

ON THE VALUE OF FOREIGN PHDS IN THE DEVELOPING WORLD: TRAINING VERSUS SELECTION EFFECTS IN THE CASE OF SOUTH AFRICA

MORITZ MÜLLER¹, ROBIN COWAN^{1,2,3}, AND HELENA BARNARD⁴

ABSTRACT. When seeking to improve science in emerging economies, uncertainty exists whether PhD training in an emerging economy can yield comparable results to PhD training in the developed world. Scientific achievements may vary because of excellent training at good universities, but also because excellent students select (and are selected by) good universities. This paper compares the career effects of overseas and domestic PhD training for scholars working in an emerging economy, South Africa. We differentiate between and examine both selection and training effects for PhDs from three tiers of South African and two tiers of foreign universities. Those with PhDs from universities in industrialised countries generally achieve greater career success than those with local PhDs, but training by universities in industrialised countries is not necessarily better than local training. Our results suggest that the perceived superiority of foreign PhD training stems from selection rather than training effects, and pure selection effects in fact explain career outcomes better than training effects. Focusing on training rather than selection, PhDs from top South African universities produce a similar quantity and quality research output compared to those trained by the leading universities in the developed world. From the perspective of an emerging economy with limited resources wishing to advance science, the development of local universities should thus be stressed, although it is clear that individuals who are able to study for a PhD abroad gain personally when they return.

JEL codes: F63; H52; I2; O15; O20; O30

keywords: Scientific mobility; Doctoral studies; University evaluation; Developing country; South Africa; Technological upgrading

⁽¹⁾ BETA, UNIVERSITÉ DE STRASBOURG

⁽²⁾ INSTITUT UNIVERSITAIRE DE FRANCE

⁽³⁾ UNU-MERIT, MAASTRICHT UNIVERSITY

⁽⁴⁾ GORDON INSTITUTE FOR BUSINESS STUDIES, UNIVERSITY OF PRETORIA, JOHANNESBURG

Date: March 10, 2018.

Acknowledgement: Financial support for the project was provided by the Institut Universitaire de France, and by the Chair in Economic Policy, ECON, KIT, Germany.

1. INTRODUCTION

The importance of universities in technological and economic development has long been recognised (Mowery and Rosenberg, 1999; Murmann, 2003; Rosenberg and Nelson, 1994) and in the context of increasingly knowledge-intensive economies, it is likely that universities will play an even greater role in the future upgrading of less developed countries (Altbach, 2013; Brundenius et al., 2011).

Although a common policy goal is to develop the research capacity of their country, local institutions are often not in a position to train young scientists in sufficient number and quality. Upgrading relevant local institutions is a lengthy and costly process and limited by the existing scientific work force (Heitor et al., 2014; Horta, 2010). Not only is it not clear whether developing countries have the capacity to provide quality training to PhDs, it also takes years before the contribution of a scholar to his or her field becomes clear. This makes it hard to judge whether scholars can be trained locally, or whether they should rather seek PhD training abroad.

In this paper, we focus on the scientific contribution of academics working in a developing country, comparing the scholarly success of academics who received their PhD training locally with those who were trained abroad. We examine two key explanations for quality differences in the scientific achievement of academics across PhD granting universities, selection effects (since universities with better reputations are likely to attract better students) and training effects (the impact of doing a PhD in a certain university).

The limited available scholarship about the capacity of universities in developing countries is often not particularly optimistic. In particular, the scientific competence of local research faculty is thought to stem mainly from advanced training abroad. The statement by Nelson (2005:27) exemplifies such a view:

Indigenous universities will play a key role as the source of students who take advanced training abroad, and as the home of faculty who have been trained abroad.

This paper contributes to the literature on universities in developing countries by examining scholars working in a developing country, South Africa, exploring whether there is a systematic relationship between where scholars have received their PhD training (locally or abroad) and the quality of their scholarship as judged by their peers. We thus examine the relative contribution of local and foreign scientific training to scientific excellence in local research faculty.

We do not address the issue of the academic diaspora, or brain circulation. Rather we explicitly focus on researchers working in their home country, and ask what contributes to their differential career success. In particular, we seek to understand for those who stayed in or returned to their home country, to what extent their different career paths can be explained by where they did their PhDs.

We (unsurprisingly) find a quality hierarchy of universities both locally and abroad. Foreign universities demonstrate a very strong selection effect with the top local students preferring study at top global universities. Even the foreign universities of a somewhat lower rank exert a strong selection effect. However, in terms of impact on subsequent scientific achievements, for those who eventually have careers in the home country, the leading local universities have a training effect that is equivalent to that of globally recognised universities.

This finding offers an alternative to the sometimes narrow view of the relative contribution of local and foreign scientific training to the quality of local research faculty in the context of a middle income economy (e.g. Nelson, 2005). The main policy implication for universities in developing countries is that real benefits can be achieved by investing in improved PhD training, and we conclude with some recommendations about how that might be achieved.

2. THE VALUE OF QUALITY SCHOLARSHIP AT DEVELOPING COUNTRY UNIVERSITIES

There is by now an extensive body of work on the role of universities in upgrading (Mowery and Rosenberg, 1999; Murmann, 2003; Rosenberg and Nelson, 1994). Moreover, work on the growing importance of scientific knowledge to economic growth suggests that the role of universities will continue to increase in importance (Conceição and Heitor, 1999; Deiacò, Hughes and McKelvey, 2012). Scholars (Altbach, 2013; Brundenius et al., 2011) have thus highlighted that universities in developing countries are likely to play an especially important role in the future.

Implicit in the body of work on the role of universities in upgrading, whether through university-industry linkages (De Fuentes and Dutrenit, 2012; Giuliani and Arza, 2009), spinoffs (Benneworth and Charles, 2005; Chen and Kenney, 2007; Kroll and Liefner, 2008), or simply the creation of a skilled workforce (Bell and Pavitt, 1997), is that the university is a key source of new and useful knowledge. This suggests that the quality of scholarship conducted there is foundational to the societal benefits that universities can provide.

A large body of research exists examining the research behaviour (especially productivity) of foreign-trained scholars in scientifically peripheral countries (Delicado, 2010; Gibson and McKenzie, 2014; Jonkers and Cruz-Castro, 2013; Li, Miao

and Yang, 2015), and there is also a substantial body of literature on the return intentions (or not) of those scholars (Grogger and Hanson, 2013; 2015; Kim, Bankart and Isdell, 2011; Soon, 2012). Another part of the literature highlights the potentially virtuous interaction between locally-trained and foreign-based scholars of developing countries (Baruffaldo and Landoni, 2012; Jonkers and Tijssen, 2008).

Although much is known about the publication and work patterns of scholars based on where they have studied, less is known about how this translates to the systemic level (see Zhang, Patton and Kenney, 2013, for an exception). In developing our understanding of universities in developing countries, we simultaneously consider scholars working there who have been trained abroad and those who have received their training locally. If a systematic difference can be seen between these two groups, it should help policymakers in deciding whether to prioritise one group over the other.

PhD students get selected at universities in a process where the preferences of students to join a certain doctoral programme must be matched with the requirements of PhD admission committees. This results in a process of (self-)selection such that characteristics of PhD students vary with the PhD granting university. Once students have been accepted, they undergo training, which is likely to also vary by university.

The close relationship between selection and training has long been recognised: In his article on the “Matthew effect”, Merton (1968) pointed out a virtuous effect by which many Nobel Prize winners had themselves been trained by Nobel Prize winners. But selection and training are conceptually distinct: Selection reflects the innate ability (“raw material”) of a student who selects and is selected by a given university, and is likely to be influenced by the perceived mutual desirability. Training reflects the transformation that the student undergoes while being at that university, and is more likely to be influenced by available resources such as equipment and supervisory capacity.

Our question is thus not only whether there are differences in the career success, measured over time, of academics working in developing countries who received their PhD locally versus in industrialised countries, but also how selection and training effects respectively explain those differences. The distinction matters, because training effects will vary based on the human and organisational resources of the university, whereas selection effects reflect a perceived view of quality, both the student’s view of the university, and vice versa.

2.1. Selection effects. The available scholarship about the research capacity of universities in developing countries is often not particularly optimistic, with Nelson (2005), for example, suggesting that the main role of universities in developing countries is to absorb foreign-trained faculty who can prepare students for advanced training at leading academic centres abroad. In other words, in much of extant work, cutting-edge science is not perceived as happening at developing country universities. Instead, in terms of their contribution to the advancement of science, the most likely contribution of scholars working at universities in developing countries is argued to be in preparing promising students for academic success elsewhere.

A similar idea is elaborated via the notion of a “social technology” (Nelson, 2004). Nelson argues that a primary purpose of the university in a developing country context is to establish the “social technologies” that support a given scientific field. Nelson and Sampat (2001) argue that social technologies, i.e. the customs and codes of practice in a field, are an essential complement to physical technologies and enable institutions to be effective. They give the example of gaining mastery of the research methods of organic chemistry, and describe the relevant social technology as “the system of training young chemists in the relevant physical technology” (2001:50). Likewise, Bell and Pavitt (1997) view the university as part of the “infrastructure” within which firms operate.

All of these views are quite limiting. They suggest that the value of universities in developing countries should be measured in terms of their contribution to a generally more enabling environment for knowledge-intensive activities, rather than in terms of scientific advancement per se. This literature is silent on the role of developing country universities in the training of PhDs, because it is understood that “serious scholars” will seek to go to universities where knowledge creation is central, i.e. abroad. This kind of (largely implicit) view strongly suggests that a selection effect will drive top students to go abroad for PhD training.

Indeed, there is a large literature documenting that there is a strong self-selection effect, resulting in gifted students going abroad for PhD study (Bertrand-Cloodt, Crvers and Heijke, 2016; Gaule and Piacentini, 2013). Given the evidence that the best local students choose to study at universities in industrialised countries, it may appear that universities in developing countries are not very effective in training PhDs. However, such a view would be inaccurate, in case the apparent difference in quality should be ascribed not to training, but to selection effects, i.e. the reasons why certain students (and not others) become students of certain universities (rather than others).

Gifted local students may not always be able to go abroad, but they can be expected to nonetheless select what they perceive as the best possible university. In this case, however, they can only select from a range of local universities. Apart from suggesting that a hierarchy of local universities should become apparent, extant evidence suggests that selection effects may well explain some part of the subsequent career success of scholars trained at the “better” universities, especially the foreign universities, and to a lesser extent, the best local ones.

2.2. Training effects. While there are many reasons to expect universities in a developing country to be disadvantaged by selection effects, training effects are likely to function differently. This is because of the global nature of knowledge creation and the importance of a thorough understanding of local conditions.

A central theme of research on upgrading is the importance of cross-border contact and learning between people, although that point has been made far more commonly about other forms of connection than about universities. A concern with how cross-national connections take place has occupied scholars focusing on the nation as their unit of analysis (e.g. Lall, 2001; Ozawa, 1992), the firm (e.g. Athreye and Cantwell, 2007; Narula and Dunning, 2000; 2010) and the individual (e.g. Almeida, Phene and Li, 2014; Saxenian and Hsu, 2001).

Scholarship has confirmed that the very process of knowledge creation is a global one (Altbach, 2013) and PhD training in a developing country is unlikely to be parochial. On the contrary, given the dominance of universities from the leading economies in knowledge creation, it can be assumed that PhD training takes place with a strong awareness of what is being done elsewhere. Indeed, it is known that the leading scholars from developing countries are themselves strongly globally connected (Barnard, Cowan and Müller, 2012).

Yet scholars also benefit from being locally connected. Local conditions in aspects such as income, weather, demography, and epidemiology (Bernardes and Albuquerque, 2003:870) require the development of new technologies to improve on often inadequate solutions that have tended to originate from high income countries. There is general consensus that the best placed universities to develop such solutions are typically the leading universities of the developing world. Moreover, as evidence of the wider benefit of something like “frugal innovation” (Zeschky et al., 2011) increases, the wider scientific community is showing a greater interest in what was previously seen as the local challenges of developing countries. The application of solutions to local problems (of developing countries) is starting to be seen globally, implying integration at one end of the innovation process (Romijn

and Caniëls, 2011). Integration further upstream may be increasingly seen as valuable.

Considering the extensive evidence about the value of both global connectedness and being locally anchored, we anticipate that local universities in a developing country may well be able to offer training at a comparable level to that of universities abroad. Of course, there is likely to be variance in both populations. But perhaps more importantly, although there is ample evidence of the value of a deep understanding of local challenges and conditions in global knowledge creation, this may not necessarily translate into effective PhD training. For example, locally connected supervisors are not necessarily globally connected, and extant evidence suggests that the real value for developing countries occurs when scholars are both locally and globally connected (Barnard, Cowan and Müller, 2012). Thus there is no reason to expect that universities in developing countries will necessarily out- or underperform their peers in industrialised countries. However, we are confident that training effects are likely to play a role in the subsequent career success of academics working in a developing country, and that such effects need to be considered separately from selection effects.

3. THE CASE OF SOUTH AFRICA

Our empirical analysis examines academic achievements of South African researchers active in South Africa. As background, this section provides a short historical account of the South African research system.

South Africa had passed through several major political regimes: The colonial period, the Republic of South Africa with an apartheid government from 1961 to 1994, and finally a democracy with ongoing, major social transformations. The colonial period imprinted South Africa with an English and Afrikaans (science and university) culture, and laid the foundation for scientific collaboration — most importantly with the UK and later the US, throughout its history (Sooyarmoothy, 2010). South African science in the early 20th century had been at the periphery to European and US science, focussing on applied research and indigenous science (Plug, 2009).

Under apartheid, universities were racially delineated, with the primary distinction between the well-resourced “white” universities and the rest. Each of these main categories had further divisions. The “white” universities had either English or Afrikaans as language of instruction. The eleven Afrikaans universities had strong ties to the ruling party of the Apartheid government (Bunting, 2006). The four English universities, less tied to the government, tended to have greater autonomy, and tended also to have strong international ties, mainly to the UK, and

their scientific reputation was thus generally stronger than those of the Afrikaans universities. The various universities for people of colour reflected Apartheid categorisations, and there were ten (generally English language) universities targeting students who were Indian, “coloured” (i.e. mixed race) and from various black ethnic groupings (Jansen, 2004). In addition, there were also “technikons”, i.e. applied technology universities, eight “black” and seven “white”. University reform began in 2000 involving a series of mergers, and repositioning “technikons” as universities of technology (Jansen, 2004). The 23 universities that emerged were categorised as “traditional” (offering theoretically oriented qualifications), “universities of technology” (focusing on vocational qualifications) and “comprehensive” (offering both).¹

South Africa experienced increasing international political isolation during the apartheid regime, including an “academic boycott”. Precisely what its effects on international collaboration were remains under debate.² However, the government responded with increased efforts to become more self-dependent generally, and advanced a national science and technology agenda. Research capacity was to be fostered through a shift away from project-based research funding towards the funding of individuals demonstrating scientific excellence at the global science level. This strategic shift was implemented in 1984 with the founding of the National Research Foundation (NRF), formerly the Foundation for Research and Development.

There was no particular national policy regarding foreign PhD training during apartheid. NRF funding of foreign PhDs was limited, with about 40 foreign PhDs between 1980 and 2000, while the major student exchange and mobility programmes financed from abroad, e.g. the South African Education Program of the Institute of International Education, focused on bachelor and masters students and introduced doctoral scholarships only in 1986 (IIE, 1986). Thus, foreign PhDs have been largely organised and financed by the students themselves, often facilitated by dual citizenships and contacts to a large diaspora of South Africans concentrated in English speaking countries.

Current South African policy considers high quality and quantity doctoral training as key to expanding research capacity. That both local upgrading and foreign

¹A 2010 study by the Council for Higher Education, using input and output indicators, found that universities clustered into three distinct groups, with five “research-intensive”, eight that were clearly preparing poor students for work, and nine that seemed to fall somewhere in the middle (Cloete, 2010). In their study, Kruss and Visser (2017) used four categories, research-intensive, comprehensive, universities of technology, and “rural universities”. While no definitive categorisation emerges, all the studies emphasise the high variability among universities.

²See for example Hyslop et al., 2006, Shaw, 1986, Sooryamoorthy, 2010.

PhDs have an important part to play is commonly understood by all stakeholders. However, how much each should be stressed is heavily debated. Herman (2011) for example, favours local upgrading over foreign PhD support. His main argument is the need to produce PhDs in sufficient quantity. On the other hand, ASSAf (2010) recommends supporting foreign PhDs to inject high quality into the research system. That doctoral training abroad is of higher quality compared to local training is taken as given by both camps. This opinion however seems not to be based on any analysis. The current paper addresses that lacuna directly.

4. DATA

4.1. Sample. We compare foreign-trained with the locally-trained PhDs active in the South African academic community, and use the rated researchers of its National Research Foundation (NRF) as our sample. The NRF (www.nrf.ac.za) is a state agency that has as its mission the promotion of research and the development of national research capacity. One of its key roles is to facilitate the “rating” of researchers at universities and other public research institutions such as museums. The rating process is similar to the tenure process at North American universities, but it is focused only on research (not teaching or service to the institution), centrally administered (not by the institution) and valid only for a set period (the exact period has changed over the years, but is around five years).

Although the NRF rating system at first glance may seem to differ from tenure in that it appears to be a voluntary process, for research-active faculty it is essentially compulsory. This is because NRF rating system is arguably the main mechanism through which universities and policymakers alike differentiate between faculty members who only teach, and the (much more desirable) faculty who are research-active. Most South African academics do not do any research, and although a number of universities are attempting to shift that, in 2015 less than 44 percent of full-time faculty in South Africa were doctorally qualified. At perhaps the leading South African university, the University of Cape Town, only two-thirds of faculty held doctoral degrees.

Getting an NRF rating not only renders researchers eligible for NRF grant funding, but ratings are also used by researchers’ universities, for example as the basis for promotion decisions or in deciding whether/when a sabbatical should be granted. Because an NRF rating translates into substantial career benefits for research-active faculty members, virtually all active researchers in the South African academic research community are NRF-rated. Just as this means that the NRF database cannot be used to map the careers of South African academics

per se, it does offer an ideal way to map the post-PhD career trajectory of South African academic researchers.

As part of the rating process, researchers provide the NRF with evidence of research outputs. NRF filings are detailed (including not only extensive personal detail but also information about outputs as diverse as peer-reviewed journal articles, books, conference presentations, patents, policy or technical reports, and publications in the public press) and suffer very little missing information. Researchers also provide the NRF with demographic information, including the institution where they have obtained their degrees.

NRF ratings are a useful indicator of the quality of researchers in the academic community in South Africa because of the rigour of the review process. A specialist review panel selects six (local and foreign) reviewers who are asked to read at least the self-identified most significant papers of the candidate. Each reviewer assesses the research outputs, and both an independent assessor and the specialist review committee consider the referee reports to assign a rating.

Our sample is based on the complete, digital files of the NRF spanning the years 2002 to 2012, comprising 4960 researchers with 7530 ratings. The NRF maintains two rating systems; one for established and the other for emerging researchers. The “emerging” category exists as a once-off option for researchers with recently completed PhDs, and acts mainly as acknowledgement and encouragement for new scholars. Because we are mostly interested in the long-term effects of PhD training, and because the number of rated emerging researchers is too small for meaningful analysis, our analysis focusses on established researchers. A second restriction is that we include only researchers whose first academic degree is observed in the data, i.e. bachelor or master, and whose last master degree prior to the PhD was obtained in South Africa. This way we ensure that we only consider students with South African education up to the master level making the decision to go or not to go abroad. Third, we include only scientists having obtained a (first) PhD between 1960 and 2003. The lower boundary is not restrictive as it excludes only one scientist with PhD obtained in 1955. The upper boundary is set because in 2004 a major restructuring of the university landscape was put in place. This second restriction excludes 73 scientists. Finally we remove another 34 individuals with missing information from the sample.

The main sample used to obtain the main results is thus restricted to 1) senior researchers active in South Africa between 2002 and 2012, 2) researchers who did their studies until masters level at a South African institution, and 3) researchers who obtained their PhDs between 1960 and 2003. In total these are 1137 scientists experiencing 2464 rating events.

4.2. **Variables.** Our main focus is how being a PhD graduate of a certain university relates to a scientist's subsequent scientific performance.

Scientific performance of scientists is measured using the **NRF ratings**. The rating system categorises researchers according to three broad categories, i.e. from C (established researchers), over B (internationally acclaimed researcher), to A (leading international researcher). Within each broader category a fine-grained rating is established, again, by excellence and internationality. Obtaining even a C-rating is considered a success, because becoming rated allows a South African academic to claim research-active status with its associated benefits. Ratings may also be unsuccessful, and we therefore consider an unsuccessful rating as a separate, and the lowest category. Thus, we arrive at the following categories (ordered from lowest to highest): Rating Unsuccessful, C3, C2, C1, B3, B2, B1, A2, and A1.³

Ideally, we would like to investigate training and selection effects for each university individually, but in practice we are forced to create groups of universities in order to combine a sufficient number of observations for the analysis. Our results then pertain to averages over universities within a group. Information loss due to grouping is smaller, the more homogeneous are the universities within a group and the more heterogeneous across groups, in terms of selection and training effects. Interpretation of the results, on the other hand, is facilitated when groups represent well established categories; in our case the idea of research excellence.

In total we construct five **PhD university categories**: three SA university categories and two foreign university categories. Table ?? shows the university groupings used for the main analysis together with the number of observations falling in each category. Note that the number of ratings exceeds the number of PhDs because researchers get re-rated roughly every five years.

³An alternative measure of scientific performance would be the number of (citation weighted) peer-reviewed articles. The rating and number of papers are strongly, but not perfectly, correlated. This is consistent with the idea that peer-based evaluation conveys more information about scientific excellence than simple article counts. Yet, there remains the possibility that peers providing the rating are influenced by the reputation of the PhD granting institution. Ideally, we therefore would perform a robustness analysis using publication counts. Unfortunately, the nature of our data are such that the publications of older researchers are under-reported. This is during the period in which Web of Science data are also under-reported, which makes it impossible to check our results by using publication or citation counts as dependent variable, as it would contain a systematic downward bias for higher rated researchers who participated in the rating system earlier.

TABLE 1. University categories

<i>University category</i>	
University	Shanghai ranking
<i>English SA university (251 PhDs, 543 ratings)</i>	
University of Cape Town	201-300
Witwatersrand University	201-300
University of Natal	401-500
Rhodes University	-
<i>Afrikaans SA university (409 PhDs, 871 ratings)</i>	
University of Stellenbosch	401-500
University of Pretoria	-
University of the Orange Free State	-
<i>Other SA university/technikon (334 PhDs, 667 ratings)</i>	
Remainder SA category	-
<i>First tier foreign university (46 PhDs, 132 ratings)</i>	
Top 30 foreign universities	1-30
<i>Other foreign university (97 PhDs, 251 ratings)</i>	
Remainder foreign category	-

Notes: University of Pretoria dropped out of the AWRU in 2008.

First tier foreign university PhDs from 3 UK and 15 US universities.

Other foreign universities ranked 31-100 (35 PhDs), 101-500 (43 PhDs), and non-ranked (9 PhDs).

We consider foreign universities ranked among the top 30 in the Shanghai ranking (ARWU, 2011) as first tier foreign universities – all of them renowned throughout the last century.⁴ The rest enters the remainder group, “Other foreign universities”.

In order to identify SA universities susceptible of providing research excellence, we use citation counts of individual papers. In detail, we downloaded from ISI web of science the 2000 most highly cited research articles with South African affiliation published in 1985, about in the middle of our PhD formation period. Seven South African universities stand out by having contributed highly cited publications. Taking into account the specificity of the South African research system, we further distinguish Afrikaans from English research universities. Interestingly,

⁴The Shanghai ranking has often been criticised as an indicator of reputation stemming from past research rather than indicating current research performance. In our context, this can be seen as an advantage, since we are interested in the quality or reputation of universities in the period 1960-2003. For SA universities the Shanghai ranking seems less appropriate as they saw some considerable development in their research performance over the period, and have a very specific history.

the Shanghai ranking of SA universities is consistent with our grouping (see Table ??).

Table ?? shows the latest ratings of scientists for each university category where the PhD was granted. Across university categories, the majority of scientists obtained a C-rating (58 to 70 percent), higher ratings (B) are relatively common (10 to 30 percent), and the highest ratings (A) are rarely awarded (up to seven percent).

TABLE 2. PhDs by university category and their latest ratings, fractions in brackets

	R.U.	C	B	A	Total
Other SA university	72 (0.22)	225 (0.67)	32 (0.10)	5 (0.01)	334
Afrikaans SA university	50 (0.12)	288 (0.70)	69 (0.17)	2 (0.00)	409
English SA university	30 (0.12)	145 (0.58)	68 (0.27)	8 (0.03)	251
Other foreign university	8 (0.08)	56 (0.58)	29 (0.30)	4 (0.04)	97
First tier foreign university	1 (0.02)	28 (0.61)	14 (0.30)	3 (0.07)	46
Total	161 (0.14)	742 (0.65)	212 (0.19)	22 (0.02)	1137

Notes: R.U. denotes Rating Unsuccessful. Fraction of PhDs from university category obtaining a certain rating in brackets.

In general there is considerable heterogeneity among SA universities, and the chance of obtaining a higher rating increases with having a foreign PhD. For example in Table ??, a B-rating or highest A-rating is obtained by about 11 percent of the graduates from “other” SA universities, 17 percent of SA Afrikaans university graduates, 30 percent of SA English university, 34 percent of graduates from the “other” foreign university categories, and 37 percent of first tier foreign universities. As expected, graduates from top-tier foreign institutions outperform all others, being more likely to receive A ratings. The econometric model is estimated on multiple rating observations per individual⁵, and fine rating categories as outcomes of scientific performance. These follow the same pattern.

If the rating of established scientists would coincide only with the quality of their PhD training, the analysis could stop here. However, as we have previously argued, the ordering of scientific achievements by PhD university category may result not only from training effects, i.e. differences in the impact of PhD training, but also from selection effects, i.e. heterogeneity of students’ potential before their PhD training.

Support for the selection effect explanation is given by the fact that first-tier foreign universities attract in particular young students who obtained their master degree with distinction. **Age at master** is the age when the first master degree was

⁵For rating observations by university category see Table 7.

obtained. Overall, fifty percent of students obtained their master degree between 25 and 30 years (interquartile range) with a median age of 26 years. PhD students at first-tier foreign universities obtained their master between 24 and 27 years with a median age of 26 years. In the estimation, we count years in decades from age 20. Furthermore, sixty percent of PhD students at first-tier foreign universities received a master degree with distinction, indicated by **Master distinction**, compared to an average of about fifty percent in the sample.

Table ?? shows the transition from master university to PhD university category for focal South African students. **Master university categories** are constructed exactly as the PhD university categories. Each row in the table captures students from one type of master institution, a cell in a row indicates where students from that type of master institution did their PhDs. Master students from research intensive English SA universities are more likely to go abroad than master students from research intensive Afrikaans SA universities or other SA universities.

TABLE 3. Transition from Master to PhD university category, frequency (row percentage).

Master University	PhD SA university			PhD foreign university		
	Others	Afrikaans	English	Others	First tier	Total
Others	259 (0.74)	41 (0.12)	17 (0.05)	26 (0.07)	6 (0.02)	349 (1)
Afrikaans	57 (0.12)	342 (0.73)	22 (0.05)	29 (0.06)	16 (0.03)	466 (0.99)
English	18 (0.06)	26 (0.08)	212 (0.66)	42 (0.13)	24 (0.07)	322 (1)
Total	334 (0.29)	409 (0.36)	251 (0.22)	97 (0.09)	46 (0.04)	1137 (1)

A further pattern in Table ?? is that most master students stay at “their” university (category) for the PhD. There are several explanations. One relates to ability, in that there may be a match between the ability of a student obtaining a master from a certain university with the requirements of that university regarding the ability of their future PhDs. Further explanations have nothing to do with scientific ability as such. Firstly, master studies orient the student to subjects treated within the faculty and thereby create a thematic fit. Secondly, uncertainty regarding both the doctoral program and the student may be reduced due to the master experience. Finally, students may have established favourable social and economic conditions during their master studies which increases their opportunity costs of choosing another university. The variable **Inertia value** therefore indicates which PhD university category in the choice set equals the student’s master university category.

The distribution of where scientists received their PhDs does vary over time. There may be many causal processes beneath this variation, among which the academic boycott, but in general we might expect a temporal effect on the decision

to go abroad (or the propensity of foreign universities to select South African students). The variable **Period of PhD** indicates the five year interval in which the PhD has been obtained.

Increasing seniority allows for higher potential scientific achievements, with a decreasing marginal effect (see e.g. Bayer and Dutton, 1977). **Experience** captures the post PhD experience of the scientist at the time of rating. We measure experience in decades after PhD, and also include its square term into the regression.⁶ Experience ranges between 5 and 48 years, interquartile range is from 10 to 22 years and a median of 15 years. The increasing trend of PhD training in SA relative to PhD training abroad creates a difference in their median of 2 years.

Finally, **Race**, **Gender**, and **Scientific Domain** are three control variables that are likely to be correlated with the rating outcome and the effect of a foreign PhD.⁷ Out of the 1137 scientists, most are white (1000), male (778), in natural sciences (that is not in social sciences) (680).

5. THE MODEL

5.1. Econometric strategy. Our concern is with the career effects of PhD study, or the added value of doing a PhD in particular types of universities. The simple estimation of the impact of the PhD-granting university on a person's career is confounded by a selection effect. Something, commonly referred to as "ability", will affect both which university selects a person, and that person's future career regardless of where the PhD was granted. Ability is broadly defined, including intelligence, motivation, or even socio-economic background. The econometric literature refers to such a setting as treatment with (self-)selection.

Simply ignoring ability in the regression and leaving it in the error term creates an omitted variables bias; with upward biased estimates expected for universities of higher reputation. Introducing a valid measure of ability into the regression yields unbiased and consistent parameter estimates. While we do have some variables which are arguably related with individual ability prior to PhD — Age at Master, Master distinction, and Master university category — they are unlikely to capture individual ability in a broad sense. A second option would be to instrument PhD university category, thus purging it from variation due to ability. Period of PhD does create some external variation in PhD university categories but is certainly not a strong instrument.

⁶We ran additional regressions with a set of dummies indicating whether the rating takes place one to five, six to nine, etc. years after PhD and obtained very similar results.

⁷Econometric results do not hinge on inclusion or exclusion of these variables in either of the equations.

A third option is to “learn” about individual ability by adding further equations where it is likely to play a role (Griliches, 1977). This strategy is the most appropriate given our setting and data at hand. More particularly, we add two auxiliary equations to the main equation. One auxiliary equation models explicitly the (self-) selection of master students into PhD university categories, based among others on their latent ability. This combination of a selection and an outcome equation within a system of equations is well known as the Roy-model, made popular by Heckman as the latent factor model (Heckman, 1976).

A second auxiliary equation allows for a relation between the master student’s latent ability and observed, potential signals of ability. This measurement equation of ability is meant to pick up correlations, rather than being a causal model of ability, and is an approach applied in Heckman et al. (2014).⁸

5.2. Model set-up.

Outline of the model. The model is comprised of three equations which we estimate as a system. The basic structure contains equations for three variables: individual ability; PhD institution selection; and rating. Individual ability is latent through being unobserved. Selection and rating are modelled as outcomes of standard discrete-choice models, i.e. a multinomial and ordered logit respectively (Train, 2009). The time ordering from ability (before PhD) to PhD institution selection and finally to subsequent ratings naturally results in a triangular system of three equations. We discuss each of the three equations below. Table ?? summarises notation and indexing of model variables.

Rating equation — ordered logit. Our observed measure of academic achievement is the rating outcome, $R_{i,t}$, of researcher i in rating event t . This is an ordinal variable, for which an ordered logit model is generally appropriate. Suppose the observed rating results from an underlying, real valued latent rating, $R_{i,t}^*$, following:

$$(1) \quad R_{i,t}^* = \mathbf{X}_{i,t}\boldsymbol{\beta} + \boldsymbol{\tau} \mathbf{D}_i + \nu_i + \epsilon_{i,t},$$

where $\mathbf{X}_{i,t}$ includes exogenous factors, i.e. experience, experience squared, race, gender and scientific domain. \mathbf{D}_i is a vector with p th entry one if the researcher’s PhD university category is p and zero else, ν_i is the individual’s ability, and $\epsilon_{i,t}$ is an iid extreme value distributed error term. PhD granting university \mathbf{D}_i is assumed to be independent of the error term $\epsilon_{i,t}$, but likely to be correlated with ability, ν_i . The latent rating, $R_{i,t}^*$, results in observed rating $R_{i,t}$ depending on which

⁸See also (Lee and Song, 2012) for measurement equations in structural equation models in combination with (Aakvik et al., 2005) for factor structure models in treatment effect evaluation.

TABLE 4. Model variables overview

Indexing:	
$i = 1, \dots, N$	Individuals in the sample
$t = 1, \dots, T_i$	Rating calls of individual i
$p = 1, \dots, P$	PhD university categories
Observed variables:	
$R_{i,t}$	Rating of individual i at application t (ordered from C3 to A1)
D_i	University (p) where PhD obtained (unordered)
$\mathbf{X}_{i,t}$	Factors affecting rating outcomes
\mathbf{Z}_i	Factors affecting PhD university choice
\mathbf{V}_i	Factors related to innate ability
Latent variables:	
$R_{i,t}^*$	Latent rating outcome
\mathbf{D}_i^*	Latent PhD university
ν_i	Unobserved individual ability
Parameters:	
ζ_p	PhD selection effect of university category p
τ_p	PhD training effect of university category p
β	Rating equation
\mathbf{c}	Hurdles in rating equation
α_p	Selection equation
γ	Ability equation

thresholds (\mathbf{c}) the latent outcome has crossed. Model parameters, β , τ , and \mathbf{c} , are freely estimated; except for the first threshold being set to zero for identification. Thus a latent rating below zero becomes a ‘‘Rating Unsuccessful’’. Latent ratings that are positive but close to zero, become a ‘‘C3’’. All the remaining 7 thresholds are then estimated freely. The main interest is in estimating τ , the effect of PhD university categories on academic achievement.

PhD university selection equation — multinomial logit. We treat selection using a multinomial logit model — any individual i has a probability of being selected into each of the P university types. Thus the dependent variable is a latent choice vector \mathbf{D}_i^* with P elements. Each element $D_{i,p}^*$ can be interpreted as the fit between individual i and a university of type p , based on ability and other controls:

$$(2) \quad D_{i,p}^* = \mathbf{Z}_i \boldsymbol{\alpha}_p + \zeta_p \nu_i + \epsilon_{i,p},$$

where \mathbf{Z}_i is a (L)-vector of observed factors which vary over individuals, the corresponding coefficients $\boldsymbol{\alpha}_p$ vary over institutions. Exogenous factors are Inertia value for staying at the Master university category, PhD period, Race, Gender, and Scientific domain. $\epsilon_{i,p}$ is an iid extreme value distributed error term.

The selection effect is modelled through the factor ζ_p . Selection is two-sided, in the sense that while it is based on individual ability, different types of universities will be differently selective on that variable. (We might assume that the top universities select only highest ability students, lower-ranked universities are more catholic.) What $D_{i,p}^*$ captures is the jointly determined desirability of agent i attending university (of type) p . If we consider $\mathbf{Z}_i\boldsymbol{\alpha}_p + \zeta_p \nu_i$ (dropping ϵ for the moment) we can think of this as the term defining the probability that an agent with ability ν_i attends a university of type p . Thus, selection probability is jointly determined by the attractiveness of universities of type p to agent i , and the attractiveness of agents with ability ν_i to universities of type p . When ϵ is realised, agent i attends the university having the largest entry in the vector \mathbf{D}_i^* . Since the alternative with the highest 'utility' is chosen, only the ranking of alternatives can be modelled. Ranking of alternatives is then with respect to a base category whose 'utility' is set to zero. Our base category are South African technikons and universities.

The vectors \mathbf{D}_i^* are not inherently interesting. Rather they are a means of extracting ζ_p and ν_i , which are otherwise unobserved.

Initial ability — linear model. Our concern is with the career effects of PhD study, or the added value of doing a PhD in particular types of universities. Thus for our purposes, “innate ability” is less interesting than an individual’s “ability” level just before starting the PhD. We assume a linear model with normally distributed error term.

$$(3) \quad \nu_i = \mathbf{V}_i\boldsymbol{\gamma} + \epsilon_i,$$

where \mathbf{V}_i includes our indicators of ability: master university category, master distinction, and age at master.

The expected unobserved ability is anchored at zero for students obtaining their master from a 'typical' SA university (our university reference category), being of average age, and obtaining no master distinction. Ability is then freely estimated for master students deviating from this reference case. The error term is standard normal distributed. This way the variance of the combination of error and ability term is fixed in subsequent equations.

Thus the model considers master studies separately from subsequent PhD studies. This is in accordance with the South African higher education system. In South Africa, differently from the US education system for example, holding a master degree or equivalent postgraduate qualification is a prerequisite for being

admitted as a PhD student (Department of Education, 1998, Du Toit, 2012). Nevertheless a problem may arise in practice when the master degree is seen as the first step in the PhD process, and so the master and PhD are considered two parts of one educational programme. When this is the case, selection takes place before rather than after the master degree. We investigate how that may affect our estimations in the sensitivity analysis by restricting attention to individuals who have master and PhD degrees from different institutions.

5.3. Identification and Estimation.

Identification. Identification rests on both the triangular form and exclusion restrictions.⁹

Table ?? shows the interdependence of the variables across the three equations. The first column lists all endogenous and exogenous variables used in the model. The subsequent three columns correspond to the three equations which model the endogenous variables, i.e. Ability, PhD university category, and Rating respectively. Whenever a variable enters an equation the table entry is marked with a cross. Here we can see the system as a whole. We note that endogenous variables form a triangular system, and each equation includes at least one exogenous factor which is excluded from all other equations.

TABLE 5. Variables by model equation

Variables	Equation (??) Ability (ν) measurement eq.	Equation (??). PhD univ. cat. (D) selection eq.	Equation (??) Rating (R) outcome eq.
<i>Endogenous</i>			
Ability (ν)	-	x	x
PhD univ. cat (D)	-	-	x
Rating (R)	-	-	-
<i>Exogenous</i>			
	V	Z	X
Age at master	x	-	-
Master distinction	x	-	-
Master univ. cat.	x	x	-
PhD period	-	x	-
Race	-	x	x
Gender	-	x	x
Scientific domain	-	x	x
Experience	-	-	x

Neither age at master nor master distinction is assumed to influence directly subsequent selection or performance equations. Clearly, distinction in a master's

⁹Only relying on structural form and error term distribution assumptions does not allow for reliable identification in practice.

degree does enter the deliberations of a PhD admissions committee, but we assume that it enters as a signal of innate ability and so should enter our ability equation. Furthermore, it seems unlikely that these characteristics exert a strong influence on an international expert committee which judges scientific excellence of an established researcher twenty years later.¹⁰

We assume that selection, training, and individual effects are all stable over time. Since it might well be the case that the comparative advantage of South African universities has improved against foreign universities over time, allowing for a trend in the selection and training effects would be desirable. Unfortunately, the data do not provide sufficient information to estimate a reliable time trend.

Estimation. The structural model constituted by equations ??, ?? and ?? is formulated in classical terms, and given above assumptions, is identified under classical conditions. We estimate the model however within a Bayesian framework, using the Markov Chain Monte Carlo process and the No U-turn sampler.¹¹ This approach is due to computational efficiency and has been already followed in prior econometric work such as (Hansen et al., 2004) for example. The Bayesian framework however necessitates to fix some prior distributions of our parameters and unobserved factors. We use flat, uninformative priors.¹²

Proper convergence of the MCMC estimation is ensured by following the suggestions of Gelman and Rubin (1992). In particular we run multiple Markov Chains starting from different initial values and verify that the chains do not diverge but instead describe the same distribution (are 'well mixing'). Furthermore we make sure that the potential scale reduction is estimated to be close to one for all scalar estimates. Finally, estimation results of the complete model are found to be reasonably close to single equation estimations.

¹⁰Age might also influence the student's decision of PhD university category. Older students are more likely to have a family for example and therefore less willing to leave the country than younger students. The year the PhD decision was made is however not available, and this decision is not necessarily made directly after master. The best available proxy is the age when PhD degree has been obtained. Age at master is significantly correlated with age at PhD (with a correlation coefficient of about 0.85). Introducing that variable into the PhD selection equation however is doubtful because the estimated coefficient of age at (end of) PhD is particularly strong and significant for top-tier foreign universities but not for other foreign universities. This signals that age at PhD is either a selection criterion or an outcome of the PhD program. Therefore we do not introduce this variable.

¹¹The results are not sensitive to the sampler used.

¹²These are the default prior distributions implemented in RStan, and in fact worked best of all those that we examined.

6. MAIN RESULT

Detailed discussion of the model and coefficient estimates is presented in sections ?? and ?. Before turning to those details, though, we discuss how well the model fits the observations. We compare predicted values and observed values for two variables: where students did their PhD training; and, given where the PhD was granted, their rating outcomes.

6.1. **Model fit.** Before we turn to estimates of specific coefficients, we assess how the model fits the data more generally. Table ?? compares the fitted probability of selecting a PhD university category with the observed choice. For example, of the students who received their PhDs from “other SA” universities, 61 percent were predicted by the model to receive their PhDs from other SA; 20 percent were predicted from SA Afrikaans universities, 7 percent from SA English, and so on. What is striking here is that looking at any column, which captures where the model predicts students will enrol as PhD students, the heaviest entry is on the diagonal — when the model predicts that a student receives a PhD from a university of type x , this is in fact where we are most likely to see him or her in the data.

TABLE 6. Fitted probabilities of PhD university category, averaged over observed PhD university. To be read as: of those observed to graduate from type 1, (other SA), 61 % were predicted to graduate from type 1, 21 % from type 2, 7% from type three etcetera.

Observed choice	Fitted Probability				
	1	2	3	4	5
Other SA (1)	0.61	0.21	0.07	0.08	0.03
SA Afrikaans (2)	0.16	0.64	0.09	0.07	0.04
SA English (3)	0.13	0.15	0.59	0.09	0.04
Other foreign (4)	0.24	0.28	0.29	0.14	0.06
First tier foreign (5)	0.16	0.30	0.33	0.12	0.09

In fact, the probability assigned by the model, of attending a particular PhD university (type) is throughout highest for those which did in fact graduate from that type. In particular for the SA PhD university categories, we find a high probability mass of around 60 percent on the categories which in fact have been observed (see values on the diagonal of Table ??). Fitted probabilities to attend a foreign university category are highest for those that actually went abroad (seen by reading down the last two columns.)

Table ?? compares for each group of PhD university category graduates the actual and fitted average rating outcomes. Fitted shares follow the observed ones

relatively closely.¹³ χ^2 tests are not able to reject the hypotheses that observed and fitted ratings stem from the same distributions.¹⁴

TABLE 7. Average rating outcome probability for PhD university categories with χ^2 statistics

	R.U.	C	B	A	χ^2	p-value
<i>Other SA university (667 ratings)</i>						
Observed	0.1634	0.6672	0.1439	0.0255		
Fitted	0.1618	0.6636	0.1579	0.0167	3.9138	0.2709
<i>Afrikaans SA university (871 ratings)</i>						
Observed	0.1022	0.6854	0.2032	0.0092		
Fitted	0.1034	0.6761	0.2027	0.0178	3.7292	0.2922
<i>English SA university (543 ratings)</i>						
Observed	0.0755	0.5893	0.2891	0.0460		
Fitted	0.0586	0.6002	0.3007	0.0405	3.4119	0.3324
<i>Other foreign university (251 ratings)</i>						
Observed	0.0717	0.5418	0.3307	0.0558		
Fitted	0.0634	0.5836	0.3079	0.0452	2.0711	0.5578
<i>First tier foreign university (132 ratings)</i>						
Observed	0.0152	0.5758	0.2955	0.1136		
Fitted	0.0289	0.5122	0.3761	0.0828	5.7075	0.1267

6.2. Regression results. Table ?? shows the coefficient estimates of the model. The first two lines of Table ?? show average ability and average training for each PhD university category — capturing the selection and training effects in latent ratings outcomes.

Average ability is derived from the estimates and is the expected ability of PhD students in each university category. Universities of higher reputation in general are found to attract more able students. In South Africa, average ability

¹³The (slight) over- and under-fitting in some cases can be traced back to the restrictions which we impose on the distribution of the latent rating outcome (R^*). In particular, we model the mean of latent rating outcomes and not the variance and skewness; be it directly through the error term in the rating function or indirectly through the ability equation. We experimented with expanding the model in that direction, but relaxing the distributional assumption comes at the cost of weaker identification, lower estimation precision, and potential over-fitting which we found to be too high.

¹⁴The χ^2 -statistics are calculated as follows: $\chi^2 = \sum_i \frac{N(p_i - \hat{p}_i)^2}{\hat{p}_i}$, where the sum is over ratings i , p_i and \hat{p}_i denote observed and fitted probabilities respectively, and N denotes the number of ratings observed for the respective PhD group. The degrees of freedom of the χ^2 distribution is three in all cases.

is estimated to be lowest at 'other SA universities' and highest at English SA universities. 'Other foreign universities' attract students of the same quality as the best SA universities. 'First tier foreign universities' host the most able students, as they have an average ability which is significantly higher than that of any other university. Overall, the estimates of ability over universities are relatively precise, though not always significant.

Average training is estimated for each university category using a dummy variable in the rating equation, with "other SA university" being the reference category fixed to zero. The training effect estimate of "Afrikaans SA university" is not significantly different to "other SA university". "English SA university", however, provides a significant training effect, estimated to be on the same level as 'first tier foreign university'. The training effect estimate of 'other foreign university' is somewhat lower but not significantly so. Relatively high standard errors suggest that there is considerable uncertainty in scientific careers of PhD graduates of any university category.

Comparing ability and training, we can observe that both follow roughly the same pattern. What is striking, however is that across university categories heterogeneity, i.e. the spread of coefficient estimates, is larger for average ability than for average training; and that estimates for ability tend to be more precise than training estimates.

We turn now to the estimates of coefficients in the three equations of the model. The section 'ability equation' in Table ?? shows that ability, at least as judged by PhD admissions committees, is particularly high for young students who received a distinction on their master degree. These effects are fixed over university categories. Although coefficients are positive for Afrikaans SA universities, as might be expected, they are not significantly different from zero due to the relatively high standard errors. English SA universities issue (signals of) ability significantly higher than all other SA universities.

Section 'Selection equation' in Table ?? contains the regression results pertaining to the PhD selection equation. The first column provides the variables, and each subsequent column corresponds to one type of PhD university. Coefficients of factors that vary only across individuals and not over alternatives are set to zero for the baseline category 'other SA university'.

In detail, inertia value contains four dummy variables which indicate whether the PhD university (category) in question was the same as the student's master university (category). The tendency of master students to stay at 'their' university for their PhD is reflected in the positive and significant coefficient estimates for

TABLE 8. Estimation results

	SA university		Foreign university			
	Other	Afrikaans	English	First tier		
Selection and training effects in latent rating (R^*)						
Avg. ability	-0.28 (0.146)	0.04 (0.209)	0.47 (0.247)	0.43 (0.221)	1.02 (0.302)	
Avg. training	0	0.16 (0.212)	0.74 (0.265)	0.54 (0.279)	0.7 (0.366)	
Ability equation (γ)						
Master university	0	0.21 (0.175)	0.93 (0.219)	-	-	
Master distinction	0.75 (0.104)	0.75	0.75	-	-	
Age at master	-0.77 (0.12)	-0.77	-0.77	-	-	
Selection equation (α_p, ζ_p)						
Inertia value	2.15 (0.217)	1.95 (0.201)	2.76 (0.244)	-	-	
Non-white	0	0.17 (0.364)	0.33 (0.367)	1.39 (0.362)	1.03 (0.514)	
Female	0	-0.14 (0.231)	-0.02 (0.267)	-0.14 (0.307)	-1.4 (0.555)	
Social sciences	0	-0.57 (0.205)	-0.93 (0.25)	-0.19 (0.266)	-0.42 (0.365)	
PhD period dummies	yes	yes	yes	yes	yes	
ζ	0	0.03 (0.148)	0.16 (0.17)	0.3 (0.18)	0.65 (0.24)	
Rating equation (β)						
Intercept	0.45 (0.25)	0.45	0.45	0.45	0.45	
Experience	1.9 (0.192)	1.9	1.9	1.9	1.9	
Experience ²	-0.23 (0.045)	-0.23	-0.23	-0.23	-0.23	
Non-white	-0.62 (0.182)	-0.62	-0.62	-0.62	-0.62	
Female	-0.37 (0.116)	-0.37	-0.37	-0.37	-0.37	
Social sciences	-0.33 (0.11)	-0.33	-0.33	-0.33	-0.33	
Rating equation hurdles (c)						
C3	C2	C1	B3	B2	B1	A
0	1.75 (0.076)	3.75 (0.095)	4.8 (0.105)	5.6 (0.113)	6.75 (0.133)	8.25 (0.172)
No. individuals: 1137			No. observations: 2464		log-Lik: -5324.71	

Notes: i) Std.dev in (). Coeff. without () are either fixed across universities to balance model complexity and data fit or are fixed to zero for identification.

ii) - if coefficient not applicable.

all four SA university categories. These coefficients can be seen as capturing the value of inertia in PhD university choice.¹⁵

Students of colour (non-white students) are particularly likely to study abroad. Female students are less likely to enter first-tier foreign universities. Social science PhDs are less sought at universities in the English SA university group. PhD period dummies are included in the regression but not displayed for brevity. They have some influence on the probability that a student is selected into a particular class of PhD university, but changes from one period to the next are not overly strong when considering the standard deviation of coefficients.

As expected, in general the aggression with which universities examine ability at the end of the master degree increases as we read the estimates for ζ from left to right on the bottom row. Roughly speaking this accords with the intuition that higher ranked universities will apply stricter standards for PhD admission. This however does not mean that English SA universities fall short in talented master students, as they simply recruit (perhaps “excessively”) their own master students. Average ability, the first statistic presented in the table, in combination with Inertia value, suggests that the strong tendency of English SA master students to stay for PhD offsets the selection factor ζ .

The two subsequent parts of Table ?? pertain to the rating equation. The coefficient estimates of the rating equation, β , represent the estimation of the real-valued, latent R^* . The bottom panel gives the threshold values that convert from R^* to the letter rating given in the raw data. To interpret the β coefficients it is necessary to make reference to the thresholds c .

For example: The average career trajectory is captured by a monotonically increasing function with decreasing slope, given the estimate of Experience (1.9), measured in decades since PhD, and its square term (-0.23). The experience curve obtains its maximum of 4.37 after around 40 years, which is most likely after the end of the career. For white, male, natural science researchers the trajectory is shifted upwards by an intercept (0.45). Ignoring ability, PhDs from “other SA university” can be expected to obtain a B rating just after 40 years (since $4.37+0.45 = 4.82 > 4.8 = B3$). A PhD graduate from a first-tier foreign university adds a training effect in average of 0.7 at each rating event, which lets us expect a B after about 30 years experience. Average ability at first-tier foreign university adds another 1.02, and thereby decreases the time needed for a B rating to 18 years.

¹⁵Recall that the sample includes only students who completed their masters degrees in South Africa, so there can be no inertia for foreign universities.

Finally, being female reduces one’s latent rating, as does being a social scientist. Both effects are of similar magnitude and statistically significant at 10 percent significant levels.

6.3. Training and selection effects. Speaking loosely, a scientist’s rating is affected by his or her innate ability and his or her training. Where ability stops and training starts is in general an open question. We have defined ability as measured at the end of a master degree. Training then refers to training received during the course of completing a PhD and further gains in terms of signalling effects or social capital. In this section we identify the magnitudes of those two effects separately.

The econometric model identifies treatment and selection effects through the joint estimation of the parameters β and ζ respectively. Given the model and data, we obtain the conditional probability of obtaining a certain rating for each individual:

$$P(R_{i,t} = j | X_{i,t}, \nu_i, \beta_p),$$

where X is the vector of controls, ν_i individual ability, and β_p training effects of the university of type p .

The probability is conditioned on both ability and training (among other things), so taking one marginal distribution (conditional on either ability or training) will give us the distribution of one or the other effects. The expected value of that marginal distribution will be the average size of the effect.¹⁶

Put another way, what we refer to as the training effect captures the following thought experiment. Send a student randomly chosen from the population to different universities to do a PhD. Different universities give different training. Depending on which university he or she goes to, we would make different predictions regarding his or her rating.

¹⁶More formally, the training effect is defined here as

$$P(R = j | D = p) = 1/N \sum_{i \in N} 1/T_i \sum_{t \in T_i} P(R_{i,t} = j | \mathbf{X}_{i,t}, \nu_i, \beta_p).$$

We define a selection effect similar to the training effect, but this time we integrate out the training effect. The selection effect enters through the fact that the distribution of unobserved ability ν , $\Phi(\nu_p)$, can be expected to vary across PhD universities.

$$P(R = j | \Phi(\nu_p)) = 1/N \sum_{i \in N} 1/T_i \sum_{t \in T_i} P(R_{i,t} = j | \mathbf{X}_{i,t}, \nu_p, \beta_s),$$

where ν_p is normally distributed with mean average ability at university p , and empirical standard deviation of ability in the population. Training β_s is from a randomly drawn university. Averages are obtained in each simulation/estimation iteration within MCMC.

Table ?? shows these probabilities (estimations, with confidence intervals next to each estimate). Each column represents the probability of having a certain rating outcome (or higher) given that a randomly chosen student attends a particular type of PhD institution. Reading along any row, the probability decreases by construction as the event “ $\geq C$ ” includes obtaining a C as well as the higher B, and A ratings. We present here estimates for broader rating categories for ease of exposition.

TABLE 9. Training effect estimates, 90 percent confidence interval ()

	$\geq C$		$\geq B$		A	
Other SA	0.879	(0.858, 0.9)	0.219	(0.186, 0.255)	0.021	(0.015, 0.027)
Afrikaans SA	0.892	(0.874, 0.908)	0.237	(0.215, 0.262)	0.023	(0.018, 0.03)
English SA	0.929	(0.91, 0.946)	0.313	(0.274, 0.353)	0.038	(0.029, 0.048)
Other foreign	0.917	(0.892, 0.938)	0.284	(0.24, 0.335)	0.032	(0.023, 0.043)
Top foreign	0.927	(0.894, 0.953)	0.311	(0.248, 0.381)	0.038	(0.024, 0.054)

Students least likely to receive a successful rating (that is, $\geq C$) are those who receive their PhDs from a generic South African university/technikon. Most likely are those who go to a top university, be it in South Africa or abroad. The probability of obtaining a B rating or higher conditional on training effects varies between 22 percent for the generic SA university and 31 percent for top universities. Chances to obtain an A rating are small for all university categories but follow a similar pattern as B ratings.

To illustrate, focus on the second column of Table ??, B rating or higher, which signals research excellence by international standards. 'Other SA' and 'Afrikaans SA' do not yield significantly different training effects, as each point estimate falls into the confidence interval of the other coefficient. In the same column, the three university categories, English SA, Other foreign, and Top foreign, are not significantly different from each other. Between these two broader groupings we find some difference in training effects (below 10% significant levels) but differences are not very stark, about 10 to 5 percentage points. This suggests that “foreign” does not in any simple sense dominate “South African”. University quality rather than location is what matters.

The sorting parameter ζ creates differences in expected ratings conditional on PhD university ex ante, or net of, the treatment. As for the rating effect, the selection effect corresponds to the expected value of the marginal distribution, but this time we integrate out the training effect of PhD universities.

Perform a similar thought experiment: What would I achieve with having just an average kind of training but the brain of a Harvard/other foreign/English SA/...student?

Selection effects displayed in Table ?? follow the same pattern as training effects. For example, obtaining a C rating or higher is, again, least likely for PhD students that obtained their PhD from the base category 'other' SA universities. This time however, as probabilities are based on 'average training effects', so the reason for the pattern lies in the lower estimated ability of the students selected into those universities. Overall, we estimate that selection effects introduce considerable heterogeneity of scientific performance across PhD university categories.

First tier foreign universities attract the brightest students. This result is also statistically significant: the 90 percent confidence intervals of first-tier foreign university students do not contain the point estimates of other university categories.

TABLE 10. Selection effect estimates, 90 percent confidence interval ()

	$\geq C$		$\geq B$		A	
Other SA	0.872	(0.844, 0.899)	0.19	(0.16, 0.223)	0.013	(0.009, 0.017)
Afrikaans	0.899	(0.878, 0.917)	0.23	(0.2, 0.26)	0.018	(0.013, 0.023)
English	0.925	(0.91, 0.94)	0.286	(0.248, 0.325)	0.025	(0.019, 0.034)
Other foreign	0.924	(0.906, 0.939)	0.282	(0.245, 0.322)	0.025	(0.018, 0.033)
Top foreign	0.95	(0.935, 0.963)	0.367	(0.309, 0.426)	0.04	(0.027, 0.057)

A further noticeable result is how the ranking of 'other' foreign university changes for the selection effect compared to the training effects. While the training effects of 'English SA universities' have been estimated higher than training at 'Other foreign universities', selection effects appear to be on the same level for both. Estimated differences however are not significant.

Finally, comparing Tables ?? and ?? we observe very similar patterns. Examining any column ($\geq B$ for example) in the two tables we observe that the magnitudes of the coefficients are very similar. In addition, the pattern across the rows, which indicates the relative value of training or selection at the different university types, is also similar in the two tables — except that the probability of obtaining a higher rating across university types is slightly more spread and more precise for selection than for training. What this suggests is that the effect of a foreign PhD on rating arises from both selection and from training, and that the magnitudes of the two effects are roughly equal: selection is at least as important as training.

6.4. Sensitivity analyses. To check the sensitivity of the results we performed additional estimations on simpler as well as more elaborate models and on variations of the sample and coding of variables. Table ?? provides one essential insight gained from each estimation, i.e. the probability of obtaining a rating B or higher due to selection and training effects as defined above.¹⁷ The information in Table ?? thus corresponds to the second column ($\geq B$) in Tables ?? and ??.

We estimated a simple ordered logit model with rating outcomes as the dependent variable, both with and without pre-PhD characteristics (Master distinction, Master university, Age at master). Model 1 is based on the assumption that the PhD university category is exogenous and thus likely to suffer an omitted variables bias. This bias is somewhat reduced in Model 2 with the introduction of pre-PhD variables as a proxy of individual “ability”. The weakness of the ordered logit is that it does not distinguish between selection and training effects, and thus is open to the objection that the better outcomes of foreign PhDs may be purely due to selection, and nothing to do with the “location” of the PhD. Model 3, the Heckman selection model without the ability equation (without pre-PhD characteristics), is also not able to identify selection effects; all ex-post heterogeneity is attributed to training and nothing to selection. Thus, Model 1 to 3 provide higher training effect estimates than those presented in the main analysis. We also find however that identification of selection effects hinges on pre-PhD characteristics, which is made possible with the specification of the model presented in the main analysis.

Model 4 goes one step further in that it allows for ability-dependent training effects.¹⁸ This more complex model provides essentially the same insights as the main model we present in the text above. Models 5 to 8 provide variations on the sample and variables. Model 5 restricts the sample by including only researchers which obtained their PhD between 1980 and 2003 and not before. Model 6 restricts to students switching university from master to PhD. In both models, results are less precise due to the smaller sample size, but we observe again the same pattern as in the main analysis — selection effects are more pronounced than treatment effects. Results are also consistent as the confidence bands of alternative regressions and those presented as main results have strong overlap. Model 7 considers all foreign universities within the top 100 of the Shanghai ranking as first tier foreign, which reduces both training and selection effects of that category compared to the main results. Model 8 differentiates two groups of English South African

¹⁷This limited presentation is due to space constraints but we refer the interested reader to the working paper, Barnard et al. (2016), for a thorough discussion of the models and results.

¹⁸This model is based on the idea that good training actually fits the capacity of the student. In that model training effects are freely estimated not only across university categories but also across bins of “ability”. Barnard et al. (2016) provides a complete description of that model.

universities based on their Shanghai ranking (i.e. grouping University of Cape Town together with Witwatersrand University on the one hand, and University of Natal with Rhodes University on the other hand, see Table 1).

TABLE 11. Probability of a rating higher or equal to B due to selection and treatment effects from alternative regressions (individuals N, observations NT, log-Likelihood LL)

	Other	SA university Afrikaans	English	Foreign university Other	First tier
1) Ordered logit without pre-PhD variables (N=1137, NT=2464, LL=-4331.69)					
Training	0.18 (0.162, 0.199)	0.223 (0.206, 0.24)	0.363 (0.336, 0.391)	0.322 (0.283, 0.36)	0.429 (0.371, 0.491)
2) Ordered logit with pre-PhD variables (N=1137, NT=2464, LL=-4265.09)					
Training	0.207 (0.184, 0.233)	0.218 (0.199, 0.24)	0.345 (0.313, 0.376)	0.308 (0.272, 0.345)	0.392 (0.339, 0.445)
3) Heckman selection without pre-PhD variables (N=1137, NT=2464, LL=-5849.97)					
Selection	0.297 (0.26, 0.337)	0.295 (0.264, 0.328)	0.292 (0.256, 0.329)	0.288 (0.247, 0.333)	0.28 (0.226, 0.335)
Training	0.177 (0.152, 0.203)	0.216 (0.194, 0.238)	0.37 (0.331, 0.411)	0.331 (0.273, 0.389)	0.451 (0.364, 0.549)
4) Ability dependent training effects (N=1137, NT=2464, LL=-5594.07)					
Selection	0.173 (0.155, 0.191)	0.233 (0.211, 0.255)	0.29 (0.266, 0.314)	0.281 (0.258, 0.305)	0.354 (0.319, 0.386)
Training	0.235 (0.21, 0.26)	0.216 (0.196, 0.235)	0.295 (0.27, 0.32)	0.286 (0.252, 0.317)	0.302 (0.259, 0.343)
5) PhD obtained year period 1980-2003 (N=978, NT=1774, LL=-3911.3)					
Selection	0.142 (0.115, 0.173)	0.16 (0.13, 0.194)	0.203 (0.163, 0.244)	0.199 (0.157, 0.244)	0.292 (0.23, 0.355)
Training	0.123 (0.1, 0.149)	0.176 (0.15, 0.205)	0.244 (0.201, 0.29)	0.228 (0.168, 0.294)	0.204 (0.136, 0.278)
6) Switching sample (N=445, NT=983, LL=-2379.43)					
Selection	0.197 (0.162, 0.234)	0.217 (0.182, 0.253)	0.283 (0.237, 0.326)	0.313 (0.268, 0.361)	0.41 (0.346, 0.472)
Training	0.267 (0.219, 0.319)	0.305 (0.26, 0.352)	0.348 (0.295, 0.405)	0.298 (0.255, 0.343)	0.314 (0.254, 0.377)
7) Shanghai top 100 in first tier foreign (N=1137, NT=2464, LL=-5337.52)					
Selection	0.193 (0.165, 0.226)	0.229 (0.199, 0.263)	0.284 (0.248, 0.323)	0.266 (0.222, 0.313)	0.341 (0.299, 0.383)
Training	0.217 (0.187, 0.248)	0.238 (0.212, 0.264)	0.315 (0.277, 0.354)	0.313 (0.252, 0.376)	0.284 (0.235, 0.336)
8) English split into first and second tier (N=1137, NT=2464, LL=-5440.79)					
Selection	0.185 (0.157, 0.217)	0.228 (0.197, 0.26)	0.315 (0.263, 0.366)	0.28 (0.242, 0.318)	0.381 (0.325, 0.438)
Selection			0.311 (0.263, 0.356)		
Training	0.23 (0.199, 0.262)	0.242 (0.216, 0.269)	0.275 (0.218, 0.337)	0.291 (0.247, 0.338)	0.304 (0.238, 0.374)
Training			0.305 (0.257, 0.358)		
9) University category of first SA employment after PhD (1129/2448/-5292.93)					
Selection	0.196 (0.168, 0.229)	0.235 (0.206, 0.266)	0.277 (0.241, 0.316)	0.278 (0.243, 0.316)	0.358 (0.299, 0.416)
Training	0.232 (0.201, 0.264)	0.234 (0.211, 0.261)	0.314 (0.275, 0.355)	0.281 (0.238, 0.325)	0.311 (0.245, 0.388)

Selection and training effects are estimated to be very similar for both groups. Finally, the estimated training effects in the main analysis include both: gains realized during the PhD period, and opportunities created due to the PhD but realized thereafter. For example, we may expect that, everything else equal, PhDs from high reputation universities are more likely to get a position in a resource rich environment than PhDs from lower reputation universities. Model 9 addresses this issue by including the university category of first South African employment after PhD in the rating equation. Although first employment at an English South African research university is estimated to be significantly positive, estimated training and selection effects stay close to those from the main model. One explanation, supported by the data, is that foreign PhDs tend to work at the same institutions as local PhDs from top South African universities; with both groups having a similar profile after master.

Which insights remain from the main analysis given the sensitivity checks? Two statements found broad support throughout: First, young students obtaining their master degree from leading SA universities with distinction tend to do their PhDs at universities of higher reputation. This sorting explains heterogeneity of career achievements as much or more than differences in PhD training across universities. That universities of a strong, world-wide reputation attract the most able students found ample support. Second, SA universities do not necessarily provide worse training than foreign universities in Europe or the U.S., at least as regards the effect on those pursuing a career in South Africa. We should note, though that in some estimations, training effects of foreign universities appear unreasonably low. This might be caused by our sample which includes only scientists returning from a foreign PhD, creating a negative bias. Therefore, we need to interpret our results from the perspective of the developing country. We are not attempting to identify the quality of training at different universities; we are identifying the value and the sources of value of sending students abroad as opposed to training them at home. With that firmly in mind we can conclude that sending students abroad rather than supporting local PhD programs does not necessarily inject higher scientific quality into research faculty.

7. DISCUSSION AND CONCLUSION

Scholarship about the role of universities in development has seen something of a hiatus in recent years as scholars struggled to disentangle the functioning of the manifold local and global institutions and influences shaping indigenous universities. In this paper, we focus on universities as providers of scholars — people who advance knowledge. We ask the question to what extent universities

in developing countries are capable of developing scholars who do cutting-edge research, and to answer that question, we focus narrowly on the scholarly success of academics. We compare the scholarly achievements of academics working in a developing country who have received their PhD locally with that of their (local) colleagues who trained abroad. To better understand the differences we observe, we differentiate between selection and training effects.

The PhD can be seen as an “entry gate” to a certain transforming trajectory. Using a unique dataset, we have been able to differentiate between different trajectories, and specifically the training and selection effects of receiving a PhD from different institutions. Our econometric model is an extension of a model that has often been used in the evaluation literature (notably in the schooling effects literature where it was introduced by Heckman) in order to account for a confounding selection effect: the factor structure model with a selection and outcome equation, where unobserved individual characteristics (typically interpreted as individual ability) are factored into a selection and an outcome equation. We expand this model by an additional measurement equation which allows for including proxies of individual ability which is not possible in the more traditional approach.

We find that the quality of the PhD granting institution is indeed correlated with future career success as an academic researcher in South Africa. Somewhat unexpectedly, the magnitudes of the effects of training and selection on career success are very similar, although selection seems to exert a somewhat stronger effect. However, our evidence suggests that the fairly common view that foreignness equates to quality does not capture the nuances of university training in a developing country, at least not in terms of research produced by academics.

Our finding that selection effects matter is hardly surprising. The perception that foreign universities offer superior training has a strong effect in terms of the (self-)selection of students. Students are particularly attracted to the universities that would be found at the top of any quality ranking — Oxford, Harvard, Stanford and so on — but this is not unexpected. Students from advanced economies also strongly prefer to study at those institutions. However, even the foreign universities that are not as highly ranked exert a very strong selection effect. This seems to be because the perceptions of the country rather than of a given institution tend to dominate assessments of quality by potential foreign students. As Marginson (2006) puts it: “for the foreign graduates returning home to Thailand or Tajikistan, all reputable foreign degrees provide positional value”. Our evidence also suggests that for individuals, study abroad carries a positive payoff. Individuals from a developing country may well decide to seek out that payoff, even given the substantial expense and disruption to them and their partners and families. The

implication for policymakers is that they cannot substantially influence *selection* preferences for those students.

However, once selection effects are controlled for, we also find very clear evidence that the leading local universities are “world-class” in the training that they offer. We emphasise here that this statement does not apply to PhD training *tout court*, but rather reflects the effect of PhD training on those who return to and pursue academic careers in South Africa. The matter of whether or not academics return to developing countries is non-trivial. Many developing economies sponsor the PhD studies of their citizens in developed economies (see Delicado, 2010 re Portugal; Güngör and Tansel, 2008 re Turkey; and Song, 1997 re South Korea), but a recurring observation of scholars is that only a subset of those whose studies abroad were sponsored by their government is likely to return.

This implies an important limitation in the interpretation of our results, as we only observe PhD students who stayed in South Africa or came back to South Africa. One could imagine that the most promising recipients of PhDs from South African universities could compete on the international job market, and leave the country to pursue their careers abroad. While perceptions about the quality of PhD training in developing countries make such scenarios relatively unlikely, to the extent that it does happen, the training and selection effect of South African universities would be underestimated. By contrast, if it is primarily scientists who did not “make it” abroad who return to South Africa, then the training and selection effects of foreign universities are underestimated. These dimensions cannot be controlled for with the current data. Again though, the over- or under-estimation of effects is about PhD training *per se*, whereas our fundamental interest is in PhD training as it affects those who return to or stay in South Africa. When that restriction of the issue is applied, these specific concerns about over or under estimation do not apply.

Even with those caveats, we provide evidence of a high quality of scholarly training at a number of the local universities. This finding may appear counter-intuitive at first; the understanding has long been that universities in developing countries are likely to develop “technicians” of research rather than scholars who can meaningfully advance their field. However, in an increasingly globally connected world the challenges faced by developing country universities are likely to attract more attention than before. Scholars from those universities who possess the fundamental skills needed to do meaningful research in the field may be able not only to join the global networks where knowledge in their fields is advanced, but also to find that their particular contextual perspective is deemed valuable.

This means that the greatest lever for policymakers relates to the PhD *training* done by developing country universities. Developing country universities can offer PhD training that is comparable to that offered at universities in industrialised countries. Future research is needed to understand the mechanisms that allow universities in a developing country to offer PhD training that allows its graduates to become research leaders in their own right, but our work already sheds light on some of the enabling conditions. The negative effect of underinvestment is evident from the poor performance of the historically black (“all other”) South African universities, and clearly, universities need adequate financial support in order to be able to provide quality doctoral training.

But although the historically English and Afrikaans universities received a similar level of financial support, our evidence also shows that the historically English universities outperform the historically Afrikaans universities in the development of research leaders. The historically English universities had the dual advantage of greater connectedness with the leading English-speaking universities worldwide and an arm’s length relationship with the government of the day. Policymakers are well advised to pursue both.

In particular, our evidence suggests that the close relationship Afrikaans universities had with government did not translate into scholarly excellence. Academic freedom and the autonomy of universities — rather than a privileged relationship with government — work in favour of robust research (Altbach, 2004; 2015). In South Africa under apartheid, English universities with their more limited ties to the government tended to have greater autonomy than the Afrikaans universities that in many ways were tied to the ruling party (Bunting, 2006). Indeed, Grogger and Hanson (2015) find that democratisation of the home country is associated with greater return rates by foreign-trained PhDs, and the step to achieve greater university autonomy was an important milestone in the reforms of the Thai university system (Liefner and Schiller, 2008). Ensuring that universities are buffered from political pressures is an important policy lever for developing countries wishing to improve the quality of scholarship domestically.

Finally, there is a substantial body of literature documenting how scholars from countries behind the technological frontier are connected with scholars, often from their own home country, in industrialised countries, and of the benefits that accrue to all parties (Baruffaldi and Landoni, 2012; Jonkers and Cruz-Castro, 2013; Jonkers and Tijssen, 2008; Li, Miao and Yang, 2015). But in addition to using the diaspora wisely, practical measures like support for researcher mobility through

streamlined visa regulations (for both incoming and outgoing researchers) and internal arrangements that enable scholars to travel to scientific events are also likely to be important.

Our results permit some optimism regarding possibilities for universities in emerging economies to catch up with their advanced economy competitors. Some local universities are good, and improved support might allow the best universities in emerging economies to become recognised worldwide for the quality of science and scientific training conducted there. The top global universities will always attract the best students, but as the country and its education system gains confidence in its (best) universities, the best local universities should be able to attract better students than the competent but non-leading foreign ones.

The study is not without limitations. Methodologically, in addition to the already-discussed matter that we do not observe non-returning South Africans who pursued their PhD training abroad, there is also the issue of how we measure selection and training. The disentanglement of selection from training effects is only as good as the instrument and observed ability measures. Our main instruments, “period of PhD” and the ability measures “Master institution” and “distinction in Master degree”, are good predictors of university selection but of course less than perfect. Thus it is possible that treatment effects in particular of foreign top tier universities are somewhat overestimated. This is an issue shared with any study of our kind and in general little can be done to formally address it. Overall, however, our results are very robust, and there is no evidence that the highest ranked universities of the world do not have a positive effect on their bright students.

We are considering historical data within an evolving system. In the process of inferring on the past, we have had to assume some stability in the system and that the effects we are measuring did not change over the period of our sample. In the context of South Africa, this is not true. In particular, in addition to relatively small adjustments after the end of apartheid in 1994, the university system underwent extensive changes from 2003 to 2005. We address this partly by our “period of PhD” variable and partly by limiting our analysis to PhDs that were obtained by 2003 or earlier. Unfortunately the data are not strong enough to include this possible instability in the analysis. Also, given the extensive changes to the system, our findings should not be used to infer anything about the current quality of South African universities.

This limitation is particularly of concern because our data span both the apartheid and post-apartheid era. Although we believe that the general processes of “selection” and “training” are valid throughout, different racial groups likely experience

them differently. For example, given that less than 10 percent of the South African population is classified as white, it is noteworthy that the vast majority of rated and highly rated scholars still are white decades after Apartheid formally ended. Teasing out how race affects scholarly success is an important avenue for future research. In addition, scholarly success is only one metric of the impact of a university. We believe that the NRF data provide excellent insight into the scholarly success of South African academics, which is what this paper is about. But the bulk of South African academics do not do research at all, and their work is not captured in that database. Understanding the functioning of that part of the university system is a very different but nonetheless important research question.

Our work does suggest that in addition to the selection effects of foreign versus local universities, there is a local selection effect. The top local universities seem to benefit from a virtuous cycle where the top local students who do not wish (or are not able) to go abroad will prefer study at the leading local institutions. Policymakers may respond to local preferences by nurturing a very small group of select institutions so that they indeed become globally acknowledged for the quality of scholarship conducted there, or may try to increase the number of locally recognised quality institutions by spreading financial and other support more widely. The likely developmental implications of these choices are very different and it may be “just as important for academic institutions to be ‘national’ or ‘regional class’” (Altbach, 2015:6). Further research is needed to understand how development is best served.

Another avenue for further research relates to social rather than human capital development. PhD training allows scholars to develop not only research skills but also relationships that can provide them with collaborators or at least access to the informal scientific networks in a field. It seems likely that PhD training abroad provides scholars with a richer social network than local PhD training, which would increase the relative value of foreign training. We believe that further research is needed into how local versus foreign PhD training shapes the social capital of researchers.

In sum, our findings allow some confidence about the role of local universities in the upgrading of developing countries. It suggests that an investment in the local science system is likely to result not only in the development of fairly basic local skills, as was previously believed, but can also result in cutting-edge science. While there is value to foreign training, countries need to consider carefully to which foreign universities they send PhD students, as not all foreign PhD training is equally useful. Indeed, from a theoretical perspective, our study suggests that the discussion about the role of universities in upgrading may be better served by

differentiating universities in terms of the roles they serve in their economy rather than along the better established “local” and “foreign” dimension. The boundary experienced for students and scientists between Oxford and the University of Cape Town may well be more permeable than the boundary between the University of Buxtehude and Oxford.

REFERENCES

- Aakvik, A., Heckman, J. J., and Vytlacil, E. J. (2005). Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs. *Journal of Econometrics*, 125(1), 15-51.
- Almeida, P., Phene, A., and Li, S. (2014). The Influence of Ethnic Community Knowledge on Indian Inventor Innovativeness. *Organization Science*, 26(1), 198-217.
- Altbach, P. G. (2004). Globalisation and the university: Myths and realities in an unequal world. *Tertiary Education and Management*, 10(1), 3-25.
- Altbach, P. G. (2013). Advancing the national and global knowledge economy: the role of research universities in developing countries. *Studies in higher education*, 38(3), 316-330.
- Altbach, P. (2015). The costs and benefits of world-class universities. *International Higher Education*, (33), 5-8.
- ARWU (2011). Academic Ranking of World Universities 2011. Retrieved from: <http://www.shanghairanking.com/de/ARWU2011.html>, 02.06.2017.
- ASSAf (2010). Academy of Science of South Africa. The PhD Study: An Evidence-based Study on how to meet the Demands for High-level Skills in an Emerging Economy. Consensus Report. Retrieved from: <https://www.assaf.org.za/files/2010/11/40696-Bolddesign-PHD-small.pdf>, 27.05.2017.
- Athreye, S., and Cantwell, J. (2007). Creating competition: Globalisation and the emergence of new technology producers. *Research Policy*, 36(2), 209-226.
- Barnard, H., R. Cowan and Müller, M. (2012). Global excellence at the expense of local diffusion, or a bridge between two worlds? *Research in science and technology in the developing world. Research Policy* 41(4): 756-769.
- Barnard, H., R. Cowan and Müller, M. (2016). On the value of foreign PhDs in the developing world: Training versus selection effects. UNU-MERIT Working paper 2016-06.
- Bayer, A. E., and Dutton J. E. (1977). Career age and research-professional activities of academic scientists: Tests of alternative nonlinear models and some implications for higher education faculty policies. *The Journal of Higher Education*, 48(3): 259-282.
- Bell, M., and Pavitt, K. (1997). Technological accumulation and industrial growth: contrasts between developed and developing countries, In: Archibugi, D., and Michie, J. (Eds), *Technology, globalisation and economic performance*. Cambridge University Press, Cambridge, pp. 83-137.
- Benneworth, P. and Charles, D. (2005). University spin-off policies and economic development in Less successful regions: Learning from two decades of policy practice. *European Planning Studies*, 13(4), 537-557.

- Bernardes, A. T., and Albuquerque, E. D. M. (2003). Cross-over, thresholds, and interactions between science and technology: lessons for less-developed countries. *Research Policy*, 32(5), 865-885.
- Bertrand-Cloodt, D., Crvers, F., and Heijke, H. (2016). Ability, academic climate, and going abroad for work or pursuing a PhD. *CESifo Economic Studies*, 63(1), 119-140.
- Brundenius, C., Lundvall, B.-Å., and Sutz, J. (2011). The role of universities in innovation systems in developing countries: developmental university systems — empirical, analytical and normative perspectives, in: Lundvall, B.-Å., Joseph, K. J., Chaminade, C., and Vang, J. (Eds), *Handbook of Innovation Systems and Developing Countries: Building Domestic Capabilities in a Global Setting*. Edward Elgar, Cheltenham, pp. 311-336.
- Bunting, I. (2006). The higher education landscape under apartheid, in: Cloete, N., Maassen, P., Fehnel, R., Moja, T., Gibbon, T., and Perold, H. (Eds), *Transformation in Higher Education*, Springer Netherlands, pp. 35-52.
- Chen, K., and Kenney, M. (2007). Universities/research institutes and regional innovation systems: the cases of Beijing and Shenzhen. *World development*, 35(6), 1056-1074.
- Cloete, N. (2010). Institutional Clusters in Higher Education in South Africa. Presentation at the Higher Education Summit of the Minister of Higher Education and Training, Cape Peninsula University of Technology, Cape Town, April 2223.
- Conceição, P., and Heitor, M. V. (1999). On the role of the university in the knowledge economy. *Science and Public Policy*, 26(1), 37-51.
- De Fuentes, C., and Dutrenit, G. (2012). Best channels of academia-industry interaction for long-term benefit. *Research Policy*, 41(9), 1666-1682.
- Deiaco, E., Hughes, A., and McKelvey, M. (2012). Universities as strategic actors in the knowledge economy. *Cambridge Journal of Economics*, 36(3), 525-541.
- Delicado, A. (2010). Going abroad to do science: mobility trends and motivations of Portuguese researchers. *Science Studies*, 23, 36-59.
- Department of Education (1998) Formal Degree/Diploma/Certificate Programme Classification Structure Manual. Manual SAPSE-004 (95/10), 2nd edition, Department of Education, Pretoria, Republic of South Africa.
- Du Toit, A. (2012) The PhD and the degree structure of South African higher education: A brief and rough guide. Memorandum presented at the Center for Higher Education Transformation seminar “Knowledge Production in South African Higher Education” on 23 February 2012, accessed 29 April at <http://www.chet.org.za/files/PhD%20Andre%20du%20Toit.pdf>.
- Gaule, P., and Piacentini, M. (2013). Chinese graduate students and US scientific productivity. *Review of Economics and Statistics*, 95(2), 698-701.
- Gibson, J., and McKenzie, D. (2014). Scientific mobility and knowledge networks in high emigration countries: Evidence from the Pacific. *Research Policy*, 43(9), 1486-1495.
- Giuliani, E., and Arza, V. (2009). What drives the formation of ‘valuable’ university-industry linkages: Insights from the wine industry. *Research Policy*, 38(6), 906-921.
- Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457-472.
- Griliches, Z. (1977). Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica*, 45(1), 1-22.

- Grogger, J., and Hanson, G. (2013). The scale and selectivity of foreign-born PhD recipients in the US. *The American Economic Review*, 103(3), 189-192.
- Grogger, J., and Hanson, G. H. (2015). Attracting talent: Location choices of foreign-born PhDs in the United States. *Journal of Labor Economics*, 33(S1)
- Güngör, N. D., and Tansel, A. (2008). Brain drain from Turkey: an investigation of students' return intentions. *Applied Economics*, 40(23), 3069-3087.
- Hansen, K. T., Heckman, J. J., and Mullen, K. J. (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics*, 121(1), 39-98.
- Heckman, J. J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. In: Berg, S. V. (Ed), *Annals of Economic and Social Measurement*, 5(4). NBER Books, Massachusetts, pp. 475-492.
- Heckman, J. J., Lopes, H. F. and Piatek, R. (2014). Treatment Effects: A Bayesian Perspective. *Econometric Reviews*, 33(1-4), 36-67.
- Heitor, M., Horta, H., and Mendonça, J. (2014). Developing human capital and research capacity: Science policies promoting brain gain. *Technological Forecasting and Social Change*, 82, 6-22.
- Herman, C. (2011). Expanding doctoral education in South Africa: pipeline or pipe dream? *Higher Education Research and Development*, 30(4), 505-517.
- Horta, H. (2010). The role of the state in the internationalization of universities in catching-up countries: An analysis of the Portuguese higher education system. *Higher Education Policy*, 23(1), 63-81.
- Hyslop, J., Vally, S., and Hassim S. (2006) The South African Boycott Experience. *Academe*, 92(5), 59-70.
- Jonkers, K., and Cruz-Castro, L. (2013). Research upon return: The effect of international mobility on scientific ties, production and impact. *Research Policy*, 42(8), 1366-1377.
- Jonkers, K., and Tijssen, R. (2008). Chinese researchers returning home: Impacts of international mobility on research collaboration and scientific productivity. *Scientometrics*, 77(2), 309-333.
- IIE (1986). Institute of International Education. 1986 Annual Report. New York, USA. Retrieved from: <https://p.widencdn.net/duj5gs/1986-IIE-Annual-Report>, 20.05.2017.
- Jansen, J. (2004). The state of higher education in South Africa: From massification to mergers. In: Daniel, Habib and Southall (eds) *State of the nation: South Africa, 2003-2004*, 290-311. HSRC Press.
- Kim, D., Bankart, C. A., and Isdell, L. (2011). International doctorates: trends analysis on their decision to stay in US. *Higher Education*, 62(2), 141-161.
- Kroll, H., and Liefner, I. (2008). Spin-off enterprises as a means of technology commercialisation in a transforming economy: Evidence from three universities in China. *Technovation*, 28(5), 298-313.
- Kruss, G., and Visser, M. (2017). Putting university-industry interaction into perspective: a differentiated view from inside South African universities. *The Journal of Technology Transfer*, 1-25.
- Lall, S. (2001). *Competitiveness, Technology and Skills*. Edward Elgar Publishing, Williston, VT.

- Laursen, K., Reichstein, T., and Salter, A. (2011). Exploring the effect of geographical proximity and university quality on university-industry collaboration in the United Kingdom. *Regional Studies*, 45(4), 507-523.
- Lee, S. Y., and Song, X. Y. (2012). *Basic and advanced Bayesian structural equation modeling: With applications in the medical and behavioral sciences*. John Wiley and Sons, Ltd, West Sussex, UK.
- Liefner, I., and Schiller, D. (2008). Academic capabilities in developing countries: A conceptual framework with empirical illustrations from Thailand. *Research Policy*, 37(2), 276-293.
- Marginson, S. (2006). Dynamics of national and global competition in higher education. *Higher education*, 52(1), 1-39.
- Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810), 56-63.
- Mowery, D. C., and Rosenberg, N. (1999). *Paths of innovation: Technological change in 20th-century America*. Cambridge University Press.
- Murmann, J. P. (2003). *Knowledge and competitive advantage: The coevolution of firms, technology, and national institutions*. Cambridge University Press, Cambridge.
- Narula, R., and Dunning, J. H. (2000). Industrial development, globalisation and multinational enterprises: new realities for developing countries. *Oxford Development Studies*, 28(2), 141-167.
- Narula, R., and Dunning, J. H. (2010). Multinational enterprises, development and globalisation: some clarifications and a research agenda. *Oxford Development Studies*, 38(3), 263-287.
- Nelson, R. (2004). The challenge of building an effective innovation system for catch-up. *Oxford Development Studies*, 32(3), 365-374.
- Nelson, R. (2005). The roles of research in universities and public labs in economic catch-up, in: Santangelo, G.D. (Ed), *Technological Change and Economic Catch-Up*. Edward Elgar, Cheltenham, UK, pp. 19-32.
- Nelson, R., and Sampat, B. (2001). Making Sense of Institutions as a Factor shaping Economic Performance. *Journal of Economic Behavior and Organization*, 44, 31-54.
- Ozawa, T. (1992). Foreign direct investment and economic development. *Transnational Corporations*, 1(1), 27-54.
- Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., and Krabel, S. (2013). Academic engagement and commercialisation: A review of the literature on university-industry relations. *Research Policy*, 42(2), 423-442.
- Plug, C. (2009). Southern African science in the year 1909-100n. *South African Journal of Science*, 105, 11:12.
- Romijn, H. A., and Caniëls, M. C. (2011). Pathways of technological change in developing countries: review and new agenda. *Development Policy Review*, 29(3), 359-380.
- Rosenberg, N., and Nelson, R. R. (1994). American universities and technical advance in industry. *Research Policy*, 23(3), 323-348.
- Saxenian, A., and Hsu, J. Y. (2001). The Silicon Valley-Hsinchu connection: technical communities and industrial upgrading. *Industrial and Corporate Change*, 10(4), 893-920.
- Shaw, W. H. (1986). Boycotting South Africa. *Journal of Applied Philosophy*, 3, 59-72.
- Song, H. (1997). From brain drain to reverse brain drain: Three decades of Korean experience. *Science Technology and Society*, 2(2), 317-345.

- Soon, J. J. (2012). Home is where the heart is? Factors determining international students' destination country upon completion of studies abroad. *Journal of Ethnic and Migration Studies*, 38(1), 147-162.
- Sooryamoorthy, R. (2010). Science and scientific collaboration in South Africa: apartheid and after. *Scientometrics*, 84, 373-390.
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press, Cambridge, UK.
- Zeschky, M., Widenmayer, B., and Gassmann, O. (2011). Frugal innovation in emerging markets. *Research-Technology Management*, 54(4), 38-45.
- Zhang, H., Patton, D., and Kenney, M. (2013). Building global-class universities: Assessing the impact of the 985 Project. *Research Policy*, 42(3), 765-775.