

Can European Economics Compete with U.S. Economics? And Should It?

David Colander

Middlebury College

The European economics profession is currently in the process of changing the incentives in its research environment to more closely mimic the US economics profession so that it can perform better globally. This has led to the development and increased reliance on quantitative metrics based on “quality”-weighted journal articles and citation measures, such as the Kalaitzidakis, Mamuneas, and Stengos (2003) metric. Many top European economists are advocating that these measures be used for funding, promotion, and advancement decisions within the E.U., so that the E.U. economics profession can compete in the global economics profession.¹ As these new global metrics become built into the European institutional incentive structure, they are changing the nature of European economics.

This paper argues that the almost unidimensional focus on the quality-weighted journal article rankings and citations, which is advocated by “global” European economists, is misguided.² It further argues that, if used, these metrics will likely undermine the traditional European economics profession’s strengths and will hamper European economics’ development into a true global economics power. The paper suggests a set of alternative metrics that build on the traditional European economics profession’s strengths. It argues that the use of these alternative metrics will put European economics in a position to leapfrog the U.S. economics profession in the coming decades.³

Formal and Informal Rankings

Rankings happen. Informally, economists rank other economists and programs all the time—this is a good program; this is a bad one; this is an economist who has something to say; this is one who does not. It is only natural that these informal rankings get formalized. The questions posed in this paper are: how formal does one want to make these rankings, and how much attention should the economics profession pay to them? I argue: not very, and not much.

The economics-ranking sub-industry developed in the United States in the 1970s. Initially, rankings were not given significant prominence in funding or research, and top U.S. economists and economic programs paid little attention to it. The rankings attracted

¹ A subset of European schools has already become “globalized;” they use these measures internally for hiring and promotion decisions. A majority of programs have not; they remain more traditionally European, but they are feeling strong pressure to become “globalized” as well.

² See, for instance, Kirman and Dahl (1994), Dreze (2001), and Neary, Mirrlees, and Tirole (2003).

³ The arguments in this paper are based in part on a survey and interviews I have done with European graduate students and European economists in preparation for an article (Colander 2008) and two books that I am working on, *The Making of a Global European Economist* (2009) and *The Changing Face of European Economics* (together with Ric Holt and Barkley Rosser) (2009)).

most interest among non-top-tier programs that wanted to build their programs up to become top-tier, and thought they could do it by hiring and promoting people who scored highly in the rankings. Thus, over time, the formal rankings of programs, and of economists, became used in U.S. economics programs for promotion decisions, especially at second-tier programs.⁴ Most top economists that I talk to still do not take these rankings too seriously; they make their own decisions about who is a good economist and which is a good program. The rankings are certainly not something that the AEA even implicitly endorses in the way that the EEA has done.⁵

Given this dismissive attitude by top U.S. economists of rankings, as I turned my interest to European economics in the early 2000s, I was surprised by the strong reliance on quantitative metrics of quality-weighted journal article output that seemed to characterize my discussions with top global European economists, and with the current reforms in European economics. I argue that this concern and focus are misplaced, both because the quantitative measures being used don't measure full output of economists, and because they don't provide a meaningful measure of future innovative cutting-edge research output, which is the type of research that can move programs into the top globally-ranked set of schools.

While these top European economists advocating the use of the journal article metrics recognize the problems with formal quantitative rankings, they defend them by arguing that the need to change the institutional structure of traditional European programs is worth the costs the rankings would impose. Their goal is to make Europe competitive with U.S. economics, and they feel that the lack of pressure to do research in traditional European programs makes that impossible. They argue that introducing quantitative measures of research, and basing funding on those admittedly imperfect measures, would increase the pressure to do quality research and should be implemented.

The argument in this paper is that the changes being imposed will likely backfire, and are unlikely to help European economics compete. While I agree that the traditional European academic economics profession needs a serious shakeup, the ranking metrics being suggested for funding and advancement are not the way to do it. Specifically, much

⁴ Top-tier programs tended to give far less weight to these rankings. In their view, good work leads to high rankings, not the other way around. The sense I get from economists at top programs is that if you have to think about rankings, then you probably don't have a truly top program. Few funding decisions in the United States are directly tied to these rankings although' informally, rankings are used to justify funding requests. Where faculty unions exist, these formal rankings have been highly formalized and embedded in formal promotion procedures. At most non-union schools rankings of an economist's journal articles are just informally used, but they are nonetheless important.

⁵ In 1989, I reviewed much of the recent literature on ranking of U.S. departments for the AEA's *Journal of Economic Perspectives*. In that article I expressed a view that was shared by the top economists I spoke with, and suggested that "the ranking game has been beat to death." (Colander 1989, 141) I wrote: "Everyone knows that any ranking loses important dimensions and, among those active in the profession, the information about which schools rank where is known more precisely than the rankings disclose, especially in view of how quickly top individuals move from school to school and how quickly topics considered important change. (ibid.) Another recent survey of rankings in the U.S. came to a similar conclusion. It found that changes in rankings methods do not change significantly the top schools, but can change significantly at the medium- and lower-ranked schools. (Thursby, 2000) Thursby writes, "There's not a hill of beans difference across large groups of departments." (p. 383)

more consideration has to be given to the limits of rankings than seems to have been given. In the absence of that consideration, the introduction of formal rankings into promotion and advancement decisions will be counterproductive. It will turn European universities into a set of perpetual second-tier global programs, like the perpetual second-tier U.S. programs, rather than turning European programs into serious competitors for dominance of global economics.⁶

Full Output Ranking Metrics

In theory, appropriate formal quantitative rankings can be used in promotion and funding decisions. In practice, because of principal-agent problems, (Holmstrom and Milgrom 1991) it is not clear whether it can or not. The problem is that the appropriate formal rankings would be quite different than the existing rankings being discussed in Europe; they would have to be what I call a “full output metric.” Anything less than full output metrics will bias activity away from non- or under-measured output that often makes up a large percentage of cutting-edge research activities likely to change the profession in the future.

Were formal research output rankings the only way to gather information about economists’ output, the problems associated with the use of output rankings would be unavoidable. But they are not. As I will argue below, the economics profession’s activities are informationally small activities, where informal knowledge that active economists have of other active economists’ work provides a better measure of an economist’s and a program’s output than will any formal ranking. You know who responds to your paper in an e-mail with insightful comments; you know who made an important point at a workshop. You know which program has people working on outside-the-box research that isn’t showing up in the ranking metrics.⁷ You know who is working on a potentially major innovation in theory or in econometrics. You know who has written a paper that is insightful, but never got around to publishing it, or who published it in an obscure journal. Thus, for example, the fact that Bill Vickrey published most of his papers in obscure journals did not stop the profession from recognizing his contribution with a Nobel Prize.

My point is that for informationally small activities, such as research in economics, formal ranking studies add little to the understanding of active professionals of other professionals’ work in an operative way; they are far too imprecise for that. The measures simply create quantitative metrics that either ratify previously held views, or, if the rankings don’t match one’s prior, lead one to look for problems in the formal rankings. “In-the-know” economists’ informal ranking supersedes any formal ranking. There is obviously an interaction, and I am not arguing that formal rankings are useless.

⁶ For all the push by various non-top-tier U.S. programs to improve in the rankings, there has been little shift over the years in the overall rankings of programs; no program has moved into the top group that was not close to the top group initially. The problem is that when a star or group of stars develops at lower-ranked programs, they are hired away by the higher-ranked, more prestigious, and generally wealthier programs.

⁷ Cal Tech in the United States is an example. It ranks much higher in many top economists’ rankings than it ranks in the standard ranking metrics.

Formal rankings can be used to refine one's informal rankings, but my claim is that the information in economists' informal rankings far exceeds the information in formal rankings.

What Is Economists' Output?

Let me now consider what the standard journal article ranking metrics miss in measuring economists' output. Most U.S. universities specify economists' output as involving three components: research, teaching, and service.⁸ Any meaningful ranking of economists' output must appropriately capture and weight all these elements.

Most U.S. universities (at least those without faculty unions) leave open the relative importance of each of these measures in promotion decisions, in part because the department's weights that are used in decisions on promotion and tenure generally differ from the funder's implicit weights, especially in public universities, which are in large part government-funded. (It is an open secret in the United States that universities transfer toward research those resources given by funding agencies for teaching.) There is usually some wording in most university by-laws that all three elements contribute to promotion decisions, but in most U.S. graduate economics departments, the sense I get from talking to economists is that about 95% of the weight is placed on research and 5% on teaching and service.⁹

Actually, teaching plays a more important role in U.S. economics departments than the 95%/5% split would suggest. While it may not be given much weight in the formal decision, at many graduate schools economists' workmanship ethic, and their desire to be seen as good in all aspects of their job, leads professors to spend significant time on teaching—far in excess of what recognition it might provide them in the advancement process. Thus, if I had to hazard a guess about the average time spent by economics faculties at the top 50 universities on the various activities, it would be 50% teaching, 80% research, and 20% service, with much of the service coming from post-tenure faculty who have gone into government and university administration positions. (This adds up to more than 100% because most U.S. economists believe that they are putting in more than 100% effort.) Normalizing this effort to 100%, I arrive at a relative weighting of 33% teaching, 53% research (of which 80% of research effort or 42% of total effort is for journal article publication), and 13% service for U.S. economists, a weighting reflecting the true allocation of U.S. economists' time. So, for most U.S. economists, a ranking based on journal article research is capturing less than half of what they would see as their output.

⁸ Each of these sub-elements of output can include many dimensions. The most ambiguous of these three is service, but it general includes activities both within the economics profession and within the broader academic and societal community. Government, central bank, international government agency, business, and think tank economists have an even broader concept of output, and their output measure would significantly expand the areas that would fall under the service heading. Their service includes providing advice to policy makers, weighing in on public and internal debates and policy issues by writing policy briefs, and talking with, and influencing, policy makers and the general public.

⁹ Where I teach, which is a liberal arts school with only undergraduates, the implicit sense I have is that after an initial high level of teaching threshold is passed, 50% of the promotion decision depends on research, 40% on teaching, and 10% on service.

In traditional European universities, incentives have been quite different. As pointed out by Frey and Eichenberger (1992, 1993) there are two major differences. First, the non-journal article aspect of research gets far more importance in Europe; far less emphasis goes into preparing research for top-ranked journals. Second, in Europe, service is considered not only at the department and university level, but also at the national level; in Europe there has been a strong tradition of professors playing active formal and informal roles in advising governments and in public debates. They argue that compared to U.S. economists, a larger percentage of European economists play the role of public intellectual and weigh in on public debates about policy issues.

In my view, contributing to the public debate is an important output of economics—it brings economic knowledge to policy. Frey and Eichenberg’s argument that European economists do more of that is consistent with the incentives they face; such output is more likely to “count” in Europe, and can lead to advancement much more than is the case in the United States. My sense is that, until recently, in traditional European programs, the attitude was that if one published in an international journal—fine, but such activity was only one measure of many that lead to advancement in the European context. That means that European economists naturally focus less on publishing in international journals even if the economists in Europe are doing research that may be highly publishable in those journals.

My point is that these different incentives also lead to a different allocation of time in U.S. and global European programs and traditional European programs. Assuming the same 20% division of time that I gave to U.S. economists for department and university service, and adding another 10% for state-level service, I would estimate that 30% of European economists’ time goes to service. Frey also states that undergraduate teaching is more important in Europe, but to be conservative, let’s say that it is the same 50% input that I estimated for U.S. economists. Research ranks lower in traditional European programs than it does in the United States. Just to put some estimate on it, say that research generally gets 30% of their time, and that 60% of that research goes toward journal-article publication. Adding these together gives a total of 110%, the lower total amount reflecting the lower incentives for work of all kinds in Europe as compared to the United States. Normalizing to 100% gives relative time spent of 45% for teaching, 28% for research, of which 60% or 17% is for journal-article publication and 28% for service. The numbers used in these calculations are obviously debatable, but the point they make is not: these differential incentives in the U.S. and traditional European programs make any comparison of traditional European programs and U.S. programs highly questionable. That point holds for any reasonable choice of percentage allocation of time in Europe and the U.S.

If my estimates of the differential incentives for the U.S. and traditional European programs are close to correct, this suggests that lower effort accounts for only 27% of the lower journal-article research output in Europe, and that almost 75% of the difference is simply a reflection of the differential incentive structure. By that I mean that these journal-article metrics are comparing 17% of traditional European economists’ output with 42% of U.S. economists’ output, which makes it a bit like comparing European football players with U.S. football players on the basis of their ability to pass, receive,

and block a U.S. football. If one did a full measure of output, this argument would suggest that there is much more parity between the traditional European and U.S. economics professions than these ranking studies have suggested, and that they present a far less distressing picture than the commentators and advocates of a change in metrics suggest.¹⁰

What Is Economist's Research Output?

My argument against the journal-article metric is even stronger than that these metrics do not measure economists' full output. It is that they also don't measure economists' research output in an especially meaningful way. The reason they don't is that research output is a multidimensional concept, and journal-article output measures only a small part of the total; thus it biases the research done to emphasize that particular dimension. Research includes working on an idea in one's mind, taking part in an on-line debate, developing new ideas, writing a book, taking part in a workshop, commenting on a paper, editing a journal, refereeing, or developing some idea within a business. When Fisher Black worked on macroeconomic theory and the lack of a pricing anchor in an economy with advanced financial markets while at Goldman Sachs, or when Hal Varian worked on search algorithms while at Google, they were doing seminal research but little of it led to journal-article publication. Similarly, work by economists at YahooResearch doesn't get published. But that work influences ideas, real world policies, and discussions within economics.

Another dimension of research involves its purpose. Theoretical research is that part of research that best fits the journal-article outlets. But the majority of economists do not do theoretical research; they do applied research, which is meant to influence policy. By focusing the output metric on journal articles, economists are led to focus on what elsewhere I have called "hands-off" applied research, which I contrast with "hands-on" applied research.¹¹ Hands-off policy analysis is written for other economists or advisors to policy makers more than it is written for policy makers. To the degree the analysis actually comes to policy conclusions (generally it concludes with a call for more research), those conclusions are highly contingent on the implicit value judgments and goals in the models. *If* the policy maker accepts these value judgments and goals, and *if* the world works like the model, then the policy recommendations of the hand-s off research are relevant, but the hands-off research papers have no discussion of these issues, which makes the research of little value to the policy maker. Hands-off researchers must leave it to an intermediary between the economic scientist and the policy maker to do the translation.

Hands-on policy analysis is different than the policy done by hands-off applied economists. Applied policy economists do work that is more econometrically sophisticated, and is written for other economists, not for policy makers. Economists' applied policy is meant to be a contribution to the scientific debate. Hands-on policy

¹⁰ As I stated earlier in this paper, my argument is not that there are no problems with academic European economics. It is simply to argue that the ranking studies comparing the programs to the United States are not a clear demonstration of those problems.

¹¹ I develop this distinction further in Colander and Nopo (2008).

analysis is designed to contribute directly to the policy debates in a country. Whereas hands-off policy analysis concentrates on the scientific aspect of policy, hands-on policy applies scientific knowledge to policy by integrating economic knowledge and economic models into a broader framework.

Traditional European programs have given students training in both hands-on and hands-off policy analysis, but only the hands off policy analysis has a reasonable outlet in the journals. The output of hands-on policy research is affecting policy, not economic journal articles. It is not written for other economists; it is written for policy makers and its output is useful advice to decision-makers often on a time-dependent topic of immediate importance to them. They need an answer in a short time. A “better” answer provided later than that time is useless to them. With the adoption of a journal-article metric, this hands-off research output, which well may be the most important research output of economists, is discouraged. Unless the metrics are adjusted to include this hands-on applied research, one of the major strengths of traditional European programs in training students to do such research will be undermined.

The problem becomes cumulative. Individuals who do hands on applied research are driven from academic institutions and are replaced by researchers who do hands off applied research. They then teach students to do hands-off applied research, even though almost all the students’ primary role will be in doing hands-on applied research. Students’ training becomes further and further removed from what the majority of them will actually be doing, and eventually the expertise in doing hands-on applied research is lost.

Even within hands-off applied research, quality-weighted journal article metrics misses much information about research output, which involves ideas, not articles. Because it involves ideas, *economists who work in a subject area, or at a school, know who is contributing to research and who isn’t regardless of whether or not those people publish*. The reason is that journal publications are not the way academic researchers communicate among themselves. They talk; they e-mail; they listen to seminars; they read. In the United States, precisely because of its use of quality-weighted journal article biased metrics, most economists who are contributing research to an area publish in journals that match the metrics, so the metrics become a better match over time, but it takes a long time and there is much noise. For example often what researchers choose to publish has to do with many issues quite separate from the quality of their paper. When a journal article research metric is used, researchers who game and play the system have an advantage over those who simply are interested in doing good work. For traditional European programs that have not previously used the metric, the metric is an especially poor way of measuring output.

The lack of correlation between true research output and journal article publication is truer today than ever before. Today, in the United States, much of the debate about an idea happens at the pre-Working Paper stage, and the research workshops at top universities are more central to the spread of ideas in economics today than are journals. If you are reading an idea for the first time when it is published in a journal, you are out of the informal research loop, and are probably at least two years behind specialists’ thinking about the idea. In fact, I would argue that a paper has already had

much of the influence it will have on cutting-edge researchers by the time it reaches the Working Paper stage. Publication in a journal is more often than not a tombstone, marking ownership of an idea, than it is an important method of communication among cutting-edge specialists. Today, were journal publication metrics not used for promotion and tenure decisions, many existing print, and even on-line, journals would disappear because they would be unnecessary.¹²

The worst effects of the metrics occur in those institutions that have most rigidly quantified them. In the United States this happens when there are faculty unions, which require carefully classified lists of what counts (and how much it counts) for tenure and promotion decisions and what does not. These lists, which can become 40 or 50 pages long, and can include weighting factors for the number of words on a page, among hundreds of other dimensions, lead to a corollary to Gresham's Law—bad research drives out good research, since faced with such a list rational researchers choose those to publish in the “least costly” venue (the journals that require the least work for the highest measured gain in the metric) to them; they structure their research agenda accordingly. Researchers looking into long-range issues, advanced speculative work, or researchers setting off in new directions are selected out of the profession and are replaced by “game playing” researchers who are more focused on maximizing output as measured by the metric. It is that tendency that leads to the sense that much of what is published is of little value.¹³

My problem with the formal rankings being used to measure research is that they give zero weights to large numbers of research activities that are central to new ideas developing. That zero weighting leads researchers away from these unweighted activities and toward weighted activities. It leads rational researchers to focus on small journal-publishable ideas, and to deemphasize large ideas that might be more interesting, and have a larger research payoff. It also leads them to worry less about what their research is contributing to knowledge or society, and more to whether it is publishable. The publication of the paper becomes an end in itself. That has happened in the United States. Essentially, my conclusion of my most recent study of top U.S. graduate economic programs (Colander 2006, 2007) is that these graduate programs have become specialists at producing highly efficient journal-article writers, but far less proficient at producing broad-based research economists and far less proficient at teaching undergraduates.

The Traditional European Educational Environment and Research Output

Traditional European programs have been far less influenced by this drive for publications. The reason is that until recently, European economists faced a more informal, political, and social, incentive systems. Often, the incentive system was vague and idiosyncratic by school and country. Thus, Europe avoided the “publication at all

¹² I have discussed these issues in Colander and Plum (2004).

¹³ In Colander (2007) one of the students expressed this sentiment when he stated, “In a top journal like the *Quarterly Journal of Economics* I'd say at least half[(of the articles published)] are useless. Probably 20% are useful and the rest are unclear....” Fifty percent will never be cited or read again.” The other students in the interview agreed.

costs” bias that characterized the U.S. system.¹⁴ In traditional European economics, doing good work, impressing higher-level professors, measures of publications and citations, and positively interacting with other professors are all important.¹⁵ This continues to hold widely in traditional European programs even as the system moves into flux. Tono Puu (2006) captures this European sensibility in his assessment of publications. He writes: “We tend to look down on the previous generation as they published relatively little. This fact, however, does not imply that they worked little or were less creative. It might just signify that they were more choosy about what they regarded as being significant enough to merit publication.”

Since within the traditional European academic system funding and advancement decisions were often made because of social and political factors rather than research output factors, there was little push for publication. There was no central job market that even implicitly used a journal-publication metric. From what I hear from advocates for the journal-article metrics, and from traditional European economists as well, the results of this social advancement system were not always a pretty picture. Programs become inbred and out of touch with developments that are going on elsewhere. Programs and professors often developed fiefdoms, and a small coterie of professors often had control over hiring and research of younger professors, who then had no reason to develop an independent research program. Connections determined hiring decisions: once hired, researchers had limited flexibility. Programs were often focused on narrow issues, and many programs made little attempt to subject their ideas to the global economics discussion, in part because those discussions were in English, while the work in traditional European programs was often in the language of the country in which the work was being done.

The negative aspects of the traditional European programs are generally known and underlie the desire for change by almost all European economists; thus I am not denying that change is necessary. But when thinking about these changes, the negative aspects must be balanced against its positive aspects, and the positive aspects of the

¹⁴ The differential assessment system in Europe and the United States is explainable; in the U.S., there has long been an integrated labor market; in Europe, there was not, so in Europe there was less need to develop a generalized ranking system to compare researchers across borders. An important reason for this difference was that, until recently, European economics was fragmented by different languages and national borders, and a different promotion system that was based less on publications and more on subjective judgments that would develop over time. The smaller markets meant that all top economists in a country would meet. They could know one another personally, discuss with each other, and be able to come to an independent judgment about each other’s knowledge and work. Assessments based on quantitative output measures, such as publications, would matter, but they would be only one small element in the subjective judgment. With the development of a common educational policy and a common language, that is becoming less the case, and the structure of European economics is now in the process of change, albeit slowly.

¹⁵ We present just one example. In Britain, judgments about who was good, not publications, guided placements, with this done through established “old boys” networks. This meant, for example, that until recently, it was considered unnecessary to get a Ph.D. in order to get a job, and in fact the getting of a Ph.D. meant that the researcher was not considered good enough to be hired in the market. As Britain integrated its profession with the U.S. profession, that is no longer the case. Indeed, today British economics may have become more “American” than American economics, slavishly following outdated orthodoxies being abandoned in the United States.

traditional European system compared to the U.S. system, have received too little discussion.

One advantage is that the traditional European system gave researchers a larger incentive to think about big questions, and to spend more time reading others' work, and to publish only when they felt they had something meaningful to say that goes beyond what others know. Put another way, it directed researchers toward *thoughtful research* over *publishable research*. This focus on thoughtful research allowed pockets of excellence to develop in Europe that could not develop in the United States as European researchers took long-term views of research rather than short-term views. For example, important work in cointegrated vector auto regression was done at Copenhagen, important work in general to specific econometrics was done in England, important work in behavioral and experimental economics initially developed in Germany and Switzerland, important work in non-linear dynamics and complex systems took place in Germany, Sweden, the Netherlands and Italy, and important work in agent based modeling was done in Italy. Most of this research reflected a deep understanding of the limitations of the current approach in economics, and was foundational work designed to change it. Aspects of this work will likely play an important role in the future development of economics.¹⁶

A second advantage of the traditional European system is that it left European economists free to choose the type of outlet that made most sense for the idea. For example, if an idea took a book to make the point, they would write a book rather than not deal with the idea because it could not fit into a journal article. Often, ideas weren't published but simply catalogued and used in discussion and classes.

A third advantage of the European system is that it allowed for more diversity than one finds in the U.S. programs. Since there was no single measure in Europe, there were many measures of excellence, and different countries specialized in different elements. France, with the regulation school, specialized in the integration of social and economic issues. Italy specialized in the history of economic ideas. The Netherlands specialized in econometric modeling. Some German universities specialized in policy and others specialized in the implications of non-linear dynamics for economic systems. The "sameness" of graduate programs that the COGEE commission criticized as characterizing U.S. graduate programs was far less prevalent in Europe. (Krueger et al. 1991)

Despite the push toward a global blending, the most distinctively local heterodox traditions have continued to flourish most clearly within the local-language journals and outlets. Thus, when I have met European economists from traditional European programs I have often been highly impressed with their broad knowledge of the economic literature, their technical knowledge, and their sense of what issues are important. In my judgment European economics is on a par with US economics, so not only can it compete; it is already successfully competing.

¹⁶ This argument is developed further in Colander, Holt, and Rosser (forthcoming 2009).

Developing a Better Assessment System

Let me now turn to the difficult question: how to design an assessment system that will weed out the less desirable elements of the traditional European system while not destroying the good elements. I premise this discussion on the acceptance that change in traditional European academics economics programs is needed, and that the only way to implement change is likely to involve the use of some metric that replaces the current individually subjective decision-making process. The politics and vested interests are too strong to allow the current non-metric-based system work. Moreover, as Europe education expands from country-based education to a European education, the subjective information about who is good becomes harder to maintain and develop.

But to say that some metrics will likely be necessary is not to say that the current ones being used are the best. Indeed, my arguments suggest that they are far from the best, and that they are being used inappropriately. Let me discuss what implications I believe my arguments have.

1. *Begin from a measure of total output, and estimate what percentage of total output is being captured by the metric you are using.* If the output of economists is a complicated and multidimensional as I am arguing it is, any metric is going to capture only a small portion of that output. Thus, at the beginning of each metric study, the authors should be required by editors to estimate the percentage of output that their metric is capturing, in order to reduce the possible misuse of that metric. For example, if the recent metrics papers had stated in their conclusion that they are capturing only a small percentage (my estimate above was 17%) of European economists' total output and that there was no objective way to justify the weights they chose, their rankings would likely have not been used in the way they have been used.
2. *Use multiple metrics to judge quality of research.* All metrics are flawed, and thus one wants to use as many different measures as possible.¹⁷ The multiple measures take account of the arbitrariness of any ranking system and better capture the multiple dimensions of research. As Zimmerman (2008) discusses, REPEC offers numerous alternative measures, which gives it a nice flexibility, and includes working papers, measures of access on the web, and other such interesting elements. The use of multiple metrics will make it easier for subjective evaluation to enter in.
3. *If you have to choose one metric to focus on, focus on an expansive output metric rather than embodying strong implicit and explicit quality weights in the metric.* In other words, the output metric should be seen as a metric measuring only the number of articles with minimal concern for where they are published. The reason is two-fold.

¹⁷ This argument is nicely developed in Henrekson and Waldenstrom, 2007.

- a. The first reason is that the quality-weighting systems currently in use are highly biased against European economists generally. This is the case since the existing measures, based upon previous importance of journals, are highly U.S.-centric. They rate internal U.S. university journals (such as the *QJE* and *JPE*) highly, as well as rating *AEA* journals highly. This creates a bias against European economists. The problem is that the peer review process cannot be totally neutral; geographic or other proximity with the editor matters, not because of any conscious bias, but because of information flow and interests.¹⁸ Being at the same workshops with the editor puts one on the same wavelength and increases chances of knowing what the editor is looking for in a paper.

Once a journal ranks high in a metric, the rankings become self-fulfilling. People send their best papers to the journals that rank the highest, and the ranking becomes built into the institutional structure. The rankings provide an institutional structure that will make it much harder for European journals to compete and will be biased against European authors. To offset that bias, at least initially, one might even consider in European rankings giving *European journals be given a higher ranking in the metrics for European promotion and advancement*. This would enhance their ability to attract the best papers. Giving double weights for publication in European journals can help offset that existing journal bias. It can be seen as a way of generating bottom-up European journals. I justify it by what might be called an infant journal argument. This would encourage Europeans to publish in European journals and build them up as an alternative to the established U.S. journals.¹⁹

- b. A second reason is that the quality-weighted systems are highly biased against traditional European economics and will likely kill off many of the strong aspects of the traditional system as well as the weak aspects. As I argued above, strengths of traditional European economics include its diversity and its pockets of excellence. These strengths have developed precisely because a

¹⁸ Journals are subject to real-world influences and to interactions between authors and editors in defining what a problem is, and the reasonable way of approaching it are likely to be shared by researchers in the same geographic area. *JPE* articles come much more heavily from Chicago, and *QJE* articles come much more heavily from Harvard, than from other schools. The editors are at these schools and, even with the best of intentions about being open, which I think they have, cannot but be influenced by social interactions and discussions about what is relevant and what is not. Decisions on publication are seldom clear-cut; reviewers differ, and choice of reviewers can play an important role in whether a paper will be accepted. Thus, even in a peer review system the editor has much more influence than is often recognized. If Europe moves toward emphasizing a single peer reviewed journal metric, it should develop its own European metric that at least initially emphasizes European-based journals (weighing such journals much more heavily) to help offset the infant-journal problem.

¹⁹ Obviously, such an alternative weighting system should not continue forever but instead have a planned phase out over a ten-year period.

ranking system was not in use, and the use of a highly qualitative ranking system will undermine that strength.

A broad-based measure—such as Coupe’s (2003) measure of output—would be far less damaging, since it would give many more outlets to publish. Let me be clear; my argument is not that journal quality doesn’t matter, and that a publication in the *AER* is the equivalent to an article in the *Journal of Human Resources* or *Public Finance*. It is simply to argue that the quality weights are article-dependent and context-dependent, and are better made subjectively. For example, much of the initial cutting-edge research that will guide economics in the future will likely be published in non-traditional journals that are more open to new approaches; the journal article metrics will miss precisely that work that is likely to be most path-breaking.

The problem of using quality weights is that they can embody significant implicit subjective judgments under the guise of objectivity. For example, as Henrekson and Waldenstrom (2007) point out, the Kalaitzidakis, Mamuneas, and Stengos ranking, which was called “the most up-to-date set of objective journal weights available” in the introduction to a set of ranking studies published in the *JEEA* (Neary et al. 2003 p. 1247) has enormous subjective judgments hidden within it. Specifically, those authors point out that “a single article in the *American Economic Review* is valued more highly in the Kalaitzidakis, Mamuneas, and Stengos metric than 10 articles in the *Journal of Financial Economics*, or 25 articles in the *Journal of Law and Economics*, 60 articles in the *Journal of Health Economics*, or all the 400+ articles published in the *Journal of Evolutionary Economics*.” (p 5) By hiding those necessarily subjective implicit weights in the ranking metric, the subjective judgments are hidden rather than made transparent, so one can judge the metric.

4. *Develop a separate output metric to judge economists’ contribution to hands-on research.* Hands-on research does not lend itself to journal articles, and should not be judged by journal article output. Often, this research will result in internal memos. Even when it does result in publication, the best publication outlets are not economic journals. Rather they include general interest magazines, newspapers, and multidisciplinary journals. Any ranking should include some measure of this form of research output, or make it clear that it is not capturing this aspect.
5. *Develop a formal subjective metric to complement and possibly replace output metrics.* If, as I argued, economics is an informationally small activity, then more information exists within the profession than can be captured in any metric. Obviously these subjective measures will be flawed, but I argue that they will be less so than the alternatives because they are more likely to include many non-quantifiable output dimensions

of economists' performance.²⁰ There are many ways this formal subjection metric could be developed. One could set up a survey sent out to specialists in a field who rank other researchers and programs in the field in terms of the dimensions that one wants to rank. Weights could be assigned to the dimensions, allowing a composite judgment to be arrived at.

Many variations of this approach are possible. For example, you could allow each researcher to write a one-page statement of her or his research agenda along with a one-page vita, which others could use to rank that research. Programs could provide two-page summaries of their research program. Then this information could be provided to all who take the survey or to a panel of experts—say U.S. Nobel Prize winning economists—who would be asked to rank the schools based on these summaries and other knowledge of the program that they have. These formal quantified subjective measures could then be used in conjunction with the physical measures of output in the more traditional approach.

6. *Develop an input metric to supplement the output metric.* With the multidimensionality of research output, with much of it unmeasurable, even using multiple dimensions of output metrics, a lot will be missed. Essentially, what one wants in research is to have really, really bright and creative people, playing with really, really hard questions. Placing any type of measure on their output will likely hurt the process. Ideally, these really, really bright people should be given wide latitude. The problem is that all of us would like to be given that latitude. Thus some filtering device must be developed to limit those who are given that latitude. Requiring researchers to have Ph.D.s is one way in which that was done. That reduces the number of researchers significantly, and insures that they have a common core of knowledge. That filter could, and probably should, be reconsidered, and additional finer filters should be developed. For example, a set of additional difficult certification exams could be developed in specific areas, or in techniques, as is done with actuaries. Researchers could take these exams and be judged according to their knowledge, not their output. Many variations of this approach are possible; for example, before one can do research using a particular econometric technique, one could be required by a journal to have completed a difficult certification exam that demonstrated full knowledge of that technique.

Providing a Level Playing Field for Traditional European Economists

I am realistic and I recognize that the above suggestions are unlikely to change the current use of metrics, and even if the arguments are accepted, their implementation will

²⁰ One could even establish an on-line virtual economics academy, in which economists who join the on-line game can create their “dream teams of researchers.” The game could include a trading system that would essentially value various economists. Thus a “player” in the virtual academy game would be given an objective function and allowed to “draft” and “trade” players for his virtual university.

take many years. In the meantime the current quality-weighted journal-article metrics will likely dominate. Thus, even if one accepts none of the suggestions, I still believe that the EU should take a number of steps to offset the U.S. bias in the existing quality-weighted journal output metrics.

1. *The EEC should subsidize the distribution of European economic journals.* As the European Economics Association has discovered, any new economics journal today faces serious problems of finding acceptance in libraries. Whereas twenty or thirty years ago new journals in economics could easily find homes in libraries, new journals today, unless they are part of a large package supplied by a journal publisher, have almost no chance of being adopted by libraries.²¹ This puts new European journals, and existing European journals that are trying to expand their readership and global reach, at a considerable disadvantage to established journals with built-in library subscriptions. European journals face a strong infant journal problem.²² By emphasizing quality-weighted journal metrics, the EU is contributing to that bias and handicapping its own journals. To offset that disadvantage, if European economics chooses to give stronger emphasis to a journal ranking system, it would pay for the EEC to consider subsidizing existing European economic journals for a limited period of perhaps ten years, making them available free on the web. This subsidization would allow all libraries in the world access to these journals and showcase European researchers' work, thereby helping offset the infant journal problem. Subsidizing all European journals for a limited time period would give them a chance to develop and compete on a more level playing field. Using European research funds to do this would likely be more effective at advancing European economics than would the same funding of European research.
2. *Establish an invited peer review process for the new European journals.* Another method that European journals might use to establish their journals as top

²¹ The AEA journals are some of the most expensive economics journals in the world to produce, both because they pay editors well and because they have a labor-intensive quality process of review and printing. The average cost of each article is more than \$8000 per article. The journals nonetheless have a low price to libraries because so many libraries subscribe. It is this advantage that is allowing the AEA to introduce four new journals at low prices. These journals will quickly move up in the rankings. The reason is twofold. First, economies of scale—the large number of library subscriptions means that the AEA can spread out the costs much more than can other journals. (They tied in their four new journals that will begin publication in 2009 to existing journals so these will also be relatively inexpensive to libraries, since very few will opt out of the total package.) The second reason is that the AEA owns EconLit, which provides a major profit (over \$1,000,000 a year) to the organization, allowing it to heavily subsidize its journals.

²² As an example of the need for such a program consider the *Journal of the European Economic Association*. It had hoped to become one of the top six journals in the world, but it has not achieved it, not because it is of lower quality in its reviewing process, but because it started later, which places it at a permanent disadvantage. Its lower rank is recognized by researchers and its lower position becomes self-fulfilling as the first pass of the best papers are sent to the highest-ranked journals. The problem for new journals seeking to move up is amplified by the fact that journals have proliferated, and have become very expensive to libraries, meaning that libraries are not buying new journals. The result is that all new journals are at a significant competitive disadvantage relative to established journals and even to those journals that are part of the private publisher's package of journals that is sold to libraries as a whole.

innovative journals rather than as second-tier journals is to institute an “invited peer review” section of the journal. For a young scholar, journal publishing is a crapshoot—all the more so if the author doesn’t know the editor or at least is not familiar with the editor’s thinking. Lack of knowledge about publishing makes it difficult for young scholars to take chances. They too often go with papers dotting i’s and crossing t’s.

To offset that incentive, one can establish a system in which European editors choose highly promising young European researchers and invite them to publish a paper on a subject in the European journal. This could be competitive in the sense that it is judged on a one-page proposal of what is intended. The papers are still peer reviewed, but the presumption is that the paper will be published, and the writer will have a chance to respond to reviewers. This is essentially how some top-ranked U.S. based journals, the *JEP* and *JEL*, work, and it allows them to get a different range of articles than shows up in other journals. Actually, in reality, journals that rely totally on submission also have aspects of these invitations since editors sometimes “encourage” economists whose work they find interesting to submit papers to their journals. The fact that that happens is one of the reasons why the geographic home of a journal matters.

Conclusion

The largest problem of European academic economics is not that it is second rate; the problem is that it sees itself as second rate, and acts second-rate by trying to copy U.S. programs rather than building on its strengths, and not worrying about how it compares to the US programs. Despite all the governance and academic institutions problems, which are considerable, its strengths are substantial. It has some of the most interesting leading-edge work taking place, and, in certain sub areas has developed pockets of excellence that match or exceed anything found in the United States. It has strength in diversity that the United States is sorely lacking; and it integrates academics into hands-on research much better than does the United States. In many ways, my goal in this paper has been a Wizard or Oz role, advising European economics to listen to European economists such as Serge-Christophe Kolm (1998), to drop that second-rate image of itself, and see its strengths. They should not be like the lion in the Wizard of Oz who needs a medal to give it courage.

So in answer to the question: Can European economics compete with U.S. economics?, the answer is yes, absolutely. I have no doubt that Europe can match the United States in terms of the journal publication and citation metrics if it changes the incentives to do that. But my answer to the second title question: Should they compete? is no, at least not in reference to any formal ranking. Just because it can compete does not mean that it should do so. Competing in reference to a formal ranking rather than competing in ideas, is a tendency of a second rate program. That’s why the metrics of competition currently being used will likely do more harm to European economics than it will do good. It will put it in a permanently second-tier position as it reduces diversity and reduces work on cutting-edge ideas that may well become the economics of the future.

What European economics needs most is not a ranking metric, but a change in attitude. European economists are so sure they are second, because they are judging themselves by U.S.-biased standards, that they develop a second-place mentality. It is time for that to end. If Europe is to move to be a leader, it must have faith in itself, and develop the mindset of a leader. That means that it must do more than simply import U.S. metrics to Europe and design its programs after U.S. programs. They have to build on the strengths of the European system, and in reforming those institutions, see themselves as leapfrogging the United States, by developing a ranking metric underlying the promotion and funding mechanisms that recognizes that subjective judgments about what is good and bad must be made. The current changes being implemented will not do that; if implemented they will condemn European economics to being a second-tier set of programs. European economics, can, and should, do better.

Bibliography

- Colander, David. 1989. "Research on the Economics Profession." *Journal of Economic Perspectives*.
- Colander, David. 2006. "The Making of an Economist Redux." *Journal of Economic Perspectives* 19:1:175-98.
- Colander, David. 2007. *The Making of an Economist Redux*. Princeton: Princeton University Press.
- Colander, David. 2008. "The Making of a Global European Economist." *Kyklos* 61 #2.
- Colander, David. 2009 (in process). *The Making of a Global European Economist*. Cheltenham: Edward Elgar.
- Colander, David, and Terry Plum. 2004. "Efficiency, Journal Publishing and Scholarly Research." Middlebury College Department of Economics Working Paper 2004-19.
- Colander, David, and Hugo Nopo. 2008. "The Making of a Global Latin American Economist." Middlebury College Working Paper.
- Colander, David, Richard P.F. Holt, and J. Barkley Rosser, Jr. 2009 (in process). *The Changing Face of European Economics*. Cheltenham: Edward Elgar.
- Coupe, Tom. 2003. "Revealed Performances: Worldwide Rankings of Economists and Economic Departments, 1990-2000." *Journal of the European Economic Association*. 1 #6.
- Dreze, Jacques. 2001. "Economics and Universities in Europe." Working Paper.
- Frey, Bruno, and Reiner Eichenberger. 1992. "Economics and Economists: A European Perspective." *The American Economic Review* (May) 82:2:216-20.
- Frey, Bruno, and Reiner Eichenberger. 1993. "American and European Economics and Economists." *Journal of Economic Perspectives*.
- Henrekson, Magnus, and Daniel Waldenstrom. 2007. "Can Research Performance be Measured Uni-Dimensionally?" Research Institute of Industrial Economics Working Paper.
- Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7 #1.

- Kalaitzidakis, Pantelis, Theofanis P. Mamuneas, and Thanasis Stengos. 2003. "European Economics: An Analysis Based on Publication in the Core Journals." *Journal of the European Economics Association*.
- Kirman, Alan, and Mogens Dahl. 1994. "Economic Research in Europe." *European Economic Review* 38:505-22.
- Krueger, Anne. et al. 1991. "Report of the Commission on Graduate Education in Economics." *Journal of Economic Literature*. 29, #3 Sept.
- Kolm, Serge-Christophe. 1998. "Economics in Europe and the US." *European Economic Review* (January) 32:1; 207-12. .
- Neary, Peter, James Mirrlees, and Jean Tirole. 2003. "Evaluating Economics Research in Europe: An Introduction." *Journal of the European Economics Association* (December) 1:6; 1239-49.
- Puu, Tõnu. 2006. *Arts, Sciences, and Economics: A Historical Safari*. Berlin: Springer Publishing Co.
- Thursby, Jerry. 2000. "Yet Another Look at Economics Department Research" *Journal of Economic Literature*, 38:2; June 383-404.
- Zimmerman, Christian. 2008. Academic Rankings with RePEc. University of Connecticut Working Paper.