985–1986, 0.098 in 1995–1996, and 0.132 in 2004–2005.

[3] Department estimates of inefficiency and marginal rates of transformation for the stochastic frontier estimates are available from the authors upon request.

We define productivity change as changes in output that occur because of efficiency change or technical change. Efficiency change occurs when departments move toward or away from the given frontier $P(x)$ and technical change refers to shifts in the frontier. Following Färe and Grosskopf (2004) the Luenberger productivity change indicator equals the sum of efficiency change and technical change:

$$L^{k,t,t+1} = \vec{D}_o^t(x^{kt}, y^{kt}; g) - \vec{D}_o^{t+1}(x^{kt+1}, y^{kt+1}; g) + \frac{1}{2}\Big[\vec{D}_o^{t+1}(x^{kt}, y^{kt}; g) - \vec{D}_o^t(x^{kt}, y^{kt}; g) + \vec{D}_o^{t+1}(x^{kt+1}, y^{kt+1}; g) - \vec{D}_o^t(x^{kt+1}, y^{kt+1}; g)\Big] \quad (14)$$

where efficiency change equals $\vec{D}_o^t(x^{kt}, y^{kt}; g) - \vec{D}_o^{t+1}(x^{kt+1}, y^{kt+1}; g)$ and the term inside the $[\cdot]$ equals technical change. Positive values for efficiency change indicate greater efficiency and positive values for technical change indicate technical progress. Färe and Grosskopf (2004) also showed that when outputs are allocated efficiently and a common directional vector is used for all departments, an aggregate indicator of productivity change can be constructed as the sum of the department productivity indicators. The time trend coefficients in the inefficiency equation are insignificant for both sets of estimates, so there are no systematic time effects on measured inefficiency. We interpret the absence of time effects to be evidence that departments did not experience a common technological shift to or away from the production frontier over time. Therefore, any productivity gains or losses are due to efficiency change. Using the GMM estimates aggregate university inefficiency increased from 3.267 in 1985–1986 to 3.627 in 1995–1996, and then declined to 3.366 in 2004–2005. The stochastic frontier estimates indicate an increase in aggregate inefficiency from 1.452 in 1985–1986 to 4.356 in 2004–2005.

For the GMM estimates, the coefficients for the departmental indicators are not significantly different from zero. For the stochastic estimates sixteen out of 46 coefficients for departmental indicators are significantly different from zero. Departments in the Colleges of Business, Liberal Arts, Science, Industrial Technology, and University Studies, all have a negative effect on measured inefficiency, while departments in the College of Education with the exception of Special Education, have a positive effect on measured inefficiency.

In 1985–1986 only four departments (Music, Chemistry, Human Environmental Studies, and Speech) had accredited programs. That number grew to five in 1995–1996 and to thirteen in 2004–2005. Given $\hat{\theta} = -0.027$ (GMM) or $\hat{\theta} = -0.034$ (stochastic), departments that achieved program accreditation experienced a decline in inefficiency.

Our final concern is with the shadow price or tradeoff between grade points produced and the entropy index, (dy_1/dy_2). Again, we use the method of Krinsky and Robb

Table 5 Departmental estimates of technical inefficiency and the shadow price ratio (std. dev.)

Department	Technical inefficiency			Shadow price ratio		
	1985–1986	1995–1996	2004–2005	1985–1986	1995–1996	2004–2005
College of business		0.088 (.187)	0.109 (.229)		−1.914 (.810)	−1.057 (.400)
Accounting, Finance, and Business Law	0.162 (.164)	0.161 (.179)		−1.749 (612.9)	−1.463 (.250)	
Accounting and Management Information Systems			0.105 (.272)			−1.687 (.37)
Administrative Services	0.245 (.165)	0.245 (.181)		−1.279 (4.26)	−1.278 (.18)	
Economics	0.125 (.167)	0.125 (.184)		−3.425 (.71)	−2.072 (.35)	
Economics and Finance			0.101 (.270)			−1.812 (.29)
Management	0.080 (.166)	0.080 (.183)		−104.98 (6,725)	−1.567 (.20)	
Management and Marketing			0.095 (.270)			−1.035 (0.26)
Marketing	0.066 (.167)	0.066 (.184)		−4.141 (57.62)	−1.218 (.15)	
Elementary and Special Education	0.108 (.144)	0.107 (.165)	0.128 (.213)	−0.114 (.21)	−0.865 (.35)	−0.501 (.28)
Educ. Administration and Counseling	0.095 (.145)	0.094 (.168)	0.115 (.217)	0.945 (278.6)	−3.252 (159.8)	−0.487 (38.17)
Special Education	0.097 (.143)	0.097 (.164)	0.117 (.213)	−1.161 (0.41)	−0.046 (.18)	−.106 (39.84)
Physical Education	0.100 (.166)	0.099 (.184)		−0.502 (0.22)	−0.61 (.10)	
Secondary education		0.017 (.184)	0.037 (.227)		−0.398 (.08)	−.274 (.21)
Aerostudies	0.000 (.142)	0.000 (.164)	0.021 (.212)	−1.308 (.37)	−0.615 (.19)	−0.761 (.25)
Speech pathology and audiology		0.086 (.191)	0.106 (.232)		−1.115 (.57)	−1.039 (.49)
Criminal justice	0.106 (.167)	0.105 (.184)		−2.314 (.69)	−1.636 (.24)	
Criminal justice and sociology			0.099 (.272)			−1.367 (.19)
Recreation and tourism		0.035 (.244)			−1.380 (.16)	
Health and recreation			0.100 (.272)			−0.599 (.15)
Human environmental studies	0.103 (.141)	0.102 (.162)	0.123 (.210)	−1.765 (.50)	−0.576 (.11)	−0.713 (.35)
Military science		0.100 (.242)			−0.745 (.15)	
Nursing	0.057 (.141)	0.057 (.165)	0.078 (.214)	−1.413 (.79)	−0.989 (.25)	−0.533 (.27)
Social work	0.096 (.142)	0.096 (.163)	0.117 (.212)	−1.837 (.72)	−1.021 (.26)	−1.170 (.21)
Art	0.115 (.144)	0.114 (.165)	0.135 (.213)	−0.910 (.21)	−1.008 (.16)	−0.920 (.14)
English	0.078 (.142)	0.077 (.163)	0.098 (.212)	−16.877 (673.1)	−2.019 (.800)	−2.998 (70.3)
Foreign languages	0.106 (.167)	0.106 (.183)		−1.693 (.35)	−1.235 (0.22)	
Foreign language and geography			0.096 (.274)			−0.948 (.17)
History	0.072 (.142)	0.071 (.163)	0.092 (.213)	−10.927 (517.3)	−1.421 (.47)	−1.504 (.36)
Communications	0.060 (.144)	0.060 (.166)	0.080 (.216)	−2.384 (.87)	−1.274 (.34)	−1.711 (7.43)
Music	0.026 (.143)	0.026 (.165)	0.046 (.213)	0.175 (4.56)	0.364 (6.00)	0.493 (10.22)
Philosophy and religion	0.163 (.163)	0.163 (.181)		−1.688 (.47)	−1.793 (.45)	
Political science	0.120 (.164)	0.120 (.180)		−3.247 (.91)	−1.430 (0.22)	
Political science, philosophy, and religion			0.089 (.275)			−1.72 (.39)
Psychology	0.124 (.143)	0.123 (.165)	0.144 (.215)	−3.637 (25.92)	−1.356 (.43)	−1.560 (.32)
Sociology and anthropology	0.079 (.166)	0.079 (.182)		−3.533 (.75)	−1.655 (.34)	
Speech and theatre	0.135 (.142)	0.135 (.165)	0.155 (.213)	−1.123 (.57)	−1.00 (.15)	−0.643 (.21)
Biology	0.063 (.142)	0.063 (.163)	0.083 (.212)	−6.836 (51.06)	−1.639 (7.43)	−1.408 (.43)
Chemistry	0.084 (.142)	0.083 (.212)	0.104 (.142)	−3.354 (7.83)	−2.455 (51.58)	−2.583 (1.62)
Computer science	0.143 (.142)	0.142 (.165)	0.163 (.212)	−3.057 (3.01)	−1.46 (.38)	−1.433 (.34)
Geosciences	0.126 (.144)	0.125 (.167)	0.146 (.217)	−2.446 (1.70)	−1.821 (10.81)	−1.569 (.39)
Mathematics	0.111 (.141)	0.110 (.164)	0.131 (.212)	19.393 (839.2)	−38.38 (2,489)	−2.290 (608.7)
Physics	0.039 (.138)	0.038 (.161)	0.059 (.211)	−3.046 (1.04)	−1.670 (.49)	−1.449 (.26)
Agriculture	0.083 (.144)	0.083 (.167)	0.103 (.216)	−0.993 (.18)	−1.049 (.13)	−0.912 (.11)
Industrial technology and engineering	0.102 (.142)	0.101 (.163)	0.122 (.213)	−1.065 (.23)	−1.041 (.18)	−1.105 (.26)
University studies		0.058 (.187)	0.079 (.229)		−1.277 (.88)	−0.736 (.52)

Shadow price ratio equals the marginal rate of transformation = dy_1/dy_2 where y_1 = grade points and y_2 = entropy

to calculate the standard deviation of the shadow price for each department. We draw 5,000 sets of coefficients of the quadratic directional distance function from a multi-variate normal distribution with a mean equal to the estimated coefficients reported in Table 3 and covariance equal to the estimated covariance matrix. The means and standard

Table 6 Mean estimates of inefficiency and the shadow price ratio

	Estimation method	1985–1986	1995–1996	2004–2005
$\vec{D}_o(x, y; 1, 1)$	GMM	0.099 (.044)	0.093 (.044)	0.102 (.032)
dy_1/dy_2	GMM	−5.21 (18.6)	−2.24 (6.0)	−1.15 (0.72)
$\vec{D}_o(x, y; 1, 1)$	Stochastic	0.044 (0.06)	0.098 (0.14)	0.132 (0.17)
dy_1/dy_2	Stochastic	−13.06 (65.1)	−3.00 (4.51)	−8.39 (19.3)
Obs. satisfying monotonicity				
dy_1/dy_2	GMM	−10.86 (25.1)	−3.81 (15.1)	−3.03 (9.4)
Proportion[a]		0.916	0.954	0.928
dy_1/dy_2	Stochastic	−18.14 (69.5)	−3.72 (2.4)	−10.53 (18.6)
Proportion		0.848	0.948	0.909

Shadow price ratio equals the marginal rate of transformation = dy_1/dy_2 where y_1 = grade points and y_2 = entropy

[a] Proportion of observations satisfying monotonicity conditions

deviations of the tradeoff for these 5,000 draws for each department and period are reported in the last three columns of Table 5. The tradeoffs are for the index values of the two outputs, where outputs have been divided by their mean values. Table 6 summarizes the mean tradeoffs for the two estimation methods. The GMM estimates yield a mean tradeoff (dy_1/dy_2) of −5.21 in 1985–1986, −2.24 in 1995–1996, and −1.15 in 2004–2005. The stochastic frontier estimates yield a mean tradeoff of −13.6 in 1985–1986, −3.0 in 1995–1996, and −8.39 in 2004–2005. Between 85 and 95% of the observations satisfy the monotonicity conditions which correspond to strong disposability of outputs and the mean tradeoffs for these observations are also reported in Table 6.

To further illustrate the tradeoff consider a hypothetical department that produces on the frontier of $P(x)$ and has 700 students completing 3 h classes with 125 students earning an A, 200 students earning B, 300 students earning C, 50 students earning D, and 25 students earning F. The average GPA is 2.50, the number of grade points earned is $2.5 \times 700 \times 3 = 5{,}250$, and the entropy index is $E = 1.34$. Holding inputs constant, if this department's average GPA were to increase to 3.0 the number of grade points would increase to 6,300. This grade point change corresponds to a change in the indexed value of $dy_1 = \frac{1050}{14609} = 0.072$. Given the 1985–1986 GMM estimate of the tradeoff in the indexed values of grade points and entropy, we have $-5.21 = \frac{0.072}{dE/1.25}$. Solving for dE, we estimate the decline in the information content of the grade distribution as $dE = -0.017$. Using the 2004–2005 mean estimate of the tradeoff in indexed grade points and entropy yields $dE = \frac{.072}{-1.15} \cdot 1.25 = -0.078$. Thus, a given increase in grade points results in a larger loss of entropy in 2004–2005 than it did in 1985–1986.

To further check on the tradeoff between grade points and entropy we estimated a quadratic stochastic production function where grade points (y_1) are a single output that are produced from the four inputs. We model the inefficiency component as a function of time and entropy. We find that increases in entropy have a significant positive effect on estimated inefficiency which supports our estimates of the tradeoff between grade points and entropy from the directional distance function.[4]

The departments of Music, Special Education, Educational Administration and Counseling, and Mathematics have a positive estimated tradeoff between grade points and Shannon's entropy index in various years. These departments produce on a positively sloped portion of the frontier of $P(x)$. Music exhibits a positive tradeoff in all 3 years. Special Education exhibits a positive tradeoff in 1995–1996 and again in 2004–2005. Mathematics exhibits a large positive tradeoff (19.4) in 1985–1986 but the estimate has large standard deviation. For each of these departments $\partial Q\vec{D}_o(x, y; 1, 1)/\partial y_2 > 0$, which violates the standard monotonicity condition that an increase in output should not increase inefficiency. One can think of these departments as producing at a point such as A on the frontier of $P(x)$ as illustrated in Fig. 1. Although point A is efficient for the $g = (1,1)$ directional vector, holding grade points constant, an increase in entropy (y_2) would move the department inside the frontier.

Table 1 reports that faculty inputs changed during the period toward a mix with relatively more undergraduate classes taught with non-tenure track faculty. What effect did this changing relative mix have on the tradeoff between grade points (y_1) and entropy (y_2)? The average number of classes taught by non-tenure track faculty is 37.7 for the pooled sample. A full-time non-tenure track faculty person will generally teach four classes per semester. To simulate the substitution of non-tenure track faculty for tenure track faculty we increase the number of classes taught by

[4] These stochastic production frontier estimates are available upon request.

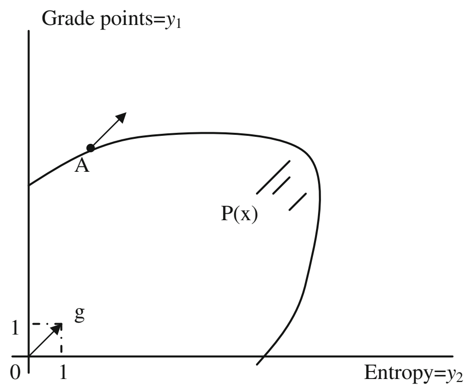

Fig. 1 A backward bending production possibility set, $P(x)$

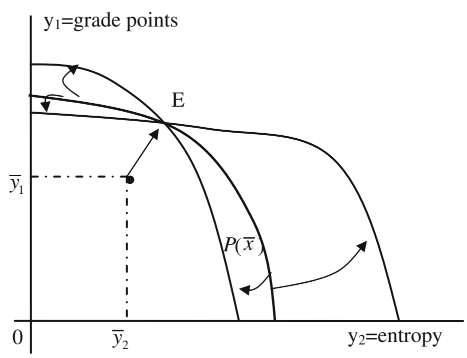

Fig. 2 Simulation of a changing input mix

non-tenure track faculty by four, which corresponds to $dx_1 = 4/37.7 = .106$. Then, we solve the quadratic directional distance function for the level of tenure track faculty (x_2) that would yield the same level of the two outputs, holding inputs x_3 and x_4 constant and reevaluate the tradeoff. In Fig. 2, a hypothetical department produces mean outputs using mean inputs with the frontier of $P(\bar{x})$ represented by the heavy line. The tradeoff for this department is evaluated at point E. If the tradeoff becomes more negative the production possibility set rotates through E and exhibits a bias toward producing more y_1 and less y_2. If the tradeoff becomes less negative, the production possibility frontier rotates through E and exhibits a bias toward producing less y_1 and more y_2. In 1985–1986, 69.7% of the departments had the simulated tradeoff become more negative due to the substitution of non-tenure track faculty for tenure track faculty in the classroom. In 1995–1996, 59% of the departments experienced a more negative tradeoff, and in 2004–2005, 54.5% of the departments experienced a more negative tradeoff. For those departments where the tradeoff becomes more negative, the simulation is consistent with Capozza (1973) and Zangenehzadeh (1988) and indicates that the substitution of non-tenure track faculty for tenure track faculty will result in more grade points and less entropy in the grade distribution.

4 Conclusions

Various researchers have argued that universities produce knowledge outputs and provide screening to allow employers to distinguish between students of different abilities and allow students to determine their comparative advantage. We used a directional output distance function to model changes in grade distributions for departments at Southeast Missouri State University. From 1985–1986 to 2004–2005 average GPAs increased at the university and the information content of those grades as measured by Shannon's entropy index fell. Estimates using a stochastic frontier and using the generalized method of moments indicated that mean departmental inefficiency increased during the period but that departments which achieved program accreditation exhibited less inefficiency. Furthermore, we found evidence of Baumol's cost disease as there were no systematic effects of time on measured inefficiency during the 20 years period. We also estimated the tradeoff between grade points earned and the information content of those grades and find that on average; increases in the number of grade points produced are associated with declines in the entropy index of the grade distribution. We also found that the substitution of non-tenure track faculty for tenure track faculty in the classroom makes substitution of grade points for entropy more likely for between 54 and 69% of departments. The empirical estimates are consistent with what has been termed "grade inflation" as departments award more grade points and the information content or screening component of education declines.

Several caveats apply to this study and suggest directions for further research. First, we were unable to obtain specific student information on student human capital that might be the cause of higher grades. For the university, average ACT scores for incoming freshmen remained constant during the last 10 years studied. If students increasingly sorted themselves into departments with students of similar ability, then increased student homogeneity might be revealed as declining entropy in the grade distribution. Although average ACT scores did not rise at the university, data on average ACT scores of students by department were unavailable to test whether students were sorting themselves by department. Second, although Southeast Missouri State University has established its primary mission as a teaching university, some of the faculty conducts research and provides community service. Although we partially controlled for faculty research by identifying departments that achieved program accreditation, more detailed data on research and service outputs were unavailable and their impact on efficiency and productivity change are unknown.

Acknowledgments We thank Gerald McDougall, Shawna Grosskopf, Kathy Hayes, Paul Wilson, and two referees for comments on an earlier version of this paper.

References

Achen AC, Courant PN (2009) What are grades made of? J Econ Perspect 23(3):77–92
Adams JD (1990) Fundamental stocks of knowledge and productivity growth. J Political Econ 98(4):673–702
Atkinson SE, Dorfman JH (2005a) Bayesian measurement of productivity and efficiency in the presence of undesirable outputs: crediting electric utilities for reducing air pollution. J Econom 126:445–468
Atkinson SE, Dorfman JH (2005b) Multiple comparisons with the best: bayesian precision measures of efficiency rankings. J Prod Anal 23:359–382
Atkinson SE, Färe R, Primont D (2003a) Stochastic estimation of firm inefficiency using distance functions. South Econ J 69(3): 596–611
Atkinson SE, Cornwell C, Honerkamp O (2003b) Measuring and decomposing productivity change: stochastic distance function estimation versus data envelopment analysis. J Bus Econ Stat 21(2):284–294
Babcock P (2010) Real costs of nominal grade inflation? New evidence from student course evaluations. Econ Inq 4:983–996
Balcombe K, Doucouliagos H, Fraser I (2007) Input usage, output mix and industry deregulation: an analysis of the Australian dairy manufacturing industry. Aust J Agric Resour Econ 51: 137–156
Bar T, Kadiyali V, Zussman A (2009) Grade information and grade inflation: the Cornell experiment. J Econ Perspect 23(3):93–108
Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empir Econ 20:325–332
Baumol WJ (1967) Macroeconomics of unbalanced growth: an anatomy of urban crisis. Am Econ Rev 57:415–426
Capozza DR (1973) Student evaluations, grades, and learning in economics. West Econ J 11(1):127
Chambers RG (1998) Input and output indicators. In: Färe R, Grosskopf S, Russell RR (eds) Index numbers: essays in honour of sten Malmquist. Kluwer, Boston
Chambers RG, Chung Y, Färe R (1996) Benefit and distance functions. J Econ Theory 70(2):407–419
Chan W, Li H, Suen W (2007) A signaling theory of grade inflation. Int Econ Rev 48(3):1065–1090
Coelli T, Rao DSP, Battese GE (1998) An introduction to efficiency and productivity analysis. Kluwer, Norwell
Cohn E, Rhine SLW, Santos MC (1989) Institutions of higher education as multi-product firms: economies of scale and scope. Rev Econ Stat 71(2):284–290
Cornwell C, Schmidt P, Sickles RC (1990) Production frontiers with time series variation in efficiency levels. J Econom 46:185–200
deGroot H, McMahon WW, Volkwein JF (1991) The cost structure of American research universities. Rev Econ Stat 73(3):424–431
Färe R, Grosskopf S (2004) New directions: efficiency and productivity. Kluwer, Boston
Färe R, Primont D (1995) Multi-output production and duality: theory and application. Kluwer, Boston
Freeman DG (1999) Grade divergence as a market outcome. J Econ Educ 30(4):344–351
Glass JC, McKillop DG, Hyndman N (1995) Efficiency in the provision of university teaching and research: an empirical analysis of UK universities. J Appl Econom 10:61–72
Glass JC, McKillop DG, O'Rourke G (1998) A cost indirect evaluation of productivity change in UK universities. J Prod Anal 10:153–175
"Grade Inflation at American Colleges and Universities" retrieved from http://gradeinflation.com September 10, 2005
Grosskopf S, Hayes K, Taylor LL, Weber WL (1997) Budget constrained frontier measures of fiscal equality and efficiency in schooling. Rev Econ Stat 79:116–124
Johnson VE (2003) Grade inflation: a crisis in higher education. Springer, New York
Kleiner C (2004) Tuition keeps rising: will new efforts by schools ease the burden? *U.S. News and World Report* USNews.com
Kohn A (2002) The dangerous myth of grade inflation. Chron High Educ 49(11):B7
Krinsky I, Robb AL (1986) On approximating the statistical properties of elasticities. Rev Econ Stat 68:715–719
Luenberger DG (1992) Benefit functions and duality. J Math Econ 21:461–486
Mansfield HC (2001) Grade inflation: it's time to face the facts. Chron High Educ, April 6, B24
O'Donnell CJ (2007) Estimating the characteristics of polluting technologies. In: 51st Annual conference of the Australian agricultural and resource economics society, Queenstown, New Zealand
Ostrovsky M, Schwarz M (2003) Equilibrium information disclosure: grade inflation and unraveling. Harvard University, Mimeo
Paul CJ, Johnston WE, Frengley GAG (2000) Efficiency in New Zealand sheep and beef farming: the impacts of regulatory reform. Rev Econ Stat 82(2):325–337
Roodman D (2009) A short note on the theme of too many instruments. Oxf Bull Econ Stat 71(1):135–158
Sabot R, Wakeman-Linn J (1991) Grade inflation and course choice. J Econ Perspect 5(1):159–170
Shannon CE (1948) A Mathematical theory of communication. Bell Syst Tech J 27:379–423 and 623–656
Spence M (1973) Job market signaling. Q J Econ 87(3):355–374
Stiglitz JE (1975) The theory of "Screening," education, and the distribution of income. Am Econ Rev 65(3):283–300
Zangenehzadeh H (1988) Grade inflation: a way out. J Econ Educ 20:217–223