

Pluralism at Risk? Heterodox Economic Approaches and the Evaluation of Economic Research in Italy

By MARCELLA CORSI, CARLO D'IPPOLITI, and FEDERICO LUCIDI*

ABSTRACT. We analyze Italy's recent research evaluation exercise (VTR) as a salient example in discussing some internationally relevant issues emerging from the evaluation of research in economics. We claim that evaluation and its criteria, together with its linkage to research institutions' financing, are likely to affect the direction of research in a problematic way. As the Italian case documents, it is specifically economists who adopt unorthodox paradigms or pursue less diffused topics of research that should be concerned about research evaluation and its criteria. After outlining the recent practice of economic research in Italy and highlighting the relevant scope for pluralism that traditionally characterizes it, we analyze the publications submitted for evaluation to the VTR. By comparing these publications to all the entries in the EconLit database authored by economists located in Italy, we find a risk that the adopted ranking criteria may lead to disregarding historical methods in favor of quantitative and econometric methods, and heterodox schools in favor of mainstream approaches. Finally, by summarizing the current debate in Italy, we claim that evaluation should not be refused by heterodox economists, but rather that a reflection on the criteria of evaluation should be put forward at an international level in order to establish fair

*Marcella Corsi and Carlo D'Ippoliti are at the Department of Social, Economic, Actuarial and Demographic Studies of the University of Rome "La Sapienza," Viale Regina Elena 295/E, 00161 Rome, Italy. Federico Lucidi is at the Fondazione Giacomo Brodolini, Viale di Villa Massimo 21, 00161 Rome, Italy. Contacts: marcella.corsi@uniroma1.it, carlo.dippoliti@uniroma1.it, lucidi@fondazionebrodolini.it.

This work is part of the project "*La valutazione della ricerca nelle Università italiane: criticità dei metodi adottati e proposte alternative*" (coordinator: Marcella Corsi), financed by the Sapienza University of Rome. We are grateful to the other members of the research team for their fruitful comments and to Andrea Salvatori for his constructive criticism. Usual disclaimers apply.

American Journal of Economics and Sociology, Vol. 69, No. 5 (November, 2010).

© 2010 American Journal of Economics and Sociology, Inc.

competition among research paradigms, thus, preserving pluralism in the discipline.

Introduction

This article aims at presenting the recent experience of Italy's first research assessment exercise (VTR, *valutazione triennale della ricerca*) as an internationally relevant example in highlighting a neglected aspect of the evaluation of economic research, that is, the impact of research evaluation on research practice itself.

It will be shown that particularly (but not exclusively) when financial resources are linked to the outcome of the evaluation, procedures and criteria of assessment may create strong incentives for researchers and research institutions to modify their original aims and strategies. Thus, it is crucial to set clear principles and objectives for economic research and to conduct any research assessment on the basis of these objectives. As the case of Italy shows, when pluralism is not explicitly among these goals, the assessment exercise may result in a marginalization of minority approaches, which instead, may be deemed worthy of survival and cultivation, both by policymakers and the scientific community.

The case of Italy is especially suited for our aims for two reasons: on one hand, pluralism of methods and topics within economics is traditionally well established in the Italian academia (if not in absolute terms, in an international comparison). Therefore, Italy's case is convenient for exemplification but is also relevant *per se*, at least from the perspective of certain economic approaches. On the other hand, the recent research assessment exercise in Italy (VTR) exhibits certain characteristics that clearly highlight the risks as well as the potentiality of research evaluation, with the aim of preserving and developing heterodox economic approaches, along with providing the *stimulus* for a lively and healthy debate within the mainstream.

Our findings support the view that if research institutions are encouraged to engage only in the lines of research that are likely to receive the highest rating according to the evaluation criteria adopted within the VTR, a convergence process is to be expected within economics, resulting in a potential disregard of heterodox schools

and historical methods, and in favor of mainstream “Anglo-Saxon” approaches and quantitative methods. Ultimately, research pluralism may be harmed. These objections have been highlighted by Lee and Harley (1998), Lee (2007), and Lee and Elsner (2008) with reference to the U.K. Research Assessment Exercise. These works show that evaluations based on the criteria of closeness to mainstream economics, by means of the subsequent allocation of funds, may shape economic research in the middle-to-long run toward the disappearance of non-mainstream research fields. Thus, a critical reflection about the rating and ranking criteria adopted in the evaluation exercise is necessary.

Specifically, we conduct a statistical analysis of the publications evaluated within the VTR, contrasting them to a comparable subsample of the EconLit dataset. Our aim is to highlight systematic patterns in the selection of the publications submitted for evaluation. The underlying hypothesis, attaching relevance to this analysis, is that research institutions in the future will discourage the development of research topics (or approaches) that they deem unsuitable for evaluation because they are less likely to be positively ranked and thus, given the link between evaluation and funding, to contribute to the institutions’ budgets.

Our main point is that, if evaluation is implicitly based on the criterion of proximity to the mainstream, as it was done in Italy’s case, such behavior on the side of institutions may negatively affect the financing of research projects by nonmainstream economists as well as their hiring and career prospects. On the contrary, we claim that it is advisable and indeed possible to conduct research assessments that prove rigorous in assessing quality and at the same time are respectful of pluralism. For these purposes, evaluation should be based on the principles of accountability of the evaluators, transparency of aims and processes, and fair competition of research approaches and of institutions.

This article is organized as follows: the next section briefly outlines the scope and relevance of pluralism in economic research in Italy, highlighting the historical origin and the current diffusion of “heterodox” approaches to economics in Italy. The following section describes the mechanism and procedures of the VTR. The fourth section presents the results of our statistical analysis, and our

conclusions summarize the ongoing debate on the evaluation of economic research in Italy and place our contribution in this context.

On Pluralism in Economic Research

Italian Context

It would be well beyond the aim of the present work to provide a complete picture of all the topics and approaches to economic research currently pursued in Italy. However, it appears possible to highlight their variety and scope, in terms of a lively competition between geographical locations, public and private sector, single research centers and institutions, and most notably among alternative methodological and theoretical approaches and research fields.

As Pasinetti and Roncaglia (2006) highlight, this plurality may be considered partly as a result, and partly as a reaction, to the long period of dictatorship that Italy experienced in the 20th century.

Indeed, Italy has been at the frontier of economic research since its inception. As Roncaglia (2005) points out, we could even date it back to the Middle Ages with the Scholastic writers, or the 17th century with Bernardo Davanzati and Antonio Serra. Since the tradition of moral philosophy and humanistic studies (which political economy was a part of) was largely fostered during the Enlightenment period, it should not come as a surprise that between the end of the 18th and the beginning of the 19th century, Italian writers heavily contributed to the early development of the marginalist approach to economics, for example, with Vilfredo Pareto, Maffeo Pantaleoni, and Enrico Barone.

The advent of fascism affected the development of economic research in Italy in three ways: a) it required academics to make a vow of loyalty to the Fascist Party; b) it imposed autarchy and a corporatist philosophy of the economic system; c) it promulgated racial laws. As a consequence, eminent economists (such as Piero Sraffa) decided to move out of the country, not to be involved in the totalitarian regime, or they were forced to move to avoid persecution because of their faith (as Franco Modigliani was).

The economists who stayed in Italy were isolated from the international debate and frequently focused on narrow topics such as

monetary issues or business cycles, or applied issues, which afforded greater intellectual freedom from the cultural yoke of the regime. After World War II, it was rightfully decided to avoid a cleansing of these scholars, limiting the democratic reaction to the dismissal only of those few academics who were most actively involved with the dictatorship.

The survival of the old school generated a favorable environment for further development of applied economics within and outside universities. This development occurred in governmental agencies aimed at forecasting (ISCO, the Istituto di Studi sulla Congiuntura), planning (ISPE, the Istituto di Studi per la Programmazione Economica), and/or supporting policy making, for example, within ministries (as in the case of the SVIMEZ, the agency for the development of the *Mezzogiorno*) or the Bank of Italy. These agencies gained a certain reputation in the cultural and political debate, as did private research centers such as that of Confindustria (Italy's largest entrepreneurs' association), and within trade unions.

Overall, this institutional plurality corresponded to a certain plurality of points of view, especially concerning policy implications.

However, a crucial boost to the reprise of internationally relevant economic research came from the many scholarships and grants aimed at allowing brilliant students to spend periods of study and research abroad. Partly due to the presence of the mentioned personalities of Sraffa and Modigliani, and partly because these were already attractive gravitation centers for Italian researchers, Cambridge (U.K.) and the Massachusetts Institute of Technology (MIT) became crucial learning centers for Italian economists, along with Oxford with John Hicks, and to some extent Harvard with J. A. Schumpeter.

The youngest generations were thus confronted with approaches rather different than that of Friedman's Chicago and the Monetarist School, being that they were more acquainted with the Austrian School and the neoclassical synthesis *à la* Hicks and Modigliani, or with more radical critiques and alternatives to static marginalism, especially the Keynesian and Sraffian approaches in Cambridge, and Schumpeter's evolutionary approach.

Along these lines, with the crucial contribution of students and researchers returning from periods abroad, the largest universities

became autonomous centers of research and training, which developed a lively and even heated debate, though they usually did not develop their own “schools.”¹

The wide scope of methods and topics pursued by these masters determined the variety and pluralism of the subsequent generations of scholars working in Italy today. Without any pretensions to completeness and without specifying the affiliation of any scholar to a certain school of thought, it is possible to identify peculiar traits, foreign to the mainstream, and rather close to post-Keynesian and Sraffian traditions (as practiced, for example, by Luigi Pasinetti, Pierangelo Garegnani, Augusto Graziani, Claudio Napoleoni, Sergio Parrinello, Alessandro Roncaglia, Neri Salvadori, Luigi Spaventa, Mario Tonveronachi, or by international scholars who worked in Italy for certain periods, such as Jan Kregel among others); to the feminists (Tindara Addabbo, Elisabetta Addis, Francesca Bettio, Marcella Corsi, Daniela Del Boca, Antonella Picchio, Annamaria Simonazzi, Paola Villa); as well as to the evolutionists, experimentalists, and behavioralists (Giovanni Dosi, Massimo Egidi, Mauro Gallegati, or eminent foreign scholars who work—also—in Italy, such as John Hey, Samuel Bowles, and Axel Leijonhufvud).

As above mentioned, these approaches flourished side by side with the cultivation of the neoclassical paradigm, more or less related to its neoclassical synthesis variant, by authors such as Tito Boeri, Francesco Giavazzi, Tullio Jappelli, Marco Pagano, Pietro Reichlin, and Guido Tabellini (to mention a few), and the Italian economists who after their studies kept working abroad (mostly in the United States and the United Kingdom): Alberto Alesina and Orazio Attanasio, for example, who frequently participate in Italy’s academic and political debate.

The plurality of points of view determined in Italy a habit and openness to the debate on the foundations of our discipline, greater than in Anglo-Saxon countries, in Austria and Germany, and in some sense closer (for example) to the atmosphere emerging in France, India, or Japan.

These fundamental debates are also related to, and the cause of, a widespread cultivation of the history of economic thought (with such authors as Giancarlo De Vivo, Maria Cristina Marcuzzo, and Annalisa Rosselli), a discipline that was already considered by the older

tradition of economics in Italy not as a distinct field of inquiry, but as a fundamental tool of core economic analysis.

At the same time, the survival of the old applied tradition, largely flexible concerning the underlying theory for the reasons hinted at above, and the preservation of pluralism of topics and methods within it, was particularly encouraged by the specific economic vicissitudes of Italy's reconstruction and Italy's subsequent role "at the frontier" of the Cold War. We can thus find diverse fields of research such as the studies on *Mezzogiorno* and on unbalanced local development, on international monetary systems, distribution of income, division of labor and international trade, industrial districts, and the construction of structural econometric models. Partly linked to these streams of research is the research on econometrics and quantitative methods, carried on by internationally visible scholars such as Marco Lippi or Franco Peracchi.

These applied themes are often dealt with in a perspective that is "not fully mainstream," although their characterization as "heterodox" is not obvious, for example, when looking at simple quantitative parameters (such as keywords, cited literature, JEL codes). Nonetheless, a quantitative outlook of the current composition of Italy's research practice, outlined in the next section, may help define the peculiarity of the case under study.

A Quantitative View

Beside the problems of identification mentioned in the previous paragraph, it should be remarked that the very definition of "heterodox" approaches is problematic and controversial: some authors or schools may perceive themselves as mainstream, while being considered unorthodox by others, or vice versa. For this reason, lacking an *ad hoc* sample on researchers' identification and self-identification (as was done by Axarloglou and Theoharakis, 2003), in the remainder of the article we will define *mainstream* and *heterodox* approaches according to publications' JEL codes.

This method entails a conservative bias, since authors may opportunistically choose "theory-neutral" JEL codes in order to maximize their chances of publication in mainstream journals, and because the

Journal of Economic Literature classification system exhibits a very basic aggregate classification, providing little detail on nonmainstream approaches and themes, and possibly lacking some relevant descriptors.²

Thus, in order to classify a product as “heterodox,” we require it to be characterized by at least one of the following JEL codes:

B5—Current Heterodox Approaches

B50—General

B51—Socialist; Marxian; Sraffian

B52—Institutional; Evolutionary

B53—Austrian

B54—Feminist Economics

B59—Other

E1—General Aggregative Models

E11—Marxian; Sraffian; Institutional; Evolutionary

E12—Keynes; Keynesian; Post-Keynesian

On this basis, it is possible to consider a rough classification of Italian economists. Specifically, we consider a dataset composed by the entries in the EconLit database authored in the period 2001–2003 by economists located in Italy (thus, not necessarily of Italian nationality), in order to ensure consistency with Italy’s research assessment exercise, as explained in the following section. While the dataset cannot provide a definitive answer on the number of active researchers working on certain topics, it can help to identify at least the relative dimensions of different subgroups or “schools.”

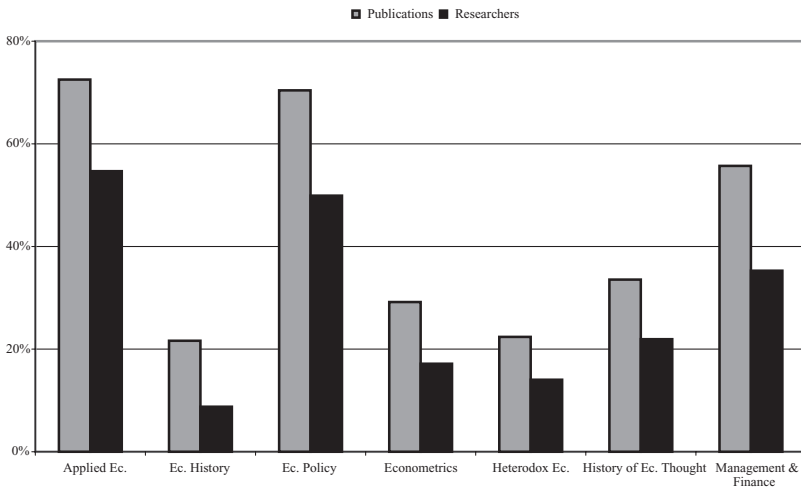
First, we classified products into eight very broad subdisciplines (fields) of economics on the basis of their JEL code(s) according to the criteria described in Appendix A. These categories are: *applied economics*; *economic policy*; *heterodox economics*; *econometrics and quantitative methods*; *economic history*; *history of economic thought*; *corporate finance and management*; and *other*. “Other” is here to be intended as a residual category, including all products whose JEL codes do not precisely fall into one of the previous categories. Thus, given our classification, it mainly includes field-specific theoretical works (for example, theoretical models of health economics, economic geography, and so on).

We then classified authors according to the subfields in which they have published. Specifically, we assigned an author to a field if he or she had published at least one product in the relevant subdiscipline in the considered period. To our aims, this criterion is superior to others equally available because it implies a convenient definition of “heterodox” economics. For example, an economist is classified as a heterodox economist if he or she manifested the nonmainstream nature of her/his work by authoring at least one publication that explicitly declares, through its JEL codes, a nonmainstream perspective.³

Figure 1 exhibits the resulting distribution of researchers and of products, for example, showing under the header “heterodox economics” the proportion of products authored by a heterodox economist (though not necessarily the number of heterodox publications, which we will deal with later). The distribution of researchers can be

Figure 1

Distribution of Italy’s Researchers and Their Publications on EconLit, Years 2001–2003



Note: Figures do not sum up to one: both publications and researchers may be classified under several categories simultaneously. The residual category “other” is not plotted to improve clarity of the Figure.

Source: our elaboration from EconLit, average values for years 2001, 2002, 2003.

used to compare the relative diffusion of certain approaches or topics in Italy, while the distribution of products is indicative of the international visibility of their publications (defined as the inclusion in the EconLit dataset).⁴

While it is likely that the EconLit dataset underrepresents the heterodox output of Italian heterodox economists, as we will discuss below, and despite the mentioned conservative bias of our method, a 23 percent share of publications and 14 percent of scholars classified as “heterodox” appear as relatively high figures when compared to other countries.⁵ Similarly, as many as 22 percent of Italian economists also write on the history of economic thought, collectively authoring more than a third of all the products in our EconLit dataset.

It should be mentioned that a quantitative analysis based on the entries recorded by EconLit (as will be carried out in the remainder of the article) bears a second conservative bias, as many relevant contributions by leading Italian heterodox economists (as well as by historians of thought) are frequently published in the form of books and book chapters, which are hardly collected by EconLit. Furthermore, heterodox economists have traditionally been very involved in local and national policy making, comparatively more than in other countries and in some periods possibly more than mainstream economists. This led many heterodox economists to focus their research on Italy-specific themes and frequently to write in Italian in order to better address the national public opinion (some examples are provided by the last two chapters of Roncaglia 2005). These factors further contribute to the underrepresentation of heterodox economics located in Italy within the EconLit dataset we employ, and thus characterize our estimates as prudential or even conservative.

Italy's Research Assessment Exercise (VTR)

The first official evaluation of Italian universities and research institutions (VTR), sponsored by the Ministry for Research and managed by an ad-hoc governmental committee, Comitato di Indirizzo per la Valutazione della Ricerca (CIVR), was set up in 2005 for the evaluation of the research output produced between 2001 and 2003. In the case of economics, this exercise focused exclusively on publications, as

authoritatively suggested by a special issue of the *Journal of the European Economic Association* published in 2003 that contained the results of a project on evaluating economic research in Europe (the project was initiated by the Council of the association in 1999). In their introduction to the issue, Neary et al. (2003) even state that “only published journal articles undergo a widely accepted process of peer review which is the essence of quality control in any scientific discipline.” However, in the Italian case, books and book chapters were considered as well.

The CIVR assessment was conducted through a qualitative peer-review process on a sample of research output selected by participating research institutions. The selection of products occurred in a “top-down” fashion, as heads of departments (and, at higher levels, of faculties) were delegated to choose which publications to submit to the evaluation (in the Italian university system, faculties are administrative units in charge of planning and managing the activities related to teaching, whereas departments organize research activities). Products’ ratings were then summed up at the level of institutions (universities or other research institutions) with the aim to construct disciplinary rankings of public research centers. The rationale of a publication’s selection, being a matter of crucial importance for the preservation of pluralism in economics, will be analyzed in detail in the next section.

The exercise proceeded as follows (see Lippi and Peracchi 2007, for details): 14 “research areas” were recognized, which represented the units of analysis. Economics, statistics, management, and business studies were grouped into Area 13, though some margins of overlapping across areas remained as the classification of products is not always straightforward (for example, in the field of statistics). As a result, neither faculties nor departments were evaluated directly, since, for example, economists may be employed in several different departments of one university, which would be evaluated for the area as a whole.

Research centers were asked to submit a number of research products equal to at least one half of the number of full-time equivalent academic staff (full professors, associate professors, and research fellows). However, this proportion was to be respected at the

level of the whole institution, not of each research area. Thus, universities were free to submit relatively more products in their specific areas of perceived excellence.

For each area, a panel of national and international experts was nominated, and submitted products were distributed among the panel's members according to their specific expertise. Each member of the panel was responsible for proposing a *rating* of the products assigned to her/him; in turn, the ratings were formulated on the basis of two or more independent referees' reports of each product. Finally, the whole panel voted on the member's proposals.

According to the norms established for all disciplines, research products were rated according to several criteria, the most prominent being the ranking of the product with respect to scientific excellence in "a value scale shared by the international scientific community." As we will discuss in the next section, the vagueness of this definition of research quality is the origin of a substantial bias in the selection of products to be submitted for evaluation.

The average of products' ratings constituted the rating of the institution itself: no consideration was taken for other variables measuring quality of research management, governance, and fund raising. Thus, it was implicit in the methodology that a unique process (leading to a single indicator) could evaluate publications, research output, research staff, and research institutions.

For each research area, universities were partitioned in four size classes according to the number of submitted products, and rankings of institutions were presented separately for each size class. In order to appear at the top of the disciplinary rankings, it was optimal for research centers to submit as few publications as possible, namely, only those they could expect to be judged as excellent by the referees, and possibly by the same authors (thus introducing a conservative and conformist bias in the process).

The introduction of size classes to partition the rankings, by preventing a direct comparison of institutions submitting very different number of publications, was an expedient to counterbalance this distortion. A crucial ambiguity characterized the rest of the process: Was the exercise meant to measure the average achievements of research in Italy by analyzing a representative sample of the research

output, or was it meant to recognize and single out excellence by assessing only the best publications? This ambiguity was never solved during the process, highlighting a fundamental lack of transparency that resulted in each institution behaving on the basis of its perceived convenience. Thus, the study of the institutions' selection of publications to be submitted is highly indicative of the *a priori* perception of a publication's or a paradigm's quality.

This issue appears to be the most relevant, since the communication of results in terms of disciplinary rankings, often with little emphasis on the methodology adopted for rating and ranking, risks being perceived as an overarching answer for whatever question concerning the quality of research institutions, as noted by Lee (2007). Due to their one-dimensional quantitative nature, rankings very easily lend themselves to many other uses beyond their original rationale, without reference to the fact that they were built on the basis of specific indicators constructed to measure some variables and not others.⁶

This lack of transparency may descend from the nature of the exercise, officially characterized as being only a pilot.⁷ However, as a consequence of the rhetoric strength of rankings, right after the exercise and until a recent decision by the central government, claims were frequently made that the resulting rankings should affect the future allocation of public funds, despite the fact that the VTR exercise was conducted as a typically *ex-post* assessment (for example, by Checchi and Jappelli 2008; Giavazzi 2008; Jappelli 2008).⁸

The results of the first evaluation of public research in Italy are undoubtedly positive, in general terms. The CIVR assessment represented a first step toward a change in the attitude of universities and researchers, introducing the principles of accountability and merit, hopefully with positive offspring in terms of research quality, competitiveness, and attractiveness of Italian universities. The outcomes of this first assessment also provide some food for thought. Its implementation brought about some criticality, concerning both the methodology and its application to the human and social sciences, economics in particular. The next section will introduce some issues by means of a statistical analysis of the products submitted for evaluation; we then draw some conclusions in light of the debate that the VTR stimulated among Italian economists.

A Statistical Analysis

Some Descriptive Statistics

The VTR evaluation exercise constitutes a good case study to highlight an issue often neglected: the impact that research evaluation may have on the development of research itself. This section argues that Italy's VTR embodied incentives for research institutions to highlight their mainstream economic research, while submitting for evaluation a small number of heterodox economic publications. This trend is particularly worrying because if an institution's budgets directly or indirectly depend on the results of the evaluation exercise, it will inevitably lead to disregard and prevent the development of research paradigms that are not likely to receive a positive evaluation. It is, therefore, of prime importance, for the sake of pluralism in economics, that these paradigms are not systematically identified with the heterodox approaches.

In order to investigate the selection of products submitted to the VTR, we will compare two datasets: one is composed of the products submitted for the fields of economics and statistics (Area 13); the other is a subsample of the EconLit dataset, containing the publications authored in the years 2001–2003 by economists located in Italy.⁹

Preliminary to the analysis, it is crucial to note that as none of the datasets is representative of the whole production of economists located in Italian research institutions, it is difficult to inquire into how they relate to the actual scientific output. The broader picture provided by the EconLit dataset, compared to the self-selected nature of the VTR dataset, allows us to consider the EconLit dataset as a benchmark to understand how faculties decided to attend the evaluation exercise.¹⁰

Our EconLit dataset exhibits 2,709 entries, authored or co-authored by 1,347 authors in the 2001–2003 period. Almost all of the products (2,509) in the EconLit database exhibit at least one JEL code. Conversely, 1,007 products are included in the VTR dataset, by a total of 842 authors. Since Area 13 of the VTR includes publications belonging to either economics or statistics, when considering only economics the dataset collapses to 597 products (22 percent of the EconLit

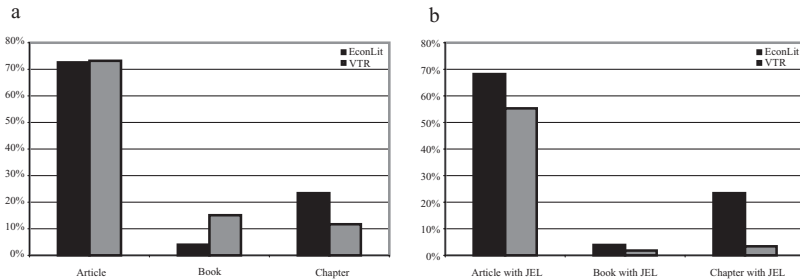
corresponding figure). By matching the datasets, we were able to assign one or more JEL codes to 361 of them.

As shown in Figure 2a, the composition of publications in the two datasets, by typology of publication, is significantly different. It emerges that research institutions decided to submit books proportionally more than their appearance in the EconLit dataset (15 percent of submitted products, *vis-à-vis* 4 percent in EconLit), while submitting book chapters in a smaller proportion (24 percent in EconLit, 12 percent in VTR). Journal articles represent more than 70 percent of both datasets.

A major source of this difference may be the partly different disciplinary scope of the two datasets, with the VTR extending over a larger number of subdisciplines within social sciences. Thus, according to Peracchi (2007), the president of the panel of experts for Area 13 (economics and statistics), Italian researchers in the fields of financial and business economics and management studies—who constitute a large share of the VTR dataset—have a higher propensity (with respect to the other researchers included in Area 13) to publish

Figure 2

Composition of the Datasets by Publication Typology: a: Composition of the Datasets; b: Composition of the Subsamples of Publications Exhibiting at Least One JEL Code



Note: a: the EconLit sample is composed of 2,709 products. For the VTR sample, only the 597 products in economics are considered; b: the EconLit subsample of products with JEL codes contains 2,509 products (92.6 percent of the sample). The VTR subsample includes 361 products (60.5 percent of products in economics). Percentage values refer to the whole sample.

books and book chapters, frequently in Italian.¹¹ It is thus likely that these typologies of publication are not overrepresented in the VTR but rather underrepresented in the EconLit dataset. Indeed, the latter exhibits a lower coverage also of interdisciplinary journals and journals focusing on subjects less related to the traditionally-defined (mainstream) topics of economics, including management and business studies journals.

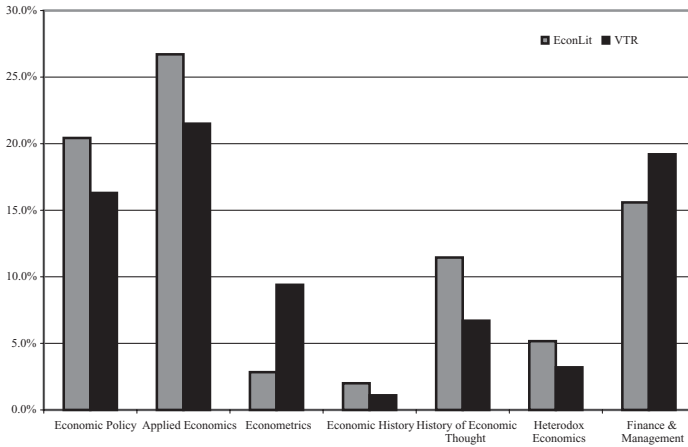
In what follows, we will focus on the subsamples of products that exhibit at least one JEL code: Figure 2b shows the relative composition of these subsamples. By comparing Figures 2a and 2b, it is evident that, when looking only at products exhibiting a JEL code, most books, as well as a considerable number of book chapters, submitted to the VTR are ignored by the analysis. There is a presumption that these are mostly the above-mentioned publications written in Italian. As a matter of fact, our sample of products denoted by JEL codes includes most products that were classified by the Area 13 panel of experts as belonging to “economics.” Conversely, more than two-thirds of products without a JEL code were submitted for the disciplines “statistics and operational research” or “business administration.”

Bearing these caveats in mind, we compare the composition of the two datasets employing the eight field categories defined above. As already noted, allocation of the same publication across several subdisciplines is allowed for products that exhibit more than one JEL code: the results are displayed in Figure 3.

As shown in Figure 3, substantial differences emerge between the two subsamples. In a number of fields, products appear in a considerable proportion in the EconLit dataset, while being drastically underrepresented in the VTR dataset: it is the case of applied economics (−5.1 percent absolute difference between the EconLit and the VTR, or −20.5 percent relative to the proportion in EconLit), history of economic thought (−4.8 percent, or −34.3 percent in relative terms), economic policy (−4.1 percent, or −23 percent), economic history (−1 percent, or −44.3 percent), and heterodox approaches (−1.8 percent or −35.4 percent). By contrast, two fields are more represented in the VTR dataset than in EconLit: financial and business economics (+3.9 percent or 26 percent of the EconLit proportion), and econometrics and quantitative methods (+6.6 percent, or +174.2 percent).

Figure 3

Composition of the Datasets by Field



Note: the field of each product is assigned on the basis of the respective JEL code(s). The EconLit subsample of products with a JEL code contains 2,509 products; the VTR subsample includes 361 products. Products with more than one JEL code are assigned simultaneously to all the relevant fields. To improve clarity, the figure does not include the category “other” (87.8 percent of entries in the EconLit subsample and 83.7 percent in the VTR one).

At a more disaggregated level, this corresponds to a reduction of the frequent (in the EconLit dataset) JEL codes A—General Economics, F—International Economics, J—Labor Economics, L—Industrial Organization, and R—Regional Economics in favor of publications denoted by JEL codes C—Quantitative Methods, D—Microeconomics, and G—Financial Economics.

The Selection of Products Submitted to the VTR

Assuming that research institutions decided to comply with the aims of the evaluation and there was no opportunistic behavior on their side, the selection of products submitted for evaluation can be studied with respect to two main hypotheses. As a matter of fact, institutions’ choice somehow reveals their degree of understanding of the aims of

the VTR, which were characterized by a substantial uncertainty and lack of transparency, as mentioned above. Thus, we consider two alternatives:

- H.1) Institutions may have chosen to submit a *representative* sample of their staff's publications, if they believed the VTR was aimed at fairly representing the state-of-the-art of research in Italy; or
- H.2) Institutions may have chosen to submit their best publications, if they believed that the VTR was aimed at signaling and rewarding *excellence*.

In the absence of a clear consensus on what the criteria are to define scientific excellence in economics, the latter hypothesis is equivalent to affirm that institutions tried to maximize their expected rating, according to the definition of excellence, that, in their opinion, the evaluating panel would adopt. As suggested in the introduction, it is the widespread adoption of the latter criterion that implies a concrete risk for the survival and development of any heterodox approach to economics.

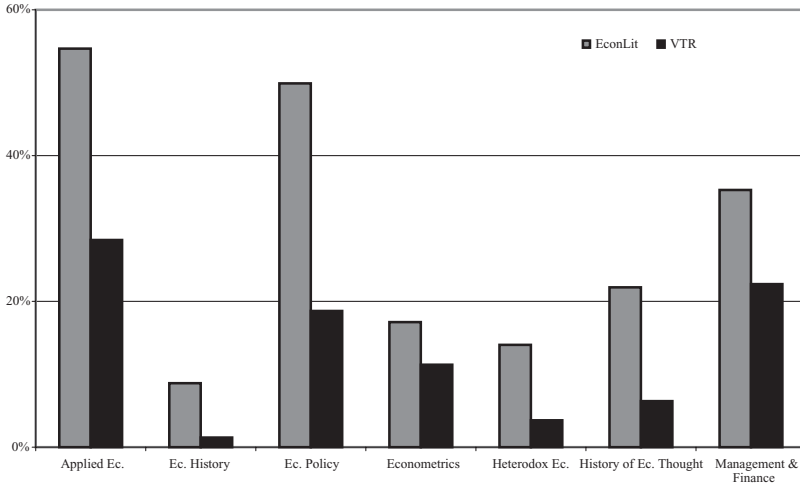
Given the absence of a complete dataset of the whole scientific production by Italian economists, it is virtually impossible to test hypothesis H.1.

However, it is at least possible to investigate the extent to which the distribution of product typologies reflects the distribution of academic staff (whose total number is known), and/or differences in the staff's average productivity. To this aim, we employ the classification of staff proposed above, based on researchers having published at least one product in a certain field in the relevant period. The composition of the research staff in the two datasets is displayed in Figure 4.¹²

Since in the VTR dataset the vast majority of authors are recorded with only one publication, the scope of each author's fields of publication is reduced, with respect to EconLit.¹³ Accordingly, in the VTR dataset all fields exhibit a lower number of active researchers recorded. However, it is clear from Figure 3 that some fields exhibit a larger drop than others. Thus, if the EconLit dataset records a 54.7 percent of authors who in the relevant period published at least once on applied issues, they are reduced to 28.4 percent in the VTR dataset. The respective figures for the other fields are: economic history 8.8 percent as opposed to 1.3 percent, economic policy 49.9 percent and 18.7 percent, econometrics 17.1 percent and 11.3 percent, financial

Figure 4

Distribution of Authors in the Two Datasets



Note. Each author is assigned a field according to the JEL code(s) of his or her publications recorded in the relevant dataset. The EconLit subsample of products with a JEL code contains 2,509 products; the VTR subsample includes 361 products. Publications with more than one JEL code are assigned simultaneously to all the relevant fields; authors are consequently assigned to more than one field. For the sake of clarity, the figure does not include the category “other” (87.8 percent of entries in the EconLit subsample and 83.7 percent in the VTR one).

and business economics 35.3 percent and 22.4 percent. After economic history, the history of economic thought, and heterodox approaches, exhibit the largest differences (relative to their value in EconLit): respectively from 21.9 percent to 6.3 percent and from 14 percent to 3.7 percent.

It is, indeed, possible that the work of these economists was less frequently selected for the evaluation because they are less productive, if they publish a smaller number of visible products with respect to their colleagues. This hypothesis is easily rejected. As shown in Table B1 in Appendix B, the average productivity of Italian economists, measured by the unweighted number of products recorded in the EconLit dataset, is not significantly different across subfields.¹⁴

Since (in total, and for each subfield) the number of researchers, times their average productivity, equals their output, under the hypothesis that the EconLit dataset provides a closer picture of reality than the self-selected sample of products submitted for evaluation, we can affirm that the allocation of products submitted to the VTR does not match the distribution of researchers' output by subfield; products were not selected with the aim to constitute a representative sample of economic research in Italy.

We are thus left with our second hypothesis, according to which research institutions tried to maximize their expected rating. The rationale of this assumption is magnified by the fact that—as was expected—results were presented in the forms of institutions' rankings. The use of “football league” rankings is relevant because evidently if someone comes before, someone else must come next. Thus, rankings introduce, beside the financial incentive to be highly ranked (not being subject to budget cuts from the central government), a further argument related to institutions' reputation (and indirectly a new financial incentive, if students' enrollment choices are affected by public rankings).

Unfortunately, by investigating the visibility of the submitted publications, we are able to collect some evidence that may corroborate this hypothesis, but not to falsify it (or to show under what conditions it would be falsified).

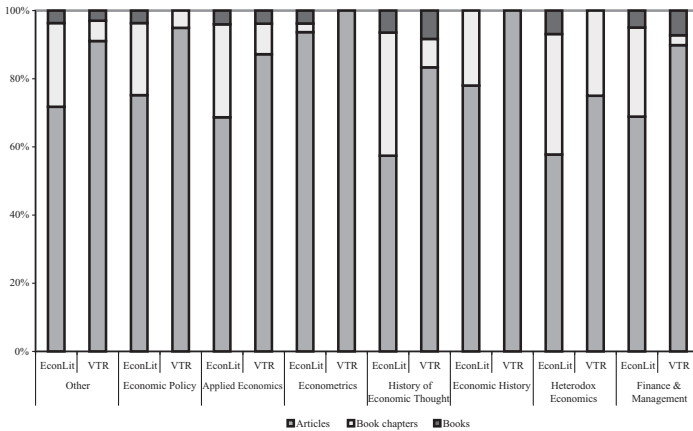
Our hypothesis H.2 respects the principle of parsimony, as it allows us to explain the observed patterns both in terms of publication typology, and of field classification simultaneously. In fact, given the *ex ante* uncertainty on the definition of “research quality,” it would have been rational for research institutions, if they wanted to maximize their expected rating, to submit the most internationally visible products, preferably already “legitimated” by previous peer-review processes (mainly journal articles) and possibly by an impact factor (IF), as a means to signaling quality. In light of the characteristics of economic research in Italy (though there is a presumption that this observation applies elsewhere), this criterion of selection ostensibly favors mainstream journals and quantitative approaches over heterodox and historical research, as well as publications written in English over any other language.

Both “demand” and “supply” reasons may explain the different publication habits across subdisciplines and between different schools and approaches. On one hand, mainstream economics enjoys larger audiences, leading to a higher demand for research products in this area. On the other hand, a nonexclusive reliance on journal articles as the best means for dissemination of results is typical of many heterodox schools, interdisciplinary approaches, and, specifically, of historians of economic thought, as mentioned above.

As it appears from Figure 5, researchers in the latter fields, indeed, exhibit a higher propensity to write books and book chapters. According to the Econlit dataset, these typologies account for more than 40 percent of the total publications in the categories history of economic thought and heterodox approaches. Beyond the quantitative dimension, there is also a qualitative argument, since these publication typologies were frequently selected for evaluation in the two subfields, according to hypothesis H.2, this means that they were perceived as being good pieces of research, liable of being positively evaluated. Indeed, although the high proportion of books and book

Figure 5

Comparison of Datasets, Product Typology by Field



Note: The subsamples include only products provided with at least one JEL code.

chapters is not perfectly respected in the VTR dataset, the history of economic thought is the only field exhibiting (relatively) more books than in the EconLit dataset, while heterodox economics is the field exhibiting the highest proportion of book chapters among the products submitted for evaluation (3.6 times the average).

The Visibility of Research Output

Thus, it appears that, under the hypothesis that institutions tried to maximize their rating, different choice rules appear to have been applied for different subfields. At the same time, institutions seem to have selected subdisciplines according to the rationale of maximizing their prospective rating, acting on the presumption that certain publications (or certain paradigms) are of less quality than others.

It appears, therefore, a sensible question to analyze the extent to which certain publication typologies, and—at the macro level—certain approaches and subfields, actually exhibit a higher international visibility than others; for example, to what extent the institutions' presumptions are correct.

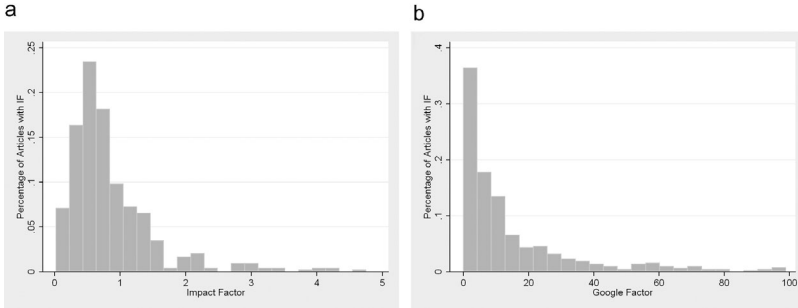
However, initially, it should be stressed that even if the mentioned criteria (peer review and IF) could be good signals of publications' international *visibility*, they would still not necessarily denote research *quality*. The rationale of using visibility as a proxy for quality only rests on the uncertainty over what the quality of economic research is, and on how to measure it. Only under this uncertainty it appears rational to rest on majority, or mainstream, criteria.

We investigate publications' visibility by counting the number of citations they received, as recorded by the Google Scholar search engine (as of August 2007): we label this indicator as the publications' Google Factor (GF). Although not as precise,¹⁵ this measure conveys more information for our purposes than the usual ISI—Web of Sciences database, insofar as product typologies other than articles are thereby considered (most books and book chapters).

Figures 6a and 6b compare the distributions of GF and IF associated with the publications submitted to the VTR. As it appears from Figure 6b, the number of individual citations as measured by the Google Factor exhibits a power-law form of distribution, which is

Figure 6

Distribution of Products Submitted to VTR, by Impact Factor and Google Factor



typical of the bibliometrics literature (compare, for example, Redner 1998; Van Raan 2005). By contrast, the distribution of the IF (for those articles published in journals that have one) exhibits a peculiar log-normal shape: with modal and median values significantly higher, and very few observations at low values of IF (which instead characterize most journals in economics). Although not conclusive, this piece of evidence suggests that the VTR sample is not random with respect to the impact factor, while being biased toward articles exhibiting high impact factors.

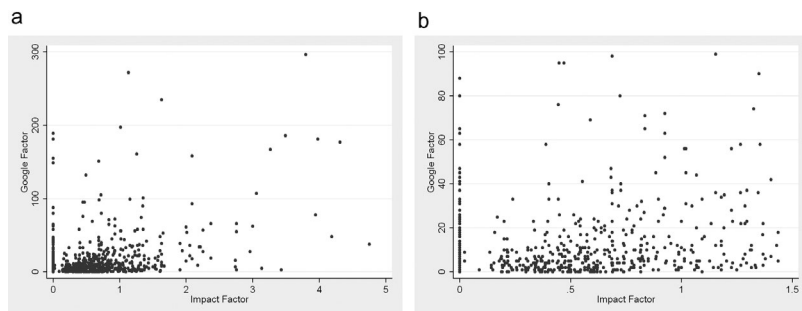
In our sample, the IF and the GF are indeed mildly correlated (the Pearson correlation coefficient is 0.48), although this correlation explains only a small fraction of the total variance. In other words, the presumption that an article's visibility can be predicted by the IF of the journal in which it is published is true on average, but is not very robust (it is frequently false in individual cases) and not substantiated by a strong correlation.

In Appendix B (Table B2) we collect some descriptive statistics on publications' GF and journals' IF by subfield. In general, it is found that there are relevant differences in the average and median IFs and GFs across subfields, as well as a relevant variance within each field.

The large variability of individual citations around the subfields' average values appears even more clearly when considering the

Figure 7

Products Submitted to VTR: Google Factor Plotted Against
Impact Factor



Note: a: The whole VTR sample is considered; b: VTR sample, products with GF < 100.

totality of publications by not limiting the analysis to journal articles. As shown in Figures 7a and 7b (where a “zero” IF was assigned to books, book chapters, and articles published in journals without an IF), a cloud of individual citations prevents us from identifying a strong correlation between the IF and the GF. Moreover, it is worth noting that many articles published in journals with an IF (37 out of 552) received no citations even after four years (the issue was emphasized also by Oswald 2007). Vice versa, we find that several books and articles published in journals without an IF were largely cited on Google Scholar.

Thus, our analysis suggests that the nonexclusive reliance on the two mentioned criteria of publications’ selection that we found as strongly present in the subfields of history of economic thought and heterodox economics are not suboptimal behavior on the side of research institutions. Instead, even if research quality could be proxied by international visibility, the criteria of peer review and IF do not appear as adequate measures of it.

Conclusions: The Current Debate in Italy

While the introduction of the principles of accountability and merit has been identified as the main improvement brought about by the

VTR, above we pointed out the following as its main drawbacks: lack of transparency on the aims and methodology of the exercise, use of “football league” rankings to disseminate its results, insufficient reference to researchers’ productivity, application of the same method indifferently for the evaluation of publications, researchers, and institutions, and finally a foggy definition of research quality.

As the analysis suggests, the last mentioned aspect—the reference to “a value scale shared by the international scientific community”—appears as crucial from the point of view of the preservation of pluralism in economics, and specifically for the survival of heterodox approaches. Luigi Pasinetti, a member of the panel of experts in charge of evaluating the publications submitted for the field of economics, deemed the topic so important to obtain the right to attach a minority report (titled “A note on points of dissent”) to the final report of the panel. He denounced that the words “shared” and “international scientific community” have been interpreted by the majority of the other members of the panel (and the majority of referees) as substantially availing the adoption of closeness to the mainstream as a fundamental criterion of judgment.¹⁶ Guido Tabellini, the chairman of the subgroup of the panel of experts with specific competence in economics, replied to Pasinetti with a further attachment to the final report. After denying that the mainstream of economic research constitutes a unique research paradigm, he stressed that reference to international standards is fundamental, and that “the real risk is that, in order to protect some sects of researchers that are dying out, we will avoid comparisons and we will abstain from discriminating the excellent research, which really moves the frontiers of knowledge, from low-quality research. Or, even worse, that by *a priori* refusing to refer to the international scientific community, we will end up evaluating according to arbitrary criteria, that represent the idiosyncrasies and the prejudices of the members of the panel” (Annex 4, p. 32, our translation).

As Pasinetti points out, the activity of this panel has been characterized by a number of points of dissent, as the members of the panel did not reach a consensus agreement on the rating of about one-third of total products (whose merit grade was decided by majority voting),¹⁷ while in some cases they decided to consult further external

experts in addition to the two referees assigned to each product. A public discussion emerged as a consequence of this disagreement, which involved economists' associations and single researchers, and took place on the Internet as well as in a number of conferences and workshops.

As of May 2009, 128 economists signed an open letter petitioning that pluralism should become a fundamental criterion of research evaluation. The open letter forcefully puts forward the argument that the evaluation of publications should be separated from that of researchers, which in turn should be different from that of institutions. Specifically, it would be necessary to evaluate publications looking at their intrinsic quality, according to the specific standards of each subdiscipline and without considering the typology (journal article, book, book chapter), language, and place of publication.¹⁸

On the other hand, considering the evaluation of institutions, it seems important to come to an agreement about what the final rankings are expected to assess, for example, if they should award the average quality of research, or if they should compare only the best research products (the "excellence"). While the second criterion could be effective in allocating funds to the most promising lines of research, it may also bring about an asymmetric information problem: in fact, if funds are allocated to faculties (and not to single researchers, nor to research projects) those unproductive researchers who work in top-ranked faculties will be awarded without merit. Conversely, if the former criterion is applied, top researchers working with mediocre colleagues could be harmed, but rankings will probably be more indicative to the stakeholders (students, academics, and government, for example) of the average quality of institutions.¹⁹ According to this criterion, the participation of all the academic personnel in the evaluation exercises should be required.²⁰

The present work contributes to this debate, while highlighting Italy as an internationally relevant case study. Beyond the interest *per se*, being the host country of a relevant number of heterodox economists, the case of Italy suggests, then, that heterodox economists in every country should not refuse the evaluation of their research and the award of merit (Tabellini clearly made the point). Moreover, both within the mainstream and within heterodox approaches it should be

clear that some pieces of research or some institutions are—from some points of view, to be strictly defined—better than others. And that the reward of merit is fundamental to set up the right incentives to researchers, as well as to provide the best environment to work in.

However, what our empirical analysis shows is that the rules governing the evaluation of research quality should respect a principle of fair competition. Specifically, although rankings were provided for research institutions as a whole, Italy's VTR did not respect a principle of fair competition among research paradigms, by providing conditions for better awarding research of distinctly mainstream character. Indeed, in the context of an ever-increasing process of multi-authorship across several “competing” research institutions, one could question the idea that only competition (rather than cooperation) takes place among institutions, while it could be claimed that competition among paradigms is more relevant for pluralism, and thus for the development of our discipline. It is, therefore, of prime importance that the criteria and the practice of evaluation do not *a priori* favor any specific paradigm or approach. The case of Italy shows that if such evaluation criteria are not explicitly stated, the large numeric majority currently enjoyed by the mainstream (even in an atypical country as Italy) risks producing a bias against minority stances.

It is consequently fundamental that minority groups (heterodox economists, as well as non Anglo-Saxon writers, and mainstream scholars specializing in less diffused topics or research areas) do not ignore the theory and practice of research evaluation, but engage in the explicit and fair *ex-ante* definition of clear rules. Only transparency and accountability of evaluation criteria and bodies can solve the century-long issue of who judges the judges.

Notes

1. However, we should recall at least Federico Caffè (whose main research fields were economic policy and public finance) and Paolo Sylos Labini (market forms, development, and technical change) in Rome, Giorgio Fuà in Ancona (development), Franco Momigliano in Turin (international economics), Siro Lombardini in Milan (monetary theory), Giacomo Becattini in Florence (industrial organization and local development).

2. Furthermore, our selection of JEL descriptors is affected by the fact that some, potentially relevant, JEL codes (for example, G01—Financial crises)

were only recently added, and are not considered here, since the publications considered in this study only refer to the period 2001–2003, in order to match the period considered by the VTR research assessment exercise.

3. When considering the single field of research in which authors have more extensively written, the following pattern emerges: researchers writing on field theoretical topics (in our classification, frequently falling into the category “other”) account for 84 percent of Italy’s academic staff, while researchers writing mostly on applied topics account for 11.5 percent and historians of thought for 7.3 percent. The rest is represented by minorities accounting for between 0.1 percent (heterodox economists, similarly to economic historians) and 5 percent (financial and business economists). Figures do not sum up to one because in the relevant period several researchers used the same number of JEL codes for different fields. This piece of evidence descends from researchers’ habit to denote their papers with an equal number of “methodological” and “theoretical” descriptors, or of “historical” and “applied” ones, in order not to qualify the paper (and themselves) as belonging to a single specialist field.

4. The differences between the two distributions cannot be imputed to differences in productivity (number of publications per author) because neither distributions sum up to one. Authors typically use more than a unique JEL code category; therefore, they are classified under more than one field; publications themselves usually exhibit more than one JEL code. Moreover, publications are often authored by two or more authors, who may not belong to the same category in our classification.

5. For example, Axaroglou and Theoharakis (2003) report that in 2002, among members of the AEA, roughly the same figure (14 percent) of their respondents self-identified with one of the schools of thought identified by the JEL codes we selected as “heterodox” (with the difference that they included all institutionalist schools, but excluded feminists). However, beside the conservative bias of our estimate, which does not include certain heterodox schools such as the experimentalists or the behavioralists, it should be noted that their respondents were (anonymously) faced with a general question on one’s overall preanalytical vision, for example, not implying as strict of a requirement as our definition does, by demanding to publicly report their vision by selecting a JEL code for at least one publication during the 2001–2003 period.

6. A typical example in the scientific literature, applied to the evaluation of Europe’s research institutions, is the well-known paper by Kalaitzidakis et al. (1999), providing for the first time a comprehensive ranking of European universities in economics, though “based on publications in a core set of highly ranked, *mainstream*, economic journals” (italics added). Astonishingly, throughout the paper the authors themselves seem to perceive their rankings as measures of unconditional research quality, thus ignoring that nonmainstream readers may embrace a different point of view (compare, for example,

the authors' conclusions: "the London School of Economics, Tel-Aviv University, and Oxford University are the three leading economic schools in Europe. . . . These results are independent of the metrics used. . . . This finding raises concerns about the current state and impact of economic research in Europe relative to North-America").

7. Formally, the CIVR was requested to assess progress toward the achievement of the political goals stated in the National Plan for Research, but it actually took the form of an internal evaluation, with little or no attention to the external impact of research, in terms of the impact of research on society.

8. Recently, the government approved a decree mandating that from 2010 up to 7 percent of the state funds yearly transferred to universities will be allocated according to the quality of institutions' teaching, research, and infrastructure (this share should subsequently increase up to 30 percent). For year 2010, 60 percent of this share was determined employing the VTR rankings.

9. In order to maintain comparability, we excluded from our analysis the few Ph.D. theses, working papers, and review articles recorded in the EconLit database because none of these typologies of products were submitted to the VTR exercise for Area 13.

10. It should be noted that a high number of scholarly economics journals published in Italy (almost 40) are included in the Econlit dataset.

11. While the VTR dataset does not entail information on products' language, we can actually confirm that the distribution by product typology of the "financial and business economics" subarea is very peculiar with respect to the other two, exhibiting as many books as 49 percent of the total, and as few journal articles as 39 percent.

12. When considering authors, the EconLit sample comes much closer to the universe than when considering publications: as Marcuzzo and Zacchia (2007) document, more than 70 percent of Italy's economists authored at least one publication recorded by EconLit.

13. For the same reason the possible second definition referred to in note 5, classifying researchers according to the field in which they have published more intensively, is unfit to analyze the VTR dataset.

14. Instead, a high variance is found within groups, significantly higher than the variance across groups. The substantially large skewness of the distributions within groups implies that, in general, average values lie far above median values, for example, within-group averages are heavily affected by single, very productive researchers. It should be recalled that the EconLit dataset encompasses all researchers working in Italy's institutions, but not all their publications.

15. In particular, it does not correct for self-citations and it does not weigh citations according to the visibility of the citing publication.

16. For example, Pasinetti reported that several referees explicitly indicated the IF as a signal of the quality of the evaluated products: “It is first hand documentary evidence from the referees’ reports I had read, before offering (or denying) my cross-panelist consensus. The referees (two for each product) had to fill in answers to 4 different questions concerning: quality, relevance, originality, and internationalization. Cases like the following were the first cause of my denying consensus (without any effect, being always in a minority). Quality of the product: ‘This paper is published in a top field journal, the IF of the journal is high, hence the paper is excellent’ or conversely (always on quality of the paper) ‘this paper is published in my opinion in a non serious journal [in the specific case of this quotation it was the *Journal of Post Keynesian Economics*], hence the quality is limited’. Notice that the evaluation I am referring to is on *quality*, not on *internationalization* of the product!” (p. 6 of Annex 4, italics in original). This is but one example for the argument for the need of a greater transparency of the process, including the publication of referees’ reports, as a fundamental requirement for keeping the evaluators accountable.

17. The Final Report of the Panel for the Area 13 (*Relazione finale di Area*), is available online (only in Italian) at http://vtr2006.cineca.it/rel_area/panel_13.pdf.

18. The petition is available online: <http://www.letteraapertavalutazionericerca.it>. The arguments that emerged in the context of the subsequent public discussion are not completely new in the literature: we refer to the introduction to this special issue and to the other articles collected here for a review of this literature. In our opinion, it is worth noting that in the specific context of peer-review processes, these critiques assume a relevant dimension even within mainstream approaches. Indeed, while the anonymity of peer-review processes is threatened, nowadays, by the availability of research databases and Internet search engines, it is likely that within specific streams of research real experts are competitors, so that referees may be subject to conflicts of interests. Moreover, reliance on evaluation criteria based mainly on international visibility may disproportionately favor consolidated research areas (Geuna 2001; Frey 2003), hindering the development of new fields by young researchers. More generally, a serious threat is posed for “periphery” countries, where such criteria may undermine the diffusion of country-specific analyses, thus reducing the (external) relevance of research in terms of impact on society and the economy (Colander 2008).

19. Two corollaries emerge from this reasoning. First, research excellence should be evaluated separately from overall university rankings (for instance, by granting funds to specific research projects, submitted to ad-hoc evaluating commissions). Second, only the application of an “average” evaluation criterion could encourage a healthy competition among universities in attracting

the best researchers in each field, rather than favoring the mobility of “top” researchers with the only purpose of maximizing expected rankings.

20. The drawbacks of such an approach, when applied to a peer-review evaluation, stand in the costs and in the complexity of the procedure, which in the long run may even outweigh benefits (Geuna and Martin 2003).

References

- Axaroglou, K., and V. Theoharakis. (2003). “Diversity in Economics: An Analysis of Journal Quality Perceptions.” *Journal of the European Economic Association* 1: 1346–1366.
- Cecchi, D., and T. Jappelli. (2008). “Come dare risorse ai migliori”. www.lavoce.info, September 4, 2008.
- Colander, D. (2008). “Can European Economics Compete with U.S. Economics? And Should It?” Paper presented at the “IZA Workshop on Research in Economics: Rewards, Evaluation and Funding,” Bonn (DE), May 26 and 27, 2008.
- Frey, B. S. (2003). “Publishing as Prostitution? Choosing Between One’s Own Ideas and Academic success.” *Public Choice* 116: 205–223.
- Geuna, A. (2001). “The Changing Rationale for European University Research Funding: Are there Negative Unintended Consequences?” *Journal of Economic Issues* 35(3): 607–632.
- Geuna, A., and B. Martin. (2003). “University Research Evaluation and Funding: An International Comparison.” *Minerva* 41: 277–304.
- Giavazzi, F. (2008). L’inizio della fine dei concorsi. www.lavoce.info, November 18, 2008.
- Jappelli, T. (2008). Largo al merito! Ma è solo l’inizio. www.lavoce.info, December 3, 2008.
- Kalaitzidakis, P., T. P. Mamuneas, and T. Stengos. (1999). “European Economics: An Analysis Based on Publications in the Core Journals.” *European Economic Review* 43: 1150–1168.
- Lee, F. S. (2007). “The Research Assessment Exercise, the State and the Dominance of Mainstream Economics in British Universities.” *Cambridge Journal of Economics* 31: 309–325.
- Lee, F. S., and W. Elsner. (2008). “Publishing, Ranking, and the Future of Heterodox Economics.” *On the Horizon* 16(4): 176–184.
- Lee, F. S., and S. Harley. (1998). “Peer Review, the Research Assessment Exercise and the Demise of Non-Mainstream Economics.” *Capital and Class* 66: 23–51.
- Lippi, M., and F. Peracchi. (2007). “Il primo esercizio italiano di valutazione della ricerca: una prima valutazione.” *Rivista italiana degli economisti* 2: 267–276.

- Marcuzzo, M. C., and G. Zacchia. (2007). "L'ECONLIT e gli strumenti per la valutazione della ricerca economica in Italia." *Rivista italiana degli economisti* 2: 277–306.
- Neary et al. (2003). "Evaluating Economics Research in Europe: An Introduction" *Journal of the European Economic Association* 1: 1239–1249.
- Oswald, A. J. (2007). "An Examination of the Reliability of Prestigious Scholarly Journals: Evidence and Implications for Decision-Makers." *Economica* 74(293): 21–31.
- Pasinetti, L. L., and A. Roncaglia. (2006). "Le scienze umane in Italia: il caso dell'economia politica." *Rivista Italiana degli Economisti* 11: 461–499.
- Peracchi, F. (2007). Closing session of the conference "Funziona! Ricetta per Riformare un Sistema Universitario Pubblico. Dal Regno Unito all'Italia." Rome, November 25, 2007, Fondazione IRI.
- Redner, S. (1998). "How Popular Is Your Paper? An Empirical Study of the Citation Distribution." *European Physical Journal B* 4: 131–134.
- Roncaglia, A. (2005). *The Wealth of Ideas*. Cambridge: Cambridge University Press.
- Van Raan, A. F. J. (2005). "Statistical Properties of Bibliometric Indicators: Research Group Indicator Distributions and Correlations." *Journal of the American Society for Information Science and Technology* 57(3): 408–430.

Appendix A

Research products were classified according to the following criteria, based on the JEL classification system.

Applied Economics: C9, C90, C91, C92, C93, C99, L6, L60, L61, L62, L63, L64, L65, L66, L67, L68, L69, L7, L70, L71, L72, L73, L74, L78, L79, L8, L80, L81, L82, L83, L84, L85, L86, L87, L88, L89, L9, L90, L91, L92, L93, L94, L95, L96, L97, L98, L99, O5, O50, O51, O52, O53, O54, O55, and the whole categories R and P.

Corporate Finance and Management: all the JEL codes included under the letters G and M.

Economic History: all the JEL codes included under the letter N.

Econometrics and Quantitative Methods: all the JEL codes included under the letter C, with the exception of C7, C70, C71, C72, C73, C78.

Economic Policy: D18, E5, E50, E51, E52, E58, E59, E6, E60, E61, E62, E63, E64, E65, E66, E69, F13, F33, F34, F35, F42, I18, I28, I38, J18, J28, J38, J48, J58, J68, J78, J88, L4, L40, L41, L42, L43, L44, L49, L5, L50, L51, L52, L53, L59, O2, O20, O21, O22, O23, O24, O29, O38, Q18, Q28, Q38, Q48.

Heterodox Economics: B5, B50, B51, B52, B53, B59, D57, E11, E12.

History of Economic Thought: all the JEL codes included under the letter B, excluding B5, B50, B51, B52, B53, B59.

The residual category **Other:** C7, C70, C71, C72, C73, C78, L1, L10, L11, L12, L13, L14, L15, L16, L17, L18, L19, L20, L21, L22, L23, L24, L25, L26, L27, L28, L29, L3, L30, L31, L32, L33, L34, L35, L36, L37, L38, L39, L3, O10, O11, O12, O13, O14, O15, O16, O17, O18, O19, O1, O3, O30, O31, O32, O33, O34, O35, O36, O37, O38, O39, O4, O40, O41, O42, O43, O44, O45, O46, O47, O48, O49, as well as the whole categories A, D, E, F, H, I, J, K, and Z.

Appendix B

Table B1

Output of Researchers in Economics: Number of Publications Recorded in EconLit, 2001–2003

	Applied Ec.	Ec. Policy	Other	Heterodox Ec.	Econo- metrics	History of Thought	Ec. History	Finance Manag.
All Publications								
<i>Mean</i>	6.6	6.6	5.8	7	7.1	6.9	9.1	7.3
<i>Median</i>	5	5	4	6	5	6	8	6
<i>Std. Var.</i>	5.5	5.3	5.1	5.3	5.7	5.3	5.4	5.8
<i>Skewness</i>	1.6	1.6	1.9	1.4	1.4	1.2	0.6	1.5
<i>Min</i>	1	1	1	1	1	1	1	1
<i>Max</i>	27	27	27	22	22	22	22	27
<i>5%</i>	1	1	1	1	1	1	2	1
<i>95%</i>	18	17	17	20	20	20	22	20
Journal Articles								
<i>Mean</i>	4.1	4.3	3.7	3.6	4.3	3.4	5	4.6
<i>Median</i>	3	3	3	3	4	3	4	4
<i>Std. Var.</i>	3.7	3.8	3.4	2.6	2.7	2.5	3.4	3.9
<i>Skewness</i>	2.4	2.3	2.5	0.7	0.9	0.7	0.7	2.3
<i>Min</i>	0	0	0	0	0	0	0	0
<i>Max</i>	22	22	22	10	12	10	13	22
<i>5%</i>	0	1	0	0	1	0	1	1
<i>95%</i>	11	11	10	8	8	8	12	11
Book Chapters								
<i>Mean</i>	2	1.9	1.7	2.9	2.5	2.9	3.6	2.3
<i>Median</i>	1	1	1	2	1	2	4	1
<i>Std. Var.</i>	2.8	2.4	2.5	3.4	3.7	3.2	3.4	2.9
<i>Skewness</i>	2	1.6	2.3	1.8	1.8	1.6	1.5	1.9
<i>Min</i>	0	0	0	0	0	0	0	0
<i>Max</i>	14	10	14	14	14	14	14	14
<i>5%</i>	0	0	0	0	0	0	0	0
<i>95%</i>	8	8	2	10	10	10	14	8
Books								
<i>Mean</i>	0.4	0.4	0.4	0.5	0.4	0.6	0.5	0.5
<i>Median</i>	0	0	0	0	0	0	0	0
<i>Std. Var.</i>	1	1	0.9	1.3	0.9	1.2	0.7	1.1
<i>Skewness</i>	3.8	3.8	4.1	4.2	2.3	3.7	1.1	3.6
<i>Min</i>	0	0	0	0	0	0	0	0
<i>Max</i>	8	8	8	8	3	8	2	8
<i>5%</i>	0	0	0	0	0	0	0	0
<i>95%</i>	2	2	2	3	3	3	2	2

Source: Our elaboration on EconLit dataset, years 2001–2003.

Table B2

Impact Factor and Google Factor: Descriptive Statistics

	Publications	Mean	Median	Std.Dev.	Skewness	Min	Max	5%	95%
Impact Factor									
<i>Applied Economics</i>	56	0.875	0.6865	0.643	3.051	0.217	4.312	0.236	2.087
<i>Economic Policy</i>	47	0.898	0.723	0.757	2.165	0.089	3.795	0.2	3
<i>Economic Theory</i>	230	0.887	0.7045	0.716	2.344	0.089	4.756	0.2	2.196
<i>Heterodox Economics</i>	5	0.501	0.444	0.124	0.699	0.403	0.688	0.403	0.688
<i>Econometrics</i>	27	0.78	0.62	0.531	1.995	0.222	1.315	0.24	1.315
<i>Economic History</i>	4	0.644	0.674	0.264	0.373	0.297	0.929	0.297	0.929
<i>History of Ec. Thought</i>	10	0.48	0.301	0.477	1.111	0.022	1.333	0.022	1.333
<i>Finance & Management</i>	49	1.101	0.806	0.823	1.443	0.135	3	0.272	3.494
Google Factor									
<i>Applied Economics</i>	78	25.846	9	43.138	3.431	0	272	0	93
<i>Economic Policy</i>	59	27.203	10	47.764	3.808	0	296	0	80
<i>Economic Theory</i>	302	23.139	9	40.431	3.704	0	296	0	93
<i>Heterodox Economics</i>	12	3.25	2.5	2.8	1.134	0	10	0	10
<i>Econometrics</i>	34	17.206	7.5	20.095	1.475	0	69	1	69
<i>Economic History</i>	4	12.75	4.5	18.209	1.138	2	40	2	40
<i>History of Ec. Thought</i>	24	3.5	1.5	6.84	3.551	0	33	0	9
<i>Finance & Management</i>	69	27.623	12	39.149	2.356	0	186	0	101
Google Factor—Articles Only									
<i>Applied Economics</i>	68	27.132	10.5	44.817	3.41	0	272	0	93
<i>Economic Policy</i>	56	28.464	11	48.719	3.714	0	296	0	80
<i>Economic Theory</i>	275	23.96	10	40.959	3.729	0	296	0	95
<i>Heterodox Economics</i>	9	3.333	2	3	1.339	1	10	1	10
<i>Econometrics</i>	34	17.206	7.5	20.095	1.475	0	69	1	69
<i>Economic History</i>	4	12.75	4.5	18.209	1.138	2	40	2	40
<i>History of Ec. Thought</i>	20	2.5	2	2.856	1.147	0	9	0	9
<i>Finance & Management</i>	62	27.258	12	37.232	2.449	0	186	0	98
Google Factor—Articles with Impact Factor Only									
<i>Applied Economics</i>	56	31.179	14	48.237	3.103	0	272	0	151
<i>Economic Policy</i>	47	29.404	12	46.478	4.187	0	296	1	76
<i>Economic Theory</i>	230	26.504	11	42.542	3.567	0	296	1	95
<i>Heterodox Economics</i>	5	3.2	3	1.924	0.396	1	6	1	6
<i>Econometrics</i>	27	19.593	8	21.855	1.16	0	69	1	69
<i>Economic History</i>	4	12.75	4.5	18.209	1.138	2	40	2	40
<i>History of Ec. Thought</i>	10	3.7	3	3.268	0.664	0	9	0	9
<i>Finance & Management</i>	49	33.327	19	39.68	2.169	0	101	1	186

Note: Only publications provided with at least one JEL code are considered.