

Monitoring, Measuring and Assessing the Italian Ph.D. System

Antonella D'Agostino, Giulio Ghellini, Laura Neri

Abstract This paper deals with several issues concerning the monitoring, measurement and assessment of the Ph.D. system in Italy. The lack of an adequate monitoring system stimulates a meaningful discussion on the need to create a national information system on Ph.D. courses and students, that has to be based on a longitudinal survey design enriched by an efficient and homogeneous set of administrative information. A general framework for performance indicators with some examples are also presented.

Keywords: *Statistics in educational research, Ph.D assessment, longitudinal design*

1 Introduction

The international dimension, in particular at the European level, of the Higher Education (HE) systems, connected with the great financial crisis of the western economy, imposes on each country the need to build up tools and procedures for a transparent accountability of training processes, both from the efficacy point of view and from the efficiency one.

In fact, it is clear that the foreseen scenario of the fulfilment of the European Higher Education Area (EHEA), started in the 2000 with the Bologna process through an increase of the cross national competition between institution, asks for improved evaluation and assessment system (as an example of that, see the European Standard

¹ Antonella D'Agostino, University of Naples "Parthenope": antonella.dagostino@uniparthenope.it

Giulio Ghellini, University of Siena: ghellini@unisi.it

Laura Neri, University of Siena: neri@unisi.it

This paper is supported by PRIN 2008 "Misure, modelli statistici e indicatori per la valutazione del sistema universitario" coordinated by Prof. V. Capursi (Univ. Palermo).

and Guidelines proposed by ENQA, 2006), useful not only as quality assurance tools, but also for making feasible homogenous comparisons across institutions and countries.

In this regard, in Italy, it is evident that the availability and transparency of information and indicators for the whole HE system is very poor and heterogeneous across institutions. This drawback can be extended to secondary school and unfortunately seems to be a typical trait of Italian informative system on Education. If that is true for the education system as whole, it is true *a fortiori* for the third level of the HE, the Ph.D. In particular, it seems rather worrying that also for this selected population - characterized by the higher level of public investment both in terms of time and resources - the collection of comprehensive set of information regarding the path from the training process to the labour market entrance has been so lightly pursued.

First of all, the existing set of indicators for Ph.D. assessment at national level (see for example, Comitato Nazionale per la Valutazione del Sistema Universitario - CNVSU (2007), (2009)) is not fully informative and pertinent. In addition, measures are obtained with heterogeneous procedures, are often not completely reliable and are affected by a certain degree of subjectivity. Furthermore, data on graduates' transition to labour market are very poor and irregular. Only in the last year we can find some partial attempts by Istat (Indagine Statistica sull'Inserimento Professionale dei Dottori di Ricerca degli anni 2004 e 2006), and by the CNVSU to collect more homogenous information on labour experience of Ph.D. graduates via cross institutional surveys. What is still missing also in these interesting attempts, in our view, is a wider and systemic approach to the Ph.D. seen as a whole. We think that the need of an adequate monitoring system, thought out for measuring both the training process development and assessment, and the output produced in terms of labour market entrance and quality of graduates jobs, has to be put in the agenda. With this paper we wish to start a useful path for going towards the creation of a systemic framework for Ph.D. courses assessment, which is essentially based on longitudinal design possibly enriched by an efficient and homogenous set of administrative information.

To do this, the paper starts with a synthetic overview of the existing data and indicators on the Ph.D. system (Section 2), followed by a definition of the main aspects and contents of the information system that we wish to propose for allowing an adequate assessment of the whole system. In particular, we identify the main dimensions that could be useful to investigate, the distinction between different levels of potential stakeholders of the produced data, and the features of the longitudinal design that is needed for achieving plausible data and indicators (Section 3). After an outline of the general framework for indicators definition with some explicative proposals (Section 4), some brief final remarks end the paper (Section 5).

2 What we know on Ph.D. system in Italy

In Italy, the available information existing on Ph.D. programmes show many shortcomings which are mainly due to an information system not able to monitor, measure and assess the educational process as a whole. Of course, many data on it are

available, but often such data are characterized by a low information content and a great deal of heterogeneity across institutions.

The central institution (MIUR) requires a final report from each university on the results of the evaluation of the Ph.D. courses. On the basis of these reports CNVSU prepares a comprehensive document regarding the annual state of the Ph.D. programmes and educational evaluation procedures adopted. Indeed the resulting assessment of Ph.D. courses has a great policy importance, since MIUR financial resource allocations are related to it. The evaluation process refers mainly to the six criteria/indicators defined by CNVSU (D.M. n° 290 10/10/2003), which seem not to be well specified for having a complete and homogeneous assessment of the courses. For example, the sixth criterion states that each Ph.D. course has to satisfy the course objectives, also in relation to professional outlets and to the achieved skills obtained during the Ph.D. training; such preposition presumes that each university can be able to relate information on the Ph.D. process and outcomes. However criteria/indicators referring to outcomes are often missing. In other cases, universities have a local monitoring system on Ph.D. outcomes and training; however, often based on *ad hoc* procedures and surveys not well documented, which tends to make them non-comparable in terms of aims and contents, collection methods and questionnaires, reference times (see for example experience reported by University of Pavia, Milano Bicocca-Trento and Siena).

Recently, some efforts have been made, trying to create a more homogeneous information system at national level. We particularly refer to two initiatives started at the end of 2009. One is the survey conducted by Istat to respond to the request for information on Ph.D. working careers formulated by OECD in the framework of the CDH (Careers of Doctorate Holders) Project; the other is promoted by the Italian CNVSU. Both are going to collect names and addresses of Ph.D. graduates from all the Italian universities for three different cohorts, with the aim of undertaking a national retrospective survey on their occupational situation.

As regards the information about applicants and the selection process for entering the Ph.D. programmes, the available data are seriously deficient. On the official website of MIUR one can find only the total number of students who apply and the number selected by gender and nationality (Italian vs not Italian). Also for the information regarding the training process the picture is quite fuzzy. In fact, the existing indicators required by the MIUR for the evaluation of the Ph.D. training are more related to a formal and bureaucratic exercise than to an effective assessment of the training supply and of the students achievement.

3 Monitoring the Ph. D.

In this section we wish to underline the need for a comprehensive and homogeneous information system able to provide consistent indicators, starting from the selection process of candidates to their job experiences. To address this proposal we introduce the following aspects: a) dimensions which fit the Ph.D. system; b) reference stakeholders for the indicators on Ph.D. system; c) information system useful for

computing indicators or making performance analysis. Each of them are described in details below.

a) Dimensions: for the definition of them we start from the idea of Ph.D. system seen both as a process and as a product (Ghellini *et al.*, (2009)). Considering firstly the Ph.D. as a process, we need to have detailed information not only on the training process but also on the input factors.

This statement leads us to define the first dimension, as “*WHO studies for a Ph.D.*” (*WHO*). The building of a national database on Ph.D. applicants and students allows many multipurpose analyses useful for different stakeholders of the system. For example, background information like age, gender, educational patterns/curricula, family characteristics and also initial motivations/expectations for Ph.D. training can be considered as control or explanatory variables for performance indicators. The collected information can also be used for studying the following two populations related to the WHO dimension of the Ph.D. process.

- i. *Applicants*. It is necessary to have detailed and comparable information on Ph.D. applicants in order to know their profiles and the reasons (motivations/expectations) supporting the decision to enter the third level of education (see Green and Powell (2007) and Salmon, (1992), as examples of Ph.D. population profile in UK). Such information can also be used for monitoring and measuring different aspects such as: changes of applicants' profiles over time or across scientific fields and universities to assess “attractiveness” both in terms of quality and quantity; mobility between scientific fields or universities to assess “openness”² perceived by the potential alumni of each doctorate; dynamics of motivations and expectations across cohorts.
- ii. *Selected Candidates*. Information on them, already known from applicants, is useful not only for having their profiles but, in particular, for understanding how the selection process works. This last aspect can be useful to study the actual “openness” of each doctorate and its capacity to select the best candidates.

The second dimension is defined as “*HOW Ph.D. training performs*” (*HOW*). This dimension regards measures for the assessment of the whole training period that can be viewed from two different points of view. The first one regards the organizational context with objective information derived mainly from administrative archives. Of course, this means the setting up of a common set of rules, accordingly defined among all the HE institution involved, for achieving a common set of information. The other one relates to the evaluation of the specific training experiences by Ph.D. students mainly collected from *ad hoc* surveys.

- iii. *Functioning of the Training*. Information should cover the available facilities, public and private grants and funds, course curricula, teaching supply and its *modus operandi*, research activity and production, international research network, rules adopted to verify periodically student skills, etc.. Such information can be very useful for analysing and describing the training process and for assessing the Ph.D. research product.
- iv. *Students Perception*. It is important to study how the Ph.D. students perceive the organizational context and the activity during the training. With these aims *ad hoc* surveys, for measuring the students' satisfaction on the aspects mentioned in the

² By “openness” we mean the predisposition of each doctorate programme and selection procedure to attract students coming from other scientific fields, universities or foreign countries.

above paragraph, should be undertaken. Moreover, in this context it can be useful to collect information regarding the students' subjective evaluation of the acquired skills (critical thinking, computing, information retrievals, writing, etc.), of research capability and of changes in motivations/expectations.

The last dimension concerns the outcomes of the Ph.D. system, that can be labelled "*WHERE Ph.D. graduation leads to*" (*WHERE*). With this dimension we want to summarize what a Ph.D. programme is able to produce in terms of increasing human capital. Also for this dimension it appears reasonable to consider, on one side the job history (objective point of view), on the other the individual perception and evaluation of the post-Ph.D. career path (subjective point of view). Consequentially the following two aspects of the *WHERE* dimension can be considered.

- v. *Job history*. Data collected can regard information on school to work transition, such as occupational status, waiting time to obtain the first job, employment episodes and their characteristics (duration, professional position, type of employment contract, degree of the use of acquired skills, wages, etc.). That should make possible to compute indicators on job placement, returns of education, job mismatch, etc. in order to assess efficacy of the Ph.D. programme.
 - vi. *Satisfaction of career paths*. For measuring job satisfaction individual data on both monetary and non monetary side have to be collected. In particular, for the non monetary one it is important to have available indicators on how Ph.D. training impacts on several aspects of the job activity (for example, involvement in research activity, contribution to innovation, job security and career path).
- b) Reference stakeholders: measures and indicators derived from the dimensions described above can form a shared frame on Ph.D. education performance for different stakeholders. Of course, for each of them the interest can focus on a subset of measures computed at different levels of aggregation and only partially overlapped. Referring to the Italian Ph.D. system, let us aggregate the stakeholders as follows.
- i. *Central HE Institution*. It should be interested mainly in a more efficiently allocation of resources among local institutions, using benchmark indicators and defining standards of excellence.
 - ii. *Local HE Institution*. Its interest should be focused mainly on efficient management of resources. This means the availability of indicators not only for optimal financial allocation of resources but also for monitoring the Ph.D. organization, teaching supply, research activity and outcome.
 - iii. *Users*. Their interest should be addressed mainly to the outcomes offered by the Ph.D. system. In particular, Students need indicators useful for making their choices related both to training/job satisfaction and job placement, while firms and institutions (Employers) are more interested in measures on the level of graduate excellence and on acquired skills.
- c) Informative system: in order to analyse the dimensions introduced above and to provide the stakeholders proper indicators for addressing their own interests and strategy, it is necessary to create an integration between administrative data and *ad hoc* surveys based on a longitudinal design that collect information at micro level before, during and after the Ph.D. Prerequisites of such design are the definition of a unique ID for the micro unit (students) and for all Ph.D. courses.

As regard to the dimension *WHO*, needed data can be obtained using the application forms that all the candidates must complete. The layout of this application form, defined by the central HE institute, should be common for all the universities, at least

for a core part of it, containing the background and motivation/expectation variables quoted before. Moreover, at this step all the contact addresses (telephone number, e-mail, etc.) should be recorded.

Referring to the dimension HOW, we suggest that information regarding *Functioning of the Training* should be recorded in administrative archives according to a common format, aiming at describing the structure, the organization and the training characteristics of each Ph.D. course. *Students Perception* data should be collected by mail/web questionnaires, during the training (for example at the end of the first year and at the end of the curricula period), asking Ph.D. students to evaluate the different phases of their experience. In addition some administrative data can be collected, developing consistent mechanism to continually record the alumni activities (publications, project participations, conference participations, research activities abroad, research awards), and for updating the contact addresses.

Finally concerning the dimension WHERE, a comprehensive perspective survey has to be designed with the aim of tracking post-doctoral career developments. For the proposed design we suggest the use of CAWI technique for data collection and three waves respectively at 1, 2 and 5 years from the degree (Ghellini, (2009); Ghellini *et. al* (2009)). At each wave of this survey, besides variables on job history and satisfaction described above, information on relevant changes in family status (marriage, children, etc.) should be collected for analysing the impact of demographic events and choices on work careers, in particular for gender analysis.

4 General framework for indicators definition

Besides the official criteria and indicators defined by CNSVU for assessing the Ph.D. programme (Section 2), an effort at defining a wider set of indicators can be found in Bianchetti *et al.*, (2000, 2004). The main deficiency of these proposals pertains to the use of data collected for administrative purposes by universities, often not comparable in content and accuracy. On our view, this situation highlights the lack of a reference framework useful for having a more precise picture of the Ph.D. system as a whole.

Starting from this statement, we think that before the definition of a relevant set of indicators, it is necessary to introduce a cross reference frame between the dimensions and the stakeholders, for clarifying the different cognitive purposes which the indicators have to serve. Such cross reference frame is presented in Table 1, where the grades of grey indicate the relevance of each dimension with respect to each stakeholder: darker cells indicate higher levels of relevance.

In each cell different indicators can be defined considering the specific purposes of stakeholders; that implies the introduction of both a WITHIN analysis, i.e. comparisons of Ph.D. courses within the same university, and a BETWEEN analysis, i.e. comparisons of Ph.D. courses belonging to the same scientific field among different universities. For example, the WITHIN analysis could be an useful tool for *local institutions* for making more informed strategic choices. By contrast, BETWEEN analysis could be used, in particular by the *central institution* and *users* for detecting best practices and excellences. In general, the indicators have to be computable at different levels of aggregation.

Table 1: Cross reference frame by Dimensions and Stakeholders

Dimensions		Stakeholders			
		Central HE Institution	Local HE Institution	Users S*	Users E*
WHO (W)	Applicants (A)				
	Selected Candidates (S)				
HOW (H)	Functioning of the Training (T)				
	Students Perception (P)				
WHERE (WH)	Job History (J)				
	Career Satisfaction (S)				

*S= Student, E= Employers

Formally, for each indicator let be l the index for the l -th scientific field ($l = 1, \dots, L$), k the index for the k -th university ($k = 1, \dots, K$), j the index for the j -th Ph.D. course ($j = 1, \dots, J_{kl}$), and i the index for the i -th student ($i = 1, \dots, n_{jkl}$). Some examples should clarify the approach. Considering the WHO dimension for Applicants (WA) and the local institution as stakeholder, for each k -th university, at least the following set of measures should be considered for assessing the attractiveness of Ph.D. courses.

- 1) A measure of the *general attractiveness* (WAA), defined as the ratio between the number of applicants (nA) and the number of Ph.D. grants.
 - 2) A measure of the *excellence attractiveness* (WAEA), defined as the ratio between the number of applicants with maximum degree mark (nA^*) and the total number of applicants.
 - 3) The *national attractiveness* of the k -th university (WANA), defined as the ratio between the number of Italian applicants having achieved the degree out of the k -th university and the total number of applicants.
 - 4) The *international attractiveness* (WAIA), defined as the ratio between the number of foreign applicants and the total number of applicants. Starting from these measures, specific indicators can be derived using convenient standardization of them.
- For example, for excellence attractiveness the following two indicators can be derived.

$$a) \text{ WAEA}_{jkl}^{(k)} = \frac{nA_{jkl}^*}{nA_{jkl}} \bigg/ \frac{nA_k^*}{nA_k} \quad b) \text{ WAEA}_{jkl}^{(l)} = \frac{nA_{jkl}^*}{nA_{jkl}} \bigg/ \frac{nA_{kl}^*}{nA_{kl}}$$

The former can be useful for ranking Ph.D. courses within the k -th university and for identifying excellences independently from scientific field, instead the latter for ranking Ph.D. courses conditionally to scientific area.

Other interesting examples can be provided for the WHERE dimension - Job History (WHJ). A preliminary list of measure can includes the following.

- 1) Waiting time for the first job (WHJJ1), defined as the number of months to obtain the first job after the Ph.D. degree.
- 2) Employment rate (WHJE) at different point in time (for example, $t=12, 24, 60$ months).
- 3) Non permanent job rate (WHJNP), defined as the ratio between the employed with term contract and the total number of employed.
- 4) Median annual labor income (WHJLI).

Considering, in particular, the Users Student stakeholders, indicators can be defined standardizing the previous measures at course level with respect to the

equivalent measure at national level for the same type of Ph.D. course. For example, for the employment rate, the following indicator can be defined:

$$WHJE_{jkl}^{(k)} = \frac{nE_{jkl}}{nD_{jkl}} \bigg/ \frac{nE_k}{nD_k}$$

where nE is the number of employed and nD is the number of Ph.D. graduates.

5 Final remarks

In this paper we address some general issues related to the assessment of the Italian Ph.D. system. In particular the central idea concerns the development of a systemic framework in which a comprehensive longitudinal approach has been adopted. In fact, in our opinion only in this context pertinent indicators can be detected. We know that this could be a long and not an easy path to pursue, but at the same time we think that, this is the best way to go towards a real Ph.D. performance measurement. If it were possible to find an agreement on this approach, next step of the research will pertain to the operative definition of shared survey instruments and to the proposal of a complete list of core indicators.

References

1. Bianchetti M., Lanzoni M., Usai M.C., Proposta di un insieme di indicatori per la valutazione dei dottorati di ricerca nell'ambito della rilevazione CNVSU Nuclei 2000, ADI - Associazione Dottorandi e Dottori di Ricerca Italiani (giugno 2000), 2000
2. Bianchetti M., Della Ratta F., Lanzoni M., Pischedda V., Rizzo R., Usai M.C., ADI - Associazione Dottorandi e Dottori di Ricerca Italiani, Progetto CNVSU per la ricognizione, raccolta e analisi dei dati esistenti sul dottorato di ricerca e per l'indagine sull'inserimento professionale dei dottori di ricerca – Rapporto intermedio: proposta di un insieme di indicatori e di relazione standard per il monitoraggio e la valutazione dei dottorati di ricerca. Prima ricognizione, raccolta e analisi dei dati esistenti sui dottorati di ricerca.”, 28 febbraio 2004
3. CNVSU, Decimo Rapporto sullo Stato del Sistema Universitario, 2009
4. CNVSU Relazione sullo stato della didattica nei corsi di dottorato di ricerca dell'a.a. 2005/06 e proposta di criteri per la ripartizione delle risorse finanziarie, Doc. 8/07, dicembre 2007
5. CNSVU, D.M. 10 ottobre 2003 “Criteri per la ripartizione tra le Università e gli Istituti Universitari delle risorse destinate al finanziamento delle borse di studio per i corsi di dottorato di ricerca e per i corsi post-laurea, per l'anno finanziario 2003”, G.U. n. 290 del 15 Dicembre 2003
6. ENQA, European Standard and Guidelines, 2006, www.enqa.eu/pubs_esg.lasso
7. Ghellini G., Indagine sulle esperienze lavorative dei dottori di ricerca dell'Ateneo di Siena, La Formazione Universitaria, UnifePress, 2009, 121-137
8. Ghellini G., Neri L., D'Agostino A., Towards a longitudinal survey design for Ph.D. evaluation, Quaderni di Statistica, vol.11, 2009, 127-143
9. Green H. and Powell S. Doctoral education in the UK, in Powell, S. and Green, H. (eds.) *The doctorate worldwide*, Maidenhead; Society for Research into Higher Education and Open University Press, pp. 88-102, 2007
10. Salmon P., *Achieving a Ph.D.: Ten Students' Experiences*, Stoke-on-Trent: Trentham Books Limited, 1992