Globally Bred Chinese Talents Returning Home: An Analysis of a Reverse Brain-Drain Flagship Policy

Giulio Marini and Lili Yang

1Quantitative Social Science, Social Research Institute (IoE) UCL, 27–28 Woburn Square, London WC1H 0AA, UK and 2Centre for Global Higher Education, Department of Education, University of Oxford, 15 Norham Gardens, Oxford OX2 6PY, UK

*Corresponding author. E-mail: lili.yang@education.ox.ac.uk

Abstract

China has launched a series of talent-recruitment policies in the last years, in order to attract back Chinese nationals who stayed abroad. Yet, little is known about the effect of such policies. This paper examines whether researchers recruited in one of the Chinese flagship talent-recruitment policies—the ‘Young Thousand Talents’ policy (Y1000T)—had, in the following years after recruitment, better research performance. We compare these recipients against other Chinese nationals who got PhDs in equally prestigious non-Chinese universities but continued to work abroad (mostly in the USA). Results of difference-in-differences regressions show that returning to China has an effect of positioning returnees both at the bottom and at the very summit of the distribution of quality of publications. Nevertheless, some Y1000T researchers seem to have prioritized the quantity of outputs, arguably to the detriment of quality. This is probably due to certain research evaluation criteria in place until recent times.

Key words: policy effect; talent mobility; China; the USA; early- and mid-career researchers; research performance

1. Introduction

China has become a competitive player in the world system of science (Altbach and Salmi 2011; Hayhoe 2011; Van Der Wende 2015; Zhou and Leydesdorff 2006). Among China’s important strategies to strengthen its research capacity is to recruit top-notch Chinese talents who have been educated and/or worked in world-leading higher education institutions (globally bred talents) (Breschi et al. 2020; Li et al. 2018b; Miao and Wang 2017). It is argued that by recruiting the globally bred talents, China can improve its capacity to produce high-quality research and educate their own new generations of talents domestically (Li et al. 2018a,b). However, it remains unclear whether or not these returned researchers can sustain their high-quality work after coming back. In other words, there is a surprising dearth of evidence-based knowledge about whether or not such talent-recruitment policy may de facto promote China’s capacity to catch up with top Global West standards. To evaluate the research performance of recruited globally bred Chinese talents is an important way to unpack the effect of such talent recruitment policies. Moreover, investigating China’s research performance, especially that of early and mid-career researchers, is enlightening to the future development of science in China.

This paper aims to analyze the research performance of Chinese researchers who were recruited back to China under the specific program of the ‘Young Thousand Talents’ (Y1000T). The program aims to attract top Chinese early- and mid-career researchers who have competitive publication records, and have studied or worked for at least 3 years in leading global research institutions outside mainland China. Specifically, the paper compares the returnees’ research performance with other globally bred Chinese researchers who obtained similar PhDs abroad, but for any reason continued to be based in the USA. This choice of comparison is one of the many possible to meet the research aim. Table 1 provides a list of possible comparisons if one wants to individuate Y1000T recipients for a comparative analysis. We focus on other international PhD-bred Chinese who continued to work abroad to answer the question: is working under this generous scheme in the Chinese system more conducive to research than continuing to work abroad? Our choice is dictated by feasibility in terms of data collection and statistical robustness, with the aim to unpack whether the Chinese system is already capable to provide early- and mid-career researchers with similar conditions for top research performances in comparison with US research-intensive universities. Notably, this research design is non-replicable for future study, for the cohorts of recipients are no more publicly available on official Web sites, making the current dataset a unique one.

Specifically for this paper, the research question we address is: are Chinese early- and mid-career researchers who received their
Table 1. Possible options in devising a comparison surrounding the Y1000T recipients.

<table>
<thead>
<tr>
<th>Possible comparisons</th>
<th>Possible rationales/questions for comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1000T Chinese-bred recipients</td>
<td>Do returnees have better research performance than domestically bred researchers?</td>
</tr>
<tr>
<td>Other Chinese academics of similar career stage</td>
<td>To what extent this policy is effective in boosting the research performance?</td>
</tr>
<tr>
<td>Other non-Y1000T returnees of similar career stage</td>
<td>To what extent returnees with the support from the Y1000T outperform those without?</td>
</tr>
<tr>
<td>Other Talents policies recipients based in Mainland China</td>
<td>By recruiting talents at what career stage can a system maximize the return of investing in talents?</td>
</tr>
<tr>
<td>Other non-Chinese Y1000T-like recipients (e.g. European Research Council or Marie Skłodowska-Curie Fellowship recipients)</td>
<td>Which policy is better devised overall? Which national system is best to host a recipient?</td>
</tr>
<tr>
<td>Other globally bred Chinese PhD holders who continue to work abroad</td>
<td>Is working under this generous scheme in the Chinese system a better choice than remaining abroad?</td>
</tr>
</tbody>
</table>

PhDs from world-leading higher education institutions and are recruited back to China under the scheme of the Y1000T better performing in terms of research quality, in comparison with those who have similar education and research experience but continue to be affiliated in the US research-intensive universities? By approaching this research question, we argue that the generous working conditions offered by the Y1000T policy are essential and beneficial, but do not necessarily include other equally important contextual conditions such as research culture or international connectedness. Working conditions in fact do not only include salary, project funding, infrastructure, and teaching loads. Working conditions reflect the comprehensive environment in which researchers do their work, such as research culture, connections to other national research systems, and other non-material conditions. We assume that within similar working conditions, researchers with similar characteristics in terms of education, working experience, and publication record would perform similarly.

The paper is organized in the following way. The next section describes Chinese policies of recruiting talents who are at the early- or mid-stage of their careers. It also provides further details about Y1000T policy in particular. Section 3 describes the state of the art in analyzing this topic. Section 4 exposes how the dataset has been designed and the respective variables. Section 5 shows some basic information by both groups (Y1000T recipients and the ‘control group’) before and after the ‘treatment’ (the Y1000T recruitment year). Section 6 presents main findings. Section 7 highlights explanations of the results for the topic and policy implications.

2. Talents policies and the young thousands talents

China’s talent attraction policies were introduced to deal with China’s brain drain problem—an increasingly preoccupation in the late twentieth century. The relatively low return rate of Chinese students and scholars in the 1990s drew concerns among governmental officials. Researchers argued that China was suffering from a large amount of brain drain (Cao 2008; Hu et al. 2017; Wang and Bao 2015). As a result, from the 1990s onward, Chinese governments successively issued a series of policies aiming at attracting overseas Chinese talents and dealing with the brain drain problem. Numerous national policies and schemes have been promoted by the Party Committee, the Central Government’s Ministries and Commissions, central public institutions (including the Chinese Academy of Sciences), and main scientific organizations such as the National Natural Science Foundation. The strategic issue of ‘talents’ also has a political connotation that justifies the epithet of flagship (Zweig and Wang 2013). In such contexts, the Young Thousand Talents Program (Y1000T) was introduced in 2011, attempting to attract relatively young talents who have potential to become leading figures. The scheme of Y1000T is open to talents in any discipline, but it predominantly recruits those in STEM disciplines.

Y1000T aims at recruiting scientists below the age of 40 years who normally have at least 3-year overseas working experience. Successful recipients shall have ‘engaged in scientific research, with formal teaching and research positions in overseas prestigious universities, institutions or enterprises’, as the policy states. The Y1000T researchers consist of those holding a Chinese university degree as well as a non-Chinese university degree, although the latter group outscores the former (Wang 2011). Existing discussions about the policy are descriptive and basic (Li et al. 2018a; Yang 2015; Zha 2016), lacking in-depth quantitative analysis of the whole program and its implications.

This policy raises multiple implications that deserve a brief exposition, as outlined in Table 1. For instance, it is relevant to know if China is able to attract back some of its brightest diaspora researchers once they have completed their education abroad. If the recruited researchers are not among the best possible candidates, or there exists bias during the recruitment process, the research performance of Chinese returnees and of overseas-based Chinese talents would be incomparable. The paper does not face such problems connected with the decision of applying, nor the issue of recruitment processes. The reason is that in terms of prestige of PhD-awarding institutions and their previous publication records, the Y1000T successful applicants (or Y1000T recipients) are clearly among the top researchers (see Table 2).

We compare the Y1000T recipients’ ex post research performances against other Chinese researchers who choose not to return.

Table 2. Detailed descriptive statistics of normalized citations by group and period.

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>Y1000T 2.419</td>
<td>1.441</td>
</tr>
<tr>
<td>Control group 2.416</td>
<td>1.609</td>
<td></td>
</tr>
<tr>
<td>St.Dev.</td>
<td>Y1000T 0.2631</td>
<td>0.5076</td>
</tr>
<tr>
<td>Control group 0.4819</td>
<td>0.6562</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>Y1000T 2.3043</td>
<td>2.7575</td>
</tr>
<tr>
<td>Control group 4.8129</td>
<td>2.7032</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>Y1000T 0.2763</td>
<td>−0.5913</td>
</tr>
<tr>
<td>Control group −1.0976</td>
<td>−0.2645</td>
<td></td>
</tr>
</tbody>
</table>
This choice excludes issues regarding whether or not some Chinese diaspora researchers who remain abroad ever applied to Y1000T or would apply to similar schemes. Previous studies have admitted that ex post research designs aiming at assessing specific policies are difficult to devise. For example, by analyzing the institution origins and tenure conditions of the same policy (Y1000T), Sun et al. (2017), for instance, argued that some US-based researchers tended to keep their positions and decline the opportunity to be a returnee under the Y1000T policy, if they already had a stable position in the USA. Though this research reveals an important finding, nevertheless, it does not assess who is more likely to produce high-quality research if both returned and diaspora researchers have secured good positions, either in China or in overseas.

3. Literature review

Talent attraction implies the possibility of mobility from one side (e.g., Chinese going abroad) and the capacity to attract back on the other side (e.g., Chinese nationals returning home). The mobility of PhD holders is key to trace leaders of the global competition in higher education and research (Lepori et al. 2015). As has been empirically found (Shen et al. 2016), the overall size and quality of international PhD students are still misbalanced between China and the Global West. Moreover, Lundh (2011) predicted that the US supremacy in talents attraction should have lasted for another 10 years at least. However, Lundh’s research is not an evidence-based forecast and does not account the more recent rise of Chinese research. Whether or not China may pair the USA in research, and if yes, when that will happen, remain unanswered.

Mobility is self-reinforcing, exacerbating differences among destination and origin countries. Classical studies (e.g. Khoury 1977) highlighted that PhDs bred in prestigious places were much less likely to work in places where higher education was less prestigious. Further, talents who get their PhDs from leading overseas universities usually perform well even before obtaining doctoral degrees (Veugelers and Van Bouwel 2015). While researchers have pointed out that mobility is reciprocally beneficial for the USA–China relationship (Chen 2015), the return of top-notch Chinese researchers to China cannot be taken for granted and needs specific encouraging policies.

Researchers return to their home countries for many reasons. The mobility of Chinese scholars who may return to China has, especially in the last decade, attracted international attention. However, the topic is still understudied, with a rather small amount of literature. In addition, most of the previous studies on the topic are already outdated (Huang 1997; Zweig and Chen 1995). Some empirical research (see for example Wang and Bao 2015) informs that the increasing reputation of Chinese higher education institutions is conducive to boost attractiveness to its own diaspora. Nevertheless, returnees may still face problems in terms of readapting to the country (Ma and Pan 2015). Apparently, expectations by young researchers may go at odds with returnees’ culture acquired overseas, rendering China’s pathway toward establishing world-leading universities more challenging (Song 2018).

Intriguingly, Wang et al. (2015) pointed out that Chinese universities may only be attractive to some of its diaspora who have not hitherto particularly productive. This finding implies that while some Chinese research-intensive universities have gained international reputation, they might still not be attractive enough for the best talents to return (Wang et al. 2015). As noted, such finding exposes a formidable problem of self-selection bias on empirical studies, which in this research, we argue, is overcome by the stringent conditions put in place by the Y1000T policy.

Institutions, and their prestige in particular, also play an important role in attracting returnees and supporting their career development. In terms of research collaboration, Li et al. (2015) found that researchers often face significant differences of cultural framework and assumptions. As researchers are often influenced by the academic norms of the country where they received PhDs, various productivity patterns of researchers trained in different countries may emerge accordingly. Tangible examples are differences in terms of publication strategies, publication behaviors, and propensity to publish in international journals (Jonkers and Cruz-Castro 2013). In addition, Lu and McInerney (2016) claimed that Chinese scholars often use their own social capitals according to their background: domestically educated ones use social capital in the closure mode, and ones with more international experience turn on their structural holes to maximize opportunities. This is a relevant pattern when valid proxies of such social capital are under consideration. For this sort of phenomenon, international co-authorship is the best proxy to grasp patterns of collaborations and outcomes of social capital dynamics (Wagner and Leydesdorff 2005). Different types of international co-authorships are in fact consistent with the idea of comparing returnees with non-returnees who still work abroad. The difference between Chinese globally bred returnees versus US-based Chinese globally bred researchers may also reflect whether returnees are more or less able to stay connected and involved in international networks.

Arguably, the comprehensive working conditions that researchers have can make a difference in their performance. Publications largely depend on available financial resources, especially for STEM disciplines (Zhang et al. 2016). While it is acknowledged that there exist problems in China’s domestic academic culture that are not conducive to high-quality academic output, as Cao (2008) has pointed out, we argue that at least regarding financial support for research and certain aspects of research infrastructure, the Y1000T recipients are often in a more advanced place in comparison with overseas diaspora. For example, the Y1000T recipients are more likely to have less teaching load and are better supported in terms of research funding and the establishment their own research teams made up of doctoral and postdoctoral researchers. Correspondingly, the main focus of many talent-recruitment policies is on both quality and quantity of research outputs (Cai 2012). Reduced teaching loads, and possibilities to establish one’s team, are also critical advantages because although preferable time to be dedicated to research depends primarily by one’s stage of research (Bentley and Kvyk 2013), certainly exceptions from consistent teaching loads at the early- and mid-career stage are favorable.

Previous studies have pinpointed that academics with qualifications attained abroad are more productive than those who are inbred (Xian 2015). However, Xian (2015) did not consider the possible effect of policies such as talents attraction ones. To this regard, Cheung and Xu (2015) demonstrated that China’s ability to narrow the gap between its research capacity and the global excellence largely depends on dedicated policies of talents attraction. According to Cheung and Xu (2015), without dedicated policies China couldn’t attract the best talents, and returnees would probably be those who are less likely to get good jobs in the USA or in other Global West countries. Talent-recruitment policies make a big difference in terms of quality of returnees, although few studies attempt an empirical analysis (Marini and Yang 2021; Yang and Marini 2019).
Studies often do not examine publications authored by Chinese people affiliated in China (regardless of their possible mobility experiences) with Chinese nationals active abroad. Some research compares Chinese citizens based in the USA with other foreign researchers active in the USA (Van Holm et al. 2019). The issue of mobility and productivity is considered in the literature, either in similar policies, or about other countries; Shin et al. (2014) and Payumo et al. (2018) and Horta (2013) for the USA. However, these studies usually do not adopt a quasi-experimental research design.

Another relevant factor in analyzing mobility of talents is the prestige of PhD-awarding institutions and the prestige of the institution where a person works afterward. To this regard, scientometric indicators at the institutional level can be used as proxies. These indicators reflect the standing of infrastructures and overall the notorious Mertonian self-reinforcing mechanism of prestige. In turn, they may affect supportive organizational climate for the goal of publishing (Uslu 2017)—this latter being a not less important working condition. Arguably, working conditions also entail the possibility to nurture well-prepared PhD students.

The aim of talent-recruitment policies is not only to reduce brain drain, but foremost to trigger some brain gain (Ma and Pan 2015). Not surprisingly, countries that are more likely to attract possible returnees are those that spend more resources in higher education (Gribble 2008). It is expected that without the presence of such policies, the number of returnees would be arguably lower for a country like China—at least this was the situation a decade ago or more (Finn 2010). The most available recent data show a quite high rate of returnees in the last available years (Kim et al. 2011), leading to the necessity to explore this phenomenon in terms of returnees’ performance after their return. Some literature has highlighted that since the ‘opening doors’ times in the 1980s, China’s science and research destiny cannot be separated from staff policies (Cao 1991).

Although China is becoming a core country in STEM disciplines at global level, there are misbalances between disciplines. This disciplinary imbalance is discussed in literature as a by-product of specific policies. For example, most of China’s talent recruitment policies focus on attracting talents in STEM areas. Underestimating the role of humanities is a possible harbinger of some undesired side effects (Oleksiyenko 2014; Song 2018). There are also certain biographical features that can help to predict the likelihood by talents to return to China (Sun et al. 2017). One of the factors is age: the older a person is, the more likely she or he wants to spend the sparing stint of her/his professional life in their motherland (Sun et al. 2017).

All in all, Chinese diaspora who have the best research performance tend to remain abroad (Sun et al. 2017). An explanation is that Chinese researchers have perceived China in the past as a yet not ready context to let them have similar (or even better) conditions to conduct research in their homeland (Sun et al. 2017). In fact, top researchers often emphasize conditions such as the possibility of networking for dense collaborations, on top of ‘hygienic’ (e.g., salary, budget, grants) conditions such as competitive salaries or steeped careers (Canhiano and Wooley 2015). The Y1000T program overcame the obstacle of attractiveness, if we look at the quality of returnees. Thus, it is necessary to explore further if the Y1000T recipients perform similarly regarding research quality compared with those who remained active in leading US research-intensive universities.

4. Data
The dataset of this study is original. It is a combination of individual curriculum vitae (CV) information plus respective Scopus publications. Individuals’ publications afford to generate further variables. Using CV as a source of data has a long-standing tradition, and a specific application in Chinese geographical field (Lu and McInerney 2016). If the person collecting data is proficient in the language, and CVs are available in official institutional Web sites, CV can reveal information such as age, institution of graduation and PhD attainment, and mobility trajectories.

The list of researchers who were recruited under the Y1000T policy was of public domain at the moment of data collection. We considered the first two waves of this policy (2011 and 2012 years) in order to have a reasonable span of time to compare scientific productions before and after Y1000T recruitment, which identifies our ‘treatment’ moment. The first two waves of Y1000T list around 350 people, among whom some biographical data are available. This number drops to around 200 (see Table 3) once we excluded those not employed in universities (some people work for public research institutions or companies). These researchers have been scanned individually to check if they at the current stage still work in mainland China and if they got their PhD in China or abroad. We include only those who got a PhD abroad (see Table 1 to appreciate that also Talents with a PhD from Chinese universities, although sporadically, might become Y1000T recipients). This group is the ‘treated’ group, meaning that they underwent the policy.

A ‘control group’ has been extracted manually from official Web sites of some US research-intensive universities. This sample was devised in a stratified way. Universities have been chosen in order to represent proportionally US institutions from the top 500, according to the Academic Ranking of World-class Universities 2018 (see Supplementary Data for the full list). Sampling of universities is proportionate to their places in rankings. By name, we individually collected people who belonged apparently to the same generation of the first two waves of Y1000T (this is possible as information about the ‘treated’ researchers’ birth year and the year of PhD attainment is available). The search was pursued in staff directories in, predominantly, STEM disciplines. Afterward, we filtered in only those who completed up to the master level education in mainland China and obtained the PhD in leading Western universities (e.g. the USA, Canada, some European countries, and few other Australasian-developed countries). We pruned manually outliers by age and prestige of PhD-awarding institutions against known ‘treated’ information in order to maximize the comparability. In total, the dataset comprises around 550 selected people. Researchers in both the ‘treatment group’ and ‘control group’ are very similar by age, career stage, doctoral background, and publication record (see Section 5 for full details and data display, plus Figs 1 and 2). The dataset comprises single scientific output as observations (around 37,000 in total), which can be referred to univocal authors, who in turns belong to either the treatment group or the control group. More information about data retrieval is available upon request.

4.1 Biographical information
Age, year of PhD attainment, gender, current scientific affiliation, and PhD-awarding institution (the latter two also indicate respective countries) are the information grabbed from CV or institutional official Web sites.

4.2 Dependent variables
From lists of publications, citations per publication (recoded as zero if missing) are available. Considering the obvious time-scaled feature of accrued citations, citations have been normalized dividing the gross
number of citations by the natural logarithm of the years elapsed from present plus one (\(\ln_{tcit}\)). We prefer this dependent variable instead of gross number of outputs as other researchers did in a similar study (Zhao et al. 2020) in order to highlight the issue of quality of research.

4.3 Independent variables

There are several independent variables in the dataset, listed as follows.

4.3.1 Type of publications

Publications can be grouped by further characteristics available from Scopus:

- Open access (OA) (binary variable): 1 is OA publication; 0 is non-OA.
- Language of the publication—(lingua) (binary): (1) is in English only; (0) is when any another language is considered by either journal or publisher.
- Typical output (Type_p) (binary): articles (1) are typical publications; any other type of publication is considered non-typical (e.g. chapters, proceedings) (0).

4.3.2 Social capital

Social capital of authors is relevant (Lu and McInerney 2016; Van Holm et al. 2019). From the list of authors of each publication and co-authors’ affiliated countries, some different proxies of social capital have been computed. In particular, the dataset includes the following variables:

- the number of co-authors by each output (no_aut);
- degree of internationality (int_coll), expressed with an indicator of international co-authoring signature out of the seventy most frequent countries found in the list of co-authoring countries. This variable does not appear in tables. int_coll2 is a respective binary variable defining ‘1’ for international publications and ‘0’ for non-international publications;
- heterogeneity of co-authorships (heter) defines the ratio between int_coll over no_aut. This variable measures the extent to which publications are heterogeneous in terms of multiple affiliations in multiple countries. It is relevant to discern whether people publish with large and pluri-international co-authorship networks, grasping a further dimension of internationality.
- institutional mobility (mob): this is a binary variable describing if scientists changed institutions from PhD awarding one into the current one (by definition, all treated Y1000T are mobile as the dataset includes only Chinese researchers who got a PhD overseas).
- Kept collaboration with China (KC): this variable disentangles productivity patterns before treatment, when all researchers were based abroad. It has value 1 when a publication was co-authored with scholars affiliated in Chinese institutions, and 0 when a single publication had no co-authors from China. All publications after treatment are computed with another value (2). This variable is a proxy to check possible self-selection bias not otherwise observable (see below—self-selection bias subsection).

4.3.3 Standings of universities

Standings of universities, both in terms of PhD attainment and of current employment, are essential. As Taylor and Cantwell (2015)
demonstrated, American universities, especially the best ones including private universities, attract PhD students at a global ray. Nevertheless, alma mater and current employer may determine the extent to which researchers can thrive. These variables are confounding ones to keep parity of contexts. Arguably, researchers benefit from their own affiliation in proportion of institution performativity. To have reliable proxies of this factor, the dataset comprises some indicators about the average quality of publications by institution. Moreover, it is consistent with literature (David and Motala 2017) that Chinese research universities are somehow still underestimated in their positions in the global rankings, if positions in rankings are compared against mere bibliometric indicators. Hence, we prefer bibliometric indicators in order to have an objective scientific measure of quality in research at institutional level. In particular, we extracted information from InCites Web of Science for any institution where talents received their doctoral degrees and where they worked in the period of 2014–2018. The list of institutional qualifiers includes the following ones (see Table 3):

- Average of Categorical Normalized Citations Indicator (CNCI) as computed by Clarivate Analytics;
- average of Journal Normalized Citation Index (JNCI);
- average of percentile of articles (av_percentile), inverted for comparability;
- percentage of top 10 percentile articles (perTOP10);
- percentage of 1 percentile articles (perTOP_1);
- percentage of cited documents (% documents cited).

Considering the limited array of disciplines under examination, we do not disentangle within universities’ disciplines, preferring a list of indicators that may grasp different aspects of institutional quality.

4.4 Determination of ‘period’
‘Period’, meaning happening before or after a given ‘treatment’, is essential to set up difference-in-differences tests. Since the dataset is based on single publications as records, it is possible to keep year of publication (year) of each publication as an information to compute further variables. For whoever was recruited by Y1000T, period before (‘0’) and after (‘1’) are simple to be computed. Although the researchers in the control group are similar by age to those in Y1000T, there is no a priori ground to believe that they should have been recruited in a certain moment such as, say, a tantamount of years after PhD attainment. We computed the most reasonable moment of hypothetical Y1000T recruitment for control group by means of using the median (and mean) of years elapsed from PhD attainment and recruitment in Y1000T policy. This results in being 5 years after PhD attainment. This referral splits publications authored by researchers in control group between a hypothetical pre- and post-treatment. Figures 1 and 2 illustrate the distribution of publications by Y1000T and control group, respectively, splitting in each chart by period—publications to be considered before (‘0’) treatment and publications to the considered after (‘1’) the actual (or potential for control group researchers) Y1000T recruitment.

Another analytical choice would have been possible—propensity score matching (PSM). This technique, however, when accounting for different ‘treatment’ moments in personal career trajectories, does not yield particularly different outputs (Veugelers and Van Bouwel 2015). Moreover, PSM would undermine consequential analysis, according to some literature (King and Nielsen 2019). PSM would also require dropping some precious observations. For this reason, ‘period’ for the control group is determined by solely the above-mentioned median of years passed from PhD Attainment and Y1000T recruitment.

4.5 Treat and period simple interaction
In order to run a difference-in-differences test, both treatment and period binary variables are needed. Treated (treat = 1) are simply Chinese who are treated by the policy and are working currently in mainland China (as noted we only considered overseas PhD holders); not treated (treat = 0) consists of the control group. ‘Period’
indicates whether any publication occurred before or after the treatment, as already defined. Table 2 shows detailed distribution with *prima facie* evidence about a reduced average of normalized citations (see Fig. 3) by Y1000T against control group, along with a more skewed on the right distribution for Y1000T—suggesting that overall negative effects are mixed with some positive ones.

As Table 2 shows, the logarithm of normalized citations before the treatment was very similar between the treatment and control group (2.419 versus 2.416), whereas after the treatment the control group has a higher average (1.61 versus 1.44).

### 4.6 Previous relationships as proxy of self-selection bias

There is no public access to the recruitment criteria and recruitment procedure of the Y1000T policy, nor the aim of the paper is to understand whether Y1000T policy was effective in recruiting its own recipients out of the larger group of applicants (see Section 1 and Table 1). It is also impossible to get access to the list of unsuccessful Y1000T applicants. Therefore, once we extracted a control group with similar characteristics, it is useful to consider the most probable bias out of observable features.

---

**Figure 2.** Distribution of publications according to years elapsed from person’s PhD attainment by period (ytreat) of treatment groups: ‘0’ equals before treatment; ‘1’ equals after treatment (both treated and control group).

**Figure 3.** Distribution of dependent variable (ln_tcit) by group and period.
Self-selection bias among returnees may occur. Potentially such self-selection bias might be at the base of poorer performance by Y1000T recipients, although a link between inbreeding and poorer performances is not proved (Jiang et al. 2020). Good productivity during PhD studentship periods instead is more likely to predict success at later stages of one’s career (Horta and Santos 2016), which is grasped in this analysis as period 0. We tried to take into account possible self-selection bias by checking one of the most determinant factors in making a Chinese scholar decide to return to China: whether these Chinese scholars had kept any scientific liaison with their mother Country before they were recruited by Y1000T (of before the moment they might have been recruited under Y1000T for control group researchers). This is also consistent with the literature (Baruffaldi and Landoni 2012). To this regard, we computed a variable called KR (having kept any scientific relationship in China via co-authorships during period ‘0’) to account for possible self-selection biases in congruence with the literature (Jonkers and Tijssen 2008). In the Chinese context, it is in fact reasonable to assume that some scholars are more likely to be recruited back to China if they had kept persistent contacts with senior scholars in mainland China although the PhD of the younger scholar was attained abroad. This feature might interfere with some aspects of one’s performance, such as quality of the publications and one’s international embeddedness. Results (Table 4) provide a test of endogeneity to this regard, by using this variable as a possible instrumental one in defining the treatment group.

5. Descriptive statistics

Table 3 shows descriptive statistics of the original dataset built by combining official information about Y1000T recipients and the manual extraction of control group researchers who are active in US research-intensive universities. The variables are those introduced in the previous section. Table 3 considers a dataset with publications collapsed by authors. Table 3 splits statistics by the two groups (the treatment and control group) and by the two periods used to run a difference-in-differences test, as exposed in the previous section (Table 2).

As seen, the treated group is relatively small, comprising less than 200 valid persons, whereas for the control group we have around a double number of persons at almost the same age, and holding PhDs awarded by institutions with similar standings. Citations for Y1000T recipients decrease more than control group researchers do, as already seen. The control group is more balanced in terms of gender, although there is no difference in terms of disciplines, as the Y1000T policy predominantly considers some specific disciplines (life sciences, engineering and materials sciences, chemistry, mathematical and physical sciences, informational sciences, environmental and earth sciences, medicine, and public health and preventive medicine) and that the control group consists of faculty affiliated in departments belonging to these same disciplines. Year of PhD attainment is very similar and normally distributed in both groups. The respective average of years of publication is also very similar. On average, the time of publication is around +2 years after year of PhD attainment, computing also publications occurred before PhD attainment.

International collaborations (\textit{int\_coll}) decrease for both groups by periods. In particular, Y1000T recipients seem less used to co-author internationally. Interestingly, heterogeneity of co-authorships (\textit{heter}) was higher among Y1000T recipients. This value dropped more for them after recruitment to China, suggesting a slightly less cosmopolitan environment for returnees, if compared with control group still active in the USA.

The frequency of publishing in Scopus-indexed non-English journals is very sporadic (\textit{lingua}). For returnees the percentage does not grow, indicating that they keep themselves focused on international scientific production. However, returnees might choose to publish in, for example, non-Scopus-indexed Chinese journals, which are not considered in the paper. Number of co-authors is also similar in averages between these two groups (from 3.28 to 6.87 for control group; and from 5.3 to 6.92 in treated). Also, the percentage of publications in OA mode is similar in growth between the period 0 and period 1. In terms of the type of publications, Y1000T researchers seem to be attentive to publish in the main form of output—journal article, albeit this prevalence rises in both groups.

KR variable reflects big differences between the returnees/treated researchers from one side and control group on the other side: observing the pretreatment groups, 60 percentile of publications of returnees are co-authored with Chinese-based researchers, whereas only 12 percentile of control group co-authored with China-based scholars up to 5 years after PhD attainment (which is the definition of pretreatment period for them). These figures give a solid ground to use it as an instrumental variable in checking possible selection bias that in turn might have engendered publications patterns also in post-treatment periods.

The list of indicators defining the standings of the respective institutions refers to the prestige of the institutions that awarded the PhD title and the current affiliation. Values may differ by period modes: before treatment means PhD-awarding institutions and current one for post-treatment period (we discard the effect of other possible affiliations that might have occurred meanwhile). Values show clearly that there is no massive difference among the four groups deriving from the combination of treatment and period. Only the percentage of top_1 percentile articles is lower for post-treatment-treated scholars. The percentage of the cited documents over the total number of publications per institution is higher for the post-treatment Y1000T group and slightly lower for control group scholars. These few differences may imply that for treated scholars their institutions are less heterogeneous, whereas for the control group there is a higher degree of variety caused probably by regular labor market dynamic. In other words, our 500 global ranking threshold results in being slightly more heterogeneous than the standing of Chinese institutions where Y1000T recipients find affiliation. Notwithstanding, many other Chinese who got a PhD in the USA might have found a job in further slightly less prestigious US universities that we exclude.

6. Results

The main hypothesis is that Y1000T-treated Chinese researchers who are globally bred are better off in terms of quality of their scientific production in comparison with other Chinese globally bred researchers who continued to be active in the USA. Since Table 2 fosters indications of nonlinear distribution, we proceeded testing possible curvilinear effects. This is tested via six models of difference-in-differences regressions resonating inter-quartiles distributions of publications by normalized citations. Model 1 predicts if there is any treatment effect on observing papers within the first quartile (most cited, normalizing by time of publication) against the rest. Models 2 and 3 check treatment prediction of the middle of the
Table 4. Difference-in-differences tests for influence of papers (normalized citations) to predict treatment in belonging to first lower quartile (Model 1), top 10% (Model 5), and top 1% (Model 6).

<table>
<thead>
<tr>
<th></th>
<th>(1) ln_tcitQ1</th>
<th>(2) ln_tcitQ2_3</th>
<th>(3) ln_tcitQ2_3a</th>
<th>(4) ln_tcitQ4</th>
<th>(5) ln_tcitTOP10</th>
<th>(6) ln_tcitTOP_1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treat</td>
<td>-0.0298</td>
<td>0.0965***</td>
<td>0.0066</td>
<td>0.0206</td>
<td>-0.137***</td>
<td>-0.0473***</td>
</tr>
<tr>
<td>(-1.09)</td>
<td>(2.24)</td>
<td>(0.10)</td>
<td>(0.37)</td>
<td>(-3.11)</td>
<td>(-3.06)</td>
<td></td>
</tr>
<tr>
<td>1.ytreat</td>
<td>0.273***</td>
<td>-0.262***</td>
<td>0.4184***</td>
<td>-0.457***</td>
<td>-0.254***</td>
<td>-0.3011***</td>
</tr>
<tr>
<td>(16.25)</td>
<td>(-8.81)</td>
<td>(13.31)</td>
<td>(-15.73)</td>
<td>(-10.47)</td>
<td>(-6.69)</td>
<td></td>
</tr>
<tr>
<td>1.treat#1.ytreat</td>
<td>0.246***</td>
<td>-0.268***</td>
<td>0.2034***</td>
<td>-0.171***</td>
<td>0.0673</td>
<td>0.0365***</td>
</tr>
<tr>
<td>(9.22)</td>
<td>(-6.11)</td>
<td>(3.94)</td>
<td>(-3.57)</td>
<td>(1.65)</td>
<td>(3.06)</td>
<td></td>
</tr>
<tr>
<td>CNCI</td>
<td>0.0309</td>
<td>-0.0419</td>
<td>-0.0582</td>
<td>0.0213</td>
<td>-0.00722</td>
<td>0.0113</td>
</tr>
<tr>
<td>(0.37)</td>
<td>(-0.43)</td>
<td>(-0.38)</td>
<td>(0.17)</td>
<td>(-0.08)</td>
<td>(0.49)</td>
<td></td>
</tr>
<tr>
<td>perDOC_cit</td>
<td>-0.000775</td>
<td>0.00412</td>
<td>0.0053</td>
<td>-0.00324</td>
<td>0.00215</td>
<td>-0.00129</td>
</tr>
<tr>
<td>(-0.22)</td>
<td>(0.96)</td>
<td>(0.74)</td>
<td>(-0.56)</td>
<td>(0.51)</td>
<td>(1.53)</td>
<td></td>
</tr>
<tr>
<td>perTOP10</td>
<td>-0.00861</td>
<td>0.0135</td>
<td>0.0034</td>
<td>-0.00113</td>
<td>0.00606</td>
<td>-0.00347</td>
</tr>
<tr>
<td>(-0.77)</td>
<td>(1.05)</td>
<td>(0.14)</td>
<td>(-0.06)</td>
<td>(0.45)</td>
<td>(1.21)</td>
<td></td>
</tr>
<tr>
<td>av_percentile</td>
<td>0.0112</td>
<td>-0.00905</td>
<td>0.0216</td>
<td>-0.0137</td>
<td>-0.0138</td>
<td>0.00168</td>
</tr>
<tr>
<td>(0.45)</td>
<td>(-0.30)</td>
<td>(0.47)</td>
<td>(-0.38)</td>
<td>(-0.48)</td>
<td>(0.26)</td>
<td></td>
</tr>
<tr>
<td>JNCI</td>
<td>0.0615</td>
<td>-0.0425</td>
<td>0.2510</td>
<td>-0.163</td>
<td>-0.0791</td>
<td>-0.0238</td>
</tr>
<tr>
<td>(0.58)</td>
<td>(-0.32)</td>
<td>(1.31)</td>
<td>(-1.04)</td>
<td>(-0.65)</td>
<td>(-0.74)</td>
<td></td>
</tr>
<tr>
<td>no_aut</td>
<td>0.000233</td>
<td>-0.000215</td>
<td>0.0003</td>
<td>-0.000308</td>
<td>-0.000271</td>
<td>-0.0000416</td>
</tr>
<tr>
<td>(0.46)</td>
<td>(-0.35)</td>
<td>(0.50)</td>
<td>(-0.52)</td>
<td>(-0.86)</td>
<td>(-0.68)</td>
<td></td>
</tr>
<tr>
<td>int_coll2</td>
<td>0.00179</td>
<td>-0.00126</td>
<td>0.0282</td>
<td>-0.0230</td>
<td>0.00637</td>
<td>0.00642</td>
</tr>
<tr>
<td>(0.13)</td>
<td>(-0.09)</td>
<td>(1.16)</td>
<td>(-1.41)</td>
<td>(0.51)</td>
<td>(1.66)</td>
<td></td>
</tr>
<tr>
<td>Heter</td>
<td>-0.00873***</td>
<td>0.00992***</td>
<td>-0.0094***</td>
<td>0.00924***</td>
<td>0.00443***</td>
<td>0.00156***</td>
</tr>
<tr>
<td>(7.68)</td>
<td>(6.07)</td>
<td>(4.93)</td>
<td>(5.09)</td>
<td>(2.72)</td>
<td>(2.88)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>-0.0136</td>
<td>0.00562</td>
<td>-0.0334</td>
<td>0.0319</td>
<td>0.0273</td>
<td>0.00462</td>
</tr>
<tr>
<td>(-0.85)</td>
<td>(0.29)</td>
<td>(1.23)</td>
<td>(1.32)</td>
<td>(1.53)</td>
<td>(0.95)</td>
<td></td>
</tr>
<tr>
<td>OA</td>
<td>0.0102</td>
<td>0.00496</td>
<td>0.0965***</td>
<td>-0.0706***</td>
<td>-0.0431***</td>
<td>-0.00195</td>
</tr>
<tr>
<td>(0.65)</td>
<td>(0.28)</td>
<td>(5.03)</td>
<td>(-4.97)</td>
<td>(-4.39)</td>
<td>(-0.41)</td>
<td></td>
</tr>
<tr>
<td>Mob</td>
<td>0.000527</td>
<td>-0.0203</td>
<td>-0.0300</td>
<td>0.0164</td>
<td>0.0373</td>
<td>0.00224***</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(-0.46)</td>
<td>(-0.45)</td>
<td>(0.27)</td>
<td>(0.65)</td>
<td>(4.30)</td>
<td></td>
</tr>
<tr>
<td>Lingua</td>
<td>-0.007550</td>
<td>0.0198</td>
<td>0.1604***</td>
<td>-0.124***</td>
<td>-0.137***</td>
<td>-0.0244***</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.50)</td>
<td>(6.30)</td>
<td>(-5.15)</td>
<td>(-5.38)</td>
<td>(-2.59)</td>
<td></td>
</tr>
<tr>
<td>type_p</td>
<td>0.0458***</td>
<td>-0.0631***</td>
<td>-0.0083</td>
<td>0.00280</td>
<td>0.0205***</td>
<td>0.0160***</td>
</tr>
<tr>
<td>(5.05)</td>
<td>(-5.57)</td>
<td>(-0.52)</td>
<td>(0.19)</td>
<td>(1.98)</td>
<td>(4.30)</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>0.269</td>
<td>-0.0249</td>
<td>-0.9794</td>
<td>1.473</td>
<td>0.0630</td>
<td>0.331</td>
</tr>
<tr>
<td>(0.31)</td>
<td>(-0.02)</td>
<td>(-0.57)</td>
<td>(1.06)</td>
<td>(0.06)</td>
<td>(1.63)</td>
<td></td>
</tr>
<tr>
<td>N (publications)</td>
<td>36,130</td>
<td>25,452</td>
<td>27,243</td>
<td>36,130</td>
<td>36,130</td>
<td>36,130</td>
</tr>
</tbody>
</table>

Test of endogeneity \( F(1,604) = 0.129991 \) \( F(1,599) = 0.001294 \) \( F(1,583) = 0.939535 \) \( F(1,604) = 1.49078 \) \( F(1,604) = 0.545283 \) \( F(1,604) = 2.61144 \)

(\( P = 0.7186 \)) \( P = 0.9713 \) \( P = 0.3328 \) \( P = 0.2226 \) \( P = 0.4605 \) \( P = 0.1066 \)

Tests of endogeneity assume variables are exogenous.
*** Indicate statistical significance at the 0.001 level. ** For the 0.01 level. * For the 0.05 level.

As Table 4 shows, Y1000Ts are more likely to produce papers with citations within the first quartile (Models 1). Models 2 and 3 show that treated researchers are less likely to fall in the middle league of the distribution. Model 4 still detects more likelihood, for the treated group, to result no better than non-returnees. Y1000Ts are instead more likely to produce top-1 percentile papers in comparison with control group people (M6), although there is no traceable effect of Y1000T policy for falling in the top-10 percentile of papers in terms of citations (M5). In all these cases, there is no rejection of the hypothesis of endogeneity, which means that some self-selection bias might have occurred. All in all, the treatment effect on citations appears to be curvilinear. The policy, if tested in comparison with non-returnees with similar characteristics, recruited researchers who eventually tended to publish either less successful publications (up to the first three quartiles in the distribution of scientific outputs by normalized citations) if the publications do not belong to the very top best 1 percentile.

In terms of covariates, all models show no influence caused by institutional indicators of standings, nor by gender and number of authors in publications, and not even by level of international collaboration in co-authorship. ‘Heter’, in contrast, is statistically significant, showing that higher degrees of internationality are fostered in self-selection bias (KR variable). Covariates are the same variables (publications) 36,130 25,452 27,243 36,130 36,130 36,130.
associated with having more high-quality papers in terms of citations. The type of publications (either articles in journals or not) and use of English (lingua) are both as significant as giving scattered indications, resulting in being not necessarily determinant of poor or extremely successful performances. Last, institutional mobility plays a positive role for the top-1 percentile of publications, indicating that inbreeding is not effective in this sample.

7. Discussion and conclusions

The effort made by the Chinese government to feed its research-intensive universities with worldwide globally bred researchers is remarkable, and worth receiving attention in order to understand the extent to which China is capable to get the most from its policies of reverse brain drain. It also sheds light on how emerging research systems can address brain drain and catch-up with leading research systems.

The question the paper aims to answer is are these returnees better off in terms of research performance compared with those who are equivalently globally bred but continued to work in the Global West, still in research intensive universities? This question does not only assess the success of individuals, but also reflects the Chinese system and its stage in achieving a leading global role in science. This quasi-experimental research leads to the conclusion that there is no simple yes or no answer to the question—the answer is more nuanced. Specifically, the Y1000T ‘treatment’ is more likely to predict publications within the bottom or the top-1 percentile. In other words, Y1000T returnees either publish less successful outputs or publish in the very top quartile—a minority of outputs though. These mixed findings may pave the way to multiple interpretations.

An interpretation might highlight the evaluation mechanism of the Chinese universities/institutions. Y1000T researchers are affiliated to top Chinese universities or public research institutes that run evaluation exercises at an institutional level (Wei and Zhao 2014). Most of these researchers start their careers with a tenure-track contract. So their academic performance and research outputs are guided by the evaluation measurements at institutional level, which are largely determined by the country’s national strategy. While Chinese universities highlight generally the number of publications and the impact factors of the publications, they are less concerned with the ‘actual’ quality of the research—meaning that the threshold of quality tended to be slightly lower in recent years, and not necessarily meeting top-1 percentile sort of requirements. This might explain the prevalence of bottom distribution for Y1000Ts’ publications. As a result, some researchers might have tended to pursue some research that is easier to be published—provided the threshold of quality is kept above decency, not necessarily these latter ones always meeting other considerations that may yield by time more citations, for instance. Y1000T may also avoid to endeavor relatively risk-taking research, which is the most likely one to secure marginal gains in recognition in the mid- and long run. This is very likely associated with citations capture by time. We might argue that this latter point is not due to the Y1000T policy itself, but as a result of the broader Chinese academic evaluation system, as has already been critiqued in the existing literature (Li et al. 2018a,b; Shao and Shen 2012; Wei and Zhao 2014; Zhu et al. 2004). Borne this latter factor in mind, it is even more striking that Y1000T is able to favor top-quality research, whereas previous assessment of this policy brought to the conclusion that the best-performing Chinese researchers do live abroad and that they are less likely to accept to go back to their country, also if under compelling and encouraging conditions (Sun et al. 2017).

Setting the Chinese case into international and comparative perspectives, the study suggests that while generous financial support is conducive to researchers’ scientific performance, the soft environment including academic culture also makes a difference, echoing findings by Scaffold and Berman (2011). This points to approaches for emerging research systems in attracting and nurturing talents—that is to focus on establishing a conducive soft research environment in addition to financial investment. According to Li et al. (2019), returnees often have a short-term decrease in research performance, and therefore, it is important to give researchers, especially early- and mid-career researchers, more autonomy and independence upon return. Possible policy implications include turning toward a less quantitative-oriented and shortsighted evaluation system of research—for example, extending the evaluation period for returnees (Li et al. 2019). Further, for the sake of boosting brain gain(-back), it seems important to consider not only financial resources, career opportunities in terms of steeped ranks ascension, or other material benefits. For providing further marginal advantages, it is also critical to assure a way to nurture researchers’ social and academic capital. To maximize one’s network of collaborations is a possible way-to-go in this case.

Indeed, the analysis reveals that higher degrees of heterogeneous collaborations are able to predict, to a certain extent, citations obtained by publications. This implies that establishing international networks of collaboration can be an important approach for individual researchers or national research systems to increase their influence. As Schott (1988) and Csomós (2018) argue, the centrality of a national science system in the world science system relies on international research collaboration and this condition is a key indicator of the system’s global influence. In this aspect, returnees are key resources in establishing and maintaining the connection (Cao et al. 2020). According to Jiang et al. (2020) and Shen and Wang (2018), researchers with international research and learning experience can carry their good learning and relevant work experiences back home—bringing positive impacts in fostering international research collaboration. This re-emphasizes the importance for countries to encourage domestic talents to gain international experience and attract international talents (Jiang et al. 2020). It also makes a point for future policies in the topic of talents and excellence, especially for emerging research systems.

Overall, the comprehensive physical and ‘soft’ infrastructures for research offered by Y1000T policy appear to have produced mixed consequences when Chinese globally bred researchers go back home. Yet, research upon the same or similar policies, including those in other countries, deserves more attention and ought to unpack the outcomes deriving from the interplay of more policies, all aiming at building a new world-leading research-intensive higher education system.

This paper also has limitations, opening the door to further questions. First, the reasons of curvilinear outcomes are not probably all observed. Second, social capital might be measured in a more fine-grained way, considering not only country level, but also institutional and individual ones. To know teaching and administrative loads across time by each person (including control group ones) would also be useful. The effect of access to funding agencies—which is available sometimes at level of grants recipients and/or single output funded by one or more agencies—on top of recruitment programs like Y1000T, also can help to understand differences in conditions and respective performances. Yet, other unobserved factors can...
determine further reasons of possible self-selection bias that may interfere with overall effectiveness. Also, comparisons between returnees and other Chinese-bred researchers under the same program may contribute to unpicking outcomes of Chinese policies on relevant topics. Other possible lines of research might be sketched in Table 1.

**Supplementary data**

**Supplementary data** are available at Science and Public Policy online.

**Conflict of interest statement.** None declared.

**References**


