Numeracy, so what?

The relationship between numeracy and literacy

Ralph Hippe[1]

– Preliminary version, please do not cite –

Abstract

Recent research has increasingly employed numeracy to approximate human capital. In particular, the age heaping strategy allows to obtain numeracy values. These values are in many cases derived from census data. For this reason, this method can often be applied not only for larger time spans but also for smaller territorial units than other historical indicators of human capital. But is this approach complementary to other more standard proxies of human capital? To answer this, I compare numeracy with another major indicator of human capital: literacy. The results show that numeracy correlates fairly well with literacy. However, different patterns of this relationship exist, particularly depending on the level of ABCC and literacy values because both variables are both lower and upper bounded..

---

[1] *University of Strasbourg and University of Tuebingen, hippe@unistra.fr*

Introduction

There has been a recent surge in research on human capital in general, as portrayed by Unified Growth Theory (e.g., Galor 2005), and on measuring human capital adequately today and in history. Concerning historical evidence, approaches focusing on literacy or signatures rates have been complemented by other proxies, such as book production or numeracy (e.g., Baten and van Zanden 2008, Crayen and Baten 2010). In particular, numeracy has been approximated by the age heaping method in a range of recent studies. This method allows to obtain valuable information on basic levels of human capital which were characteristic for mot historical societies. In this way, it is also possible to calculate earlier estimates on human capital than other methods. But to trace back the history of human capital in the long-run, it appears crucial to connect or to make the link between human capital indicators to assess the long-term implications of human capital on the economy and the society. However, evidence on the relationship between numeracy and other human capital proxies has been restricted to a few studies mostly at the national level or at the regional level for one particular country (e.g. A'Hearn *et al.* 2009, Crayen and Baten 2010).

For this reason, in this paper I use part of a recently constructed large data set on numeracy, which covers most of the European regions in the 19th century, to advance the understanding of the relationship between literacy and numeracy. This dataset is complemented by additional regional data for developing countries outside Europe in the 20th century. By using these data I am able to compare numeracy and literacy data at a regional level in different parts of the world at different points in time. This enriches the existing literature in both space and time dimensions.

The paper is organised as follows: first, I review the literature on literacy and numeracy and the historical educational context. Then, I present the data used and the methodology to compare the different proxies of human capital. In particular, the age heaping method is used to approximate numeracy and the ability to ‘read and write’ measures literacy. The results indicate that there is a high correlation between numeracy and literacy indicators. However, this relationship becomes less evident as one indicator approaches its lower or upper bound. A conclusion sums up the results of the paper.

Literacy

There is a broad range of human capital proxies which are used today. Woessmann (2003) lists some of them. He names variables such as education-augmented labour input, adult literacy rates, school enrolment ratios, level of education attainment and average years of schooling. Still, not all of them can be used in an historical perspective since data are often not sufficiently available. Therefore, the most important historical proxies for human capital in Europe are literacy (taking the form of the ability to ‘read and write’ and signature rates), and, more recently, numeracy. Other indicators include book production, school enrolment ratios, the number of schools or the number of teachers.

In the next sections, I review first some of the literature on literacy, before taking a closer look at the relationship between literacy and numeracy. Finally, numeracy itself and the broader context are highlighted, thus allowing a broader understanding of the issues at stake.

The capacity of writing has a very long tradition, even though it was only open to the elites of society during most of history. Still in 1750, more than 90 % of the worldwide population was not able to write and did not have access to institutions teaching it (Cipolla 1969). Therefore, the broad majority of the population was excluded from literacy. However, literacy is very important since it is a “tool for enabling individuals and social groups to extend their understanding of themselves and their world” (Vincent 2000, p. 24). In this way, individuals are more receptive to new ideas and adapt themselves faster to the changing demands of their work and their environment. This is a crucial point because the advances in technology necessitated this flexibility in different areas of the job market throughout history.

In general, literacy is a very popular measurement method of human capital. However, interest in its historical development had languished for a long time. Pioneering works by Cipolla (1969) and Stone (1969) initiated this research field. Since then, research has been conducted much more widely on historical literacy (Graff 2009). As Graff (1991) points out, one can classify research on historical literacy during the last decades in three broad generations. The first generation was constituted by research during the end of the 1960s, in particular by Schofield (1968), Cipolla (1969) and Stone (1969). Nevertheless, already some works in the 1950s were predecessors of this research line (Webb 1955, Fleury and Valmary 1957). The first generation created a foundation for upcoming studies by stressing and demonstrating the importance of literacy. Furthermore, it indicated future research possibilities with respect to more extensive numerical sources and to broader research themes. Based on these results, a second generation began its work in exploiting even more detailed quantitative data and distinguishing historical patterns of literacy. Finally, the third generation has been, among other things, opting for more interdisciplinary research between different fields to advance the knowledge on literacy. In this context, the study of literacy combined with the one of numeracy appears to be a logical and valuable addition to the existing literature.

Considering data availability, literacy rates in the form of reading (and writing) ability are available for most European countries only from the middle of the 19th century onwards. Accordingly, Cipolla states that “for the periods preceding the second half of the nineteenth century the information [on illiteracy] is very poor” (Cipolla 1969, p. 15). Therefore, many studies on early literacy developments use signature rates of conscripts or newly married couples (e.g., Schofield 1981, Mitch 1993, Reis 2005).[2] In the case of the latter, marriage contracts or other official documents had to be signed by the eligible person. However, this person was not always able to sign the contract. For this reason, taking the share of people who were able to sign with respect to the entire population might be employed as an indicator of literacy.

The idea of taking signature rates of marriage contracts is nothing new. For instance, this proxy was already employed in *Statistique générale de la France* from 1854 onwards (Furet and Ozouf 1977). Interest in reconstructing educational levels is nothing new either. In 1877 Louis Maggiolo, ancient rector of Nancy Academy, began his work in recollecting marriage signatures all over France for the years 1686 to 1690, 1786 to 1790, 1812 to 1816 and 1872 to 1876.

Still, there are potential biases and disadvantages of this method. For example, did the individual sign the corresponding document or was it done by another person, such as the bride or the priest? In France, new married couples were obliged by law to sign their marriage contract since 1647, before it was often the priest who signed them. Yet this example may not be confounded with other countries less stringent on marriage signatures. For instance, it was not obligatory for a couple to sign the registers in Italy (Cipolla 1969). Moreover, even when laws were passed, their application was often a different matter, with biasing effect on the data. The degree of spatial and time coverage is thus very different from country to country in Europe.

More globally, when measuring literacy a range of methodological and conceptual issues have to be addressed. What is illiteracy? There are several possible answers which make it not always easy to define literacy in a unique way. One can define a literate person by his capacity to read and write and an illiterate one by the lack of this capacity. However, there

[2] Other literacy measures include book-ownership, books borrowed or direct tests. Sources for literacy data include wills, petitions, criminal records, applications, depositions, inventories, deeds and others. For an overview see Graff (1991).

is still a group of 'semi-illiterates' (Cipolla 1969) who can read but not write. In some censuses of the 19th century data were collected explicitly for those semi-illiterates, in others they were not. Another issue that arises is the question of the quality of reading and writing. More specifically, people may be able to read but not necessarily capable of understanding the content. On the other hand, a person who is able to sign may not be able to write anything else than the name he has been trained to write in some way in order to fulfil certain minimum requirements for contracts.

Literacy and numeracy

Compared to the literature on literacy, research on numeracy is still in its infancy (see e.g., Thomas 1987, Emigh 2002, Netz 2002). Why is there such a gap between these two research fields? A major problem for researchers has long been the quantification of numeracy (A'Hearn *et al.* 2009). A coherent measure was lacking because few statistics were collected on numeracy in the past (Thomas 1987, Vincent 2000). Some even believed that it was impossible to construct one (Cohen 1982). As we will see later on, this has been overcome by the age heaping strategy (A'Hearn *et al.* 2009).

Nevertheless, research on both indicators appears to be helpful to understand the development of human capital due to several reasons. First, anecdotal evidence already suggests that there is a relationship between the two proxies. For example, young children who are good in literacy (here: reading) are often also high performing in numeracy (arithmetic) (Bulcock and Beebe 1981).

Second, literacy and numeracy have been closely intertwined throughout history. Accordingly, Netz (2002, p. 323) points out that "there is no difference between the history of numeracy and the history of literacy". In ancient cultures, the use of numerical symbols paved the way for verbal symbols. For instance, by analysing the emergence of writing in

Mesopotamia, Schmandt-Besserat (1992) comes to the conclusion that "in early cultures, numeracy drives literacy rather than the other way around" (Netz 2002, p. 323).

An example of the linkage of literacy and numeracy are also the 'Arabic numerals' widely used around the world today. In fact, these Arabic numerals should rather be called 'Indian' numerals since they were invented in India. Their first recorded images on pillars date from around 250 BCE (Woods and Woods 2000). Subsequently, they were passed on to other neighbouring cultures which traded with India. In this way, peoples from the Middle East adopted and adapted this numerical system. It took quite some time for its breakthrough in Europe, though. Only beginning from 976 CE onwards it became known in Europe because Europeans traded with Middle Easterners (Woods and Woods 2000). As Europeans did not know the origins of the numerals, they named them 'Arabic numerals'. Later on, the invention of the printing press led to a process of standardisation of the numerals and increased the acceptance among Europeans. Finally, Europeans brought this system to other parts of the world by means of their trades and conquests.

In consequence, Arabic numerals eventually replaced the old Roman system in Europe. They did so because they have several advantages. First, Roman numerals do not allow calculation as easy as Arabic ones and large sums cannot be expressed in a short manner. Second, there is no means to obtain fractions and a symbol representing zero does not exist. Third, it is possible to record transactions, in contrast to calculations which had to be done with counters such as an abacus (Thomas 1987). This, however, is a key feature of Arabic numerals. It combines numerical and verbal practices so that arithmetic can be put down by the use of paper and pen, and not by counters. In this way, both practices are linked to each other and their histories are intertwined.

Numeracy

Numeracy (and particular arithmetic) may be influenced by different additional factors, most importantly education, state bureaucracy and capitalism (Emigh 2002). Population statistics such as censuses and tax assessments carried out by the state may oblige individuals to keep records and to correspond to the described requirements. Nevertheless, Emigh (2002) argues that the chain of causation runs the other way round because states reacted to a rise in numeracy by collecting more thorough data. There is no point in collecting data if the individuals do not have the necessary knowledge to provide it. Otherwise the role and the power of the state on the population would be overemphasised (Tilly 1999).

Moreover, market capitalism furthered numeracy due to the necessity to keep records and to be able to calculate. Accordingly, arithmetic was also perceived as being at the core of trade (Hodder 1671). Tradesmen, state bureaucrats and plenty of other occupations needed arithmetic (Thomas 1987). Nevertheless, numeracy was also important in local economies without market capitalism (Emigh 2002). Therefore, numeracy already played an important role in everyday life before the Industrial Revolution. As the Catasto of 1427 in Florence highlights, the ability to work with numbers was needed for transactions such as property sales, payment for whatever service, testaments or dowries (A'Hearn *et al.* 2009).

By contrast, calculating was not accessible to and not perceived to be needed by everyone. Taking the example of early modern England, most grammar schools did not teach arithmetic before 1660 and afterwards only in form of an extra. As a result, still during the last decades of the 17th century, "fewer than four hundred men could be said to be mathematically minded" (Cohen 1982). The focus lay on literary classes at school, also because mathematics was still disregarded in society. The attitude by the public was rather negative, mathematics being perceived as an anti-social object (Thomas 1987). Moreover, arithmetic was considered to be only important for certain occupations and not for the general training and education of everyone. These occupations were to be found in commerce and trade. By contrast, higher society despised such occupations to be beneath its status. On the other hand, the bottom of the society had no access to educational facilities. Only at the end of the 17th century did this slowly change and arithmetic was included in the curricula of more grammar schools. But this was not always the case. For example, arithmetic became obligatory at Eton only in 1851 (Houston 2002).

The importance for improved navigation and rising overseas trade helped to spur this formation of arithmetic skills alongside commercial developments. In addition, reformers favoured the inclusion of arithmetic because "it disciplined the mind, encouraged inductive thought, and developed habits of precision, attention to detail, and a love of factual knowledge" (Houston 2002, p. 164). Finally, as Thomas (1987) resumes the research on this point, this development led to the acknowledgement of mathematics being a fundamental part of the education of a gentleman in England in the 18th century. However, the change in attitude towards a gentleman's education was not equivalent to the one of a lady. Mathematics was not deemed to be appropriate for women. Accordingly, it was mostly not taught to girls. Not surprisingly, women are mostly less numerate than men in many historical numeracy studies in Europe and elsewhere (e.g., Manzel and Baten 2009).

Other factors influencing literacy and numeracy

The evolution and relationship of literacy and numeracy has also to be seen in a broader social and economic environment. Because this is a too large domain, it is more appropriate to focus on a few important issues in this context.

During the first phase of the Industrial Revolution in England, the basic school system was not substantially contributing to economic growth. By contrast, economic growth rather accelerated despite the education system. Only during the further advancement of the economy became education increasingly important because some basic knowledge in different disciplines was required for more and more occupations. This need for skilled personnel after mid-19th century finally evoked a surge for educational facilities and education in general. Education became much more widely available than before.

Social developments encouraged this spread of education. For example, illiteracy became to be deemed to be a national disgrace (Cipolla 1969). Moreover, education was perceived as a requirement and a manifestation of the state's authority (Green 1990, Vincent 2000). More schools were set up and thus school attendance rose. These facts illustrate the new importance given by the state to the education of the people. The state became the principal promoter of schooling and replaced in some ways the church in enforcing education, even though the latter still played an important role. In fact, no institution other than the church had the necessary facilities, the means and experience of managing the bureaucratic requirements in order to provide large-scale teaching. Therefore, as Vincent (2000, p. 7) puts it, "the capacity of […] structures of power [other than the state], the family, private philanthropy, the market place and the church, to realize this objective was called into question". In consequence, the church joined more vigorously the efforts of the state later on.

Nevertheless, even if the church and the state endeavoured to improve the quantity and the quality of education, parents still had a preeminent influence on the education of their children. Until the final compulsion by the state (and afterwards), there were at the beginning very convincing reasons not to let one's children go to school. In addition to fees, clothing and stationary, parents had to do without the contributions that the work of the children made to the income of the family. Moreover, the training given by schools was not always regarded as important for the future work of their children either. For this reason, authorities aimed at meeting the preferences and demands of parents in order to increase the effective enforcement of schooling laws. For example, they adapted the school calendar more to the needs of the parents. Still, school attendance was not granted itself by a law.

Therefore, “across Europe as a whole, formal schooling cannot be taken as a proxy for the distribution of reading and writing skills” (Vincent 2000, p. 57; Houston 1985). For this reason, direct literacy values as stated by individuals during censuses appear to be better suited for this purpose. This underlines once more the meaningfulness to use literacy data (to be able to ‘read and write’) in the present study.

Data

I analyse numeracy and literacy by using historical census data from Europe in the 19th century and more recent data from other continents during the middle to the end of the 20th century. Almost all data stem from official census publications.[3] Historical numeracy estimates are derived from the database by Hippe and Baten (2011). This database covers almost all European countries at the regional level in the 19th century. Literacy data have been added to this source, mostly taken from the same official publications.

Second, I use the Integrated Public Use Microdata Series (IPUMS) database which includes microdata for many countries in the world. Clearly, age heaping and illiteracy have still to be sufficiently present in these censuses to compare both human capital indicators. Crayen and Baten (2010) and Hippe and Baten (2011) have already demonstrated that age heaping was already low in many of today’s industrialised countries at the country and at the regional level at the end of the 19th century. For this reason, I use available IPUMS data from developing countries in the Americas, Africa and Asia for the second part of the 20th century. Because literacy and numeracy are both upper bounded[4], I always use the earliest census data to avoid biases as best as possible. By using these criteria, I selected data from countries such as Mexico, India and Kenya where both age heaping and illiteracy were still observable. Tables 1 and 2 give additional information on both datasets.

[3] All data except the data from India. These are derived from an employment survey (see Table 2).
[4] And also lower bounded but this is not an issue concerning the data used here.

Methodology

What are the advantages of using regional data instead of national ones? Cipolla (1969) explicitly points out that *intra*national differences in human capital can be equally important as *inter*national ones. The importance of interregional variation is confirmed for many countries in the 19th century by Hippe and Baten (2011). The data of that paper suggest that regional variation appears to be quite significant in numeracy, allowing a closer comparison of these regional numeracy data with literacy data.

Taking regions as the standard unit of analysis can considerably improve and complement the existing results obtained at the national level. This reasoning has also recently been promoted by other economic theories as, for example, by New Economic Geography and its Nobel prize winning initiator, Paul Krugman. The regional perspective has the advantage that inherent cross-country differences do not bias the results as might be the case when performing pure country comparisons. This is particularly helpful for literacy comparisons because literacy was not always measured in the same way in each country and other time. In contrast, numeracy values are directly derived from total census outcomes and do not rely on changing definitions throughout history. This appears to be an important advantage of numeracy for means of comparisons. By contrast, early signature rates typically rely on rather small samples and do not cover the whole population but only a specified fraction such as married people at different ages, military recruits, etc. The majority of literacy data in this study are derived from the ability to 'read and write' as stated by all individuals in censuses and not only specific categories of the population. Therefore, it is a more modern and more complex interpretation of literacy and is still used in developing countries today.

Because the majority of age heaping studies are still rather recent (e. g., A'Hearn *et al.* 2009, Manzel and Baten 2009), it appears necessary to explain the underlying methodology. In general, the age heaping method takes advantage of the fact that in many historical official documents the ages of the concerned group of people are listed. This is notably the case in population censuses but also other material can be used.

More concretely, individuals were asked their ages by a census taker. These statements are available either in individual or aggregate form. However, individuals did often not know their exact age. This is why a heaping is discernible on certain ages, i.e. the so-called age heaping. For example, an individual was 33 years old but told the census taker that he was 35. This means that there are clear rounding effects because the individual was not able to count correctly. In consequence, a heaping on '0' and '5' is visible in many historical cases.

Evidently, reasons other than human capital might be attributable to this age heaping effect. These include the role of administrations in public affairs and when conducting the census and false age declarations on purpose. However, Crayen and Baten (2010) have shown that the influence of human capital is the most important factor. Moreover, earlier studies on this numeracy proxy have shown a high correlation between numeracy and literacy on a national and in some cases at the regional level (Crayen and Baten 2010, Hippe and Baten 2011). These findings underline the significance of the human capital effect. The inherent characteristics of this method are also quite advantageous. Age statements are available for (almost) all time periods during the last couple of centuries and beyond. The long-term measurement of human capital thus becomes possible. Moreover, these data are often quite more spatially available, meaning that the analysis can be brought to smaller spatial units than before.

Nonetheless, it is clear that the age heaping method only captures very basic numerical abilities. In this sense, it is a proxy for very basic human capital values. These, however, persisted in today's industrialised countries until the 19th century and in many developing countries until today.

In comparison, literacy as measured by the ability to 'read and write' concerns already higher human capital levels. A more basic indicator would be the ability to 'read only'. This indicator is given only in few historical censuses, the majority preferring both capacities, i.e. to read *and* to write. Exemptions prove the rule, as one of the early leaders in literacy, Sweden, was actually a leader in *reading*, not necessarily in *writing*. The aim was to enable the believers to read the bible. Writing, in contrast, was not strictly relevant for this purpose. Still, the higher competencies required for reading and/or writing generally lead to general the observation that numeracy values as measured by age heaping are lower than literacy values. This is important to know in the further analysis of the data.

In addition, both literacy and numeracy proxies used in this study rely on output values, i.e. they measure the performance of individuals in given tasks such as reading. In contrast, other human capital measures such as enrolment rates, the number of schools or teachers are input indicators. They describe, for example, how many children go to school but it is apparent that the attendance of school is more beneficial for some students than for others. The learning success is not equal among all individuals. In contrast, the human capital proxies analysed here measure the actual acquisition of some basic component of human capital. This makes them better comparable than using input proxies.

The calculation of the ABCC index, which is able to capture this age heaping behaviour of individuals, is as follows. The most intuitive way is to begin with the Whipple Index (which is also used by the United Nations). It is defined as

$$WI = \frac{\sum_{i=5}^{14} n_{5i}}{\frac{1}{5}\sum_{i=23}^{72} n_i} \times 100,$$

where i stands for years of age and n for the number of observations. Values range from 100 to 500. A value of 100 means that age heaping is not present and 500 all age observations end on ‘0’ and ‘5’. Because this range is not very intuitive, A’Hearn *et al.* (2009) propose a new index, the ABCC Index, which is a linear transformation of the Whipple Index.

$$ABCC = \left(1 - \frac{WI - 100}{400}\right) \times 100$$

The ABCC Index has the advantage to be handier than the Whipple Index. Here, values range from 0 to 100, where 100 is the maximum numeracy level and 0 the lowest. This makes the analysis much easier because literacy rates are commonly defined by the same value range. Thus, the ABCC Index achieves a higher comparability between literacy and numeracy values than the Whipple Index. Therefore, the following analyses are performed by using the ABCC Index.

Ideally, literacy should be defined in this study as

$$Literacy = \frac{\sum_{i=23}^{72} rw_i}{\sum_{i=23}^{72} n_i} \times 100,$$

where *rw* stands for the number of individuals able to ‘read and write’[5]. This would give the same age range as for ABCCs and would arguably allow a maximum of comparability. Unfortunately, this perfect standardisation is not always possible because literacy is often defined as a share of individuals able to read and write above a certain age, often the age of 7 years. Therefore, the ABCCs are compared with the available literacy definition in each case. Details on the literacy definition of each country can be seen in Tables 1 and 2. Moreover, Table 3 presents some descriptive statistics for the historical European dataset.

[5] Including if a person is only able to ‘read’.

Results

Figures 1 to 7 present scatter plots for the relationship between ABCCs and literacy at the regional level for the different countries under study in 19th century Europe.

Literacy and numeracy appear to be well correlated at the regional level in these countries. However, there are some apparent outliers. These outliers are mostly highly urbanised areas which are characterised by (far) higher literacy rates than other regions. This is notably the case in Greece (Attica, the greater Athens region; Figure 2), in Russia (e.g. the largest cities Saint Petersburg and Moscow; Figure 6) and Serbia (the largest cities Beograd and Nis; Figure 7). But also in Italy (Figure 5) the regions of two of the three largest cities, Naples and Rome, are more advanced in literacy than their rather average numeracy level would suggest (between an ABCC of 80 to 85). A possible reason might be that bureaucracy was important in the capital and that literacy was thus very important. Moreover, access to schools was much higher and easier in these urbanised regions and the focus was possibly still more on reading and writing ability than on numerical capacity.

In addition, the relationship between the ABCCs and literacy is not always the same. When the ABCC values are already quite high, literacy rates range a lot even though the range of ABCCs is quite limited (e.g. in Spain). This changes when the ABCCs are on a lower level. A similar observation can be made when literacy rates are fairly low. In this case, literacy rates do not vary a lot but ABCCs do, as for example in Serbia.[6] What might cause these differences in the slope? Clearly, ABCCs and literacy rates are bounded by 0 and 100 (%). This means that as ABCCs get closer to the upper limit, literacy rates are wider ranging and the relationship becomes less important. Therefore, when the mean values of the ABCC are higher, the slope is less high. This is a common issue for bounded variables and so for the two variables under analysis too.

[6] Not taking account here of the apparent urban outliers.

Figure 8 displays the relationship in some selected developing countries where the ABCC did not yet attain its upper bound at the time the censuses were taken. Similar tendencies appear as in the former dataset. However, the literacy values are generally not as low as in some European countries in the 19th century. Correspondingly, the slope is less steep. More concretely, Chile has on average the highest ABCC and literacy. Bolivia and Kenya are characterised by the highest ranges in literacy. In contrast, in India the deviation from the mean ABCC value is the most important and the scatter points are more dispersed.[7] Still, the relationship also appears to hold in these cases.

However, a clearer econometric analysis has still to be advanced in more detailed research to specify more concretely the foundations of this relationship.

Conclusion

This paper has analysed the relationship of numeracy and literacy. Numeracy and literacy have been historically intertwined so that, at least theoretically, a link between the two variables should be at hand.

This link was further investigated by using historical census data from Europe in the 19th century and more recent data from developing countries outside Europe in the 20th century. As the literacy proxy was (mostly) used the ability to 'read and write' and as the numeracy pendant the ABCC Index, a linearly transformed Whipple Index with the same value range as the literacy proxy. This allowed a better comparability between the two human capital proxies. For the same reason, similar age ranges of the individuals were used, in most cases individuals between the ages of 23 to 72 years.

The results show that numeracy and literacy indicators are well correlated both historically and in more recent data. The slopes become less high when ABCCs come closer to its maximum value. On the other hand, very low values of the ability to 'read and write'

[7] The fact that the data come from an employement survey might be related to this.

lead to a much more important range in ABCC values. Thus, in this dataset the problem of an upper-bound is crucial for the ABCCs and basic numerical capacities are not able to be traced any further, whereas the lower bound inhibits to gain more information in the case of literacy abilities.

Moreover, the historical data of several European countries show some outlier regions. These are (mostly) attributable to comparatively highly urbanised areas where literacy values are often much higher than numeracy ones. This characteristic should be taken into account when analysing in more detail the relationship between several human capital proxies at subnational level. Still, both indicators of literacy and numeracy appear to be well linked to each other. Clearly, neither the ABCC nor literacy may appropriately capture human capital in more advanced countries today. But particularly for historical purposes these may deliver valuable information on human capital formation in Europe and elsewhere.

References

A'Hearn, B., Crayen, D. and J. Baten (2009). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital, *Journal of Economic History* 68 (3): 783-808

Baten, J. and J. L. van Zanden (2008). Book production and the onset of modern economic growth, *Journal of Economic Growth* 13: 217-235

Bulcock, J. W. and M. J. Beebe (1981). Some Common Causes of Literacy and Numeracy, *Canadian Journal of Education* 6 (3): 19-44

Cipolla, C. M. (1969). Literacy and development in the West, Baltimore: Penguin Books

Cohen, P. C. (1982). A calculating People: The Spread of Numeracy in Early America, Chicago: University of Chicago Press

Crayen, D. and J. Baten (2010). Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth, *Explorations in Economic History* 47: 82-99

Emigh, R. J. (2002). Numeracy or Enumeration? The uses of numbers by states and societies, Social Science History 26 (4): 653-698

Fleury, M. and P. Valmary (1957). Les Progrès de l'instruction élémentaire de Louis XIV à Napoléon III d'après l'enquête de Louis Maggiolo (1877-1879), Paris: INED

Graff, H. J (1991). The Literacy myth. Cultural Integration and Social Structure in the Nineteenth Century, New Brunswick: Transaction Publishers

Graff, H. J. (2009). Themes and Questions, in: Graff, H. J., Mackinnon, A., Sandin B. and I. Winchester (eds.). Understanding literacy in its historical contexts: Social cultural-history and the Legacy of Egil Johansson, Lund: Nordic Academic Press

Green, A. (1990). Education and State Formation: The Rise of Education Systems in England, France and the USA, Basingstoke: Palgrave

Hippe, R. and J. Baten (2011). Regional Inequality in Human Capital Formation in Europe, 1790 – 1880, AFC Working Papers Nr. 2011-7

Hodder, J. (1671). Hodder's Decimal Arithmetick, sig. A4

Houston, R. A. (1985). Scottish Literacy and the Scottish Identity: Illiteracy and Society in Scotland and Northern England, 1600-1800, Cambridge: Cambridge University Press

Houston, R. A. (2002). Literacy in Early Modern Europe, Harlow: Pearson Education Limited

Manzel, K. and J. Baten (2009). Gender Equality and Inequality in Numeracy: The Case of Latin America and the Caribbean, 1880-1949, *Revista de Historia Económica* 27 (1): 37-73

Mitch, D. (1993). The Role of Human Capital in the First Industrial Revolution, In: Mokyr, J. (ed), *The British Industrial Revolution: An economic perspective,* Boulder: Westview Press, 267-307

Netz, R. (2002): Counter culture: Towards a history of Greek numeracy, *History of Science* 40: 321-352

Ozouf, J. and F. Furet (1977). Lire et écire. L'alphabétisation des Francais de Calvin à Jules Ferry, Les editions de minuit.

Reis, J. (2005). Economic Growth, Human Capital Formation and Consumption in Western Europe before 1800, In: Allen, R. C., Bengtsson, T. and M. Dribe (eds.), *Living Standards in the Past*, Oxford: Oxford University Press, 195-225

Schmandt-Besserat (1992): Before writing: From counting to cuneiform, Austin: University of Texas Press

Schofield, R. S. (1981). Dimensions of illiteracy in England 1750-1850. In: Graff, H. J. (ed.), *Literacy and social development in the west: a reader*, Cambridge: Cambridge University, 201-213

Stone, L. (1969). Literacy and Education in England, 1640-1900, *Past and Present* 2: 69-169

Thomas, K. (1987). Numeracy in Early Modern England, The Prothero Lecture, *Transactions of the Royal Historical Society* 37: 103-132

Tilly, C. (1999). Power – top down and bottom up, *Journal of Political Philosophy* 7: 330-352

Vincent, D. (2000). The Rise of Mass Literacy, Cambridge: Polity Press

Webb, R. K. (1955). The British Working Class Reader: 1790-1848; literacy and social tension, London: Allen & Unwin

Woessmann, L. (2003). Specifying Human Capital, Journal of Economic Surveys 17 (3): 239-270

Woods, M. and Woods, M. B. (2000): Ancient computing: from counting to calendars, Minneapolis: Runestone Press

Appendix

Data on numeracy and literacy in Europe: see Hippe and Baten (2011)

Other data on numeracy and literacy: Minnesota Population Center (2011). Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]. Minneapolis: University of Minnesota

Table 1 Details on historical European data

Country	Code	Census	Sexes	ABCC definition	Literacy definition
Greece	GR	1907	Both	23-32 yrs.	23-32 yrs.
Hungary	HU	1869	Both	23-72 yrs.	7+ yrs.
Ireland	IE	1841	Both	23-72 yrs.	26-75 yrs.
Italy	IT	1871	Both	23-72 yrs.	23-72 yrs.
Russia	RU	1897	Both	23-72 yrs.	23-72 yrs.
Serbia	SR	1895	Both	23-72 yrs.	7+ yrs.
Spain	ES	1887	Males	23-72 yrs.	21-70 yrs.

Table 2 Details on data from developing countries

Country	Code	Census	Density	Sexes	ABCC definition	Literacy definition
Bolivia[1]	BO	1976	10 %	Both	23-72 yrs.	23-72 yrs.
Brazil	BR	1960	5 %	Both	23-72 yrs.	23-72 yrs.
Chile	CL	1960	1 %	Both	23-72 yrs.	23-72 yrs.
Colombia	CO	1964	2 %	Both	23-72 yrs.	23-72 yrs.
Ecuador	EC	1962	3 %	Both	23-72 yrs.	23-72 yrs.
India[2]	IN	1983	0.091 %	Both	23-72 yrs.	23-72 yrs.
Kenya	KE	1989	5 %	Both	23-72 yrs.	23-72 yrs.
Mexico	MX	1960	1.5 %	Both	23-72 yrs.	23-72 yrs.
Panama	PA	1960	5 %	Both	23-72 yrs.	23-72 yrs.
Tanzania	TZ	1988	10 %	Both	23-72 yrs.	23-72 yrs.

Note : [1] Excludes 11 states in the north; [2] Employment survey

Table 3 Descriptive statistics for ABCC and literacy in Europe

Code	Obs.	ABCC				Literacy			
		mean	*sd*	*min*	*max*	*mean*	*sd*	*min*	*max*
ES	50	93.34	4.70	83.90	100.00	0.53	0.17	0.27	0.84
GR	26	61.43	8.17	50.64	79.83	0.37	0.08	0.27	0.64
HU	18	89.97	3.29	82.23	94.73	0.49	0.09	0.31	0.63
IE	32	73.24	4.13	65.39	81.88	0.50	0.16	0.21	0.82
IT	69	88.62	6.93	72.62	99.01	0.29	0.15	0.11	0.65
RU	35	80.06	6.39	66.85	91.78	0.24	0.09	0.16	0.60
SR	18	60.08	9.30	41.77	82.75	0.17	0.14	0.06	0.61

Figure 1 ABCC and literacy in Spain[8]

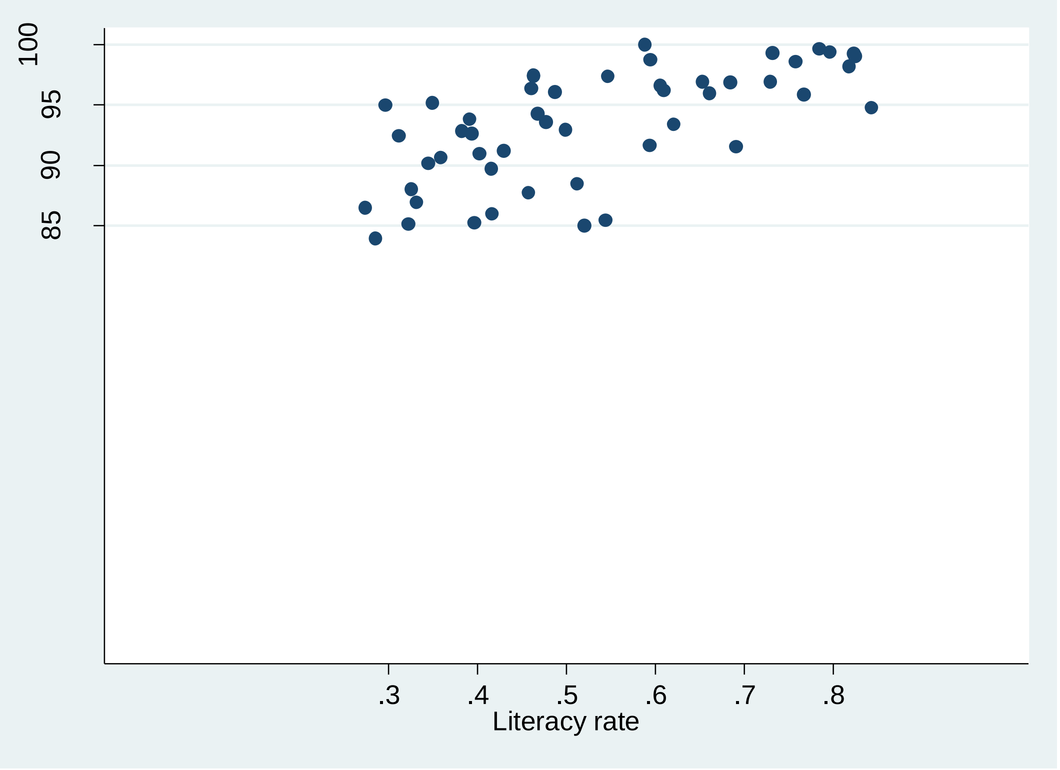
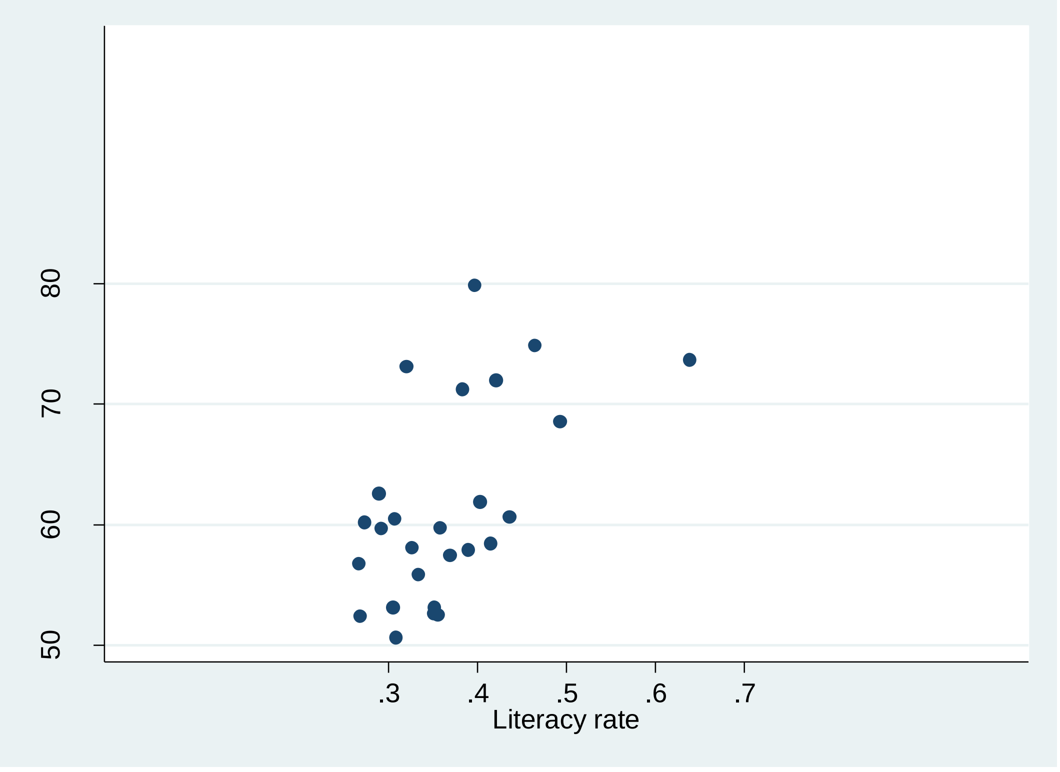


Figure 2 ABCC and literacy in Greece[9]



---

[8] Only data for the male population available.
[9] Only data for the ages 23 to 32 years is available.

Figure 3 ABCC and literacy in Hungary[10]

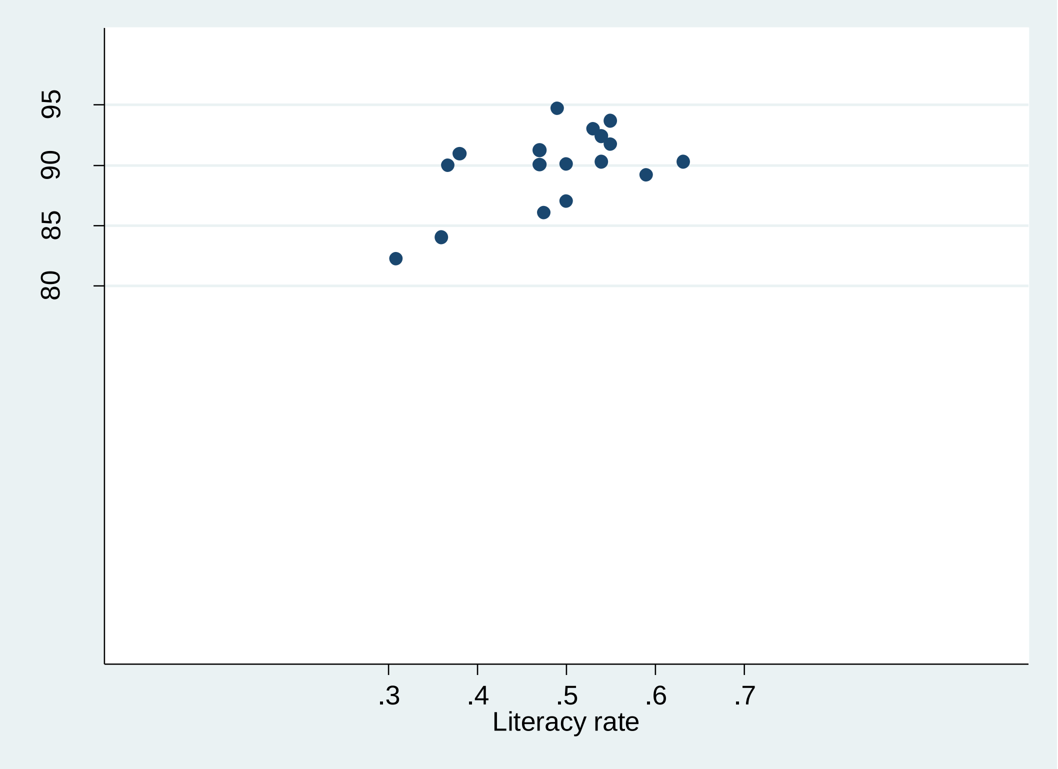
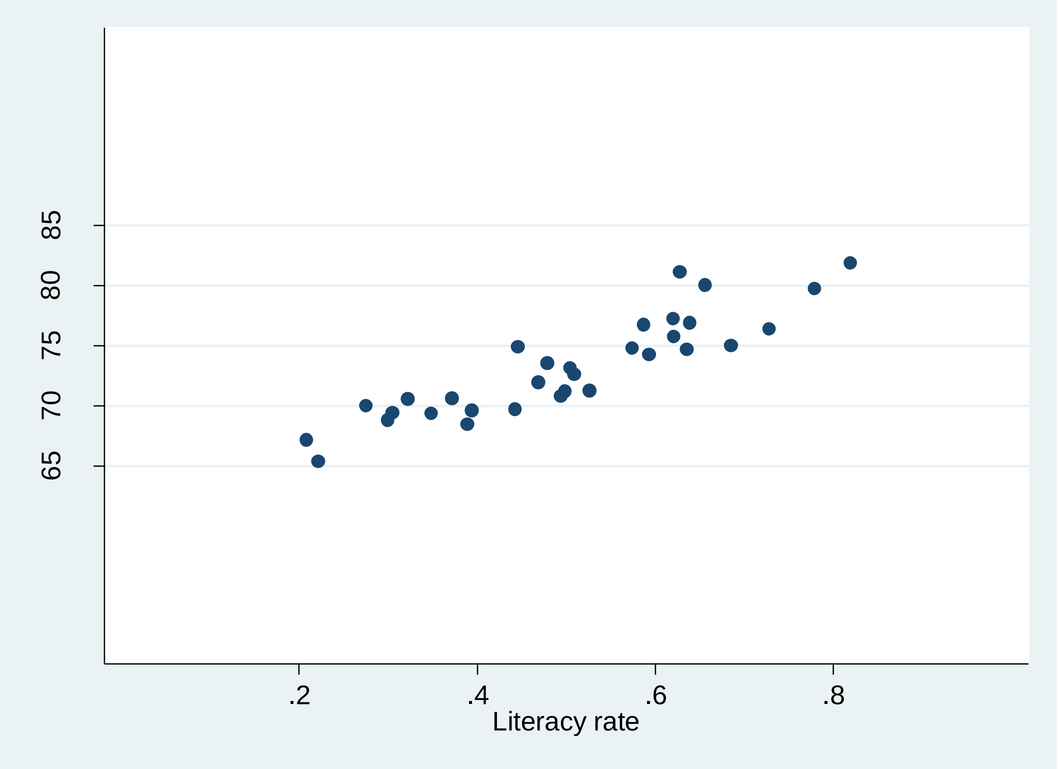


Figure 4 ABCC and literacy in Ireland



[10] Hungary within today's borders.

Figure 5 ABCC and literacy in Italy

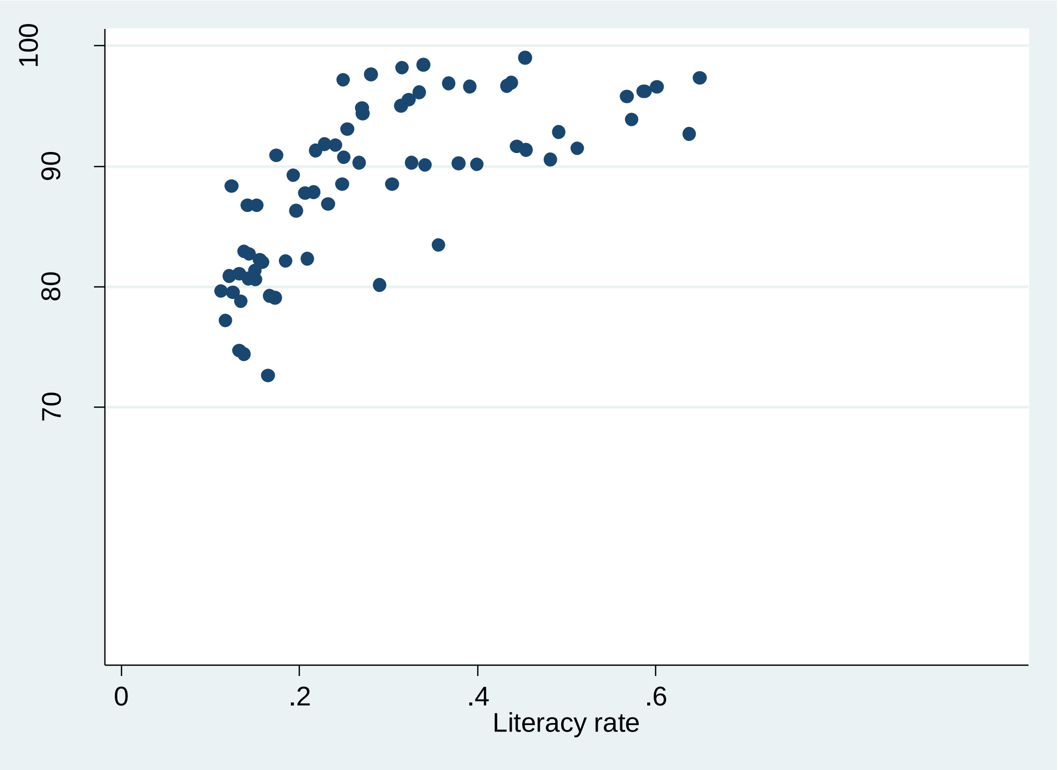
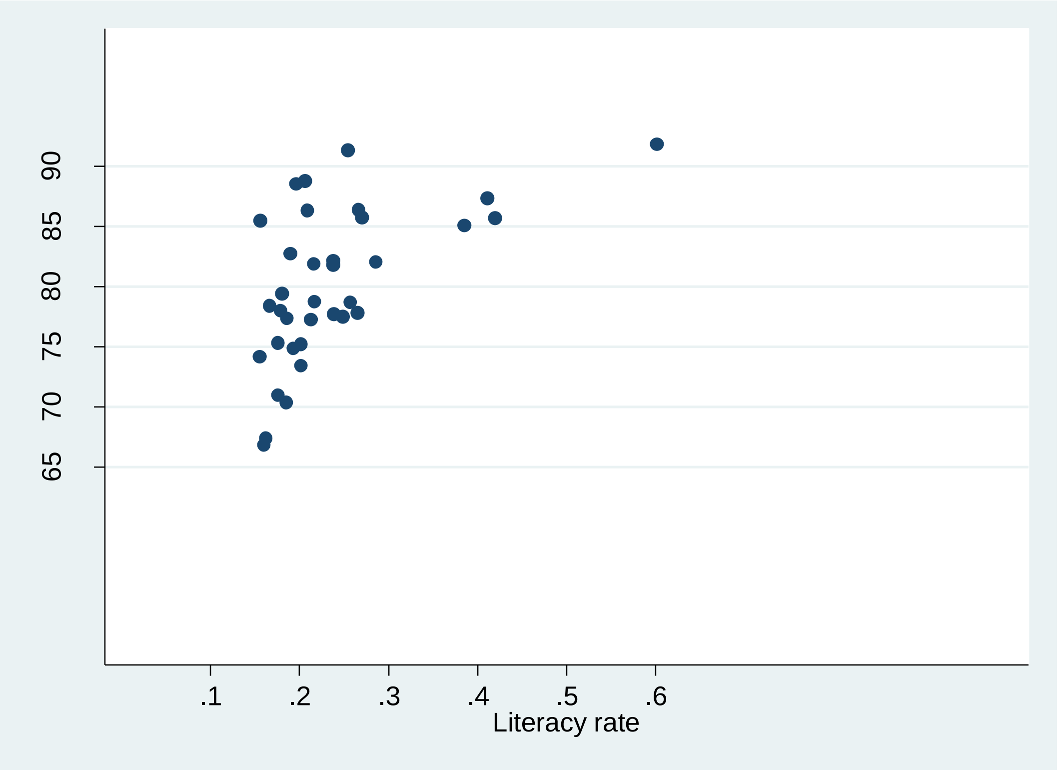


Figure 6 ABCC and literacy in Russia[11]



[11] Russia comprises Russia's European part in today's borders.

Figure 7 ABCC and literacy in Serbia

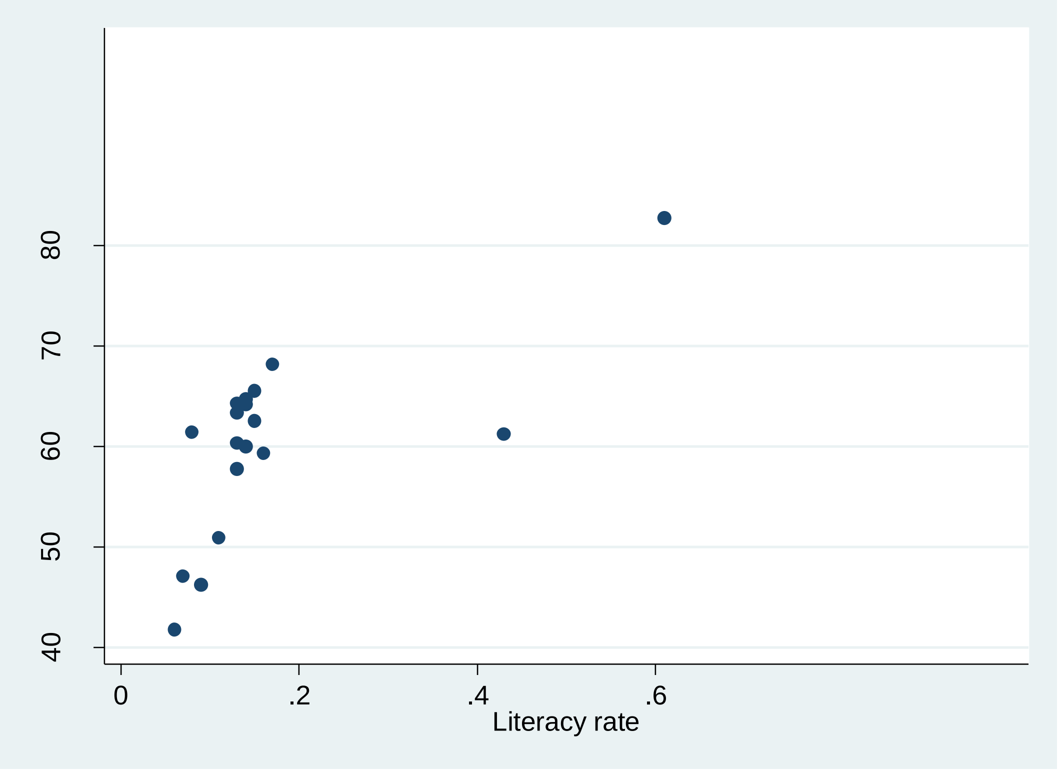


Figure 8 ABCC and literacy in developing countries

Note: bo = Bolivia, br = Brazil, cl = Chile, co = Colombia, ec = Ecuador, in = India, ke = Kenya, mx = Mexico, pa = Panama, pk = Pakistan, tz = Tanzania, ug = Uganda