

Comparing Universities: A Case Study between Canada and China

Daniel W. Lang and Qiang Zha

Department of Theory and Policy Studies, Ontario Institute for the Study of Education, 252,
Bloor Street West, Toronto, Canada M5S 1V6.
E-mail: dan.lang@utoronto.ca

Diversity among institutions or, at least, institutional types, is a policy objective that most systems of higher education pursue although to different degrees. North American colleges and universities, for example, are among the most diverse in the world. In the East, Chinese universities have until recently tended to be highly homogeneous. At the same time, both systems are concerned about equity of access, the quality of educational opportunity, and the role of higher education in worldwide economic competition. Individual institutions, for a variety of reasons, ranging from accountability to the allocation of scarce resources, attempt to compare or 'benchmark' themselves against other institutions. Although diversification and benchmarking involve measurement, classification, and the selection of peers, they often work against one another. League tables and rankings, which are the most evident and accessible manifestations of benchmarking, tend to encourage isomorphism instead of diversification among universities. This irony becomes a particular problem when comparisons are attempted between Western and Eastern systems of higher education. This study is an effort to move beyond or at least refine benchmarking by devising a new paradigm based on peer selection. The study is a collaborative venture involving the University of Toronto, in Canada, and Tsinghua University in the People's Republic of China.

Higher Education Policy (2004) 17, 339–354. doi:10.1057/palgrave.hep.8300061

Keywords: Higher Education; methodology; International Comparison; Techniques of Comparison; Institutional level.

Introduction

Since the 1990s, Chinese higher education has been experiencing a large-scale and systemic structural reform. Different from the higher education system reorganization in early 1950s, this wave of restructuring features increases in institutional autonomy and enhancements the institutional capacity for self-development. Since late 1990s, the cost-sharing and cost-recovering policies

within Chinese higher education, and the consequent rapid increase in its aggregate size, have attracted public attention to performance of the institutions. This would require not only diversification of higher education governance at a macro-level but also efficiency and effectiveness of internal management at a micro-level. Without doubt, these are exactly the goals of the current higher education reform that would need to make comparisons between/among institutions when implementing the reform. Put in another way, the comparison is necessary no matter for the sake of measuring diversity of the system, for assessing performance of the institutions, and for providing models or paradigms for system and institutional reform.

Meanwhile, the Chinese government has committed to build the world-class universities by launching the Project 21/1¹ and 98/5,² which, however, results in much attention and controversy as to methods and indicators used to assess the performance of these high-profile universities as well as fund them. The question that followed is: what is a world-class university? How do we know one when we see it? These projects are themselves a reflection of diversification of Chinese higher education system, and would require a world vision of university comparison.

In this regard, the presidents of the elite universities of China have probably the truest sense. While they target their universities to provide elite education, or becoming comprehensive, research-intensive, and open institutions, they suffer from a lack of systematic criteria by which they can define world-class in terms that are applicable and relevant to their institutions. They have few means of making comparative assessments from system to system or from country to country. In other words, it is easy to reach a consensus as to the notion that universities need to be compared, but it is not easy to implement the comparison in a way that leads to diversity. Academic researchers tend to contribute to the controversy. While some scholars argue that the gap between the world-class universities and best universities in China is widening, others believe the opposite. There are also scholars who are suspicious of comparisons between Chinese universities and their counterparts elsewhere.

The authors and their participating universities believe there is a need for better means of making comparisons and benchmarking universities, given the goal of building world-class universities in China. It is significant also in terms of raising the management efficiency of Chinese higher education system and its institutions. Based on this understanding, this article provides a brief account of theories and methods with respect to a specific form of university comparison, peer selection, in Western higher education literature, then focuses on a case study of the peer selection practice of the University of Toronto in Canada, and attempts to depict the implications for Chinese universities.

The Significance of Peer Selection

Strategic planning

Comparison and emulation are components critical in institutional strategic planning. Peer comparisons can provide a basis for the rational evaluation of differences and of similarities among institutions, and of identifying relative strengths, weaknesses, and possible opportunities or niches.

Mission statements are often vague or abstract statements about institutional goals and priorities (Lang and Lopers-Sweetman, 1991). Comparative analysis can help institutions delineate their own identity in more concrete terms. In this regard, such comparisons can be a helpful antidote to external funding and coordination efforts that, deliberately or inadvertently, blur useful distinctions among institutions within a given jurisdiction.

Strategic planning is about a higher education institution's future aspirations and realistic possibilities. Throughout the research literature on strategic planning there are frequent references to environmental scanning (Bryson, 1988) for the purpose of identifying opportunities, challenges, and the best fits between what the institution is and what its sponsors, users, or beneficiaries wish it to be. Logically, the environment to be scanned for any given institution could have wide and quite indefinite boundaries, so broad and so uncertain as either to defeat scanning or to render it meaningless. By determining its peers, an institution can give shape to its environmental scanning exercise. The concept of environment becomes extremely complex when strategies involve comparisons among institutions in systems that have deeply different cultural roots, as in the case of institutions in Canada and China.

Just as some mission statements are vague and abstract, others are about aspirations, which may or may not be realistic or practicable (Lang and Lopers-Sweetman, 1991). One might think of this means of expressing an institutional strategy as definition by association, whether or not there is a sound basis in fact for the association. The key, then, to an aspirational approach to determining institutional strategy is to confine or direct aspiration to institutions that, on the basis of comparative data, seem to share a given college or university's mission generally, but appear to be more successful in achieving it. Aspiration, however, is often a core purpose of comparison between different national systems of higher education.

Alternatively, a given institution could postulate a different role for itself in the future by defining a 'desired institution' containing targets for factors that are potentially controllable by it in the long term. Those factors might include, for example, total enrolment, graduate share of total enrolment, a balance between part-time and full-time balance, library size, instructional program mix, and targets for external circumstances that the college or university might try to have changed, and then use a peer selection methodology to identify

those institutions most similar to this desired institution. The institutions thus identified become a benchmark or milestone against which that specific institution can measure its progress.

Assessment of an institution's performance

In the absence of absolute standards or frames of reference in higher education for the evaluation of institutional performance, governors and administrators understandably tend to turn to the behavior of other institutions, either individually or as a group, to establish norms for guidance. This approach is called 'comparative benchmarking.' In the United States and Canada, the National Association of College and University Business Officers (NACUBO) conducted a large-scale benchmarking exercise that assembled a very extensive and detailed database that covered virtually every area of institutional activity in higher education. Participation in the NACUBO study was voluntary. The result was an array of participating institutions that was highly diverse and therefore not conducive to reliable comparison. In other words, there was a peer selection problem that made the benchmarks problematic.

The Theoretical Foundations of Peer Selection

Peer selection, at the core, is unlike benchmarking in that peer selection asks why universities are different from one another while benchmarking relies on either their similarities or their desire to be similar. What explains institutional differentiation?

Resource dependency theory

The resource dependence approach stresses that each actor needs to interact with (other actors in) the environment to reach goals. 'In social systems and social interactions, interdependence exists whenever one actor does not entirely control all of the conditions necessary for the achievement of an action or for obtaining the outcome desired from the action.' (Pfeffer and Salancik, 1988, 40). It relies heavily on a political view of inter- and intra-organizational interaction. Yet, the theory departs from earlier open-system theory in its emphasis on how organizations act strategically and make active choices to manage their dependency on those parts of their task environment that control vital resources. This perspective introduces several factors that sever a deterministic link between an organization's resource dependencies and its actions.

First, organizations are usually in a position of interdependencies. Second, organizations have other options apart from complying with external demands.

They can manage and manipulate their dependencies in several ways. Third, environments are not treated as 'objective realities'. They become known through the process of enactment. 'Rather than taking the environment as a given to which the organization then adapts, it is considerably more realistic to consider the environment as an outcome of a process that involves both adaptation to the environment and attempts to change that environment.' (Pfeffer and Salancik, 1988, 222). An important aspect of the resource dependence perspective is its emphasis on intra-organizational factors to understand how organizations react and interact with their environments, in other words, collective responses to environmental pressures (Scott, 1995).

Neo-institutionalism theory

Many of the studies within neo-institutionalism emphasize the survival value of organizational conformity to institutional environments (Tolbert and Zucker, 1983), that is, the effect of institutional isomorphism on organizational forms (DiMaggio and Powell, 1983). Isomorphism is a constraining process that forces organizations to resemble other organizations that face the same set of environmental conditions. According to DiMaggio and Powell, there are three forms of institutional isomorphism, all leading to an increasing similarity in organizational behavior and producing a decrease of system diversity. Coercive isomorphism results from the pressures applied by other organizations in the environment, on which the organization is dependent (e.g. governmental policies and laws). Mimetic isomorphism stems from uncertainty caused by the symbolic environment, which induces organizations to imitate the behavior of perceived successful organizations. Normative isomorphism stems from professionalization or professionalism.

A Typology of Peer Selection Methodologies

There are several methodologies available for determining peer groups among colleges and universities. In the United States, for example, the American Association of University Professors (AAUP), the Carnegie Commission for Higher Education, the National Center for Higher Education Management Systems (NCHEMS), and a few individual states, for example, Washington and Kansas, have developed formal methodologies. Others, like the *Maclean's* magazine survey in Canada, are less definitive but aim for a similar result. Each uses different criteria but usually includes some subset of the following variables: enrolment, numbers of degrees awarded, programs offered, professional staffing, average salaries, and research expenditures, among others. Some take local geography and demographics into account.

It is important to understand that there can be significantly large differences between methodologies that organize individual institutions into groups or categories, and then make comparison among the groups or categories, and those methodologies that aim actually to measure the differences or similarities among individual institutions so that they can be compared one to another. With very few exceptions, the existing methodologies are of the first type: they construct groups of approximately similar institutions according to relatively short lists of characteristics. Once the groups are constructed, the institutions that they comprise are assumed to be identical. These methodologies can assist in comparing jurisdictions in order to measure diversity, but they are unhelpful and even misleading in making other comparisons.

A typology of approaches to developing institutional peer groups is presented in Table 1. The top half of the table summarizes the major approaches. The bottom half displays a continuum of options for emphasis which ranges from and entirely judgment free, purely statistical approach (Cluster Analysis) to and approach (Panel Review) that relies wholly on the professional judgment experts.

Cluster Analysis

Cluster Analysis is a set of statistical procedures that are designed basically to calculate statistical distance. Alternative ways of making the calculation distinguish alternative clustering methods. Clustering algorithms ensure that the institutions in a given cluster will be more similar to each other, with regard to the variables being evaluated, than the institutions in any other cluster. The approach relies heavily on multivariate statistics and computer processing to manipulate large quantities of institutional descriptors. Other statistical techniques may be used in conjunction with the cluster analysis procedures. Factor analysis is sometimes used as a step preliminary to *Cluster Analysis* as a means of incorporating a large amount of data in the peer selection process. Discriminant analysis is used to examine the results of the clustering techniques.

Hybrid Approach

The *Hybrid Approach* incorporates a strong emphasis on data and input combined with custom-designed statistical algorithms for manipulating data.

Table 1 Typology of approaches to developing institutional peer groups

Technique	Cluster Analysis	Hybrid Approach	Threshold Approach	Panel Review
Emphasis	Data plus Statistics	Data plus Statistics plus Judgment	Data plus Judgment	Judgment

The *Hybrid Approach* also involves a degree of professional judgment in selection of data and the construction of algorithms. Thus, the *Hybrid Approach* usually involves fewer data than *Cluster Analysis* because of the pre-selection of data. Various forms of this approach are conceivable. One such approach is that used by the Kansas Board of Regents to identify peer groups for the six 4-year institutions under its jurisdiction (Teeter and Christal, 1987). This methodology was revamped in the fall of 1980 to revise earlier peer selections made by the Kansas Board of Regents, which used these selections as aids in developing finding formulas for institutions in Kansas.

Threshold Approach

The *Threshold Approach* relies primarily on thresholds and raw data, and depends little, if at all, on statistical methods. It is useful to think of it as a procedure for reducing the universe of institutions until a residue of acceptable ones remains. Although not a pure threshold approach, the National Center for Higher Education Management Systems (NCHEMS) uses a methodology that comes close in practice to such an approach. The *Threshold Approach* is essentially historical in that it accepts and reinforces data based on fixed performance.

Panel Review

In the *Panel Review* approach, peer groups are developed primarily through informed judgment and are based on a consensus of knowledgeable individuals. Data are used only informally. This approach is commonly used, although descriptions of this approach are difficult to find because of its simplicity and unscientific foundation. Reputational surveys are often used either to inform the *Panel Review* approach or to confirm its results.

Each of the approaches offers advantages and disadvantages. Some are more appropriate in certain circumstances than others. For example, one advantage of the *Cluster Analysis* approach is that it does not require arbitrary judgments to be made in advance about the appropriate cut-off points for interval variables as required by the *Threshold Approach*. Considerable judgment is still required to decide both how and where group boundaries will ultimately be drawn, and how to assign weights to the variables entering the analysis. *Cluster Analysis* and the statistical techniques that support it, on the other hand, are complex and sophisticated, and require more than a basic understanding of statistics. *Cluster Analysis* raises other statistical concerns. The manner in which data are standardized can cause problems whereby variables that have the largest variance will have the largest impact on the cluster results, regardless of whether that makes sense substantively. Factor analysis based on

samples of fewer than 300 cases may only have fair reliability. The technical complexity and abstractness of *Cluster Analysis* makes it less practical to implement, explain, and understand. Non-statisticians generally have to accept on faith that this approach is appropriate for the selection of peer institutions, and that the human interventions required by these procedures have been reasonable.

Both the *Threshold Approach* and *Panel Review* approach are simple in the use, but have notable weakness. A major weakness of the *Threshold Approach* is that it ignores the extent to which institutions miss the value range for a given variable selected by the home institution. The price of this enhanced credibility is a higher degree of logistical complexity. The *Panel Review* approach could inevitably avoid subjectiveness, with the likelihood of mistakenly selecting an 'aspirational' institution as a peer. Such erroneous *Panel Review* classifications could jeopardize the credibility of comparisons.

The *Hybrid Approach* strikes a deliberate and reasonable balance between having statistical integrity and utilizing professional judgment. It is not so heavily reliant upon judgment that it runs the risk of selecting aspirational institutions as peers or of creating the perception that data have been manipulated to promote institutional self-interest. The major area of subjective judgment — the assignment of selection variable weights — is clearly visible, and thereby open to further review and discussion as necessary. The *Hybrid Approach* is not so statistically intricate that it is incomprehensible. It is, however, sufficiently elaborate and thorough to discourage the manipulation of results. It permits extensive examination of institutions, particularly with respect to degrees awarded by degree level and instructional program area, and incorporates information on state and provincial characteristics.

Of all of the peer selection approaches, the *Hybrid Approach* is the only one that explicitly takes into consideration the characteristics of the nation, state, province, and city in which the candidate institutions are situated. This is desirable because environmental factors are important elements of comparative analyses, for example, ability to pay or cost structures that are based on local costs of living. This recommends the *Hybrid Approach* to select peers across borders, and in this case across national histories and cultures.

An Illustrative Case Study of Peer Selection by the University of Toronto

Background

The peer selection practice described here grew from four similar but separate events, each involving the University of Toronto. First was the university's participation in two major data exchanges, the Canadian Universities Data Exchange Consortium (CUDEC) and the American Association of

Universities Data Exchange (AAUDE). Comparisons based on peer selection, regardless of theoretical approach, depend heavily on the availability of institutional data. These exchanges provided a wide array of data organized by mutually agreed and recognized definitions.

Second was a large-scale benchmarking study sponsored by the National Association of College and University Business Officers (NACUBO). Although NACUBO is a US organization, Canadian institutions were invited to participate in the study, and the Canadian Association of University Business Officers (CAUBO), which is NACUBO's counterpart in Canada, kept an active watching brief on the project. The University of Toronto was a full participant in the project for 2 years.

Third, in 1991, the Minister of Colleges and Universities in Ontario struck a Task Force on University Accountability. The task force's final report, which appeared in 1993, made a number of recommendations about performance indicators and how they should be properly deployed. In the task force's judgment, proper use of the indicators depended on definitive mission statements and deliberate and objective identification of peers.

Finally, an Advisory Panel on Future Directions for Post-Secondary Education was set up by the provincial government in 1995 and reported in 1997. The panel raised a number of questions about how differentiation among institutions might be measured and promoted, and how distinctive institutional missions and roles might be recognized within a single system of higher education. Responding to these queries and suggestions required some yardstick by which to express and measure similarities and dis-similarities among institutions.

The University of Toronto, like many other colleges and universities, therefore had a number of reasons to develop a process for identifying peers and had access to data on which such a process might depend. Of particular importance and relevance were the close similarities between the objectives of the 1991 task force and the 1995–1997 advisory panel in Ontario and the objectives of Project 98/5 and Project 21/1 in China. Their purposes as matters of public policy were almost identical.

Approach

After examining the several theoretical peer identification schemes, and favoring the Hybrid Approach, the University of Toronto decided that it should develop that approach further to include four different 'slates' of peers: 'Base', 'Research', 'Compensation', and 'Government Ability to Pay'. Each slate would be used in different circumstances but would be based on the same definitions and data, and organized by program as well as by institution. All data would be drawn from either AAUDE or CUDEC. In addition, data were assembled from various sources on jurisdictional (state or province) characteristics.

That there would be a 'base' slate could be taken as given. That there should be a 'research' slate was mostly explainable by the role of the University of Toronto. A 'compensation' slate was needed for several reasons. Comparisons almost always play a role in labor negotiations about salaries. Salary expense, which is any college or university's single largest cost, can vary significantly among programs. Thus, the mix of programs in a given institution can appear to overstate or understate comparative costs unless there is a specific comparison algorithm for compensation. The 'compensation' slate is in some respects an expression of costs of living in different locations. So, for example, all salaries and wages in both the public and private sectors in a large urban area might be relatively high, in which case an unadjusted comparison of higher educational costs would be misleading. A separate 'compensation' slate can provide such an adjustment. The performance of colleges and universities and the degree of diversity in systems of post-secondary education depend heavily on levels of funding. Yet, those levels often are not really the result of policies directed specifically at higher education. Instead, they are artifacts of larger policies and circumstances that affect the entire public sector, for example, the rise and fall of general revenue. Hence, there is the need for an 'ability to pay' slate.

Process

Exchange rate

As both the US dollar and the Canadian dollar float, a 'fundamental equilibrium exchange rate' was set and deployed to align all financial information among institutions. The consistent use of one exchange rate that factored out cyclical variations in currency values was especially important for time series analysis. The adjustment algorithm for Canadian and US dollars (at 0.66 for the inter-institutional methodology) would pose problems for longitudinal comparisons, so another algorithm was introduced. The Organization for Economic Cooperation and Development, which makes many inter-jurisdictional comparisons in many fields, has devised a method of adjusting currencies for comparison called 'purchasing power parity'. While the two algorithms for adjusting for the relative value of the Canadian dollar and the US dollar produced somewhat different absolute results, they did not produce different results in terms of rankings of similarity.

Cost of Government (COG) Index

Price differences among geographic areas can create significant differences in purchasing power, a condition of major importance in public finance but often overlooked in comparisons and equity considerations. Comparisons of

revenues and expenditures lose much of their value if nominal dollar amounts are not adjusted for equal purchasing power. Consequently, the financial data for each AAUDE institution were adjusted using a state COG Index developed by the US Department of Education.

The COG reports the market prices and real wages that state and local governments would negotiate for a fixed basket of goods and services purchased for the current operation of their collective public human services, excluding medical services. While not specifically designed for colleges and universities, the COG reflects theoretical minimal prices generally applicable to all public services. For all states, the COG values ranged from a high of 127 for Alaska to a low of 89 for Mississippi. For the 25 states, which contained at least one AAUDE member, the COG values ranged from a high of 115 for New York to a low of 90 for North Carolina.

Considerable effort would have to be expended to develop an individual COG value for Ontario, which would be based on the same 'basket of goods and services' as the American COG values. Alternatively, it was possible to use three variables in the peer selection model (population size — 25% weight; urbanization level — 25%; nominal per capita income — 50%) to select the five states that were most similar to Ontario, and then use the average of those states' COG values. Thus, the proxy COG value for Ontario was 98.4 based on Colorado, Florida, Michigan, Ohio, and Washington.

Addition of a library variable

The University of Toronto placed and continues to place a high priority on its library system as reflected by a formal budget policy for library acquisitions budgets against budget reductions, price inflation and currency fluctuation, in other words, measures to ensure that real purchasing power is maintained. Given that priority, two selection variables — total library volumes and total library materials expenditures — were added to the peer selection model. This is a good example of the combination of statistical analysis, professional judgment, and selection of data under a Hybrid Approach.

Selection process

Variable weight distributions

The University of Toronto deployed four slates, which are differentiated by the relative weights assigned to the peer selection or data input variables as in Table 2:

The selection variables were conceptually grouped into three categories: *State/Provincial Characteristics*, *Enrolment/Financial/Library*, and *Degrees*

Table 2 Peer selection variable weights

<i>Selection variable</i>	<i>Base slate</i>	<i>Research slate</i>	<i>Compensation slate</i>	<i>Government ability to pay slate</i>
<i>Enrolment</i>				
FTE enrolment	5.0	2.0	0.0	5.0
Head count enrolment	3.0	0.0	0.0	0.0
Graduate and first professional as % of FTE enrolment	8.0	12.0	16.0	5.0
	16.0	14.0	16.0	10.0
<i>Financial</i>				
Current fund expenditures	4.0	2.0	2.0	0.0
Tuition and fees revenue as % of current fund revenues	4.0	2.0	8.0	15.5
Instruction expenditures	4.0	2.0	2.0	0.0
Restricted funds revenues	4.0	2.0	8.0	2.0
Research expenditures	4.0	6.0	2.0	0.0
Research expenditures as % current fund expenditures	4.0	8.0	2.0	0.0
	24.0	22.0	24.0	17.5
<i>Library</i>				
Library volumes	4.0	8.0	0.0	0.0
Total library materials expenditures	4.0	6.0	0.0	0.0
	8.0	14.0	0.0	0.0
<i>State/provincial</i>				
Population size	0.0	0.0	0.0	15.0
Urban as % of population	2.0	0.0	10.0	5.0
Per capita personal income	2.0	0.0	10.0	25.0
	4.0	0.0	20.0	45.0
<i>Degrees awarded</i>				
Bachelor	12.0	7.5	10.0	13.9
Masters	12.0	15.0	10.0	7.8
Doctoral	12.0	20.0	10.0	2.5
First professional	12.0	7.5	10.0	3.2
	48.0	50.0	40.0	27.5
Total	100.0	100.0	100.0	100.0

Awarded. The total residual weight between the latter two categories was split 50:50 once the weight for the first category has been determined. For the *Base* and *Compensation* slates, the total weight assigned to the degrees awarded category was then equally distributed among the selection variables for each of the four-degree levels. For the *Compensation* slate, higher weights were assigned to the urbanization level, per capita income, graduate and first professional share of full-time equivalent enrolment, tuition and fees revenue,

and restricted funds revenue. For the *Research* slate, higher weights were also assigned to the research expenditures, graduate and first professional share of full-time equivalent enrolment, and library selection variables. For the *Government Ability to Pay* slate, higher weights were assigned to the state or provincial characteristics, and tuition and fees revenue selection variables. The degree level weights for the *Government Ability to Pay* slate reflected the actual distribution of degrees conferred in 1987–1988 by degree level expressed in government funding units.

Selection procedure

For selection steps, see Table 3.

Table 3 Peer selection steps

<i>Procedure</i>	<i>Purpose</i>	<i>Content</i>
Step one	Screening a candidate group	A 'candidate group' was screened by similarity to the University of Toronto with respect to enrolment, funding and expenditure patterns, library volumes and materials expenditures, state/provincial characteristics, and degrees awarded.
Step two	Calculating 'comparison scores'	A mean and a standard deviation were calculated for each selection variable from which a z-score (I -score = (raw datum – mean for variable)/standard deviation for variable) was generated for each institution. Each candidate's z-scores are compared to those of the University of Toronto by taking the absolute value of their differences. The results of this process are referred to as 'comparison scores'.
Step three	Standardizing 'comparison scores'	All comparison scores (c) were then standardized using the formula $X = 10 + 5c$. Since z-scores commonly range between -3 and 3 , this conversion caused the comparison scores to become non-negative with broader ranges. In the case of degrees awarded, however, only standardized comparison scores were provided for each institution, one score for each degree level.
Step four	Calculating 'similarity scores'	Weights (totaling 100) were applied to the standardized comparison scores of the selection variables. The scores thus weighted were summed to create similarity scores. The institutions were then rank-ordered by similarity score. These rankings then served as a valuable aid in selecting a final set of peer institutions.

Selection results

For selection results, see Table 4.

Table 4 University of Toronto's peer institution top 10 results (in descending order)

	<i>Base slate</i>	<i>Research slate</i>
1	Arizona	Arizona
2	California, Berkeley	California, Berkeley
3	Illinois, Urbana-Champaign	California, Los Angeles
4	Michigan	Michigan
5	Minnesota	Minnesota
6	North Carolina, Chapel Hill	North Carolina, Chapel Hill
7	Ohio State	Ohio State
8	Rutgers	Rutgers
9	Texas, Austin	Texas, Austin
10	Washington	Washington
	<i>Compensation slate</i>	<i>Ability to pay slate</i>
1	Arizona	Arizona
2	California, San Diego	Illinois, Urbana-Champaign
3	Florida	Michigan
4	North Carolina, Chapel Hill	Michigan State
5	Kansas	Minnesota
6	Missouri	Missouri
7	Ohio State	North Carolina, Chapel Hill
8	SUNY, Buffalo	Ohio State
9	Texas, Austin	Texas, Austin
10	Washington	Washington

At this point it is critically important to recognize the crucial role that the selection variable weights played in the analysis. Changes in the weighting resulted in changes in the similarity scores. The weights were the connection between the statistical dimension of the Hybrid Approach and its judgmental dimension. While this characteristic of the Hybrid Approach is not difficult to understand in theory, it is difficult to deploy in practice. The weights were in effect a missing link that solved this problem.

Conclusions and Implications

The case study suggests that university comparison or peer selection exercises should start from program comparison. This requires a special note: very few league tables and rankings function at the program level. The vast majority

function only at the institutional level. Chinese universities should adopt some indicators featuring programs. In particular in the case of international comparisons, even though the higher education systems in question might display many differences, at least programs are comparable.

The case study shows how much attention the University of Toronto gives to teaching and education, which is evident not only in terms of inclusion of such variables of enrolment, instruction expenditures, library volumes and expenditures, and degrees awarded but also in weights distribution among the variables. For example, a heavy weight is assigned to degrees awarded in all of the four slates. This contrasts the heavy emphasis given by Chinese universities to research resources and outputs. In a popular Chinese university ranking exercise, weights as high as 42.9% are assigned to research alone, that is to publications, awards and patents.

Valid and sufficient data are crucial in university comparison and peer selection. There is a need of such databases as CUDEC and AAUDE in China and, when conditions permit, some Chinese universities might think about participating in these data exchange mechanisms worldwide. Only with sufficient data, the comparative analysis concerning Chinese universities' strength and identity in the world university community can be valid, and therefore useful for them to make right decisions as to the strategic planning.

China is still in the process of modernization. This means that there is a developmental gap compared to the major developed countries, in particular, in the spheres of science and education. In this case the total adoption of Western indicators in the university comparison and peer selection exercises could be misleading, while comparison in the priority areas or selected slates become thus more significant and may set up the practical goal to achieve.

Funding may determine to a large degree the performance of a university, which contributes mostly to the gap between Chinese universities and their Western peers. However, in many resource-related comparative studies between Chinese and Western universities, the differences in operational cost and the effect of purchasing power parity have not been given adequate attention, which results in comparisons that can be misleading if not specious. This study suggests that such useful tools as the OECD's 'purchasing power parity' and the US Department of Education's COG Index probably be employed to process the financial and cost data involved in the comparative analysis between Chinese universities and the Western peers.

The 'government ability to pay' factor seems to be unique in the case study. Given the political and economic systems existing in China, the government is often dominant in many domains of public goods, including higher education, and shows strong ability and will to pay. This is particularly evident with Project '21/1' and '98/5', and the recent rapid expansion of higher education system in China (a 300% increase in enrolment from 1999 to 2002). This

should be reflected in the variable designing in the comparative study between Chinese and Western universities.

Notes

- 1 Announced in 1993, Project 21/1 is to identify for the 21st century 100 institutions and a number of disciplines of 'world standard' by preferential treatment.
- 2 This is commonly regarded as the elitist part of Project 21/1, yet it is otherwise called Project 98/5, as it was officially approved in May 1998 during the centennial anniversary of Peking University, when the Ministry of Education appealed to the top leadership for reserving China's 1% of its annual revenue in 3 consecutive years, from 1999 to 2001, for the purpose of building world class universities in China. It is estimated that the total investment on this project will be over ¥30 billion current price in RMB (Wang, 2002). The objective of the large extra investment is to promote the around 10 universities into world well-known universities.

References

- Bryson, J.M. (1988) *Strategic Planning for Public and Nonprofit Organizations*, San Francisco: Jossey-Bass.
- DiMaggio, P.J. and Powell, W.W. (1983) 'The iron cage revisited: institutional isomorphism and collective rationality in organizational fields', *American Sociological Review* 48: 147–160.
- Lang, D. and Lopers-Sweetman, R. (1991) 'The role of statements of institutional purpose', *Research in Higher Education* 32(6): 599–624.
- Pfeffer, J. and Salancik, G.R. (1988) *The External Control of Organizations. A Resource Dependence Perspective*, New York: Harper & Row, p 40.
- Scott, W.R. (1995) *Institutions and Organizations*, Thousand Oaks: Sage.
- Teeter, D.J. and Christal, M.E. (1987) 'Establishing peer groups: a comparison of methodologies', *Planning for Higher Education* 15(2): 8–17.
- Tolbert, P.S. and Zucker, L. (1983) 'Institutional sources of change in the formal structure of organizations: the diffusion of civil service reform, 1880–1935', *Administrative Science Quarterly* 30: 22–39.
- Wang, H. (2002) 'Zhongguo ming xiao zhengduo 300 yi bokuan, 10 nian yao zuo 'shijie yiliu' (High profile universities in China compete for ¥ 30 billion governmental appropriation, and to be 'world class' in 10 years)', *21 shiji jingji baodao* (21CN BUSINESS HERALD), 27 May 2002: 1 & 3.