

Establishing World Class Universities in China: Deploying a Quasi-experimental Design to Evaluate the Net Effects of Project 985

Xiaohua Zong and Wei Zhang¹

Institution of Education, Nanjing University, Nanjing, China;

School of Education, University of Leicester, Leicester, UK

Abstract

The rapid advances of China's universities in major international league tables are generally believed to be an accomplishment of Project 985. A quasi-experimental study was, therefore, undertaken to test the belief and evaluate the effectiveness of this policy intervention, using 15-year panel data between 1998 and 2013. Results from a difference-in-differences model showed that Project 985 had a positive effect on publication outputs of '985' universities. Tier 2 '985' universities gained strong momentum in publication growth in international and ISI journals. Additionally, Theil index decomposition was employed to examine the stratification effect of Project 985. Results confirmed the homogenizing trend within '985' universities. In contrast, the vertical differentiation between '985' and '211' universities was noticeably enlarged. While sustained public funding is still vital, the study findings inform policy makers and HE leaders of greater sectorial and institutional reforms to fulfil individual HEI's needs and remove bottlenecks in publication growth.

Key words: world-class university; Project 985; net effect; stratification; homogenization

¹ Corresponding author. Email: wz24@le.ac.uk

Introduction

Building world-class universities (WCUs) has become a global trend as higher education (HE) plays an important role in promoting national competitiveness in the knowledge based economy (Power, Millington, and Bengtsson 2015). Yet creating universities of international status in the People's Republic of China is not new and can be observed in national reforms and investment initiatives known as Project 211 and Project 985. Project 211 was initiated in 1995 by the Ministry of Education (MOE) aiming at raising the research standards of around 100 key universities for the 21st Century (Ministerial Office 2007). More than 17 billion RMB yuan (equivalent to £ 1.7 billion) has since then been spent on 118 universities but individual institutional impact remains insignificant due to the large number of higher education institutions (HEIs) supported (Cheng and Wang 2012). Project 985, launched in 1999, further shortlisted 39 universities from '211' institutions with a grand expectation from the MOE that limited resources could be concentrated on a batch of flagship universities that were best placed in China to realise world-class ambitions (Chen 2006; Ngok and Guo 2008).

In the intervening years, China has made huge strides towards restructuring HE, encouraging international research collaboration, funding students and academics on overseas placements and exchange, and adopting many other ways of rapidly promoting research excellence in HE (Cheng and Wang 2012; Wang, Cheng, and Liu 2012). The Chinese government also increased their share of Gross Expenditure on Research and Development (R&D) from around 0.64 percent of GDP in 1994 to over 2.08 percent in 2013 (China Statistical Yearbook 2014). Two decades into its effort in development of WCUs, China accounts for 13% of the world's total scientific research output, 25% of the world's R&D workforce (*THE*, October 14, 2013) and hence has emerged as the world's second largest publisher of research (RCUK 2014). The results of the Times Higher Education (THE) World University Rankings 2016/17 reveal that 38² Chinese universities have joined in the world's top 800 with two in the top 50 (*THE*, September 21, 2016). The rapid advances of China's universities in major international league tables are perceived to have been achieved through Project 985.

² The number does not include universities in Hongkong, Macau and Taiwan.

National and provincial governments pump billions into '985' universities mainly based on the deeply embedded assumption that political will and abundant public funding will accelerate the establishment of WCUs. As Project 985 completed its second phase in 2007, the Ministry of Finance has been increasingly concerned about its funding efficiency, particularly since the National Institute of Education Sciences Report (2013) demonstrated that nearly half of the '985' universities had been identified as inefficient between 2007 and 2011 with high level inputs but low research outputs. Recent studies have also pointed out that Project 985 may not have the expected effectiveness due to the lack of competitive mechanism and solidification of '985' identity (China Daily November 13, 2014; the State Council 2015; Liu and Fang 2016). One striking gap throughout the lifetime of Project 985 is a lack of systematic evaluation as to the impact of the intervention on participating universities. This study addressed this gap by deploying a quasi-experimental design to explore the net effects of Project 985 on the publication outputs of 32 member universities between 1998 and 2013, using 27 '211' universities as a comparison group. We also intended to examine whether Project 985 has a more significant effect on the top listed nine universities, which have attracted more funding from the public and private sectors than the rest of the '985' universities. The study addressed the following research questions:

1. To what extent has Project 985 contributed to improved research outputs among participating universities?
2. Has Project 985 led to institutional differentiation among its member universities, and between '985' and '211' universities?

The work described in this paper explicates the study design, including innovative methodology, panel data construction, data analysis, organisation and rationalisation. A second paper will offer more sophisticated engagement with a comprehensive conceptual framework by locating Project 985 in the global, national, political and cultural contexts to explain the results and to build theories.

Project 985

HE in mainland China is a differentiated system of many types of institutions (Figure 1). Public funding from the central government is allocated chiefly on the basis of merit, which

allows some institutions to prosper while others lag (Altbach 2009). Project-based funding initiatives reinforce this tendency to differentiation. The ‘211’ and ‘985’ projects have been the government’s conscious strategies to support research universities and to fund the elite sector of the system. According to figures released by the MOE, ‘211’ and ‘985’ universities received more than 60 percent of their research funds from financial allocations while non-participating HEIs receive less than 40 percent (*Xinhuanet*, November 14, 2014). Project 985 is a massive programme of selective investment (Zhang, Patton, and Kenney 2013). It has passed through three phases (1999-2002, 2004-2007 and 2009-2013). The MOE stipulated four key areas of development on which project funds should be spent (Table 1). While these are broad guiding principles, some critics have identified an absence of formal or strict rules on the flow of the project funds and their use (Cheng 2011; Wang 2011).

[Table 1 near here]

Within the set of ‘985’ universities, the state government has drawn distinctions by identifying nine top universities (Tier 1) during Phase I with disproportionate investment (more than half of the project fund) in an effort to propel them into the world top fifty (Salmi 2009; Wang 2011). 30 more research universities (Tier 2), varying in quality, have managed to join Project 985 by the end of its second phase in 2007, supported by local governments where those universities are located (Cheng 2011). Central government expected these institutions to develop into internationally well-known universities (MOE 2008). Despite relative lower requirements, a review of Tier 2 ‘985’ universities’ vision statements shows that those 30 resource-rich institutions also aspire to attain world-class standards. Understandably, the WCU notion mobilises the entire ‘985’ club for quality enhancement and inadvertently obscures institutional differences.

[Figure 1 near here]

Nonetheless, two problems have been identified in the implementation of this public funding project that lead some observers to argue that Project 985 is a politicised policy of exclusion (*Guangming Daily*, November 24, 2014; *Sciencenet*, August 20, 2015; *Huanqiumil*, February 15, 2016). First, it lacks transparent selection mechanism. Performance and meritocracy did not entirely dominate the selection and expansion procedures. The selection mechanism also took into account sectoral, regional, ethnical, ideological and defense factors (Chen 2006;

Wang 2011). Therefore, there is great variation as to the research quality among '985' institutions. Second, due to budgetary constraints at the end of Phase II, the central government announced that Project 985 would concentrate its limited resources on the existing 39 institutions and no more new HEIs would be accepted (MOE 2011). This suggests that the structure of the HE system and institutional differentiation are further solidified, with the majority of the Chinese HEIs being permanently excluded from the nation's WCU construction list (Wang 2011).

The long term funding associated with Project 985 since 1999 suggests the need for ensuring accountability for public funds through an examination of its effect on participating universities' performance. Formal government reviews took place during the gap years between each consecutive phase. But the results were only released to its participating institutions rather than to the public domain (Deng, Wang, and Liu 2010; Wang 2011). Independent studies of the policy effects in the past decade are simplifications on the ground of international rankings of universities (Deem, Mok and Lucas 2008; Shen 2015). However, the league table evidence is far from sufficient to warrant the claim that the implementation of Project 985 resulted in publication growth at Chinese research universities. The growth of academic outputs is also shaped by multiple and often complex factors, including the increase of operating budget and research funds from other sources, the enrolment expansion of doctoral students, the widening opportunities of international academic exchange and so forth (Cai 2012; Yi 2011).

Surprisingly, only one empirical study (Zhang, Patton, and Kenney 2013) was conducted to measure the effects of Project 985 by controlling for only observable variables. The truth of their claims was doubtful, arguably, since the research design was flawed by the lack of a comparison group to eliminate or control for latent and contextual variables that also influence university research performance. Clearly any claim concerning the positive effects of Project 985 demands an adequate empirical base. This necessitates a more sophisticated study of the effectiveness of Project 985 that draws on robust methodologies and techniques.

Research design

The study design and data analysis is an evolving process. For such longitudinal work, it is recommended that we use panel data to describe the problem and explain the pattern of the research universities nationwide (Hsiao 2003). Comparable studies in China and South Korea show that there are no agreed methods for evaluating the effects of research funding projects. Zhang, Patton, and Kenney (2013) pioneered their Project 985 impact study on the growth rates of publications in international journals at 24 member universities, relying on the approach of linear mixed modelling. Shin (2009) chose an interrupted time-series designed to compare research publications of all participating universities before and after the BK 21 programme, employing hierarchical linear modelling growth analysis. Their methodological implication for our study is a quasi-experimental design with a comparison group, which was impossible in Shin's study (2009) as all research universities in South Korea joined in the programme.

Quasi-experimental design controls various confounding variables and contextual factors (Muijs 2011; Gorard 2013). It is especially suited to measure the impact of policy interventions like Project 985 by comparing the differences in publication growth before and after the policy intervention between '211' universities (comparison group) and '985' universities (treatment group). The '211' universities make a good comparator because they are all research universities similar in origin to the '985' institutions. As illustrated in Figure 2, the net effect of Project 985 is the change in intercept in the intervening years, that is, the growth of academic outputs from '985' universities since the implementation of the policy compared with that from '211' universities. Notably, the effect of the intervention, nevertheless, becomes observable only after around a minimum of two years of elapsed time between research and publication (Shin 2009).

[Figure 2 near here]

Population and data

The population for this study included all universities affiliated to the Chinese MOE. Seven '985' universities under the administration of different ministries and all '211' universities that belong to local governments were not included in this study due to data inconsistencies. Small specialised HEIs with a curriculum emphasis on particular fields such as arts, language and culture, law and finance were also excluded because similar scale data comparable with

the comprehensive universities are not available. After filtration, the final dataset consisted of 59 research universities, among which 32 were ‘985’ universities and 27 ‘211’ universities.

The unique panel data of this study was carefully aggregated from three sources to extract data associated with input (independent) and output (dependent) variables covering the period between 1998 and 2013. A large proportion of the data were relatively easy to access from the public domain. HEI data extracted from the MOE database were strictly confidential and barred to outsiders. Zong’s positionality as a researcher in one ‘985’ university ensured access to the necessary information to conduct the study. One challenge we encountered was that the amount of money allocated by Project 985 to each individual institution was unavailable and difficult to pull out from the annual R&D revenue or expenditure. We therefore used R&D expenditure instead. Provincial per Capita GDPs were assembled from China Statistical Yearbooks (1999–2014) that release data with a time lag of one year. It is also noted that all data on economy and funds were converted into 2010 constant Chinese Yuan (CNY) (currency). Table 2 presents an overview of our data set and the three databases from which we extracted the variables.

[Table 2 near here]

The key advantage of this data set, compared with previous studies on Project 985 impact, lies in the expansion of both input and output variables from 59 universities to enhance validity and reliability. More specifically, we included more indicators to control research inputs such as the number of researchers, doctoral students, international academic exchanges; provincial per Capita GDP; and R&D expenditures. Meanwhile, we secured a full set of scholarly publications in domestic, international and ISI journals that cover research outputs not only in sciences-related disciplines but also in social sciences and humanities. Compiling this dataset was an intensive process to ensure data were coherent and consistent all the way through. For example, many universities experienced institutional merger and changed names during the period covered by our study. The study entailed a manual scrutiny of publication data to ensure that they mapped to institutions before and after the amalgamation, and data of those related institutions were integrated into one unit. Likewise, data on branch campuses were aggregated to form data of one observation.

Model specification

Following the seminal work by Ashenfelter and Card (1985), Difference-in-differences model (DID) has been widely used in policy research impact evaluations that employ a quasi-experimental design. This analytical approach is an optimum choice as it allows us to control for variables we cannot observe or measure like cultural factors or heterogeneity of development practices across HEIs (White and Sabarwal 2014). We estimated the DID in terms of changes in publication growth over time between ‘985’ and ‘211’ universities to investigate the effectiveness of Project 985 for supporting the development of publication outputs of participating institutions. Our assumptions were that

1. research publication growth in ‘985’ universities would be greater than that of ‘211’ universities;
2. research publication growth in Tier 1 ‘985’ universities would be greater than that in Tier 2 universities.

The standard set up of a DID model (Wooldridge 2013) can be illustrated as:

$$Y = \alpha_0 + \alpha_1 D + \delta T + \gamma(D * T) + \varepsilon \quad (1)$$

where Y is the outcome of interest, and D is a dummy variable for the treatment group, coded 1 if the unit is in treatment group, 0 if in comparison group. T is a dummy variable for the time period after an intervention, coded 1 if the observation is in the time period after the intervention and 0 prior to the intervention. Hence, $(D * T)$ is a dummy variable indicating when $D=T=1$, and its coefficient, γ , is the DID estimator, which captures the treatment effects of policy intervention.

The standard DID model can be transformed into a more advanced version with panel data. As illustrated in model (2), instead of including an independent dummy variable (D) in the standard DID model, a whole set of dummy variables (u_i) are introduced for all the observations in the panel data DID model, intending to capture the unobservable effects inherent with each case. This specification makes the model more accurate and interpretive when applied in the HE sector, because HEIs, often with a long and unique history, differ from each other in nearly every aspect, far beyond a dichotomy between those selected by the project and those not. A matrix of control variables, Z_{it} , which are assumed to affect the outcome can also be included in the model. The interpretations of coefficients in a DID

model with panel data resemble that in the standard model, with the DID estimator, γ , pinpointing the treatment effects of policy intervention.

$$Y_{it} = \alpha_0 + \beta Z_{it} + \delta T + \gamma(D*T) + u_i + \varepsilon_{it} \quad (2)$$

Grounded on model (2), we set up a panel data version of the DID model

$$\begin{aligned} \log(Y_{it}) = & \alpha_0 + \beta_1 Rescher_{it} + \beta_2 Doct_{it} + \beta_3 R \& D_{it} + \beta_4 Intexch_{it} + \beta_5 pergdp_{it} \\ & + \delta T + \gamma Project_{it} + u_i + \varepsilon_{it} \end{aligned} \quad (3)$$

where Y_{it} is the academic outputs, the subscript label i denotes an individual university and the t subscript indicates a year. Similar to model (2), T denotes time dummy for each year when there was at least one university entering Project 985. Note that T is not a singular variable in model (3) but an array of time dummies because of the different time these universities included in the project. $Project_{it}$ is the Project 985 dummy variable, assuming a value of 1 for all years after which the ‘985’ contract was signed by university i with the MOE, and 0 otherwise. If it is non-‘985’ universities, the dummy assumes 0 for all the years under study. The project dummy is equivalent to $(D*T)$ in model (2), and hence, its coefficient, γ is the DID estimator. u_i is a whole set of dummy variables for each university, indicating the time-fixed unobservable effects. ε_{it} is an idiosyncratic error term.

Five control variables are also added in the model: four input factors and one environmental variable - per capita GDP of provinces where universities are located. The reason to control for input variables is outlined earlier. Provincial per capita GDP is also incorporated as a control variable because earlier research finds that this variable has a significant influence upon research productivity of universities (Johnes and Yu 2008). Since the selection of ‘985’ universities was not a random process but took into account a variety of factors inherent within the universities, such as history, region, sector, etc., the condition of $E(u_i + \varepsilon_{it} | 985 Project_{it}) = 0$ cannot be satisfied. Therefore, this panel data DID model should be estimated by fixed effects instead of random effects to obtain consistent estimators.

As embedded in research question 2, it was also our intention to estimate the DID between Tier 1 and Tier 2 ‘985’ universities by comparing their publication growths over time with ‘211’ universities. Two interaction terms between the project dummy and tier dummies are

added to substitute for the single project dummy, which can distinguish the effects of the project on different tiers of ‘985’ universities.

$$\log(Y_{it}) = \alpha_0 + \beta_1 Rescher_{it} + \beta_2 Doct_{it} + \beta_3 R \& D_{it} + \beta_4 Intexch_{it} + \beta_5 Pergdp_{it} + \delta T + \gamma_1 (Tier1 * Project_{it}) + \gamma_2 (Tier2 * Project_{it}) + u_i + \varepsilon_{it} \quad (4)$$

The variables are specified in detail in Table 3. Considering that outputs of R&D activities usually have at least a delay of one year (Zhang et al. 2011; Zhang, Patton, and Kenney 2013), all five control variables in the models are regressed by a lag of one year. Concerning the dummy for Project 985, following the specification in Shin’s (2009) model, a 2-year lag of time is assumed in this study. It seems justifiable to set a 2-year lag for the ‘985’ dummy if the institutional context is taken into account. In China, government budgets are routinely drafted in the previous year and then audited, and approved by National People's Congress at the very beginning of the next year. Accordingly, the special funds prescribed by Project 985 would not be available after the ‘985’ contracts were signed until next year due to this statutory budgeting process, and the effects of the project on academic outputs could be perceptible even later.

[Table 3 near here]

Results of the DID models and discussion

The effect of Project 985 on ‘985’ universities

The descriptive statistics are presented in Table 4. It is clear that there were substantial gaps between ‘985’ and ‘211’ universities both in research inputs and outputs. The gaps were distinctively larger regarding publications in international and ISI journals than those in domestic journals. Figure 3 further examined the trends of publication growth separately for ‘985’ and ‘211’ universities between 1998 and 2013. A sharper publication increase in ISI journals was also observed in ‘985’ rather than ‘211’ universities, but there was a flat rise with diminishing momentum in domestic journals across both groups of universities. This is a reflection of both groups of HEIs’ different policy requirements for academic publications. ‘985’ universities, with an aspiration for international prominence, set out guidelines about

international journals, particularly ISI ones, on which most of the publications should be targeted (Vidovich, Yang and Currie 2007; Zhang 2013).

[Table 4 near here]

The results of estimations by fixed effects on six models are reported in Table 5 and Table 6 respectively. Based on model (3), the first three models (3-1 – 3-3) estimate the average effects of Project 985 on academic outputs of participating universities as a whole. The next three models (4-1 – 4-3), hinged on model (4), assume different effects of Project 985 on Tier 1 and Tier 2 universities. It is worth noting that, DID effects are estimated with a comparison group, which in this study means that the estimated policy effects are the results from comparison with the performance gain of ‘211’ universities.

[Table 5 and Table 6 near here]

Results of the first three models revealed that Project 985 had a positive and significant effect on academic outputs of participating universities across all three publication types, after controlling for various research inputs and contextual factors. This finding lends weight to previous accounts of targeted support from Project 985 effectively resulting in an increased number of publications (e.g. Cheng and Wang 2012; Zhang, Patton, and Kenney 2013). This confirmative result produces nothing new or exciting but is derived from a more reliable evidence base than those deployed by earlier studies.

By comparing these three models, it can be found that the magnitude of coefficients for the project dummy in model (3-2) and model (3-3) was much larger than that in model (3-1), indicating that the output increase in international journals induced by the project was higher than in domestic journals, though the positive effects on domestic publications were not negligible. The gradients in the magnitude of the three coefficients for the project dummy confirm that the participating universities prioritise research publications in international peer reviewed journals, seeking more desired impacts and a rise in global league table rankings.

The effect of Project 985 on both Tiers within ‘985’ institutions

When the net effects of the intervention were analysed separately upon Tier 1 and Tier 2 '985' universities, similar and different patterns of publication growth emerged. The DID estimators in the first column in Table 6 were still positive and significant at 0.01 level, consistent with the estimation in model (3-1), suggesting that both tiers have experienced significant growth in domestic publications as a result of Project 985. However, the DID estimators in the second and third columns of Table 6 were noticeably different between the two tiers, with the policy effect on Tier 1 '985' universities being not significant. The results of model (4-1) to (4-3) indicate that while the publications from Tier 2 '985' universities increased significantly after the implementation of Project 985, Tier 1 '985' universities did not experience significant growth in academic outputs except in domestic publications. Notably, this distinction was more evident in model (4-3), where the estimated policy effect on ISI publications from Tier 2 universities was not only significant but its size was also much larger than that from Tier 1 universities. This is despite the fact that Tier 1 '985' universities received preferential treatment with tremendous resources and favourable governance structure (Salmi 2009; Wang 2011). The divergence in the DID estimates between both tiers of '985' universities nevertheless implies a catching-up process. In other words, those Tier 2 universities have gained stronger momentum than Tier 1 in international publications after participating in the project.

In light of this surprising finding, we conducted a nested F-test for model (4-3). The result did not provide strong evidence for the conjecture that the DID estimator for the second tier was significantly larger in magnitude than that of the first tier, although the former was positive at 0.01 significant level in comparison to '211' universities. Therefore, further statistical analysis should be undertaken to investigate the convergence or divergence trends within the '985' universities.

Findings from the control variables

It is notable that, among the control variables, only per Capita GDP showed a consistent and significantly positive effect on these six models. Researchers, doctoral students and international academic exchanges had positive correlations with all three publication indicators but were statistically significant only with domestic articles. The insignificant effects of those three input variables on international publications were contrary to the findings of Shin (2009); Cheng and Wang (2012); Wang, Cheng and Liu (2012), all of whom

found that various WCU strategies and doctoral students potentially contribute to academic outputs. A possible explanation is that many Chinese researchers with limited linguistic capability, and doctoral students with less academic writing experience in international realm and limited time span until graduation, prefer to publish in domestic journals, especially in the fields of Arts, Humanities and Social Sciences (Flowerdew and Li 2009).

The study also found that R&D expenditure had a negative and significant effect on ISI publications but no significant effect upon the other two categories. This means that there would be no positive marginal gain in academic outputs if future WCU policy intervention were to increase investment on '985' universities. It is an important finding because it raises the question of the accountability of the public funds and draws our attention to the arguments of Cheng (2011) and Wang (2011), both of whom suggest that much emphasis should be given to improvement in research funding allocation, flow and utilization efficiency.

The vertical stratification of research universities in China

This section further investigates whether the research universities were more unequal regarding international publications after Project 985. In an earlier study on differentiation between the world's top 500 universities, Halfman and Leydesdorff (2010) reported that the overall inequality among universities regarding ISI publications decreased internationally. This convergence trend was also observed in their sample of eight Chinese HEIs, a rather small fraction of China's research universities. With a substantially improved dataset in terms of coverage and time period, we took another close look at differentiation between institutions to test Halfman and Leydesdorff's conclusion and our DID findings.

Theil Index and decomposition

The Theil index was used at the final stage of the study to examine the stratification effect of Project 985. This method provides a useful statistical tool in the development of the analysis of the inequality measure, which is one of a class of decomposable generalized entropies developed by Shorrocks (1984). By applying the Theil index, we were able to discern the within-group inequality and the between-group portion of total inequality since this technique is additively decomposable by subgroups within the population (Shorrocks 1980). The Theil

index equation was as follows:

$$T_1 = \sum_{i=1}^n f(p_i) \frac{p_i}{\mu} \ln \frac{p_i}{\mu} \quad (5)$$

where p_i is the academic output measured by ISI articles from i^{th} university; μ is the total population mean of academic outputs; $f(p_i)$ is the population share of p_i in the total population and n is the total population.

The evolution of vertical stratification of China's research universities

Based on the formula above, we decomposed the inequality in academic publication in ISI journals by the stratification between '985' and '211' universities, and between Tier 1 and Tier 2 '985' universities. The results are presented in Table 7. The decomposition results with regard to the differentiation between '985' and '211' universities showed that the between-group component accounted for 43.30 per cent of the total inequality in 1998, a year before the introduction of Project 985. It then steadily rose and peaked at 53.28% in 2012. This indicates that disparities between '985' and '211' universities were enlarged after the implementation of Project 985. This finding confirmed the implication from the earlier DID estimation that Project 985 would widen the gap between '985' and '211' universities and echoed the serious concern with equity in the HE sector (Wang 2011; *China Daily*, November 13, 2014).

[Table 7 near here]

The Theil decomposition between the two tiers within '985' universities showed that the share of between-group inequality increased from 50 % to 63 % of the total inequality as measured by the Theil index, which suggests that the disparity between the two tiers of '985' universities was remarkably large before 2005. However, the discrepancy between both tiers of '985' universities, after peaking at 62.75 % in 2005, started to decline continuously to a new lowest point of 52.52% in 2012. It is important to note that the decline began in 2006 when all Tier 2 universities were accepted in Project 985 with the majority of them having been in the project for at least two years. This result provides additional support to the speculation in the preceding section that Tier 2 '985' universities, though still lagging behind

their Tier 1 counterparts, had made significant gains in international publications, which ushered in a catching-up process between the two groups.

This is not an indication of Project 985 failing to achieve its purposes. The disparity can be explained through reference to the historical contexts of the nine top listed universities. As China's leading research institutions, with a strong history of scholarship, they have placed great emphasis on research performance long before the implementation of Project 985 since China's opening up in the 1980s (Yang 2009; Li and Fang 2012). To this end, the implementation of Project 985 in this elite group appears to merely serve as a natural extension of their aspirations for achieving international prominence rather than an intervention, which, however, has given Tier 2 '985' universities a whole new look in research productivity. Moreover, Tier 2 '985' universities, enjoying the advantages of later-comers, have emulated the leading universities and hence progressed at a quicker speed (Chandra, Lin and Wang 2013). In sum, the estimation of policy effects and the implications with regard to the evolution of stratification in the system are reconciled by the decomposition analysis which demonstrates a trend of divergence between '985' and '211' institutions after the introduction of Project 985 and a convergence process between the two tiers of '985' universities. These two trends are clearly illustrated in Figure 4.

[Figure 4 near here]

Conclusion

Project 985, as an initiative for creating WCUs, plays a developmental as well as a political role in HE in China. Preferential allocation of resources to '985' universities has long been considered by the central government to be critical to promoting research capacity and the quality of elite HEIs. The study results extend the public recognition of the significantly positive effects of Project 985: '985' universities as a whole have experienced a significant increase in academic publications. Thus, this finding adds to a more nuanced and sophisticated understanding of the Project 985 effect. It differs from previous studies in terms of the extended coverage of dataset and the quasi-experimental design. Most notably, this large scale evaluation involves sophisticated statistical models and rigorous analysis.

Project 985, together with restructuring in higher education, has contributed to China's vision of building WCUs, but the improvements hide significant disparities: our results also suggest the vertical differentiation between '985' and '211' universities has been noticeably enlarged. While '985' universities on average have made significant achievements, '211' universities and many more that exist outside of 'excellence' schemes would be demotivated to compete for quality improvement, which will potentially yield an unbalanced HE system (Altbach 2016). Therefore, the overall quality improvement of the HE sector has been outweighed by the political imperatives of developing WCUs.

The impact of Project 985 on participating institutions' research performance is more complex than policy-makers and the public seem to believe. Our study also identified 'solid micro evidence' (Bolton 2014) that Project 985 has not had the desired effect on Tier 1 universities. A long-term tendency of publication saturation was observed in those nine Tier 1 '985' universities, despite generous investment from the government. This may have been the result of premature academic 'burn out'. Beneficial effects have largely been restricted to Tier 2 '985' institutions. We argue that the initial differences between both tiers of Project 985 would diminish and homogenizing tendencies would gather force over time.

The development pathways of Tier 1 institutions enable good predictions of the trajectories Tier 2 universities will follow later. Capacity constraints at Tier 1 '985' universities make us wonder whether Tier 2 '985' universities will withstand the pressure to continuously improve their research performance. This, together with the negative effect of R&D expenditure on ISI publications, makes it imperative to provide a mature recognition that a continued relentless focus on funding is necessary but not sufficient. The study findings also inform policy makers and HE leaders that greater sectoral and institutional reforms will inevitably fulfil individual HEI needs and remove bottlenecks in publication growth. The emphasis, therefore, should be placed on fluidity in the membership of WCUs, transparent selection and evaluation process (Wang 2004; Zhou, Thijs and Glanzel 2009; and *The Economic Observer*, December 5, 2014). This emphasis also encourages an efficient flow and use of public funds, an optimal research culture, academic freedom and a conducive external environment (Altbach 2009; Cheng and Wang 2012; Yang 2016).

References

- Altbach, P.G. 2009. "One-third of the globe: The future of higher education in China and India." *Prospects* 39: 11-31.
- Altbach, P.G. 2016. "Chinese Higher Education: 'Glass Ceiling' and 'Feet of Clay'" *International Higher Education* 86:11-13.
- Ashenfelter, O., and D. Card. 1985. "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs." *The Review of Economics and Statistics* 67(4): 648-660.
- Bolton, L. 2014. *Helpdesk Report: Education Expenditure Reviews*. Health and Education Advice and Resource Team, Oxford, UK.
- Cai, H. 2012. "Deploying the Chinese knowledge diaspora: a case study of Peking University." *Asia Pacific Journal of Education* 32(2): 367-379.
- Chandra, V., J.Y. Lin, and Y. Wang. 2013. "Leading Dragon Phenomenon: New Opportunities for Catch-up in Low-Income Countries." *Asian Development Review* 30(1): 52-84.
- Chen, X.F. 2006. "Ideal Orientation Policy Making: Analysis on 985 Project Policy Process." [In Chinese.] *Peking University Education Review* 4(1): 145-157.
- Cheng, Y. 2011. "A Reflection on the Effects of the 985 Project." *Chinese Education and Society* 44(5): 19-30.
- Cheng, Y. and Q. Wang. 2012. "Building World-Class Universities in Mainland China." *Journal of International Higher Education* 5(2): 67-69.
- China State Council. 2015. "Coordinate Development of World-class Universities and First-class Disciplines Construction Overall Plan." http://www.gov.cn/zhengce/content/2015-11/05/content_10269.htm.
- Deem, R., K.H. Mok, and L. Lucas. 2008. "Transforming Higher Education in Whose Image: Exploring the Concept of the 'World-Class' University in Europe and Asia." *Higher Education Policy* 21: 83-97.
- Deng, Q.Q., Q. Wang, and N.C. Liu. 2010. "National Initiatives for Building World-Class Universities: Comparison between Asian and European Experience." *Higher Education Forum* 7: 35-51.
- Deng, H. 2014. "Post '985' '211' Period: How to Break the Boundaries that Fragment China's Higher Education System." *Guangming Daily*, November 24.

- Flowerdew, J., and Y. Li. 2009. "English or Chinese? The trade-off between local and international publication among Chinese academics in the humanities and social sciences." *Journal of Second Language Writing* 18(1): 1-16.
- Gorard, S. 2013. *Research Design: Robust approaches for the social sciences*. London: SAGE.
- Halfman, W., and L. Leydesdorff. 2010. "Is Inequality Among Universities Increasing? Gini Coefficients and the Elusive Rise of Elite Universities." *Minerva* 48(1): 55-72.
- Hsiao, C. 2003. *Analysis of panel data*. 2nd ed. Cambridge: Cambridge University Press.
- Li, J., J. Lin, and G. Fang. 2012. "Nanjing University: Redeeming the Past by Academic Merit", in *Portraits of 21st Century Chinese Universities: In the Move to Mass Higher Education*, edited by Hayhoe, R., J. Li, J. Lin and Q. Zha, 131-161. Netherlands: Springer.
- Liu, J. and X. Fang 2016. "What are the Dispute of the Existence and Abolition of '211' and '985'." *Contemporary Educational Management* 6: 20-25.
- Ministry of Education, China. 2011. *985 Project Construction Report 1999-2008*. [In Chinese] Beijing: Higher Education Press.
- Ministerial Office of 211 Project, China. 2007. *Report on 211 Project 1995-2005*. [In Chinese] Beijing: Higher Education Press.
- Muijs, D. 2011. *Doing quantitative research in education with SPSS*. 2nd ed. London: Sage.
- National Bureau of Statistics of China. 2015. *China Statistical Yearbook. 1999-2014*.
<http://www.stats.gov.cn/english/statisticaldata/AnnualData/>
- National Institute of Education Sciences 2013. "A Report on Performance Evaluation Research of Colleges and Universities in 2012." *Educational Research* 10: 55-64.
- Ngok, K.L. and W. Q. Guo. 2008. "The Quest for World Class Universities in China: critical reflection." *Policy Futures in Education* 6(5): 545-557.
- Power, L., Millington, K.A. and Bengtsson, S. 2015. "*Building Capacity in Higher Education Top Guide*." The Health & Education Advice & Resource Team
- Salmi, J. 2009. *The challenge of establishing world-class universities*. Washington, DC: The World Bank.
- Chen, B. 2015. "The Enlarged Differences in terms of Income and Expenditure in Chinese Higher Education Institutions." *Sciencenet*, August 20.
- Shen, G. 2015. "Building World-Class Universities in China: From the View of National Strategies." *Global University Network for Innovation*,

- <http://www.guninetwork.org/articles/building-world-class-universities-china-view-national-strategies>
- Shin, J. C. 2009. "Building world-class research university: The Brain Korea 21 project", *High Education* 58: 669-688.
- Shorrocks, A. F. 1980. "The Class of Additively Decomposable Inequality Measures." *Econometrica*, 48(3): 613-625.
- Shorrocks, A. F. 1984. "Inequality Decomposition by Population Subgroups." *Econometrica*, 52(6): 1369-1385.
- Siyue. 2016. "'985' Universities Research Grants Rankings." *Huanqiumil*, February 15.
http://www.huanqiumil.com/news/54374_3.html
- Vidovich, L., R. Yang, and J. Currie. 2007. "Changing accountabilities in higher education as China 'opens up' to globalisation." *Globalisation, Societies and Education* 5(1): 89-107.
- Wang, Z. 2004. "Two Points of Criticism on the 985 Project." [In Chinese.] *Impact of Science on Society* (3), 63-64.
- Wang, Q. 2011. "A Discussion on the 985 Project from a Comparative Perspective", *Chinese Education & Society* 44 (5): 41-56.
- Wang, Q., Y. Cheng, and N.C. Liu, eds. 2012. *Building World-Class Universities: Different Approaches to a Shared Goal*, Rotterdam/Boston/Taipei: Sense Publishers.
- White, H., and S. Sabarwal. 2014. "Quasi-experimental Design and Methods", *Methodological Briefs: Impact Evaluation 8* UNICEF Office of Research, Florence.
- Wooldridge J. M. 2013. *Introductory Econometrics: A Modern Approach* 5th ed. South-Western: Cengage Learning.
- Yang, R. 2009. "Enter the Dragon China's Higher Education Returns to the World Community: The Case of the Peking University Personnel Reforms." In *Higher Education: Handbook of Theory of Research*, edited by Smart, J. C., 427-461. Netherlands: Springer.
- Yang, R. 2016. "Toxic Academic Culture in East Asia", *International Higher Education* 84:15-16.
- Yi, L. 2011. "Auditing Chinese higher education? The perspectives of returnee scholars in an elite university." *International Journal of Educational Development* 31(5): 505-514.
- Zhang, D., R. D. Banker, X. Li, and W. Liu. 2011. "Performance impact of research policy at the Chinese Academy of Sciences." *Research Policy* 40(6): 875-885.

- Zhang, J.N. 2013. "Promotion criteria, faculty experiences and perceptions: A qualitative study at a key university in China." *International Journal of Educational Development* 33(2): 185-195.
- Zhang, H., D. Patton, and M. Kenney. 2013. "Building global-class universities: Assessing the impact of the 985 Project." *Research Policy* 42: 765-775
- Zhou, P., B. Thijs, and W. Glanzel. 2009. "Is China also becoming a giant in social science?" *Scientometrics* 79(3): 593-621.

List of Tables

Table 1 Project 985 funding focus across three phases

Phase	Project funding focus
Phase I (1999-2001)	<ul style="list-style-type: none">• infrastructure and other hardware construction on campuses
Phase II (2004-2007)	<ul style="list-style-type: none">• consolidating the achievement of the first place round• building platforms for scientific and technological innovation and bases for the philosophy and social science innovation
Phase III (2009-2013)	<ul style="list-style-type: none">• establishment of talent teams and improvement of the capacity of indigenous innovation

Source: MOE (2013)

Table 2 The study dataset and data sources

Variable	Source
<i>Inputs</i> <ul style="list-style-type: none">• Researchers participating in R&D activities• Doctoral students• R&D expenditure• International academic exchanges	Ministry of Education Department of National Universities data set (1998–2013)
<i>Outputs</i> <ul style="list-style-type: none">• Articles published in domestic and international journals	
<ul style="list-style-type: none">• Articles published in SCI/SSCI/A&HCI journals	Thomson Reuters (ISI) Web of Knowledge
<i>Environment</i> <ul style="list-style-type: none">• Provincial per capita GDP	China Statistical Yearbooks (1999–2014)

Table 3 Variable specifications

Variable	Specifications
$\text{Log}(Y_{it})$	The number of academic publications in logarithmic form for university i in year t , including publications in sciences, social sciences, and humanities. There are three alternatives: (1) number of articles published in domestic journals; (2) number of articles published in international journals; (3) number of articles published in ISI journals.
Rescher_{it}	The number of researchers engaging in R&D activities at university i in year t , including researchers in sciences, social science and humanities.
Doct_{it}	The number of doctoral students at university i in year t .
R\&D_{it}	The total expenditures on R&D per year, from all sources , by university i in year t .
Intexch_{it}	The international conference attendances from university i in year t , counted only when papers were accepted.
Pergdp_{it}	Per capita gross provincial product in year t for the province in which university i is located.
T	T denotes time dummy for each year when there was at least one university entering Project 985.
Project_{it}	Project 985 dummy variable
Tier1	Tier1 dummy variable
Tier2	Tier2 dummy variable

Table 4 Descriptive statistics of main variables, 1998-2013

	Total		'985' Universities		'211' Universities	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Articles in Domestic Journals	2782.50	1961.45	3586.49	2093.70	1755.74	1134.57
Articles in International Journals	932.48	1125.60	1340.07	1307.01	411.96	471.17
Articles in ISI Journals	855.97	1087.08	1312.93	1280.34	314.39	316.23
Researchers	1169.40	983.45	1578.44	1108.60	649.37	390.14
Doctoral Students	1939.65	1704.82	2747.61	1833.85	907.83	653.58
R&D Expenditure (Million CNY)	261.07	305.46	347.12	360.39	151.40	159.29
International Academic Exchanges	552.19	678.56	721.48	733.99	337.38	528.67
Per capita GDP (Thousand CNY)	34.10	22.90	33.89	22.66	34.36	23.21

Table 5 Regression results of the average effects of Project 985

	Model (3-1)	Model (3-2)	Model (3-3)
	Domestic Articles	International Articles	ISI Articles
<i>Constant</i>	7.13560*** (0.07801)	4.72614*** (0.15068)	4.35264 (0.17403)
<i>Researchers</i>	0.00004* (0.00002)	0.00001 (0.00005)	0.00002 (0.00003)
<i>Doctoral Students</i>	0.00007* (0.00004)	0.00005 (0.00005)	0.00005 (0.00006)
<i>R&D Expenditure</i>	-0.00017 (0.00013)	-0.00025 (0.00039)	-0.00046** (0.00022)
<i>International Academic Exchanges</i>	0.00006* (0.00003)	0.00004 (0.00006)	0.00003 (0.00006)
<i>Per Capita GDP</i>	0.00859*** (0.00242)	0.03446*** (0.00519)	0.04329*** (0.00492)
<i>985 Project</i>	0.24641*** (0.06815)	0.41130*** (0.14048)	0.36662*** (0.13106)
R-squared (within)	0.3950	0.6612	0.8015
Number of Obs	841	841	860

Note: *, **, *** denote significant at the 0.1, 0.05, 0.01 level, respectively. Robust standard errors are presented in the parentheses.

Table 6 Regression results of the differential effects of Project 985

	Model (4-1)	Model (4-2)	Model (4-3)
	Domestic Articles	International Articles	ISI Articles
<i>Constant</i>	7.13661*** (0.07850)	4.72687*** (0.15174)	4.34937*** (0.17385)
<i>Researchers</i>	0.00004* (0.00002)	0.00001 (0.00005)	0.00002 (0.00003)
<i>Doctoral Students</i>	0.00007* (0.00004)	0.00005 (0.00006)	0.00005 (0.00006)
<i>R&D Expenditure</i>	-0.00017 (0.00013)	-0.00025 (0.00029)	-0.00046** (0.00021)
<i>International Academic Exchanges</i>	0.00006* (0.00003)	0.00004 (0.00006)	0.00003 (0.00006)
<i>Per Capita GDP</i>	0.00854*** (0.00242)	0.03442*** (0.00520)	0.04345*** (0.00498)
<i>985 Project on Tier1 (Tier₁*Project_{it})</i>	0.29848*** (0.10466)	0.44865 (0.31201)	0.21229 (0.19825)
<i>985 Project on Tier2 (Tier₂*Project_{it})</i>	0.24321*** (0.06776)	0.40900*** (0.14320)	0.37508*** (0.13809)
R-squared (within)	0.3951	0.6612	0.8017
Number of Obs	841	841	860

Note: *, **, *** denote significant at the 0.1, 0.05, 0.01 level, respectively. Robust standard errors are presented in parentheses.

Table 7 Theil Index decomposition

Year	Between '985' and '211' universities				Between two tiers of '985' universities			
	Total	Within-group	Between-group	Share of Between-groups (%)	Total	Within-group	Between-group	Share of Between-groups (%)
1998	0.66	0.38	0.29	43.30	0.37	0.18	0.19	50.56
1999	0.65	0.37	0.28	43.10	0.37	0.18	0.19	51.50
2000	0.63	0.35	0.27	43.82	0.35	0.16	0.20	55.48
2001	0.62	0.35	0.27	44.31	0.34	0.14	0.20	58.77
2002	0.61	0.35	0.26	43.18	0.35	0.15	0.20	57.55
2003	0.58	0.32	0.26	44.40	0.32	0.13	0.19	59.49
2004	0.54	0.28	0.26	47.75	0.29	0.11	0.18	61.01
2005	0.49	0.27	0.22	45.21	0.28	0.10	0.17	62.75
2006	0.43	0.22	0.20	47.48	0.23	0.09	0.14	60.68
2007	0.39	0.20	0.19	48.97	0.20	0.09	0.11	56.12
2008	0.37	0.19	0.18	49.17	0.19	0.09	0.11	55.88
2009	0.35	0.17	0.18	52.12	0.17	0.07	0.10	57.37
2010	0.32	0.15	0.17	52.34	0.16	0.07	0.09	56.69
2011	0.31	0.14	0.16	52.80	0.16	0.07	0.09	55.56
2012	0.30	0.14	0.16	53.28	0.15	0.07	0.08	52.52
2013	0.30	0.14	0.16	52.86	0.15	0.07	0.08	54.49

List of figures

Figure 1 The structure of HE system in China in terms of administration and status

Figure 2 Policy effects in DID design

Figure 3 Publications from '985' universities (A) and from '211' universities (B), 1998-2013

Figure 4 The shares of between-group inequality in total inequality