Last updated: September 10, 2020

Updating China's Harmonized Test Scores for the Human Capital Index

Noam Angrist, Syedah Aroob Iqbal, and Aart Kraay[1]

I. Background

The Human Capital Index (HCI) measures the level of human capital that a child born today can expect to attain by age 18, given the risks of poor health and poor education that prevail in the country where she lives. The health and education components of the index are combined in a way that reflects their contributions to worker productivity. The index units represent productivity relative to a benchmark of complete education and full health, on a scale of 0 to 1.

The education component of the HCI—Learning-Adjusted Years of School—combines information on the quantity of school using Expected Years of School and its quality using Harmonized Test Scores (HTS). HTS are national average scores from major international and regional student achievement testing programs (ISATs and RSATs respectively), harmonized into common units. The HTS uses TIMSS-equivalent units, where 300 is minimal attainment and 625 is advanced attainment. HTS for the 2018 HCI were retrieved from the Global Database on Education Quality (Patrinos and Angrist, 2018).[2] The 2020 HCI update uses the January 2020 update of this database.

The HTS database harmonizes scores from seven ISATs and RSATs into HTS units by creating an exchange rate between international standardized achievement tests (ISATs) such as PISA, TIMSS, and PIRLS and their regional counterparts (RSATs) such as SACMEQ, LLECE, PASEC, and PILNA, as well as EGRAs. This exchange rate is derived by comparing average scores for countries that participate in both an RSAT and an ISAT in a given time period, schooling level (primary and secondary), and subject. This exchange rate is then applied to country-level average scores for all countries in the testing program.

II. An updated HTS for China

This note describes the process for estimating nationally-representative test scores for China for use in the September 2020 update of the World Bank's Human Capital Index (HCI). The methodology described here updates the approach described in Annex 4 of Patrinos and Angrist (2018)[3] that was used in the inaugural version of the HCI released in October 2018. The main change here is that newly-available data from the 2018 PISA round are now also included in the calculation.

With nearly 1.4 billion people, the People's Republic of China is the world's most populous country and has the world's largest education system. While special administrative regions of China, namely, Hong

---

[1] World Bank Education Global Practice (Iqbal), World Bank Development Research Group (Kraay), and Oxford (Angrist). The views expressed here are the authors' and do not reflect those of the World Bank, its Executive Directors, and the countries they represent.

[2] Patrinos, Harry and Noam Angrist (2018). "Global Dataset on Education Quality: A Review and Update (2000-2017). World Bank Policy Research Working Paper No. 8592. http://documents.worldbank.org/curated/en/390321538076747773/Global-Dataset-on-Education-Quality-A-Review-and-Update-2000-2017.

[3] Patrinos, Harry and Noam Angrist (2018). "Global Dataset on Education Quality: A Review and Update (2000-2017)". World Bank Policy Research Working Paper No. 8592. Available at http://documents.worldbank.org/curated/en/390321538076747773/Global-Dataset-on-Education-Quality-A-Review-and-Update-2000-2017.

Kong and Macao, have been regular participants in ISATs like TIMSS, PIRLS and PISA, China as a whole has yet to participate.

However, a few provinces have participated in PISA over time, providing some further insight into the quality of education in China. In 2009 and 2012, the province of Shanghai participated in PISA and was the top-performing economy in both years. Shanghai is predominantly urban with 88 percent of the residents in urban areas. The city's income per capita is also more than twice as high as the national average.[4] In 2015, four provinces, Beijing, Shanghai, Jiangsu and Guangdong (B-S-J-G) participated in PISA. These four provinces of China cover 17 percent of China's population and, with the inclusion of relatively less urban and less affluent provinces of Jiangsu and Guangdong as compared to Shanghai, provide insight into education quality in a less affluent region than Shanghai alone. However, their combined average per capita income is still well above the national average, and so test scores from these four provinces are unlikely to be representative of China as a whole. In 2018, four provinces, Beijing, Shanghai, Jiangsu and Zhejiang (B-S-J-Z), participated and for the same reasons as in 2015, are unlikely to be representative of China as a whole.

Apart from the participation of five provinces in PISA, some independent efforts have been conducted to assess education quality in other provinces in China using internationally-comparable metrics. One such effort is led by the Rural Education Action Program (REAP) at Stanford. In 2015, researchers at REAP led a reading assessment to assess education quality in Shaanxi province and rural areas of Guizhou and Jiangxi provinces.[5] The reading tests were constructed by trained psychometricians and used test items from the Progress in International Reading Literacy Study (PIRLS), thereby allowing international comparison. Results show that Chinese provinces of Jiangxi and Guizhou stood last in comparison to other countries and economies participating in PIRLS 2011. The provinces represent 5.8 percent of the population in China and give an insight into the education quality in rural, less affluent regions of China: average per capita income in these three regions is well below the national average.

III. Extrapolating to China National Average Test Scores

Information from PISA scores as provided by the OECD is combined with PIRLS scores as provided by REAP to estimate nationally-representative average education quality in China. For both programs, the fundamental problem is that the test scores are obtained in provinces that are unlikely to be nationally-representative of educational quality given their income gaps noted above, and given the observed income gradient in test scores across provinces. Even taking a population-weighted average of PISA and PIRLS scores for the seven provinces is unlikely to result in nationally-representative scores. This is because the average per capita income of these areas is still much higher than the national average, since the poorer regions covered in REAP are considerably less populous than the richer regions covered in PISA.

Therefore, national-level test scores are approximated by extrapolating the observed PISA and PIRLS scores based on per capita income. Specifically, PISA scores for Shanghai (2012), B-S-J-G (2015), and B-S-

[4] Throughout this note we measure per capita income as household per capita disposable income, as reported by NBS based on China's household survey. Available from the NBS website at http://data.stats.gov.cn/english/easyquery.htm?cn=E0103.

[5] Gao, Qiufeng, Yaojiang Shi, Hongmei Yi, Cody Abbey, and Scott Rozelle (2017). "Reading Achievement in China's Rural Primary Schools: A Study of Three Provinces". Stanford University, Freeman Spogli Institute Working Paper, available at https://fsi.stanford.edu/publication/reading-achievement-chinas-rural-primary-schools-study-three-province.

J-Z (2018) are extrapolated to the national average, separately by subject (Table 1), using a simple linear regression of test scores on province-level log per capita income. Averaging across subjects, this gives an extrapolated national PISA score of 418, as compared with PISA scores of 588 (Shanghai 2012), 514 (B-S-J-G 2015), and 579 (B-S-J-Z 2018). This corresponds to an extrapolated national HTS of 432. The same extrapolations are repeated using the PIRLS scores for reading only (Table 2). This gives an extrapolated national-level reading score of 449. Since the units of PIRLS and HTS are on an international scale, this is also the HTS.

The PISA and PIRLS-based extrapolations give reasonably consistent results. This is most apparent from comparing the PISA reading score extrapolation with the PIRLS extrapolation (Figure 1). The graph plots reading scores from the two programs against log per capita income. The small blue (orange) dots represent the data points for PISA (PIRLS). The dashed lines represent the corresponding extrapolations. Finally, the two large data points show the extrapolated values for China national-level reading scores, which are 398 and 449 for PISA and PIRLS respectively.

Figure 2 places China's extrapolated test score in international perspective. The gray dots in the graph report data from the 2015 PISA and 2016 TIMSS and PIRLS rounds for countries and territories participating in these programs. The test scores are on the vertical axis and are calculated as the average of whichever of TIMSS 2015, PISA 2015, and PIRLS 2016 are available for the country, after harmonizing to TIMSS-equivalent HLO units. The horizontal axis is log real GDP per capita. Over top of this we superimpose (a) the extrapolated China national test score (large red dot in center of graph) and (b) the actual PISA and PIRLS scores for Chinese provinces described above.[6]

IV. Use in the HCI Update

In summary, extrapolating PISA scores based on income differences across provinces suggests a national-level HLO of 432, while doing the same using PIRLS scores suggests a national-level HLO of 449. For the 2020 HCI, the HTS is calculated as an average of these two figures, 441, as the best estimate of the national-level HLO for China, with the values of 432 and 449 as lower and upper bounds.

Following best practices with index updates, the 2020 update to the HCI will include a back-calculated version of the 2018 HCI using the most recent data available now for 2018 for all countries. The updated HTS for China is based on data collected in 2012, 2015 and 2018, and for the purpose of the HCI is assigned the year 2015. This makes it the most recent measurement available as of 2018 that is used in the back-calculated 2018 HCI as well as in the 2020 HCI.[7] This updated HTS estimate of 441 is slightly lower than the HTS estimate of 456 for China reported in Patrinos and Angrist (2018) that was used in the 2018 HCI, and was based on similar extrapolations. This value should not be interpreted as a decline over time in

[6]Absent better data, for these provinces we assume that the ratio of GDP per capita to the national average is the same as the ratio of household income per capita to the national average. Since the household per capita income data across Chinese provinces is not adjusted for differences in purchasing power parity, the dispersion along the horizontal axis likely overstates what the dispersion would be if adjusted for differences in purchasing power parity.

[7] Naturally the extrapolations relied on in this note are highly tentative, as they are based on only three data points for PISA, and three data points for PIRLS. However, they are necessitated by the absence of published nationally-representative test score data for China. It is worth noting that a high-quality national level assessment with provincial-level results exists for China. Moreover, that assessment – known as NAEQ – is linked with PISA 2012. However, the results have yet to be published. The publication of such data could help validate these extrapolated scores and could be used in place of this extrapolation.

learning as measured by the HTS. Rather, this is an updated estimate of the overall level of the HTS based on the limited available internationally-comparable learning data for China based on the additional information available in the 2018 PISA assessment.

Table 1: Extrapolating PISA Scores by Subject

	Household Per Capita Disposable Income*		Actual and Imputed Test Scores				
	Yuan	Log	Math	Reading	Science	Overall	HLO
Shanghai (PISA 2012)	41919	10.64	613	570	580	588	608
B-S-J-G (PISA 2015)	33368	10.42	531	494	518	514	532
B-S-J-Z (PISA 2018)	42594	10.66	591	555	590	579	598
China National (Extrapolated to 2016)	23821	10.08	433	398	422	418	432
PISA-HLO Conversion Factor			1.05	1	1.05		

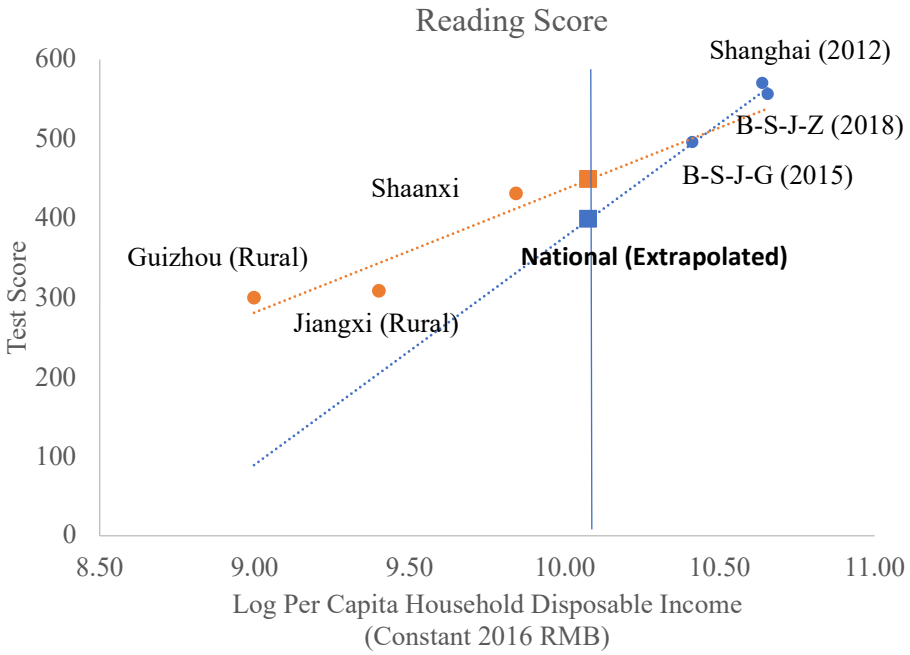
Note: measured in constant 2016 RMB in the indicated year. The PISA-HLO conversion factor for math and science is anchored to the TIMSS international assessment.

Table 2: Extrapolating PIRLS Scores

	Household Per Capita Disposable Income*		Actual and Imputed Test Scores	
	Yuan	Log	Reading	HLO
Shaanxi	18874	9.85	430	
Jiangxi (Rural Only)	12138	9.40	308	
Guizhou (Rural Only)	8090	9.00	300	
China National (Extrapolated to 2016)	23821	10.08	449	449
PIRLS-HLO Conversion Factor			1	

Note: measured in constant 2016 RMB.

Figure 1: Comparing PISA and PIRLS Extrapolated Reading Scores



Note: Orange dots represent PIRLS, blue dots represent PISA.

Figure 2: China's Extrapolated National HLO In International Perspective