QUANTITATIVE ANALYSES OF LONG-RUN HUMAN CAPITAL DEVELOPMENT:
AGE HEAPING AS AN INDICATOR FOR NUMERACY

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Wirtschaftswissenschaftlichen Fakultät
der Eberhard-Karls-Universität Tübingen

vorgelegt von

Dorothee Crayen
Marburg a.d. Lahn

2008

Dekan:	Professor Dr. Kerstin Pull
Erstkorrektor:	Professor Dr. Jörg Baten
Zweitkorrektor:	Professor Dr. Laszlo Goerke
Tag der mündlichen Prüfung:	10. März 2009

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor Joerg Baten. His dedication to research and to the mentoring of PhD students is unique, and he gave me invaluable guidance. He also made it possible for me to gain teaching experience, for which I am deeply grateful. Moreover, his encouragement to present my research to an international audience on various summer schools, workshops and conferences gave me important professional feedback and experience. I am particularly indebted to Laszlo Goerke. He has not only been one of the best teachers I have ever encountered, his unprejudiced approach to research is exemplary and made a deep impression on me. I am glad he was willing to act as my second adviser and I am thankful for his valuable comments to my research. I am greatly appreciative to Josef Molsberger for taking the chair of my disputation and for giving me valuable feedback.

I recognize, and am grateful for, the financial support of the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). Being part of the DFG Research Training Group at the Faculty of Economics and Business Administration at Tuebingen University broadened my knowledge and stimulated my research by creating an important platform for the presentation and discussion of research. I am also owing gratitude to the Marie Curie Research Training Network 'Unifying the European Experience: Historical Lessons of Pan-European Development' that granted me a pre-doctoral fellowship at the Universitat Pompeu Fabra in Barcelona. My paper with Joerg Baten and Hans-Joachim Voth originated during this research stay. Other parts of my dissertation have gained a lot from summer schools and workshops organized and funded by Globaleuronet, by the European Science Foundation, the Marie Curie Research Training Network, and by the Centre for Economic Policy Research.

Several other people have played a decisive role in my understanding of economic history, in the process of knowledge accumulation and in my ways of working. By virtue of his objectivity, diligence and kindness Brian A'Hearn acted as a role model. I enjoyed every discussion with him and his comments were always insightful. I learned a lot from Hans-Joachim Voth, who shared his ideas and experience and gave me inestimable advice. He has got a admirable intuition for asking the right questions and his creativity inspired my work. I would like to express my gratitude to all the participants of conferences, workshops and summer schools who contributed to my dissertation by their critical questions, their advice and by their ideas that supplied new input for my work. Special thanks go to Jan Luiten van Zanden, Robert Allen, Gregory Clark, David Mitch, Tim Leuning and Albert Carreras. The discussions with Peter Koudijs, Karine van der Beek and Kerstin Enflo were also invaluable to me. I thank Friedrich Breyer, Juergen Meckl and Manfred Stadler for being supportive from the beginning on, both personally and professionally. Their perpetual confidence in my person often amazed me and it means a lot to me.

I am grateful to my peers and colleagues of the Tuebingen research group for many conversations about my work, and help with the constant challenges I faced in completing this project. In particular, I wish to thank Normann Mueller, Kirsten Labuske, Mojgan Stegl, Kerstin Manzel, Dominic Behle, Aravinda Meera Guntupalli, Alexander Moradi, Nikola Köpke, Margaryta Korolenko, Sonja Rabus, Eva Rosenstock, Daniel Schwekendiek, Linda Twrdek, Luis Huergo and Dominik Ohly.

Mein ganz besonderer Dank gebührt meiner Familie, deren bedingungslose und immerwährende Unterstützung mir den Weg in die Forschung und Wissenschaft geebnet und mir die Kraft gegeben hat, meine Arbeiten auch über anstrengende Phasen hinweg weiter zu entwickeln und zum Abschluss zu bringen. Knut, ich kann Dir nicht genug danken für Dein Vertrauen, und den mentalen wie praktischen Beistand. Deine Ruhe und Geduld geben mir die nötige Sicherheit und Bodenhaftung, ohne die ich sonst nie solche Leistung bringen könnte. Ich danken meinen Eltern von Herzen, dass sie mir immer größte Freiheit gelassen, und mich stets bereitwillig und nach allen Kräften unterstützt haben. Missen möchte ich auch nicht die Unterstützung meiner Geschwister sowie von Bianca, Christoph, Regine, Horst, Iris, Paul, Ilka, Malte, Lilly, Kathrin, Nina, Tini und Tine.

TABLE OF CONTENTS

SYMBOLS AND ABBREVIATIONS

AH	Age heaping
CWR	Child-Women Ratio
DFG	Deutsche Forschungsgemeinschaft
DHS	Demographic and Health Surveys
DY	Demographic Yearbook
EQ	Equation
EU	Europe
FE	Fixed Effects estimation
IDB	International Database
IPUMS	Integrated Public Use Microdata Series
IV	Instrumental variable estimation
MEN	Middle East / North Africa
OLS	Ordinary Least Square estimation
RE	Random Effects estimation
UN	United Nations
US	United States
W, WI	Whipple Index
$\hat{W}$	transformed Whipple Index
WP	Working paper
WW2, WW II	World War 2

1 INTRODUCTION

> "Unlike literacy, for which signatures, though a poor measure, provide some baseline, there is no good measure of the level of numeracy."

Keith Thomas, 1987

I Motivation and general comments

Since the emergence of endogenous growth models in the late 1980s (e.g. Lucas 1988, Romer 1990) human capital has been recognized as one of the most essential factors determining the growth path of an economy.[1] Based on these theoretic models, cross-country and panel data regressions were conducted to assess the empiric impact of various measures of schooling on economic development and confirmed the stimulating role of human capital in the growth process (e.g. Barro and Sala-i-Martin 1995, Barro 2003, Drèze and Sen 1995, Levine and Renelt 1992, Levine and Zervos 1993, World Bank 1993).

For researchers studying the economic development of nations, the availability of human capital data is crucial. Fortunately, from the very start of their existence, the United Nations encouraged and supported their member states to collect data and to implement registration systems. By these means, the United Nations statistical division after 1945 was able to provide a broad and standardized data set for most countries. For

[1] Following Adam Smith (1776), I use the term human capital to refer broadly to skills, dexterity (physical, intellectual, psychological, etc) and discernment.

instance, the *Unesco Statistical Yearbooks* and the *U.N. Demographic Yearbooks* contain figures on total pupils, enrolment rates and current educational spending. The *World Tables* and *World Development Reports* compiled by the World Bank provide figures on literacy rates as well as primary and secondary school enrolment ratios. The *International Data Base (IDB)* made available by the U.S. Census Bureau contains literacy rates and other demographic and socio-economic data for 227 countries from 1950 to present. In addition to the quantitative indicators named above, there are also qualitative measures of schooling available, e.g., test scores, dropout rates and average teacher salaries.[2] The various human capital indicators enable researchers to determine approximations for a country's level of human capital and to analyze complex macroeconomic interrelations. The relevant data is obtained from registration systems, surveys or censuses at regional or country level. Thanks to the endeavours of the United Nations, there are only a few countries left nowadays which have not taken at least one census in the post-war period.

By contrast, economists making efforts to examine the impact of human resources on economic development before the end of World War II often face massive problems. Methods of collecting data, i.e., registration systems, population censuses or surveys did neither acquire sufficiently detailed data nor was the recorded data standardized in any consistent form. Early population counts were conducted primarily for reasons such as deriving the number of prospective warriors, to identifying the amount of resources available for military purposes and for estimating tax income. Therefore, little attention was paid to direct information about a population's level of education. Typically, population counts only report the number of inhabitants, possibly subdivided by sex. More often than not, those counts, undertaken to determine fiscal, labor, and military

[2] For a database including qualitative measures of human capital, see for example Barro and Lee (1996).

obligations did not count women, children or slaves, but were limited to heads of households, males of military age or taxpayers. Economic historians are thus often left without adequate material for their research on stocks and flows of human capital. This is particularly unfortunate in the light of the renewed interest, which the historical evolution of human capital has received by the recent development of very-long-run growth models. These models seek a unified explanation of pre-modern Malthusian dynamics and modern economic growth, typically assigning a key role to a fertility transition in which families switch from valuing quantity to quality in children, raising investment in human capital (Galor and Weil 2000, Cervellati and Sunde 2005).

Faced with these data constraints, economic historians commonly use signature ability rates to proxy human capital when schooling data dry up for today's industrial countries in the early nineteenth century and for today's developing countries in mid-twentieth century.[3] Signature ability rates, however, are subject to measurement errors (for example, few countries insisted on the bride and groom subscribing the register until the late eighteenth or nineteenth century so that often the priest rather than the couple signed the marriage registers, see Houston 2002), capture a very specific aspect of human capital only, and are limited to records requiring signatures, thus discriminating against less developed countries. The need to derive an alternative way of approximating the level of human capital is thus great in economic history.

In search for measurements of the stock of human capital for early time spans, more than any other data, demographic records are available. This thesis proposes to use irregularities in the reporting of age to estimate the people's level of education. Single

[3] Signature ability refers to the ability to sign one's name on marriage registers and legal documents (instead of making a cross or other mark).

year age data, which enables researchers to depict a population's detailed age structure, almost always exhibits irregularities in the form of heaped data, i.e., the age distribution does not run smoothly but exhibits sharp jumps and clustering at certain ages. This phenomenon is attributed to age heaping, a term which describes people's ignorance to their age or people's tendency to round their ages off. To a varying degree, age heaping exists in nearly all historical age statistics when people were asked for their age (as opposed to age records that were calculated from birth certificates or alike). By measuring the degree of age heaping one is able to derive a simple proxy for human capital covering a greater number of countries and regions as well as earlier periods than signature ability rates. Being a proxy of numerical comprehension and diligence when responding to age questions, age heaping also serves not necessarily as a substitute but as complementary human capital measure to signature ability rates to capture as many aspects of human abilities as possible. With respect to literacy, numeracy has been understudied in economic history and its inclusion in long-run growth analysis can provide new insights into the role and development of human capital.

While signature ability rates and age heaping levels generally correlate rather well, they can diverge for several reasons. For example, common tradesmen, farmers and craftsmen were often well-trained with regard to numbers and simple calculations due to their occupational profile, while they were less competent in spelling and writing. As a consequence, their signature ability rates (or more advanced literacy measures) would produce much lower estimates of their average skill level than age heaping. In this context, using a proxy for numerical knowledge adds substantial value to the analysis of historic human capital levels and trends. Moreover, literacy measures are sometimes discriminated against certain languages. In India, for example, people who were able to read and write the Latin alphabet were considered literate by the colonial government whereas Indians whose writing skills were limited to Sanskrit or other local alphabets were often deemed illiterate by officials. Because these local languages were often taught in monasteries or other religious schools together with basic mathematical skills, in this case the age heaping proxy captures educational conditions more comprehensively than literacy rates do.

In short, though age heaping is a very basic indicator for human capital, and though it faces more limitations and it is more inaccurate than modern standardized measures such as enrolment or dropout rates, it helps to overcome the scarcity of data in economic history, which generates a need to extract new information from old data. It is the purpose of this thesis to offer profound statistical and econometric analysis to assess the usefulness of age heaping indexes to capture basic numerical skills, and to apply the age heaping concept on current research questions in economic history.

II Outline

The thesis comprises five manuscripts that are intended for separate publication and of which four out of five are co-authored. Each of the manuscripts forms one chapter of my thesis. For this reason, within the chapters, I refer to the respective manuscript as ‘paper’, ‘article’ or alike.[4] Two out of these five working papers are to be submitted in the near future, one to an economic and one to a demographic journal, one paper is currently under review (resubmission) and two are accepted, one by the *Journal of Economic History* and one by the *Economic History Review*. The papers are arranged in chronological order according to the time periods under study.

Given that the concept of age heaping as a proxy for numerical skills and discipline is relatively new in economic history, the working papers compiled in this doctoral thesis could not draw from a large existing stock of literature. Instead, they offer basic research on the micro foundations of age heaping and newly constructed data sets extending over several centuries and as many as 165 countries. This work also includes first applications of the new age heaping indicator. Together with conference presentations and joint projects with other economic history departments, these manuscripts addressed a broad audience of quantitative economic historians and applied macroeconomists and were written to provide them with comprehensive information on the age heaping indicators to facilitate general acceptance of this new proxy variable for human capital.

The first paper offers a profound analysis of the statistical features of various age heaping indicators, which formed the basis for my later work (A'Hearn, Baten and Crayen 2006, 2008). It also takes a look on the determinants of age misreporting on individual level, followed by an analysis of national numeracy levels introducing a large and newly constructed data set on age heaping stretching from the 14th to the early 19th century. These data are complemented by signature ability rates to systematically examine the relation between age heaping and literacy rates.

The second paper uses the age heaping technique to construct a measure of human capital inequality in 17th and 18th century France and 19th century United States (Crayen

[4] Moreover, the numbering of subsections, tables and figures restarts in each chapter.

and Baten 2008b). We employ inequality values on provincial level to track the change of the distribution of human capital over time and to analyze its importance and its interaction with other variables in growth processes.

The third manuscript employs the age heaping method as a proxy of cognitive ability to understand determinants of low living standards in the past (Baten, Crayen and Voth 2008). Using the quasi-natural experiment of Great Britain during the Napoleonic blockade when imports were restricted and food prices soared, the paper studies the link between net-nutrition and cognitive ability. Given that the welfare system in 18^{th} and 19^{th} century Great Britain was organized at county level, relief payments for the poor varied considerably within the United Kingdom. We use the regional variation of living standards for the poor in combination with the temporary food crisis to elaborate the causal mechanism between net-nutrition and cognitive skills, as well as between cognitive abilities and wage income.

The fourth article documents the construction of a data set providing decadal age heaping estimates for 165 countries for the period of 1820-1940.[5] Since the estimates are derived from census age data only, the age heaping values can be widely considered representative for the population. We assess possible determinants of age heaping such as schooling, a strong state bureaucracy, the numbers of census enumerations conducted in the past and the Chinese calendar. First growth regressions employing age heaping as proxy for human capital suggest a statistically and economically significant and positive influence of basic numerical knowledge on subsequent economic growth (Crayen and Baten 2008a).

[5] For most of today's less developed countries, the time series only start in mid- to late 19th century.

The fifth paper exemplifies the dual purpose of single year age data that was collected for age heaping analysis (Crayen 2008). By comparison of successive age distributions of a given population life expectancy estimates are derived. Complex adjustment procedures for age heaping, migration and other data quality issues neutralize potential sources of bias. The estimates are then used to contribute to a debate in historical demography on the spatial demographic pattern in India. Chapter seven summarizes my findings and gives directions for future research.

References

A'Hearn, B., J. Baten and D.Crayen (2006). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital. Universitat Pompeu Fabra Economics Working Paper Series 996.

A'Hearn, B., J. Baten and D.Crayen (2008). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital (revised version). Working Paper Franklin and Marshall College/ University of Tuebingen.

Baten, J., D.Crayen and H.-J. Voth (2008). Poor, Hungry and Stupid: numeracy and the impact of high food prices in industrializing Britain, 1780-1850. Working Paper Universitat Pompeu Fabra /University of Tuebingen.

Barro, R. (2003). Determinants of Economic Growth in a Panel of Countries, Annals of Economic and Finance. 4: 231-274.

Barro, R. J. and X. Sala-i-Martin (1995). Economic Growth. New York: Mc Graw-Hill.

Cervellati, M. and U.Sunde, (2005). Human Capital Formation, Life Expectancy and the Process of Development. American Economic Review 95 (5), 1653-1672.

Crayen, D. and J. Baten (2008a). Global Trends in Numeracy 1820-1949 and its Implications for Long-run Growth. Working paper University of Tuebingen.

Crayen, D. and J. Baten (2008b). New Evidence and New Methods to Measure Human Capital Inequality Before and During the Industrial Revolution: France and the U.S. in the 17th to 19th centuries (revised version). Working paper University of Tuebingen.

Crayen, D. (2008). Life Expectancy in Mysore State During the Early 20th century: The Dual Purpose of Single Year Age Distributions. Working paper University of Tuebingen.

Drèze Jean and Amartya Sen (1995). India: Economic Development and Social Opportunity. Oxford University Press.

Galor, O. and D. Weil, (2000). Population, Technology and Growth: from the Malthusian Regime to the Demographic Transition. American Economic Review 110, 806-828.

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

Levine, R. and D. Renelt (1992). A Sensitivity Analysis of Cross-Country Growth Regressions, American Economic Review. 82: 942-963.

Levine, R. and S.J. Zervos (1993). What have We Learnt About Policy and Growth from Cross-Country Growth Regressions?, American Economic Review 83: 426-430.

Lucas, R. (1988). On the Mechanics of Economic Development. Journal of Monetary Economics 22 (1): 3-42.

Romer, P. (1990). Endogenous Technological Change. Journal of Political Economy 99 (5): S71-S102.

Thomas, K. (1987). Numeracy in Early Modern England: The Prothero Lecture. Transactions of the Royal Historical Society, 5^{th} Ser., 37.

2 QUANTIFYING QUANTITATIVE LITERACY: AGE HEAPING AND THE HISTORY OF HUMAN CAPITAL

Abstract

Age data frequently display excess frequencies at round or attractive ages, such as even numbers and multiples of five. This phenomenon of age heaping has been viewed as a problem in previous research, especially in demography and epidemiology. We see it as an opportunity and propose its use as a measure of numeracy in historical periods where other measures of human capital are unavailable, much as signature rates are used to measure literacy. A simulation study yields methodological guidelines for measuring and interpreting differences in age heaping, while analysis of historical datasets demonstrates the existence of a robust correlation between age heaping and literacy at both the individual and aggregate level, where both can be observed. A preliminary dataset is constructed that illustrates the wide range of primary sources from which age heaping can be estimated, and which considerably expands the frontier of our knowledge of human capital, both geographically and temporally. Our estimates reveal the medieval roots of human capital accumulation in Western Europe, which had diverged from the East and reached high numeracy levels by 1600, long before the rise of formal schooling.

This chapter is based on a working paper written jointly with Brian A'Hearn (Franklin & Marshall College) and Joerg Baten (University of Tuebingen). It is conditionally accepted by the *Journal of Economic History.* The concept for the paper was developed jointly, the analyses and writing were equally shared.

I. Introduction

Twenty-first century legislatures and educational bureaucracies creak into action at the publication of the latest international league tables of educational attainment.[6] In the U.S., such concerns have a history, and a particular focus on math, dating back to Sputnik-era anxiety about falling behind in a technological race with economic and national security implications. This led to reforms like the "New Math" of the 1960s, measures like the Education for Economic Security Act of 1984, and debates like the current "Math Wars" about classroom pedagogy. But these are only the most recent manifestations of national math anxiety. More than a century ago it was British fear of German technological ascendancy that spurred a reform movement demanding better mathematical, technical and scientific education.[7]

Concern with educational policies is not misplaced, according to a voluminous (if inconclusive) literature on the determinants of economic growth.[8] But growth regressions have yielded few insights into the importance of *different types* of knowledge, education, or skill. Large panel datasets rarely have information more detailed than gender- and level-specific enrollment or attainment rates. An oft-cited analysis indicating that engineering students raise growth while law students lower it is the exception rather than the rule.[9] In studies of individual labor market outcomes, the consistent, positive effect of human capital variables is not in question. Debate concerns only the precise mechanisms

---

[6] Influential comparative projects include TIMMS (Trends in International Mathematics and Science Study; International Association for the Evaluation of Educational Achievement), PISA (Programme for International Student Assessment; OECD), and ALL (Adult Literacy and Lifeskills; Statistics Canada and OECD).

[7] This anxiety was captured in the title of a well-known analysis of the industrial position of the country: "Made in Germany." See Sanderson (1999).

[8] A few influential studies that investigate the connection between schooling and economic growth include Levine and Renelt (1992), Mankiw, Romer and Weil (1992), Barro and Lee (1994), Barro (1997), Sala-i-Martin (1997), Bils and Klenow (2001), and Hanushek and Kimko (2000).

at work, for example whether education primarily screens ability or creates it, or whether skill- or college-premia are increasing due to rapid technological change. In this context, achievement and skill test results are often introduced as measures of ability alongside educational attainment and experience variables. Invariably, ability measured in this way matters; occasionally, it is broken down by specific skill areas such as quantitative and verbal reasoning.[10] In a study of the U.S. labor market, Murnane, Willett, and Levy (1995) found that cognitive abilities had greater predictive power than educational attainment, with mathematics skills the most highly correlated with wages. Similarly, Rivera-Batiz (1992) found that "quantitative literacy" significantly (in both senses) raised the probability of full-time employment among U.S. workers. Outside the U.S., recent studies have found numeracy to be positively associated with labor force participation, full-time employment, annual weeks worked, and income in Britain, Canada, and Australia.[11] Often numeracy dominates literacy as an explanatory factor, particularly for women and among the less educated.

Interest in the historical evolution of human capital has been stimulated by the recent development of very-long-run growth models. These models seek a unified explanation of pre-modern Malthusian dynamics and modern economic growth, typically assigning a key role to a fertility transition in which families switch from valuing quantity to quality in children, raising investment in human capital.[12] The empirical basis of such

---

[9] Murphy, Shleifer and Vishny (1991).

[10] Studies with evidence on general cognitive ability include Neal and Johnson (1996) and Heckman (1995).

[11] See Chiswick, Lee, and Miller (2003) for Australia; Charette and Meng (1998) and Finnie and Meng (2001) for Canada; Parsons and Bynner (2005) for Britain.

[12] In Galor and Weil (2000), population size eventually causes technological progress to speed up, raising the return to human capital, causing families to increase investment in human capital, lowering fertility and increasing the rate of growth. Thus is the Malthusian cycle broken. Becker, Murphy and Tamura (1990),

theorizing is shaky, however. For the nineteenth century, measures of human capital investment, such as school enrollments and attainment, or census data on literacy, are available – at least for some of today's rich countries. O'Rourke and Williamson (1997) were able to include schooling in European convergence regressions for the 1870-1913 period, for example. (Their conclusion was that globalization forces were much more important influence on comparative development than human capital.[13]) No data whatsoever are available on individual cognitive ability, or on different types of ability, however.

Pushing back into the early nineteenth century and before, even schooling data become scarce, and literacy must generally be inferred from a proxy: ability to sign one's name on marriage registers and legal documents. For the years around 1800, Reis (2005) is able to assemble such data for 15 European regions. The figures indicate that male literacy varied widely, from over 60% in northwestern Europe to below 20% in parts of Italy and under 10% in eastern Europe. Pushing back still further into the early modern era, it becomes increasingly difficult to find systematic, comparable data. Relatively plentiful data on reading ability in Scandinavia and signature ability in the Netherlands, Britain, France, and Spain allow Graff (1987) to document a considerable improvement in literacy in the seventeenth and eighteenth centuries. But there is very little information available for the rest of Europe. Allen's (2003) conclusion that human capital has no ability to explain progress and poverty in Europe between 1300 and 1800 may result more from his use of urbanization as a proxy than from literacy's actual irrelevance.

Lucas (2002), and Cervellati and Sunde (2005) are further examples of models explicitly built on such human capital foundations.

[13] Tortella (1994), using literacy data, offers a different interpretation, at least for southwest Europe.

Indeed, Baten and van Zanden (2006) reach the opposite conclusion using book production to measure human capital.

What of *numeracy* as a historical measure of human capital? For Weber, Sombart, and Schumpeter, numeracy was at the very heart of modern, rational capitalism. They traced the roots of both to the invention of double-entry bookkeeping in late medieval Italy. Carruthers and Espeland (1991) describe in some detail the process of abstraction and organization inherent in compiling a ledger, which made possible the development of concepts like capital, depreciation, and rate of profit.[14] It is no accident that the introduction of Arabic numerals into Europe (by the merchant Leonardo of Pisa, a.k.a. Fibonacci) and the earliest accounts of mathematics education date from this same time and place. Numerous *scuole d'abbaco* thrived in Renaissance Florence according to Goldthwaite (1972), where the young sons of the commercial classes studied a mathematics curriculum that would change little before the nineteenth century.[15] Italy remained the European center of publication and instruction in mathematics and accounting until at least 1500, according to Swetz (1987).

Emigh (2002) has investigated the numeracy of ordinary Tuscans in this period by analyzing their tax declarations for the famous Florentine *catasto* of 1427. She finds that ordinary citizens and peasants much more often provided too much quantitative information (rents, size of plots, yields, debts, salaries) than too little, with respect to the demands of the tax officials. This implies that causation ran from market activity to

---

[14] The authors argue that although double-entry bookkeeping truly was a superior technology, that its potential was seldom exploited by practicing merchants. It came to have a powerful rhetorical significance, as a symbol of meticulousness and probity.

[15] Routine commercial calculations in the middle ages could include conversions between non-decimal monetary systems with fluctuating exchange rates, estimation of the volume of containers, the reckoning of interest, or the division of profits between partners with different amounts of capital invested at different times.

numeracy to tax design, rather than in reverse. Cohen (1982), another historian of numeracy, finds that the expansion of market activity in the early nineteenth century U.S. was only partially responsible for making Americans “a calculating people.” A reform of math pedagogy actually de-emphasized commercial applications in favor of teaching abstract thinking in this period.[16] A sophisticated literature on the history of numeracy certainly exists, but it does not yield statistical measures. Can we quantify quantitative reasoning?

II. Age Heaping

It turns out that we can. As signature ability can proxy for literacy, so accuracy of age awareness can proxy for numeracy, and for human capital more generally. A society in which individuals know their age only approximately is a society in which life is not governed by the calendar and the clock but by the seasonal cycle; in which birth dates are not recorded by families or authorities; in which numerical age is not a criterion for access to privileges (voting, office-holding, marriage, holy orders) or for the imposition of responsibilities (military service, taxation); in which individuals who know their birth year have difficulty accurately calculating their age from the current year. Within a society, the least educated and those with the least interaction with state, religious, or other administrative bureaucracies are least likely to know their age accurately. Age awareness thus tells us something about both the individual and the society he or she inhabits, about human capital and what might be called “administrative capital.”

[16] Equally important was the triumph of “political arithmetic” as the favored tool for assessing the experiment that was republican America, and for influencing public opinion. This presupposed a certain basic level of numeracy.

Approximation in age awareness manifests itself in the phenomenon of age heaping in self-reported age data. Individuals lacking certain knowledge of their age rarely state this openly, but choose instead a figure they deem plausible. They do not choose randomly, but have a systematic tendency to prefer "attractive" numbers, such as those ending in 5 or 0, or even numbers, or in some societies numbers with other specific terminal digits. Age heaping can be assessed from any sufficiently numerous source of age data: census returns, tombstones, necrologies, muster lists, legal records, or tax data, for example. While care must be exercised in ascertaining possible biases, such data are much more widely available than signature rates and other proxies for human capital.

Age heaping is a well-known phenomenon among demographers. Already a half-century ago influential studies by Bachi (1951) and Myers (1954) investigated age heaping and its correlation with education levels within and across countries. Myers (1976) demonstrated the correlation at the individual level between age awareness and income. For others, including epidemiologists, age heaping is a problem to be solved, a source of distortion in age-specific vital rates. Zelnik (1961), for example, assessed age misreporting in the United States between the 1880 and 1950 censuses. Finding real and approximately linearly decreasing levels of age heaping for each census, he constructed correction factors to adjust age cohort sizes.[17] Development economists and anthropologists use age heaping as a measure of data quality and consistency.

Meanwhile, historians have studied age heaping as a topic of interest in its own right. A pioneering example is the study by Herlihy and Klapisch-Zuber (1978) of the

[17] Discussions of age heaping as a problem in the demography literature include Coale and Kisker (1986), Preston, Elo, Rosenwaike and Hill (1996), and Vallin, Meslé, Adamets, and Pyrozhkov (2002). Denic, Khatib and Saadi (2004) discuss the issues for medical, Crockett and Crockett (2006) for historical research. See also UN Statistics Division (2003).

Florentine tax records from the fourteenth and fifteenth centuries. In a chapter devoted to age heaping, they documented marked heaping on even numbers for children and on multiples of five for adults, to a degree similar to that reported for Egyptian census data in 1947. Among other findings, they demonstrated that age heaping diminished substantially over successive tax enumerations from 1371 to 1470, and that it was more prevalent among women, in rural areas and small towns, and among the poor. A second well known example is the study of Duncan-Jones (1990), who used grave monument inscriptions to estimate age heaping for men and women in twelve provinces of the Roman Empire. He found age heaping on multiples of five at levels not dissimilar to those for medieval Tuscany or developing countries of the 1950s and '60s, and higher for women than men.[18]

The first use of age heaping as an indicator of human capital in the economic history literature is relatively recent. Mokyr (1983) tested for positive selection or "brain drain" in pre-famine Irish emigration by comparing age heaping among migrants and in the population at large. Developing original measures of age heaping along the way, he found no support for the conventional wisdom that the best and brightest emigrated. Budd and Guinnane (1991) studied Irish age misreporting in linked samples from the 1901 and 1911 censuses. They found considerable heaping on multiples of five in the 1901 census, which was greater among the illiterate, the poor, and the aged. The introduction of state pensions for individuals 70 and older in 1908 changed incentives regarding age reporting: For those potentially eligible, there was on the one hand an incentive to exaggerate age, on the other a concern to report age accurately, since pension examiners had access to

[18] Other interesting studies of age heaping are Kaiser and Engel (1993) on early modern Russia, and Jowett and Li (1992) on contemporary Chinese ethnic groups.

census returns. On balance, age heaping declined significantly in the 1911 census, and its variation across social groups became much less clear. The longitudinal data enabled the authors to check the consistency of individual age reports. Mean elapsed time was not 10 but approximately 12 years, the positive discrepancy being larger for Catholics, illiterates, and those reporting heaped ages in 1901. In another study of Ireland, O'Grada (2006) used a higher degree of age heaping among Dublin's immigrant Jewish population to show that their lower literacy did not refer only to the English language and that their lower mortality was the result of religious practices rather than education.

In a linked census sample for Britain in 1851 and 1881, Long (2005, 2006) assessed both aggregate age heaping at the county level and, exploiting the repeated observations, individual age discrepancies. Fully a quarter of his sample of 1851 school-aged children reported ages in 1881 with discrepancies of from two to five years. While countywide age heaping had a limited impact on individual outcomes once other county characteristics were controlled for, individual age discrepancy had a significant impact on socio-economic status, wages (10% higher for 0-discrepancy individuals), and the probability of rural-urban migration. What these studies have in common is that all find evidence of significant age heaping, and that it varies across individuals or groups in a way consistent with its interpretation as a measure of human capital.

Researchers in a number of disciplines are familiar with age heaping, and it has been used as an indicator of human capital in studies of particular times and places. But several different measures of age heaping have been employed, complicating comparisons. And awareness of age heaping is somehow less than the sum of these parts.

Age heaping has an enormous and unexploited potential to yield new insights into the comparative historical evolution of human capital.

III. Measuring age-heaping

To deploy age-heaping as a useful indicator of human capital, we require a measure that allows us to track its variation over time and across groups. This is a question of *how much* age heaping is present in the sample, conceptually distinct from the question whether age heaping is present at all. For the second of these questions, the familiar Pearson chi-squared statistic would be suitable:

$$(1) \qquad H_P = \sum_{i=1}^{k} \frac{(n_i - \hat{n}_i)^2}{\hat{n}_i}$$

where *H* is chosen to stand for heaping and *P* for Pearson, *i* indexes the *k* ages to be considered, and n_i and $\hat{n}_i$ are the observed and expected frequencies, respectively. Note that the chi-squared statistic does not attempt to estimate a parameter such as, for example, the share of ages incorrectly reported. It can take on quite a wide range of non-negative values, depending on *k* and the particular pattern of expected frequencies (the null hypothesis). Like the other indices to be considered, H_P is based on comparison of actual and expected frequencies, and like them it must standardize and aggregate these deviations. The chi-squared statistic standardizes by squaring and expressing the result as a percentage of the expected frequency. It aggregates by summing with equal weights. H_P is asymptotically distributed as a χ^2_{k-1} random variable, facilitating hypothesis testing concerning the presence of heaping.

But our primary interest is in comparing the *degree* of age-heaping in different samples: the "how much" question. And for this purpose the chi-squared statistic has the drawback of scale dependence. For clarity, it is useful to distinguish *mathematical* from *statistical* or probabilistic, scale dependence. The chi-squared statistic is scale dependent in the mathematical sense that inflating both observed and expected frequencies by some common factor causes H_P to grow by the same factor. It is *not* scale dependent in a statistical sense, in that $E\left(\chi^2_{k-1}\right) = k-1$, regardless of sample size. The two statements are reconciled by noting that, under the null hypothesis, a larger sample is not expected to increase all observed frequencies by the same factor. Rather, the observed and expected distributions will tend to conform more closely on average. Mathematical scale dependence complicates comparisons, since two samples of different sizes, with identical patterns and degrees of heaping, will yield different values of H_P. Similarly problematic is the dependence of H_P on the age range considered (*k*), which varies across samples. Of course, it is possible to consider only subsamples of like size and age-range, but this either throws away valuable information or necessitates a cumbersome resampling and averaging procedure. Another potential weakness of H_P is that it weights the deviations for all ages equally, even those for which smaller frequencies and greater variability are expected.

In this paper two indices based on the chi-squared statistic are considered. Both are mathematically scale-independent, and both weight deviations by the age's expected sample share so as to reduce the influence of potentially unreliable observations. They differ in their methods for standardizing the deviations. The first, dubbed the "ABC" Index after the initials of authors, is a weighted sum of the squared percentage deviations from predicted values:

$$(2)\ H_{ABC} = \sum_i \frac{\hat{n}_i}{N}\left(\frac{n_i - \hat{n}_i}{\hat{n}_i}\right)^2,$$

where the notation is as before and N denotes the sample size. If possible discrepancies between expected and actual sample size are ignored, H_{ABC} can be rewritten as $\frac{1}{N}H_P$, so that it is in fact a scaled version of the chi-squared statistic.

Squaring the deviations has the effect of heavily weighting outliers. The second index, "Lambda," avoids this by relying on the absolute value of the percentage deviations:

$$(3)\ H_\lambda = \frac{\sum |n_i - \hat{n}_i|}{\sum \hat{n}_i} = \sum_i \frac{\hat{n}_i}{N} \frac{|n_i - \hat{n}_i|}{\hat{n}_i}.$$

Neither the ABC nor the Lambda Index embodies particular assumptions about how expected frequencies are generated. Both can take on a wide range of values depending on the number of ages k and the particular distribution of expected frequencies.[19]

Several commonly used indices aggregate observed and expected frequencies over terminal digits before calculating deviations. Bachi's index sums the differences between actual and expected frequencies of each terminal digit, considering only the positive values – intuitively similar to the use of absolute values in (3).[20] The resulting sum is expressed relative to sample size:

$$(4)\ H_B = \frac{1}{N}\sum_{j=0}^{9} \mathrm{I}(j)\cdot(n_j - \hat{n}_j), \quad \mathrm{I}(j) = 1 \text{ if } n_j > \hat{n}_j, \text{ else } \mathrm{I}(j) = 0,$$

where I(j) is an indicator variable as defined above, and j indexes terminal digits from 0 to 9. The Bachi index is not mathematically scale dependent, weights terminal digits with

---

[19] The Lambda Index was proposed by Mokyr (1983), who also considered a "Gamma" index, which can be rewritten as $\sum_i \frac{\hat{n}_i}{N^2} \frac{(\hat{n}_i - n_i)^2}{\hat{n}_i}$, showing it to be a weighted chi-squared type index. In simulations, Gamma's performance did not differ in interesting ways from that of H_{ABC} and H_λ.

positive deviations equally, and makes no assumptions about how to derive the expected frequencies for each terminal digit. If all expected terminal digit frequencies are assumed to be 10%, H_B can take values between 0 and 0.90 and can be intuitively interpreted as an approximation of the percentage of the sample reporting an inaccurate age.

Another commonly used measure is Myers' Blended Index, which differs from those considered thus far by making a specific assumption about expected frequencies, and by making an adjustment to observed frequencies. Predicted terminal digit shares are set at 10% (meaning the index must be applied to age intervals that are multiples of 10 years). This is of course not always a realistic assumption; in a sample of individuals aged 60 to 79, one would expect more ages ending in 0 than in 9, for example. Rather than adjust *expected* frequencies accordingly, H_M adjusts the *observed* frequencies using a "blending" procedure.[21] Unlike H_B, which considers only positive deviations and expresses them relative to N, the Myers Index sums the absolute values of all deviations, and expresses them relative to $2N$:

$$(5)\ H_M = \frac{1}{2}\sum_{j=0}^{9}\left|\frac{\tilde{n}_j}{N} - .10\right| = \frac{1}{2}\sum_{j=0}^{9}\frac{\left|\tilde{n}_j - \hat{n}_j\right|}{N},\ \hat{n}_j = 0.1N\ \forall j,$$

where $\tilde{n}_j$ is the "blended" observed frequency of a particular terminal digit j. The index is not mathematically scale dependent, and can vary between 0 and 0.90 like the Bachi.

The indices considered so far can be used to detect any type of heaping, and do not rely on particular mechanisms for generating expected frequencies (with the exception of the Myers Index). More specialized measures focus on a particular type of heaping. The well-known Whipple Index is designed to capture heaping on ages ending

---

[20] This is only one of three indices proposed by Bachi. The others focus on a particular age or a particular terminal digit.

in 0 or 5. H_W sums the frequencies of all ages ending in 0 or 5 and expresses the result relative to one-fifth the sample size:

$$(6)\ H_W = \frac{\sum(n_{25} + n_{30} + n_{35} ... + n_{60})}{\frac{1}{5}\sum_{i=23}^{62} n_i}.$$

The summation notation in the denominator (rather than the N used in Equations 2-5) is meant to emphasize that H_W must be defined over an interval in which each terminal digit occurs an equal number of times, such as 23 to 62. Implicitly, equal terminal digit shares in unheaped data are assumed. This would be correct for a uniform distribution of ages, but can only be approximate for typical samples in which frequency decreases with age. The Whipple index makes no adjustment to correct for this problem. H_W can range from 0 in the case of no observations on 0s and 5s, through 1 in the case of a uniform distribution of terminal digits, to 5 in the case of 100% heaping. In application, H_W is typically multiplied by 100.

An alternative index of heaping on the terminal digits 0 and 5 makes a less restrictive assumption regarding expected frequencies: that they evolve approximately linearly over any three year age range. The "Multiples of Five" index is a simple, equally weighted, average of the frequencies of terminal digits 0 and 5, each expressed relative to the average frequency of immediately adjacent ages, as in the following example:

$$(7)\ H_{M5} = \frac{1}{3} \cdot \left\{ \frac{2n_{20}}{n_{19} + n_{21}} + \frac{2n_{25}}{n_{24} + n_{26}} + \frac{2n_{30}}{n_{29} + n_{31}} \right\}.$$

[21] See Shryock and Siegel (1973) for an abbreviated description of Myers' blending procedure.

IV. Evaluating Indices of Age-Heaping

Which of the indices defined in Section III is best? The familiar criteria used to assess estimator performance, bias and variability, are not directly applicable here, as the indices under consideration are not estimators of unknown parameters. Three desirable properties of an age-heaping index are statistical scale-independence, a linear response to the degree of heaping, and the ability to reliably rank samples from populations with different degrees of heaping. Because the distributions of the indices are unknown, these properties must be investigated by simulation.

a. *Simulation details*

Three types and five degrees of heaping are distinguished in the study. The *types* are: heaping on even numbers, heaping on multiples of five, and mixtures of the two. The *degrees* are: for the pure types, 0, 5, 10, 15, and 20 percent of the sample subjected to heaping; and for the mixed type 5 percent even-heaping plus an additional 5, 10, 15, or 20 percent multiples of five heaping. The predominance in the mixed scheme of heaping on multiples of five over even-heaping conforms to the pattern found in a wide range of historical datasets. The mixed schemes are the focus of most discussion here.

The three heaping patterns were imposed on random samples drawn from distributions in which frequencies decrease with age, as is typical of data from military, census, and randomly-sampled sources. More specifically, samples of size 250, 500, 1000, 2000, and 5000 were drawn from the 23-42 range of a $N(20,10)$ distribution rounded to integer values, then heaped according to the relevant scheme. The entire exercise was repeated for the wider age range 18-57, yielding results sufficiently close to those reported here as not to merit separate discussion. 1000 repetitions were carried out for each sample size and heaping scheme.

As noted in Section III, the ABC, Lambda, and Bachi indices do not embody specific assumptions about expected frequencies. In the simulation study, expected frequencies were generated using locally weighted regression. The intuition is to use a regression of observed frequencies on age to generate predicted frequencies, allowing the estimated relationship to vary locally. For each age, the slope is estimated using only data in a window around that age (here set to include 80% of all observations), using a kernel (here the tricube) to weight them inversely with the distance from the age being considered. This approach is similar to other smoothing methods, in particular kernel density estimation. It has the advantage of being easily implemented and flexible, making no *a priori* assumptions about the form of the underlying (unheaped) distribution of ages. Its potential disadvantage is oversensitivity to (undersmoothing of) the observed, randomly- or systematically-heaped frequencies. In principle, this problem can be addressed by optimal choice of bandwidth. But this choice depends on the type and degree of heaping and on the underlying distribution of unheaped ages, all of which are unknown in practice. Hence, a uniform (non-optimal) bandwidth was applied for all sample sizes, heaping schemes, and indices. The evaluation of the ABC, Lambda, and Bachi indices is in fact a joint evaluation of the index itself and the lowess method of deriving expected frequencies. Similarly, the Myers Index is evaluated together with its blending method of adjusting observed frequencies.

As an example of the results, Figure 1 displays kernel density estimates of the distribution of the Lambda Index under the mixed heaping scheme with a sample size of

1000. The leftmost curve corresponds to no heaping, the rightmost to maximum heaping. It is evident that on average the index value increases with the degree of heaping, as it was designed to. The graph also indicates for this specific case that the increase is not linear with the degree of heaping, that the distribution is right-skewed at low degrees of heaping, that the variance of the distribution increases, and that there is substantial overlap in the distributions for successive degrees of heaping.

b. *Scale dependence*

The indices under study are mathematically independent of scale by construction. But they may be *statistically* scale dependent, in the sense that their expected value may be a function of sample size. Random sampling variation always creates deviations between observed and expected frequencies, even when the population is unheaped. Since such deviations are squared, converted to absolute values, or considered only when positive in Equations 2-5, there is no tendency for them to cancel each other out, so that the means of these indices are not zero even in the absence of heaping. When such deviations are expressed relative to expected frequencies or sample size, however, they tend to diminish in larger samples. This creates what is termed here statistical scale dependence, which creates the same difficulties as mathematical scale dependence.

Figure 2 plots the means of all six indices as a function of sample size under the 05-10 mixed heaping scheme (5 percent even-heaping, 10 percent heaping on multiples of five). Statistical scale dependence is apparent for all but the Whipple Index. The ABC and Lambda indices perform poorly on this criterion. Their decrease with a doubling of sample size (at least when starting from a small sample) is enough to offset the increase that would result from doubling the degree of heaping. Among the “all-purpose” indices, the Bachi seems to do best. The pattern of Figure 2 is representative of results under all types and degrees of heaping, though the severity of scale dependence is somewhat reduced when the degree of heaping is extreme.

c. *Response to heaping*

Figure 3 plots mean index values for increasing degrees of mixed heaping, for sample sizes of 500. All index means rise monotonically with the degree of heaping, as designed, but responsiveness varies considerably. The Multiples of 5 and Whipple Indices increase linearly, while the ABC and Lambda Indices increase at increasing rates. The sharply nonlinear response of the ABC and Lambda indices creates two related problems. Most importantly, it is difficult to distinguish low degrees of heaping from zero and from each other, especially once random sampling variability is taken into account. In addition, even substantial differences in heaping across samples become less easily interpretable. An increase of 0.2 in the Whipple Index always corresponds to an increase of approximately 5% in the share heaped on multiples of five. No analogous simple rule describes the indices with a non-linear response. They do not, for example, double when the share heaped doubles. The Bachi and Myers Indices have an intermediate response pattern, which is reasonably close to linear.

The response patterns in Figure 3 are representative of all sample sizes and of both mixed and pure heaping on multiples of five. Pure even-heaping produces somewhat different results. It is no surprise but deserves emphasis that the Whipple Index does not respond at all, actually declining somewhat as the share subjected to heaping increases in the simulation results. Similarly, the Multiples of 5 Index increases only imperceptibly. The other indices increase monotonically with the degree of heaping, in a distinctly non-linear way in the case of the ABC and Lambda Indices. The rate of increase is less than for types of heaping including multiples of five, making it difficult to distinguish varying degrees of even-heaping. As a practical matter, however, datasets from a wide variety of countries and time periods have failed to yield an example of pure even-heaping. When present, even-heaping is "outweighed" by heaping on multiples of five.

d. *Precision and the probability of ranking errors*

Accuracy in ranking samples depends not only on the response of the index mean to age heaping, but equally on its variation around that mean in repeated sampling. Figure 1 illustrated how the distribution of the Lambda Index varied with the degree of heaping, shifting steadily to the right, decreasing in skewness, and increasing in variance. The substantial overlap of the estimated densities for successive degrees of heaping in Figure 1 signaled likely difficulties in distinguishing small variations in the degree of heaping from random sampling noise. Index distributions are also a function of sample size. This is illustrated for the Bachi Index in Figure 4, which plots kernel density estimates for 05-10 heaping in sample sizes of 250, 500, 1000, 2000, and 5000. The Bachi's previously-noted scale dependence is again evident; the low, wide curve with the rightmost mode corresponds to a sample of 250, while the tall, narrow, leftmost curve is for sample size 5000. Also evident is a quite dramatic decrease in the variance of the distribution as the sample size increases, a characteristic of all the indices.

A useful gauge of reliability in discerning the degree of heaping turns out to be the probability of incorrectly ranking sample pairs. A ranking error occurs when the sample from a population with a low degree of heaping, say 05-05, yields an index value exceeding that for the sample from a population subject to a greater degree of heaping such as 05-10. We could write this as $H^{05-05} > H^{05-10}$, or alternatively $D = H^{05-10} - H^{05-05} < 0$. The distribution of differences like D can be estimated by recourse to the already simulated index realizations under various types and degrees of heaping. In many cases these differences have an approximately normal distribution. Figure 5 provides an example: the estimated density of the difference $D = H_B^{05-10} - H_B^{05-05}$ for the Bachi Index. The probability of a negative difference, hence a ranking error, is given by the shaded area, which in this case equals 0.15.

Table 1 presents estimates of the probability of a ranking error for different sample sizes under the mixed heaping type. The differences refer to "adjacent" degrees of heaping; $D1 = H^{05-05} - H^{00-00}$, $D2 = H^{05-10} - H^{05-05}$, and so on. Immediately apparent from the figures in Table 1 is that none of the indices perform reliably in small samples. In samples of 250, the "all-purpose" indices routinely misrank samples from unheaped populations as having more heaping than those from 05-05 populations: 37% of the time. The Whipple Index offers dramatically better performance for low degrees of heaping, and remains the best by a small margin even with more severe heaping. In large samples, the accuracy of all indices is much improved, the error probabilities comfortably small. Again the Whipple Index is clearly the most accurate of the group, especially in distinguishing among relatively low degrees of heaping. It bears emphasis that the Whipple is quite incapable of capturing forms of heaping other than multiples of five. Among the more flexible indices, the Bachi offers the lowest error probabilities.

Table 2 reports the analogous probabilities for samples that are two degrees of heaping apart. In other words, $D20 = H^{05-10} - H^{00-00}$, $D31 = H^{05-15} - H^{05-05}$, and so on. While all indices have trouble reliably distinguishing fine differences in the degree of heaping, the results in Table 2 indicate considerably better accuracy for larger heaping differences. Here again, the Whipple Index is clearly the most reliable – so long heaping on terminal digits 5 and 0 characterizes the data. In large samples, all indices are extremely accurate, with incorrect rankings not a practical problem.

In short, our simulation studies show that the Whipple Index trumps the competition, offering three specific advantages. First the mean of *W* is not scale dependent, meaning that Whipple Indices can be compared across samples of widely varying size.[22] Second, *E*(*W*) increases linearly with heaping, again facilitating comparisons.[23] Finally, the coefficient of variation of *W* is systematically lower than for the alternatives, at all sample sizes and for all degrees of heaping. This leads to greater reliability in correctly ranking samples according to the true extent of heaping in the underlying populations.

In this paper we employ a simple transformation of the Whipple Index that yields an estimate of the share of individuals that correctly report their age:

$$(8)\quad \tilde{W} = \left\{1 - \frac{(W-100)}{400}\right\} \times 100 \text{ if } W \geq 100\text{; else } \tilde{W} = 100.$$

V. Numeracy and literacy in the US census

Our claims are that 1) age heaping is a useful indicator of basic numeracy in historical contexts; and 2) basic numeracy is an important component of human capital more broadly. An absence of historical data on individual cognitive abilities makes it difficult to test either proposition directly. We propose to evaluate age heaping's usefulness as a measure of human capital by investigating its correlation with literacy. For the nineteenth

[22] Means of the alternative indices fall with sample size, because random deviations between observed and predicted frequencies that are large in proportional terms occur less frequently in large samples.

century, census manuscript records offer a wealth of individual level data on age and literacy. The Integrated Public Use Micro Samples (IPUMS) of the United States censuses are an especially valuable source because they offer large samples from a wide range of birthplaces in the US and abroad.

We extracted records for nearly 650,000 men and women aged 20-69 from the IPUMS samples for 1850, 1870, and 1900.[24] 1850 is the earliest available IPUMS sample, while 1900 is important because census enumerators inquired about both age and birth-year, possibly eliciting more considered responses that better reflect age awareness. The data used for this study include age, race, literacy (the ability to read and write, in any language, with no reference to proficiency), and birthplace.

We first consider the numeracy-literacy relationship at an aggregate level. For each census – birthplace – ethnic group combination with at least 100 observations, we calculate $\tilde{W}$ and the literacy rate, averaging over all age groups in the 20-69 range and over both genders. Literacy rates vary widely, from under 25% among the former slaves of the American South in the 1870 census, to over 99% among the white populations of several Northeastern states. Numeracy measured by $\tilde{W}$ displays somewhat less variation, ranging from under 65%, for 1850 free and 1870 emancipated black populations, to over 98% for whites in several US states and foreign birthplaces in 1900. The higher minima for $\tilde{W}$ suggest that age heaping captures a very basic level of numeracy.

---

[23] Several other indices respond more than proportionally to increases in heaping.

[24] The IPUMS data are available online from the Minnesota Population Center at the University of Minnesota. Steven Ruggles, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004.

Table 3 and Figure 6 illustrate the regional numeracy-literacy relationship in the IPUMS data. Overall and in every subsample there is a positive, statistically significant correlation. OLS regressions of literacy on numeracy yield slope estimates mostly in the range 1.5 and 3.0, so that small changes in numeracy are associated with larger changes in literacy. This reflects the wider range of variation in literacy.

There is considerable noise in the relationship in some samples, as indicated by low R^2 values. In particular, the correlation weakens considerably when only high-literacy, high-numeracy regions are considered. Though still positive and statistically significant, slope coefficients are less than one, while R^2's are quite low. It is worth noting that the correlation of literacy and numeracy is quite robust, emerging also in pooled and region fixed-effects models and with controls for gender balance (relevant for a number of immigrant samples). Nor does it vanish in the census of 1900, when more accurate age reporting may have been induced by separate questions about age and birth year. All in all, these results suggest that $\tilde{W}$ is reliably correlated with human capital (as proxied by literacy) in contexts where both are low, as in the southwest quadrant of Figure 6. This, as we shall see, is the range typical of pre-industrial Europe.

The IPUMS data, at the individual level, allow us to address the question of whether age heaping reflects more the characteristics of individuals or of the society they inhabit. We can also study the effects of variables we will be able to observe in our early-modern age-heaping dataset, in particular gender and age group. We model the probability of reporting a multiple-of-five age as a logistic function of the birth-region literacy rate, which captures aspects of the social environment, and of individual characteristics: personal literacy, and dummies for sex, age group, and Irish ethnicity.

Personal literacy is intended as a measure of human capital. A female dummy is included because basic numeracy skills could be acquired and maintained independently of formal education and literacy, for example through frequent market transactions. It is at least possible that such opportunities differed by gender. Age group controls are included because the tendency to heap can vary with age itself. Younger individuals are more likely to have recently passed landmarks at which their age was ascertained, such as marriage, military service, or immigration. Older individuals may be more likely to forget their ages.[25] (The extremely old, who sometimes deliberately exaggerate their age, have been excluded from the sample.) And society may provide incentives to deliberately distort true age. There is some evidence that young women tended to round their ages down, for example -- possibly in response to marriage market pressures.[26]

The estimates displayed in Table 4 indicate that the probability of reporting a heaped age depends negatively on both personal and birth region literacy. This is true of the sample as a whole and every subsample. In the full sample, the marginal effect of personal literacy is to reduce the probability of reporting a heaped age by about 4 percentage points.[27] This magnitude is fairly large compared to the 9.2 percentage point excess frequency (0.292-0.200) multiple of five ages. Estimates for individual sub-samples vary from -1 to -6 percentage points, and are statistically significant at the 1% level in all but two cases. At the individual as well as the aggregate level, then, accurate age reporting is correlated with another measure of human capital: literacy.

---

[25] See the discussions in Kaiser and Engle (1993) and Ewbank (1981).

[26] Dillon (2007), Ch. 3.

[27] This and all subsequent marginal effects are evaluated at the mean of all explanatory variables. The results are not meaningfully different when evaluated at other points, for example with all dummies set to zero. Predicted mean probabilities of reporting a heaped age are also conditional on all explanatory variables at their sample means.

An individual's probability of reporting a heaped age also depends on the society in which she is brought up. Birth region literacy has a statistically significant (at the 1% level) negative marginal effect on age heaping in the sample as a whole and every subsample. The full-sample estimated effect of a 10 percentage point rise in regional literacy is to lower the probability of a multiple of five age by about 2 percentage points. The range in individual subsamples is from -1.5 to -3.1 points.

Because we have controlled for individual literacy, these estimates mean that even an illiterate is less likely to report a heaped age in a highly-literate society. This interpretation is borne out by the smaller estimated effect of personal literacy in the sample of highly literate birth regions. A range of factors could underlie this result, from greater reliance on written records and greater spread of markets in highly literate societies, to more precise knowledge on the part of the household member interviewed by the census enumerator regarding others' ages.

That $\tilde{W}$ reflects not only individual but also social characteristics is evident in the large positive effects of Irish ethnicity in the foreign-born subsamples. The Irish, who comprise the largest share of the foreign-born, have low levels of literacy but *extremely* low levels of age numeracy, relative to other groups. Turning to gender, women are estimated to have a higher probability of reporting a heaped age in all samples. The effect is small on average, +1.1 percentage points in the full sample, and not always statistically significant in subgroups, but is consistently positive.

Finally, age group effects are large, statistically significant, and consistent across all subsamples; the probability of reporting a heaped age rose with age itself. In the full-sample estimates, individuals in their 20s were almost 8 percentage points less likely to report a heaped age than those in their 30s (the reference group), while those aged 40 and above were 3 points more likely to do so. Across subsamples, the magnitude of the aging effect varies with the degree of heaping. Among native blacks in the 1870 census the excess frequency of multiples of five is 27.7 percentage points (0.477-0.200), and the age group effects are -21 percentage points and +10 points for those under 30 and over 40, respectively. Among native whites in the 1900 census, by contrast, the excess frequency of heaped ages is only 3.2 points (0.232-0.200) and the age group effects are a similarly small -1.7 and +1.7 points. Comparison of the same cohorts across successive censuses revealed that these effects can indeed be attributed to aging, rather than a time trend. Experimentation with alternative specifications indicated that in the IPUMS data there were no discernable differences by age within the over-40 group.

VI. A European age heaping dataset, 1300-1800

To explore the potential of age heaping as a long-run indicator of human capital, we have assembled a dataset that deliberately maximizes breadth of coverage across time, place, and source type. The numeracy database covers over 130 locations in 16 European countries over birth decades in the half-millennium from 1350 to 1840. It consists of over 300 *samples*, which we define as data from a particular source type, from a specific decade, in a specific location, for one gender. The median sample size is just over 900 individuals, while at the extremes we have a handful of samples with less than 100 observations and a few with several million (from mid-nineteenth century census data covering entire countries). In this section we discuss the types of sources useful for the study of numeracy and describe the construction of our dataset.

Age distributions can be recovered from a wide range of sources in early modern Europe. The most comprehensive are of course censuses of population, which aimed at complete coverage across genders, ages, social classes, and locations. Though censuses date as far back as ancient times, early examples often did not inquire about exact age, or reported the data in aggregated form only, for example by five-year age groups. Only from the late eighteenth century do surviving census records commonly provide single year age frequencies. The accuracy of census data – even in modern times – depends on the efforts of enumerators and the details of interview procedures and questionnaire design. And these varied over time and across countries. Our discussion of the IPUMS data (not in our dataset) touched on one example: the question about birth year added to the 1900 U.S. census in hopes of eliciting more accurate age responses. While such variation must be acknowledged as a potential source of spurious variation in numeracy estimates, the criticism applies with equal force to other indicators of human capital, in particular literacy estimates based on signature rates in marriage registers. Some clergy, in some times and places, insisted on signatures, while others did not, and the choice itself was probably not random. Our dataset includes 130 samples drawn from census data, covering 15 countries, and birth decades from 1390 to 1840.[28] The database includes two precocious censuses from Pozzuoli and Sorrento in southern Italy, dating from 1489 and 1561, respectively. Our census data were assembled from previously compiled age distributions in digital or paper format, as detailed in the appendix.

Dating from as early as the sixteenth century, "soul registers" (*libri status animarum*) are another rich source of historic age data, mainly for central Europe. Soul

[28] For an analysis based entirely on census data and covering the period from 1820 to the present day, see Crayen and Baten (2008).

registers were compiled by local clergy interested in gauging the needs and revenue-generating potential of their parishes, as well as the spiritual welfare of their flocks. (Enumeration revealed such phenomena as cohabitation among unmarried couples and the presence of individuals of other faiths.) They aimed at census-like coverage of the relevant ecclesiastical unit. While soul registers were common across a wide range of time and space in both Catholic and Protestant areas, *surviving* records are available only sporadically and therefore yield less systematic evidence than censuses. Having access to birth and baptism registers, clerics sometimes counterchecked and corrected age statements in the soul registers. As a consequence, some soul registers exhibit essentially no age heaping ($W < 105$) even in largely illiterate, rural communities. When soul registers indicate extremely low age heaping, comparison with alternative sources from similar regions is prudent as a check. Our database includes approximately 70 samples drawn from soul registers, covering six central European countries and the UK, over birth decades from 1580 to 1830.[29] Our data, in the main, were originally transcribed and/or digitized by local historians and genealogical researchers, as detailed in the appendix.

The earliest efforts by temporal authorities to collect systematic data on their domains are, unsurprisingly, fiscal in nature. Tax rolls listing the number of hearths (households) were common in medieval times. In some cases, detailed information including the self-reported age of household heads was collected. Our dataset includes nearly 40 samples from fifteenth century Italy drawn from tax registers, principally the famous 1427 *catasto* of Florence and its possessions. The data, comprising over 50,000 individual observations, are taken from the work of Herlihy and Klapisch-Zuber. Tax

[29] Other ecclesiastical sources of data include monastic necrologies and the personnel records of religious orders such as the Jesuits. See Nalle (1989) for a study using the records of the Spanish Inquisition. Age

registers ordinarily provide information on heads of household only. Even within this group, there may be selection by social group if only property owners are included. For both reasons the concern is that fiscal data may suffer positive selection by social class or wealth. The same is true of civil legal documents that report the ages of litigants, witnesses, transactors, or beneficiaries. Our database includes one such sample, Dutch wills of the 1510s. Negative selection may characterize sources deriving from the criminal justice system, such as our eight samples drawn from Oxley's nineteenth century British prison record data. In principle, the occupational data often contained in fiscal and legal records permit a control for social selectivity, but this has not been pursued in this study.

Military records offer copious quantities of data, and come in two forms. Conscription lists offer nearly complete data on young men eligible for military service, a category that could cover a range of (self-reported) ages in the case of episodic levies and early conscription systems. The database includes 2 such samples, one from Southern Germany in the 1520s, the other from Hungary in the 1710s. More common are muster lists, which enumerate the actual personnel serving in military units. The database includes 30 samples from the French army, covering birth decades from 1650 to 1750. These data were drawn from Komlos' height dataset for early modern France. The French army in this period was a volunteer force, in which the lower classes were overrepresented (since non-inheriting sons and others with relatively low opportunity costs had greater incentives to enlist). On the other hand, the social selection of older soldiers, who tended to be either officers or skilled specialists such as physicians, is less clear. And minimum height requirements and other recruiting biases favoring tall (better

---

information from her data could not be used in this study due to small sample sizes.

nourished and healthy) individuals further offset negative social selection. In both types of source, numeracy cannot be assessed for the younger ages due to the distortion typically induced by a uniform age of conscription or minimum enlistment age.

The movement of people also occasioned the recording of personal data including age. By the nineteenth century passports, ship manifests, and the records of immigration authorities are useful sources. Passenger lists from oceanic voyages survive from even earlier periods, as do indenture contracts. Our dataset includes 2 samples from passenger lists, covering several hundred individuals who emigrated from Britain in the 1630s and ’40s. Though there is something of a presumption in the literature that emigrants are positively selected, it is difficult to document this because early ship manifests typically did not record occupations or wealth. Our passenger lists were taken from internet-based genealogical sources (as listed in the appendix).

Death and marriage registers are a further source of age data for the early modern period. In many marriage registers, ages were self-reported and not cross-checked against birth registers. As with military records, accurate assessment of age awareness can be impeded by concentration of betrothed partners in a narrow range of ages. The use of recorded age at death to assess numeracy was pioneered by Duncan-Jones (1990), who studied tombstone inscriptions around the Roman Empire in antiquity. Like tombstones, death registers reflect relatives’ knowledge of an individual’s age, and could thus lead to upwardly biased estimates of heaping levels. Unlike grave monuments, it is unlikely that death registers suffer from positive social selection. Our database includes approximately 20 samples drawn from death registers in Geneva, Berlin, Paris, Milan, and the Polish town of Radzionkow, over birth decades from 1650 to 1820. Sources are given in the appendix.

The extraordinary breadth of our database (8 source types, 49 birth decades, 130 locations in 16 countries) inevitably implies that coverage is also thin, and that the data are heterogeneous. A continuous run of samples over many decades, from a representative set of locations in a particular country, drawn from the same source type, is very much the exception in the numeracy database. Our samples cover men and women, urban and rural places, various age-ranges, and source types with diverse potential biases. In this exploratory study we have opted for simplicity, transparency, and uniformity in procedures for generating age heaping estimates.

Age itself proved the thorniest issue. A preliminary and fairly straightforward decision was to restrict the analysis to ages 23 to 72. The young were dropped because several of our sources, in particular the military data, have such a concentration of frequency in a narrow range of ages around 20 that age heaping measures are distorted. The old were dropped because of the tendency to deliberately exaggerate age, along with small sample sizes. More problematic was the aging effect identified in the IPUMS sample, which is also evident in our early modern data. Table 5 presents Whipple indices (*W*) calculated separately for 10-year age-groups from 23-32 to 63-72, for five typical samples with complete information, representing various countries and source types, and both genders. The general tendency of age-heaping to increase with age is apparent. Because mean age varies considerably across samples, and because in some cases just one sample represents a particular country in a particular period, some type of age-standardization is necessary if age-heaping comparisons are not to mislead.[30]

A natural solution is to regress W on age-group dummies, with controls for time, gender, location, country, and source type. Estimated coefficients could then be used to standardize each sample to a common basis, for example males aged 33-42 from census data sources. But the aging effect is far from uniform, as the figures in Table 5 suggest.

The only consistent pattern that can be discerned in the database as a whole is that the aging effect is weaker, in both absolute and proportional terms, the lower is the baseline level of heaping in the 23-32 group. Meanwhile, gender gaps vary across countries in magnitude, trend, and even sign. Source type correlates with time and place: census data come mostly from the end of the sample; military records are mostly French; tax register data are early and Italian. And of course the hypothesis that time trends vary across countries is the very motivation for the study. This heterogeneity undermines efforts to identify a parsimonious model capable of fitting the data well.

We opt instead for an approach that is simple and respects patterns specific to time and place as illustrated by Figure 7. We begin by completing all samples with missing age groups – roughly half the total.[31]

When at least two age groups are observed, missing values are filled in by linear interpolation or extrapolation, subject to the constraints that W be nondecreasing in age and be bounded by 100 and 500. For the handful of samples with only a single age-group, we impute the missing data on the basis of a regression of W on age-group dummies, a gender dummy, and a time trend, using data from a neighborhood around the sample,

---

[30] The interdecile range of sample mean age, based on frequency-weighted averages of age-group midpoints, is from 30 to 49 years.

defined as the same country within plus or minus 150 years.[32] An overall sample value for W is then calculated as a simple average of the age-group specific figures.

In aggregating samples, we take simple averages across all locations and genders[33], by country and half-century of birth. We lack data on within-country variation in urbanization, economic structure, or income that might permit the identification of geographic patterns such as urban-rural disparities. Gender differences can be identified but, as discussed, are too varied to permit any uniform treatment or standardization. Though a promising topic for further research, in this paper gender gaps disappear into a single aggregate.[34] All numeracy figures ($\tilde{W}$) discussed in the remainder of the paper are based on such 50-year, countrywide averages. Table 6 presents the database in summary form.

VII. Age heaping and literacy in pre-industrial Europe

In late nineteenth century data, age heaping-based numeracy is well correlated with literacy across US states and European countries, particularly in contexts where both were low. Can the same be said of early modern Europe? To answer this question, we culled literacy estimates for the period before 1850 from a wide range of secondary sources (detailed in the appendix). The original sources of these literacy estimates are predominantly signature rates – mostly those of spouses and witnesses in parish marriage registers, but in some cases deriving from court proceedings or other legal documents.

---

[31] Age groups with fewer than 50 observations are treated as missing in this process.
[32] For an alternative approach to age-standardization that exploits the greater regularity of aging effects in recent historical data, see Crayen and Baten (2008).
[33] For some 30 samples the sexes are mixed or no indication regarding gender is given in the source.

The limitations of signature ability as a measure of functional literacy are obvious, but the same can be said of self-reported “ability to read” in census data. In practice signature and reading ability are well correlated where both can be observed. In historical curricula, reading instruction was largely completed before writing was started. Perhaps even more than the age-heaping based numeracy figures, the literacy estimates tend to be based on small samples, the representative nature of which is difficult to judge. And when the secondary sources consulted omitted data on age or birth year, it was necessary to assume an average age based on the source type (e.g. mid-twenties for spouses) in order to assign the literacy estimate to the relevant birth half-century. For 28 country – birth half-century pairs we observe both literacy and numeracy. These points are plotted in Figure 7.

The expected positive correlation between age heaping-based numeracy and signature-based literacy is evident in Figure 8. The data points would not look out of place in Figure 6, though lower literacy levels would situate them in the lower portion of the graph. At 0.86, the slope of the fitted regression line is considerably less than in the nineteenth-century data, as is the R^2 of 0.40 (cf. Table 4). These differences are consistent with greater measurement error in the early modern data. Numeracy and literacy also move together *within* countries over time. Figure 9 again plots literacy against numeracy, connecting successive (not necessarily adjacent) birth half-centuries for the same country with arrows indicating the direction of time. In general, countries move in a northeasterly direction, indicating simultaneous improvement on both measures of human capital. Protestant Germany, France, Denmark, and the United Kingdom show this pattern. The

[34] The sources are well balanced by gender, with the exceptions of the entirely male French muster lists and predominantly male Italian tax registers.

only clear exception, Protestant Germany between 1750 and 1800, can be attributed to a change in the underlying sources.[35]

A more meaningful variation on the pattern is the dependence of the path followed on the starting point. Three countries start from unusually low levels of numeracy ($\tilde{W} < 60$): Northern Italy (from 1450), Hungary (from 1650), and Russia (for which only the 1650 value is plotted due to the small sample size in 1800). All three show dramatic improvements in estimated numeracy, with little or no improvement in literacy. Such development is consistent with the ideas that age-awareness measures a very basic level of numeracy, and that numeracy can be acquired without formal schooling. Countries that depart from higher levels of numeracy, by contrast, improve literacy dramatically, without being able to generate similar improvements in numeracy. This is consistent with age-heaping innumeracy having been reduced to a "hard core" of relatively inaccessible regions or social groups.

VIII. New estimates of human capital in the very long run

Our age-heaping based numeracy estimates are plotted against time in Figure 10. It is apparent that Western Europe already enjoyed high rates of numeracy in the early modern era. By the period around 1600, the Netherlands, Britain, Northern Italy, and probably France boasted figures of 70% or more. By 1700, numeracy rates are in the vicinity of 90% for these countries, as well as for Scandinavia. The individual human capital, collective number discipline, and administrative capacity that age heaping reflects

---

[35] Protestant Germany's numeracy estimate for 1750 is based on census data for the city of Kiel and the town of Kellinghusen, while that for 1800 is based on death registers from Berlin.

were evidently well-developed long before the spread of schooling, the rise of literacy, or the industrial revolution. In Northern Italy, we observe improvements in numeracy as far back as the Middle Ages, starting from levels in the fourteenth century that are not far removed from (and probably lower than) those that had prevailed a millennium earlier in Roman times.[36] It is natural to speculate that Italy experienced a long stagnation or decline following the collapse of the Empire, followed by turnaround associated with medieval state-building or commercial expansion, and slow but steady progress into the early modern period. The cases of Southern Italy and Ireland (within the UK) suggest the possibility of substantial and persistent regional disparities, that of Belgium a capability for rapid catch-up.

The German-speaking lands of central Europe present a varied picture (Figure 10b). Both Austria and the Protestant areas of what would later become Germany have impressive numeracy rates circa 1600, on a par with those of Western European countries at the same time. In Germany as in Italy, this seems to have resulted from a long period of improvement beginning in the Middle Ages. The Catholic areas of Germany, along with Switzerland, have lower numeracy rates when we first observe them around 1700, but display strong convergence thereafter.[37] The ethnic Germans of Prussian Silesia (in modern Poland) fit the overall pattern with high levels of numeracy in the eighteenth century. The development of basic numeracy in Eastern Europe contrasts sharply with that in the West. In the mid-1600s, Bohemia, Hungary, and Russia have numeracy rates

---

[36] Estimates of $\tilde{W}$ for Roman Italy can be derived from Duncan-Jones' (1990, pp. 86, 90) data on grave monument inscriptions primarily from the Principate period, roughly 0-200 C.E., to a lesser extent from later dates. Averaging over both men and women, the city of Rome and the rest of Italy, $\tilde{W}$ = 55%.

[37] The earliest Swiss estimate, for a town near Zürich in 1600, would put Switzerland on a par with Germany and Austria, but is based on a very small sample and therefore not plotted (but see Table 6). The somewhat lower figure for 1700 is based on Geneva death registers.

in the range from 30 to 45%. These are some of the lowest figures observed in the dataset, similar to levels observed in the eastern provinces of the Roman Empire circa 200 C.E., or to those of medieval Northern Italy.[38] It is interesting to note that East-West disparities are greatest when we first observe them, in the early seventeenth century. Any divergence between the two would appear to have taken place already in the Middle Ages and Renaissance, rather than the later era of "second serfdom" in the East. That period instead saw convergence in numeracy rates, at least from 1650 and among the Habsburg dominions (Bohemia, Hungary), possibly in Russia as well.[39] In Serbia, still a province of the Ottoman Empire in the eighteenth century, numeracy remained as low as 55% even in the decades around 1800, lagging Western Europe by more than two centuries.

IX. Conclusion

It is difficult to imagine capitalism without calculation. Rates of interest, profit, and exchange were at the heart of things long before engineering or science became important in generating technical change. So measurement of numerical cognitive abilities is a natural complement to literacy in the history of human capital. As signature rates, despite their limitations, can proxy for basic literacy, so age heaping can be used as an index of basic numeracy. Both measures offer a partial view of human capital, and both reflect not only individual but also broader social capabilities – what might be called administrative capital. They are well correlated, when both can be observed. This is true at both the

---

38 Data are again from Duncan-Jones. $\tilde{W}$ ranges from 30 to 45% in four eastern provinces corresponding roughly to Bulgaria, Croatia, Hungary, and Romania. (The Roman provinces were Moesia, Dalmatia, Pannonia, and Dacia.)

aggregate and individual levels in late nineteenth-century data. In earlier periods, numeracy estimates confirm signature rate studies in identifying a high level of human capital in the Netherlands, United Kingdom, and Scandinavia on the eve of the industrial revolution. They also yield the new insight that, starting from low levels of human capital, basic numeracy in some cases increased significantly before substantial improvements in literacy were achieved.

The primary contribution of age-heaping based numeracy estimates to the study of human capital, however, is to extend coverage to times and places where signature based literacy estimates are not available. Our new dataset illustrates the wide range of primary sources available, and takes a first step in extending human capital estimates back into the Middle Ages and east into Central and Eastern Europe. While our findings must remain tentative at this early stage, they are certainly intriguing.

In the early Middle Ages, numeracy may have remained at or below levels observed in antiquity, a thousand years earlier. In Western Europe and in the German speaking lands of Central Europe, a slow but steady improvement appears to have set in by late medieval times, bringing several countries to numeracy rates on the order of 70% by the decades around 1600. This period appears to have been the point of maximum divergence from Eastern Europe, where numeracy rates were only beginning to rise from very low levels. The development of numerical cognitive abilities and administrative capabilities thus had deep roots – long predating the spread of mass schooling – and displayed persistent variation across countries. The early and substantial rise of numeracy

---

39 Not plotted in Figure 9 is the Russian estimate of 89% for the birth half-century around 1800, based on a very small sample (Table 6). Czech estimates based on similarly small samples indicate that the improvement in numeracy in Bohemia was mostly complete by 1700 (Table 6).

in the West lends strong support to accounts that assign human capital the role of cause, rather than effect, of the industrial revolution.

References

Allen, R.C., 2003. Progress and poverty in early modern Europe. Explorations in Economic History 56 (3), 403-443.

Bachi, R., 1951. The tendency to round off age returns: measurement and correction. Bulletin of the International Statistical Institute 33, 195-221.

Barro, R.J., 1997. Determinants of Economic Growth : A Cross-country Empirical Study. MIT Press, Cambridge MA.

Barro, R.J., Lee, J.-W., 1994. Sources of economic growth (with commentary). Carnegie-Rochester Conference Series on Public Policy 40, 1-57.

Baten, J., van Zanden, J.L., 2006. The impact of human capital on early modern welfare growth, 1450-1913: book production as a human capital indicator. University of Tübingen Working Paper.

Becker, G.S., Murphy, K.M., Tamura, R., 1990. Human capital, fertility, and economic growth. Journal of Political Economy 98, 12-37.

Beloch, K.J., 1937. Bevölkerungsgeschichte Italiens I. Walter De Gruyter, Berlin.

Bils, M., Klenow, P.J., 2001. Does schooling cause growth? American Economic Review 90 (5), 1160-1183.

Budd, J.W., Guinnane, T., 1991. Intentional age misreporting, age heaping and the 1908 old age pensions act in Ireland. Population Studies 45 (3), 497-518.

Carruthers, B.G., Espeland, W., 1991. Accounting for rationality: double-entry bookkeeping and the emergence of economic rationality. American Journal of Sociology 97 (1), 31-96.

Cervellati, M., Sunde, U., 2005. Human capital formation, life expectancy and the process of development. American Economic Review 95 (5), 1653-1672.

Charette, M.F., Meng, R.,1998. The determinants of literacy and numeracy, and the effect of literacy and numeracy on labour market outcomes. Canadian Journal of Economics 31 (3), 495-517.

Chiswick, B., Lee, Y.L., Miller, P.W., 2003. Schooling, literacy, numeracy and labour market success. The Economic Record 79 (245), 165-181.

Coale, A., Kisker, E.E., 1986. Mortality crossovers: reality or bad data? Population Studies 40, 389-401.

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

Crayen, D., Baten, J., 2008. Global Trends in Numeracy 1820-1949 and its Implications for Long-Run Growth. CESifo Working Paper Series No. 2218.

Crockett, A., Crockett, R., 2006. Consequences of Data Heaping in the British Religious Census of 1851. Historical Methods 39(1), 24-46.

Denic, S., Khatib, F., Saadi, H., 2004. Quality of age data in patients from developing countries. Journal of Public Health 26 (2), 168-171.

Dillon, L., 2007. The Shady Side of Fifty: Age and Old Age in Late Victorian Canada and the United States. Montreal: McGill-Queen's University Press.

Duncan-Jones, R., 1990. Structure and Scale in the Roman Economy. Cambridge University Press, Cambridge.

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

Ewbank, D.C., 1981. Age Misreporting and Age-selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis. National Academic Press, Washington, D.C.

Finnie, R., Meng, R., 2001. Cognitive skills and the youth labour market. Applied Economic Letters 8, 675-679.

Galor, O., Weil, D.N., 2000. Population, technology and growth: from the Malthusian regime to the demographic transition. American Economic Review 110, 806-828.

Goldthwaite, R.A., 1972. Schools and teachers of commercial arithmetic in Renaissance Florence. Journal of European Economic History 1, 418-433.

Graff, H., 1987. The Legacies of Literacy: Continuities and Contradictions in Western Culture and Society. Indiana University Press, Bloomington IN.

Hanushek, E.A., Kimko, D.D., 2000. Schooling, labor-force quality, and the growth of nations. American Economic Review 90(5), 1184-1208.

Heckman, J.J., 1995. Lessons from the Bell Curve. Journal of Political Economy 103 (5), 1091-1120.

Herlihy, D., Klapisch-Zuber, D.,1978. Les Toscans et leurs familles: une étude du catasto florentin di 1427. Editions de l'Ecole des Hautes Etudes en Sciences Sociales / Presses de la Fondation Nationale des Sciences Politiques, Paris. Chap. XIII.

Jowett, A.J., Li, Y-Q., 1992. Age heaping: contrasting patterns from China. GeoJournal 28 (4), 427-442.

Kaiser, D., Engel, P., 1993. Time and age-awareness in early modern Russia. Comparative Studies in Society and History 35 (4), 824-839.

Kwasny Z., 1996. Ksiega slubów parafii radzionkowskiej z lat 1801 - 1850 / wydal i wstepem opatr. Zbigniew Kwasny. - Wroclaw : Wydawn. Uniw. Wroclawskiego.

Kwasny, Z., 1998. Ksiega pogrzebów parafii radzionkowskiej z lat 1811 - 1850 / wydal Zbigniew Kwasny. - Wroclaw : Wydawn. Uniwersytetu Wroclawskiego.

Levine, R., Renelt, D., 1992. A sensitivity analysis of cross-country growth regressions. American Economic Review 82, 942-963.

Long, J., 2005. Rural-urban migration and socioeconomic mobility in Victorian Britain. Journal of Economic History 65 (1), 1-35.

Long, J., 2006. The economic return to primary schooling in Victorian England. Journal of Economic History 66(4), 1026-1053.

Lucas, R., 2002. Lectures on Economic Growth. Harvard University Press: Cambridge, MA. Ch. 5.

Mankiw, N.G., Romer, D., Weil, D.N., 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics 107 (2), 19-36.

Mokyr, J., 1983. Why Ireland Starved. George Allen & Unwin, London. Chap.8.

Mols, R., 1956. Introduction a la demographie Historique des villes d'Europe du XIVe au XVIIIe siècle. Tome Troisieme, Annexes,Université de Louvain.

Murnane, R.J., Willet, J.B., Levy, F., 1995. The Growing importance of cognitive skills in wage determination. Review of Economics and Statistics 77(2), 251-266.

Murphy, K. M., Shleifer, A.,Vishny, R.W., 1991.The allocation of talent: implications for growth. Quarterly Journal of Economics 106 (2), 503-530.

Myers, R., 1976. An instance of reverse heaping of ages. Demography 13 (4), 577-580.

Myers, R., 1954. Accuracy of age reporting in the 1950 United States census. Journal of the American Statistical Association XLIX, 826-831.

Nalle, S., 1989. Literacy and Culture in Early Modern Castile. Past and Present 125, 65-96

Neal, D.J., Johnson, W.R., 1996. The role of pre-market factors in black-white wage differences. Journal of Political Economy 104 (5), 869-895.

O'Grada, C., 2006. Dublin Jewish demography a century ago. UCD Centre for Economic Research Working Paper Series WP 06/01.

O'Rourke, K., Williamson, J.G., 1997. Around the European periphery 1870-1913: globalization, schooling and growth. European Review of Economic History 1, 153-190.

Parsons, S., Bynner, J., 2005. Does Numeracy Matter More? National Research and Development Centre for Adult Literacy and Numeracy, London.

Preston, S.H., Elo, I.T., Rosenwaike, I., Hill, M., 1996. African-American Mortality at Older Ages: Results of a Matching Study. Demography 33 (2), 193-209.

Reis, J., 2005. Economic growth, human capital formation and consumption in western Europe before 1800. In: Allen, R.C., Bengtsson, T., Dribe, M. (Eds.), Living Standards in the Past. Oxford University Press, Oxford, pp. 195-225.

Rivera-Batiz, F., 1992. Quantitative literacy and the likelihood of employment among young adults in the Unites States. Journal of Human Resources 27, 313-328.

Sala-i-Martin, X., 1997. I just ran 4 million regressions. National Bureau of Economic Research Working Paper no. 6252.

Sanderson, M., 1999. Education and Economic Decline in Britain, 1870 to the 1990s. Cambridge University Press, Cambridge.

Shryock , H.S., Siegel, J.S., 1973. The Methods and Materials of Demography. U.S. Bureau of the Census, Washington, D.C. Chap. 8.

Swetz, F.J., 1987. Capitalism and Arithmetic: The New Math of the 15th Century. Open Court Publishing Company, La Salle.

Tortella, G., 1994. Patterns of economic retardation and recovery in south-western Europe in the nineteenth and twentieth centuries. Economic History Review 47 (1), 1-21.

United Nations Statistics Division, 2003. Nonsampling Errors in Surveys. United Nations Secretariat, New York.

Vallin, J., Meslé, F., Adamets, S., Pyrozhkov, S., 2002. A new estimate of Ukrainian population losses during the crises of the 1930s and 1940s. Population Studies 56 (3), 249 – 264.

Zelnik, M., 1961**.** Age Heaping in the United States Census: 1880-1950. The Milbank Memorial Fund Quarterly 39 (3), 540-73.

Tables

Table 1. Probabilities of ranking errors, adjacent degrees of mixed heaping

	Sample size 250					*Sample size* 1000			
Index	*D1*	*D2*	*D3*	*D4*		*D1*	*D2*	*D3*	*D4*
ABC	.37	.27	.18	.20		.11	.04	.03	.04
Lambda	.36	.27	.22	.24		.11	.06	.04	.04
Bachi	.37	.24	.22	.22		.08	.05	.03	.03
Myers	.37	.31	.24	.24		.14	.10	.05	.07
M. of 5	.23	.24	.24	.30		.05	.09	.09	.12
Whipple	.16	.16	.17	.19		.02	.02	.02	.04

Notes: Estimated probabilities that *D1* etc. < 0; $D1 = H^{05-05} - H^{00-00}$, $D2 = H^{05-10} - H^{05-05}$ etc.

Table 2. Probabilities of ranking errors, two degrees of mixed heaping apart

	Sample size 250					*Sample size* 1000			
Index		*D20*	*D31*	*D42*			*D20*	*D31*	*D42*
ABC		.17	.06	.04			.00	.00	.00
Lambda		.17	.08	.05			.00	.00	.00
Bachi		.15	.08	.04			.00	.00	.00
Myers		.20	.10	.07			.00	.00	.00
M. of 5		.09	.08	.10			.00	.01	.01
Whipple		.03	.02	.03			.00	.00	.00

Notes: Estimated probabilities that *D20* etc. < 0; $D20 = H^{05-10} - H^{00-00}$, $D31 = H^{05-15} - H^{05-05}$ etc.

Table 3. Regional literacy and Numeracy in 19th century US censuses

Sample	*Numeracy coefficient*	R^2	*N*	*Mean literacy*	*Mean numeracy*
All	1.97 (0.12)	0.68	213	80.7	87.3
Census of 1850	1.67 (0.15)	0.79	41	81.5	83.0
Native whites	2.01 (0.06)	0.88	25	86.0	87.6
Native blacks	1.58 (0.16)	0.69	7	58.5	69.6
Foreign born	1.79 (0.44)	0.77	9	87.0	80.3
Census of 1870	2.64 (0.15)	0.83	84	76.7	83.6
Native whites	1.86 (0.62)	0.69	34	89.2	88.5
Native blacks	3.18 (0.52)	0.72	20	34.9	70.0
Foreign born	1.38 (0.54)	0.48	30	90.4	87.0
Census of 1900	2.79 (0.15)	0.77	88	84.1	92.3
Native whites	1.85 (1.01)	0.36	42	93.6	95.4
Native blacks	2.72 (0.40)	0.82	17	62.0	86.0
Foreign born	2.50 (0.33)	0.65	29	83.4	93.2
Native whites	1.62 (0.26)	0.65	101	90.3	91.2
Native blacks	1.84 (0.20)	0.59	44	49.1	76.1
Foreign born	1.00 (0.24)	0.26	68	86.9	88.8
Low literacy	1.58 (0.18)	0.55	64	51.4	77.3
High literacy	0.43 (0.08)	0.15	149	93.3	91.6
Low numeracy	2.20 (0.23)	0.55	80	61.4	77.6
High numeracy	0.67 (0.20)	0.08	133	92.3	93.1

Notes: Estimated slope coefficients from OLS regressions of birthplace literacy rates on birthplace numeracy rates; birthplaces with samples of at least 100; robust standard errors in parentheses; high/low literacy and numeracy are above/below the sample means of 80.1 and 87.3, respectively; IPUMS data, ages 20-69.

Table 4. Marginal effects on the probability of reporting a heaped age, IPUMS data

	1850			1870			1900		
	Native white	*For.*	*Native black*	*Native white*	*For.*	*Native black*	*Native white*	*For.*	*Native black*
Personal literacy	-.008 (.006)	-.058 (.013)	-.007 (.008)	-.037 (.005)	-.057 (.008)	-.027 (.007)	-.032 (.006)	-.029 (.008)	-.062 (.007)
Regional literacy	-.002 (.000)	-.002 (.001)	-.002 (.001)	-.002 (.000)	-.002 (.000)	-.002 (.000)	-.001 (.000)	-.002 (.000)	-.002 (.000)
Female	.009 (.003)	-.013 (.008)	-.013 (.023)	.015 (.003)	-.005 (.005)	.063 (.005)	.002 (.002)	.003 (.004)	.015 (.007)
Age < 30	-.063 (.004)	-.166 (.009)	-.178 (.027)	-.065 (.003)	-.139 (.006)	-.211 (.006)	-.017 (.003)	-.057 (.006)	-.086 (.008)
Age ≥ 40	.032 (.004)	.073 (.010)	.098 (.028)	.007 (.003)	.073 (.005)	.105 (.006)	.017 (.003)	.029 (.005)	.074 (.009)
Irish		.113 (.012)			.123 (.007)			.062 (.006)	
Av.depvar	.285	.377	.436	.273	.350	.477	.232	.251	.334
N	73,381	15,485	2,076	124,304	45,172	43,145	136,341	42,241	20,623
Pseudo-R^2	.009	.048	.047	.007	.046	.061	.002	.011	.027
χ^2 LR	768.1	987.6	134.7	1059.1	2717.6	3632.3	250.0	535.3	701.0

	All	1850 *all*	1870 *all*	1900 *all*	*Native White*	*For.*	*Native Black*	*Low literacy*	*High literacy*
Personal Literacy	-.037 (.002)	-.019 (.005)	-.041 (.004)	-.048 (.004)	-.026 (.003)	-.048 (.005)	-.035 (.005)	-.039 (.003)	-.025 (.004)
Regional Literacy	-.002 (.000)	-.003 (.000)	-.002 (.000)	-.002 (.000)	-.002 (.000)	-.024 (.000)	-.003 (.000)	-.002 (.000)	-.003 (.000)
Female	.0107 (.001)	.0027 (.003)	.0206 (.002)	.0042 (.002))	.0085 (.002)	.006 (.003)	.047 (.004)	.031 (.003)	.003 (.001)
Age < 30	-.076 (.002)	-.085 (.004)	-.116 (.003)	-.031 (.002)	-.045 (.002)	-.108 (.004)	-.175 (.005)	-.147 (.003)	-.045 (.002)
Age ≥ 40	.032 (.002)	.034 (.004)	.039 (.003)	.026 (.002)	.016 (.002)	.042 (.003)	.096 (.005)	.069 (.003)	.019 (.002)
Irish						.123 (.004)			
av.depvar	.292	.304	.331	.246	.259	.313	.431	.372	.254
N	502,768	90,942	212,621	199,205	334,026	102,898	65,844	148,131	354,637
Pseudo-R^2	.026	.015	.040	.009	.006	.033	.058	.042	.004
χ^2 LR	15,727.2	1716.8	10,876.5	2,043.1	2,149.2	4,271.9	5,250.9	8,293.5	1,652.8

Notes: Marginal effects of estimated coefficients of logistic regression, evaluated at means of all variables; IPUMS data on adults aged 20-69 from birthplaces with at least 100 observations; marginal effects in gray are not statistically significant at the 3% or better level; degrees of freedom for the chi-squared likelihood ratio test are 5 or 6 depending on the model, but in all cases the *p*-value is zero to three decimal places.

Table 5. Age-group and age-heaping in 5 complete samples

	Sample				
Age group	3	79	279	400	464
23-32	163	105	226	125	252
33-42	191	107	288	119	247
43-52	248	111	394	138	248
53-62	245	110	388	161	252
63-72	218	107	403	149	364

Notes: Whipple Index values by 10-year age-groups in five representative samples with data for all age-groups: 3 – Soul register of Amland, Austria, 1730, women; 79 – Census of Belgium, 1860, men; 279 – Catasto of Castiglione Fiorentino, Italy,1427, men; 400 – Census of Lister og Mandal, Norway, women; 464 – Census of Tula, Russia, 1715 and '20, men.

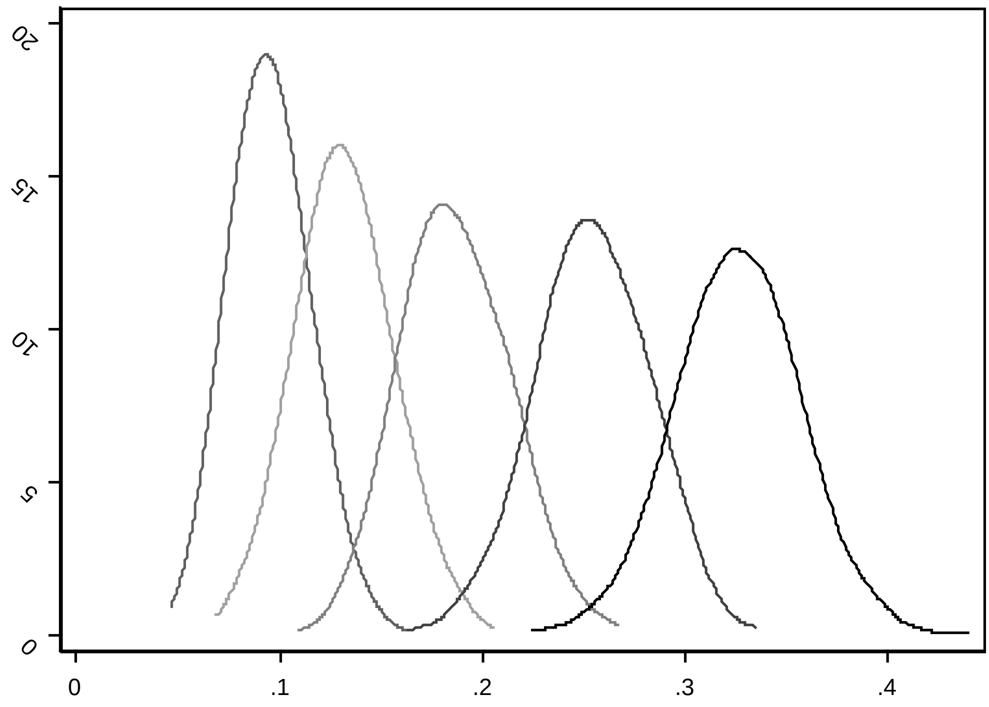
Table 6. Age-heaping based estimates of numeracy ($\tilde{W}$)

	Birth half-century centered on									
	1350	1400	1450	1500	1550	1600	1650	1700	1750	1800
Austria						84		81	86	96
Belgium								72		98
Bohemia						44		*85*	*85*	84
Denmark								90	96	100
France							79	89	93	96
Germany										
Protestant			40			*78*	*79*	87	96	88
Catholic								68	86	
Hungary							32	59	66	
Ireland										77
Italy										
North	31	46	55						89	87
South			18	26						
Netherlands			72							98
Norway									93	96
Poland									*94*	91
Russia							43			*89*
Switzerland						*87*		66		98
Serbia										55
United Kingdom						76		*93*	*93*	93

Notes: Average values of the alternative Whipple index ($\tilde{W}$) described in text, for all locations and both genders, by country and half-century of average birth year (e.g. "1600" = 1575–1624); estimates in italics are based on samples of less than 500.

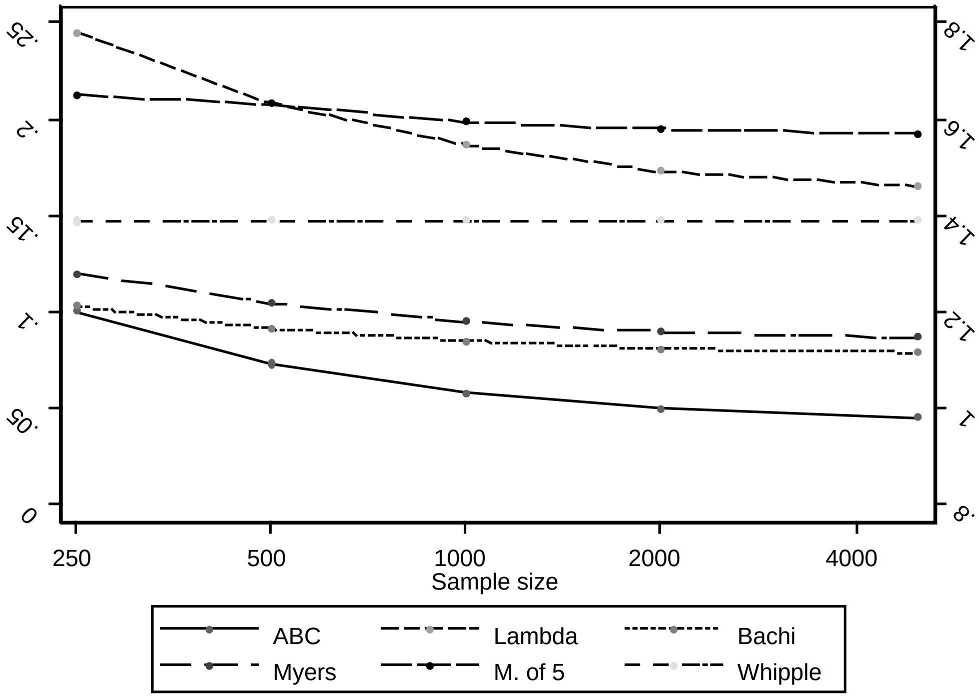
Figures

Figure 1. Distribution of the Lambda Index under increasing degrees of mixed heaping



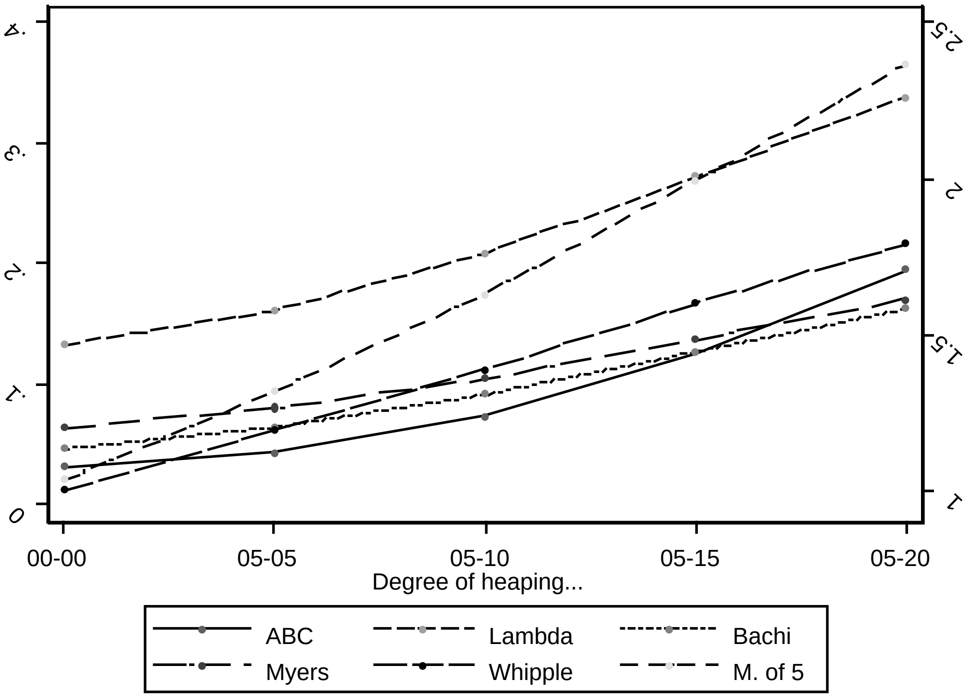
Notes: Kernel density estimates (Epanechnikov kernel, bandwidth .01); mixed heaping schemes as described in text; sample size 1000; repetitions 1000.

Figure 2. Scale dependence under mixed heaping



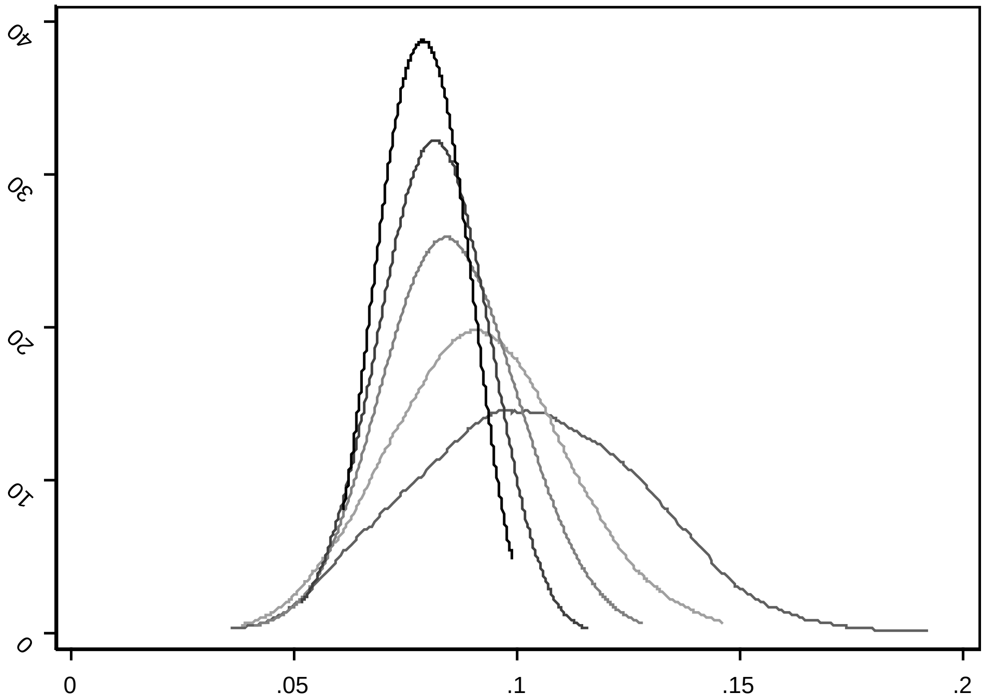
Notes: Mean values of indices under mixed heaping scheme 05-10 for indicated sample sizes; Multiples of 5 and Whipple indices plotted against right axis; log scale x-axis; 1000 repetitions.

Figure 3. Index response to degree of heaping



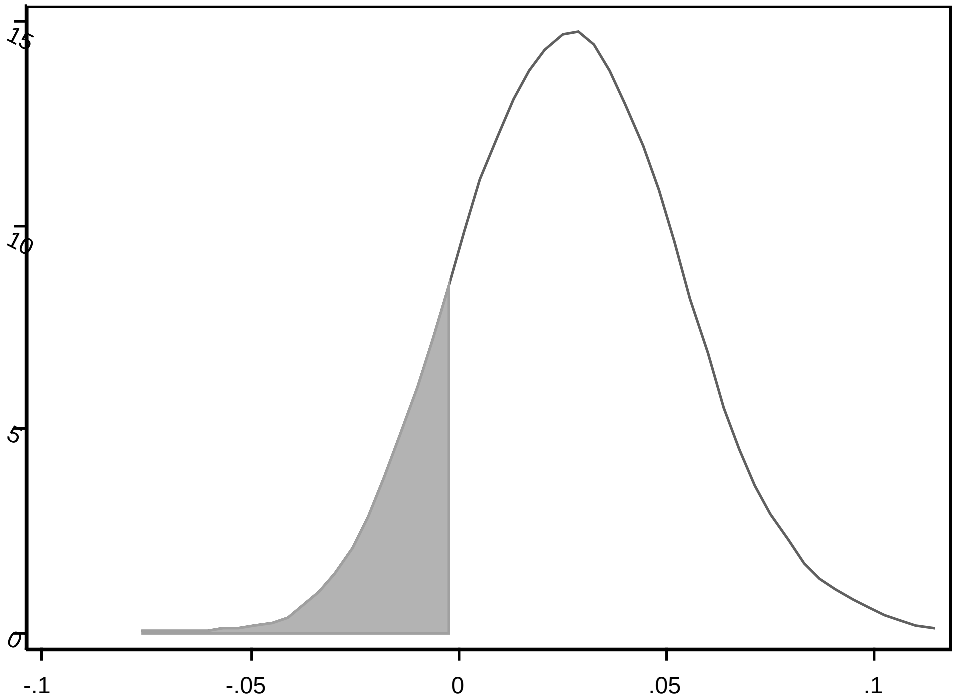
Notes: Mean values of indices under mixed heaping schemes as indicated on horizontal axis and described in text; sample size 500; Multiples of 5 and Whipple Indices plotted on right axis; 1000 repetitions.

Figure 4. Distribution of the Bachi Index for various sample sizes



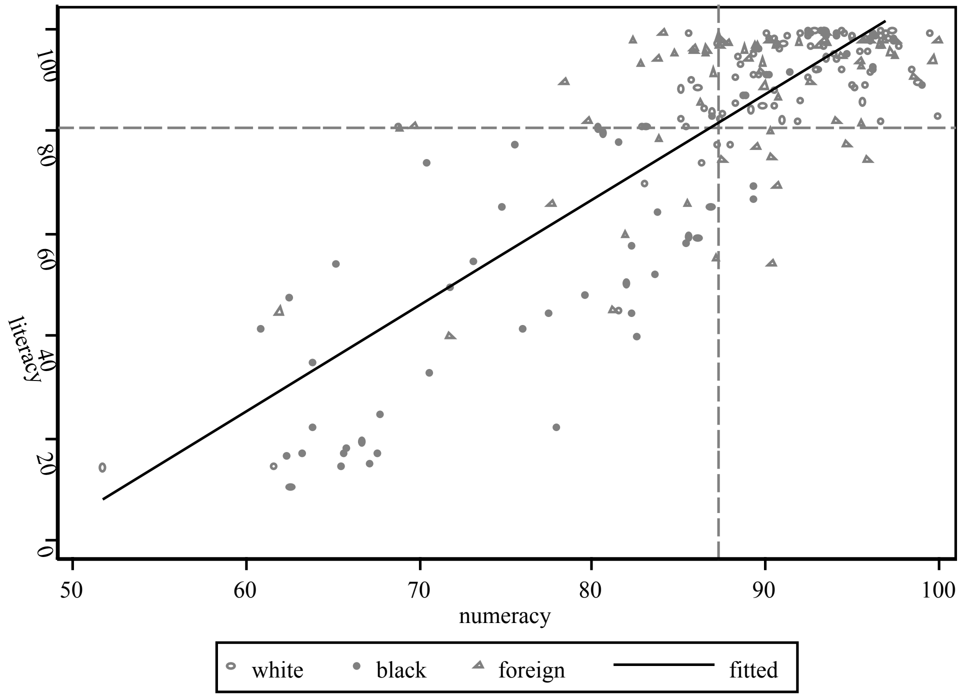
Notes: Kernel density estimates (Epanechnikov kernel, bandwidth .075); Sample sizes 250, 500, 1000, 2000, 5000; Mixed heaping type 05-10.

Figure 5. Bachi Index probability of ranking error



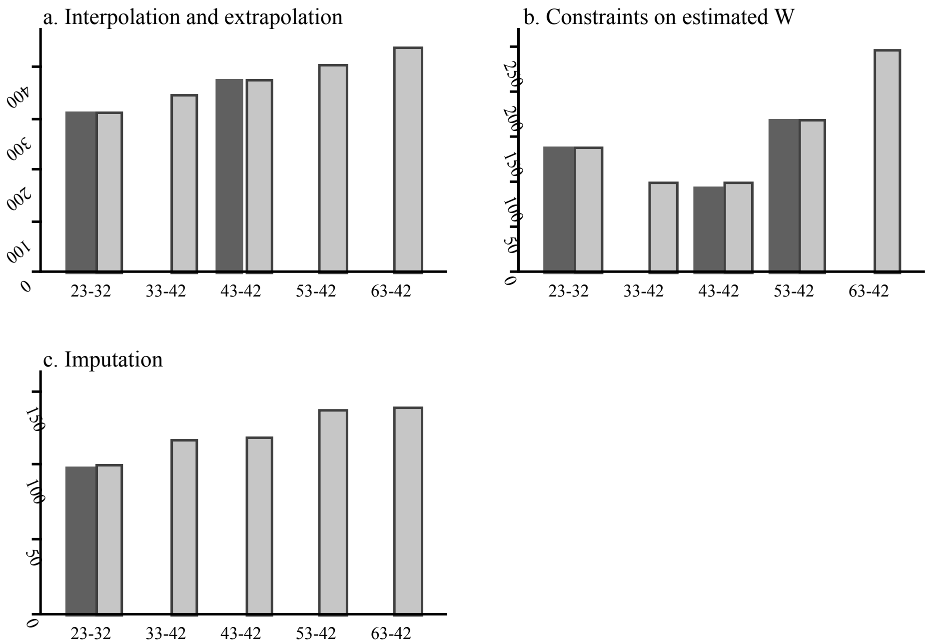
Notes: Kernel density estimate of the distribution of difference between Bachi Indices for random samples from 05-10 and 05-05 populations (Epanechnikov kernel, bandwidth .01); Sample size 500.

Figure 6. Literacy and numeracy in three U.S. censuses



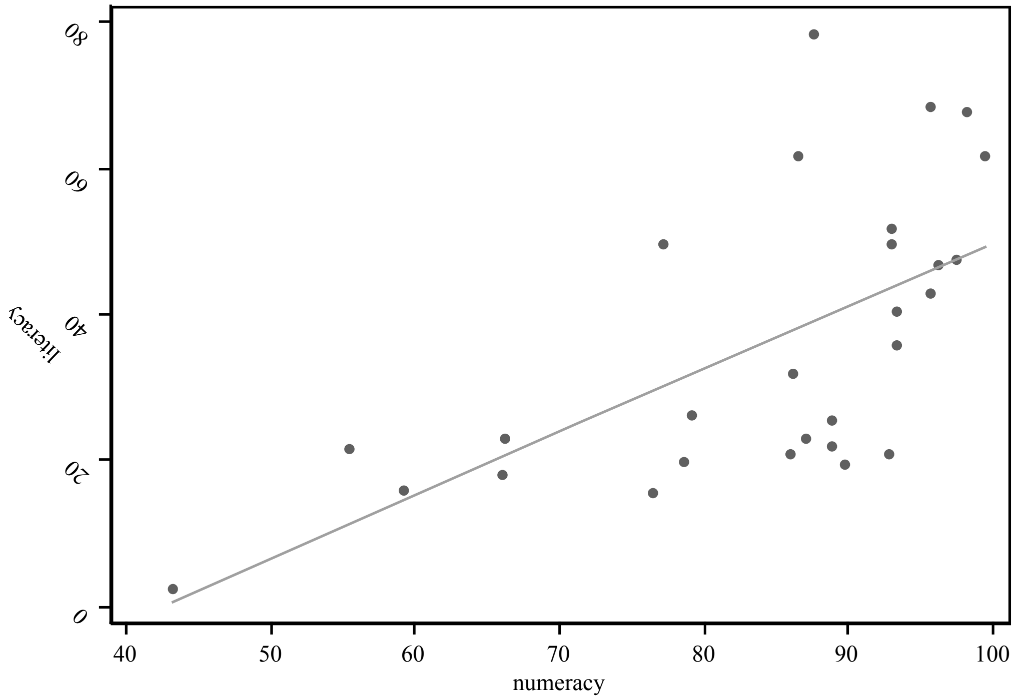
Notes: Alternative Whipple Index and literacy rate by birth regions (US states and territories, foreign countries and provinces) with at least 100 observations; IPUMS data; fitted values based on OLS regression; dashed lines indicate sample means.

Figure 7. Estimation of missing age-heaping data in three incomplete samples



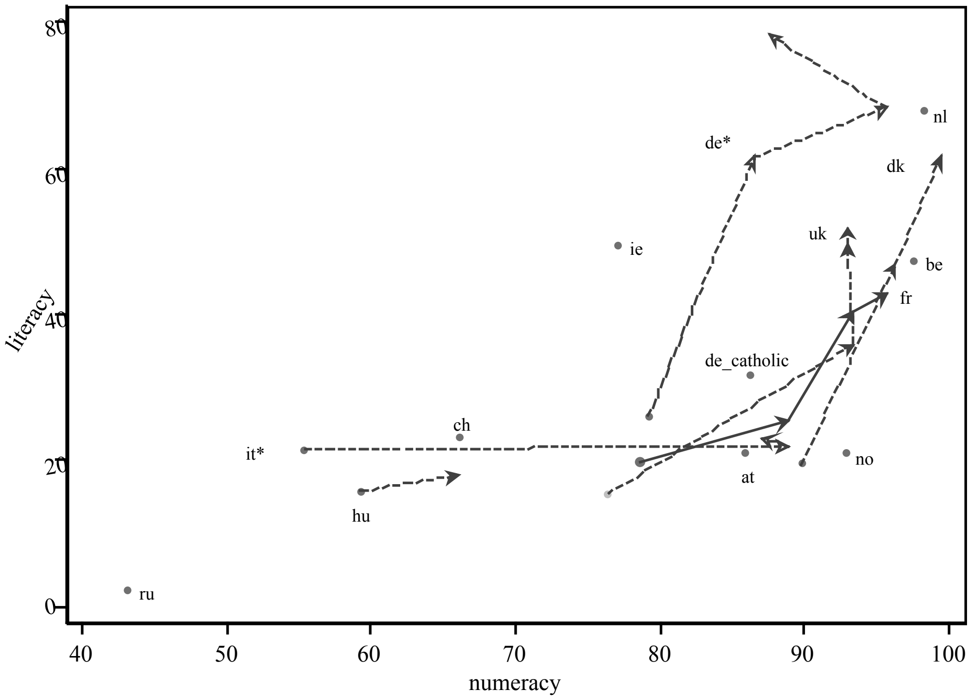
Notes: Values of the Whipple Index by age-group; dark bars observed, light bars estimated by interpolation, extrapolation, or imputation, as described in text. Panel a: soul register, northwestern Bohemia, ca. 1600. Panel b: census data for Eckernförde, Denmark, ca. 1700; the constraint that $\hat{W} \geq 100$ binds on the 43-52 age group, the constraint that $\hat{W}$ be non-decreasing in age binds on the 33-42 group. Panel c: army muster lists, central France, ca. 1750.

Figure 8. Literacy and numeracy in early modern Europe



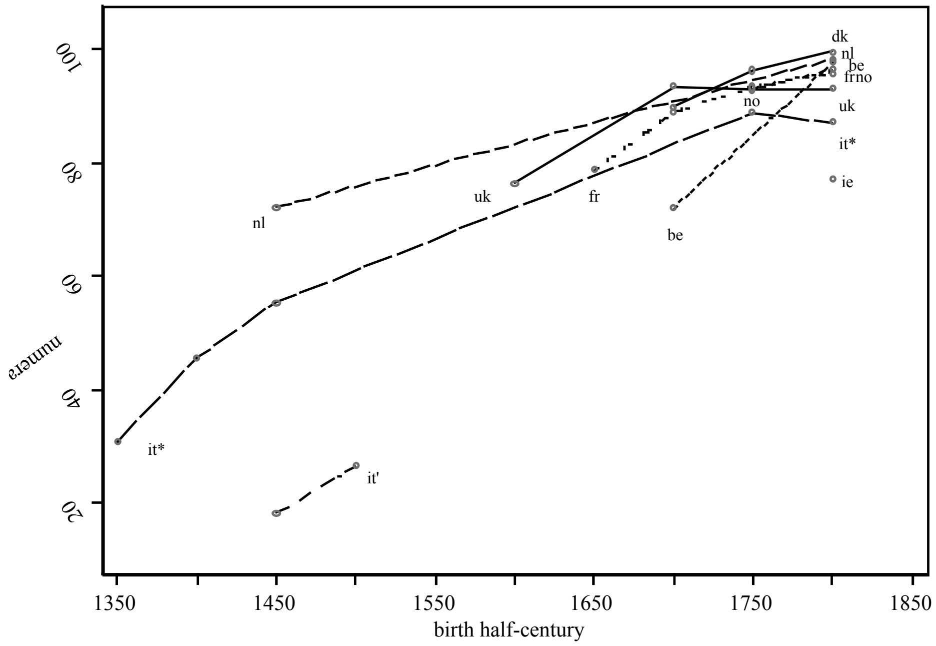
Notes: Alternative Whipple Index and literacy rate by country and birth half-century; fitted values based on OLS regression; samples with at least 100 observations.

Figure 9. Coevolution of literacy and numeracy in early modern Europe

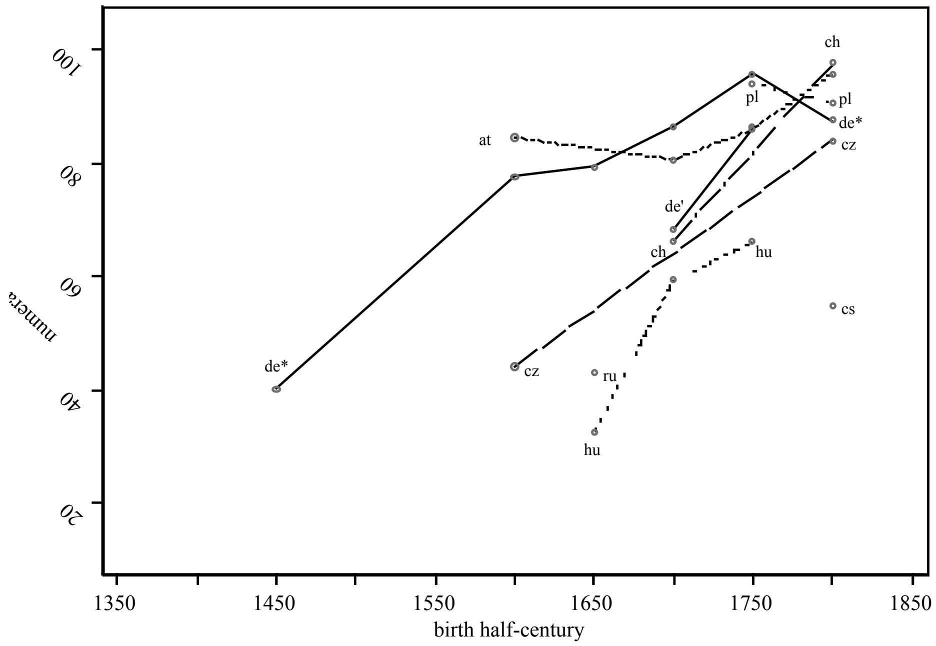


Notes: Alternative Whipple Index and literacy rate by country and birth half-century; arrows indicate direction of time; samples with more than 100 observations only; de* is Protestant Germany; it* is Northern Italy.

Figure 10. Numeracy in the long run
a. Western and Northern Europe



b. Central and Eastern Europe



Notes: Alternative Whipple Index and literacy rate birth half-century; samples with more than 100 observations only; de* is Protestant Germany; it* is Northern Italy.

Appendix

Data Sources

Country	*Source for literacy estimates*	*Source for numeracy estimates*
Austria		1647-1919 soul registers on Viennese suburbs,, see The Vienna Database on European Family History (http://wirtges.univie.ac.at/famdat/). Special thanks go to Josef Ehmer.
		1880 census of Cisleithania, see Österreichische Statistik 1882, Wien.
		1880, 1890, 1900 and 1910 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
Belgium	Ruwet, Joseph (1978). Un essai de synthese provisoire. In: Joseph Ruwet and Yves Wellemans (eds.). L'analphabétisme en Belgique (XVIIIème-XIXème siècles). Louvain: Bibliothèque de l'Université.	1740 death registers Liège, see Roger Mols S.J. (1956), Introduction a la demographie Historique des villes d'Europe du XIVe au XVIIIe siècle, Tome Troisieme, Annexes,Université de Louvain.
	Woude, A.M. van der (1980). De alfabetisiering. Algemene Geschiedenis der Nederlanden 7, 257-264.	1846, 1856, 1866, 1880, 1890, 1900, 1910 and 1920 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
Czech Republic		1651 census data for Pilsen, Beraun, Karlstein, Chrudim and Jung Bunzlau from the State Archive in Prague (signatory R 109/45). Special thanks to Katarina Hodinova for processing the data in the context of her diploma thesis.
		1657 soul register of the district of Leitmeric. Special thanks to Katarina Hodinova and Karl-Heinz

		Hospodarz.
		1689-1708 death register of the district of Leitmeric. Special thanks to Katarina Hodinova for processing the data in the context of her diploma thesis.
		1745 Bösig household data, see Kirchenbuch Bösig L3 / 3 , Hirschberger Weiß-oder Mannschaftsbuch 1745 (stored at the Hirschberg Museum), processed by Heinz Knobloch.
		1880 census of Bohemia and Moravia, see Österreichische Statistik 1882, Wien.
Denmark	Markussen, Ingrid (1990). The Development of Writing Ability in the Nordic. Countries in the Eighteenth and Nineteenth Centuries. Scandinavian Journal of History 15 (1), 37-63.	1769 census data (Duchy of Schleswig) provided by the Arbeitskreis Volkszahl-Register (AKVZ, www.akvz.de).
		1870, 1880, 1890, 1901, 1911 and 1921 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
Finland		1880, 1890, 1900, 1910 and 1920 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
France	Chartier, Roger (1985-7). Les Pratiques d'Écrit. In : Philippe Ariès and Georges Duby (eds.), Histoire de La Vie Privée. Paris : Seuil.	1740 death register Paris, see Roger Mols S.J. 1956, Introduction a la demographie Historique des villes d'Europe du XIVe au XVIIIe siècle, Tome Troisieme, Annexes,Université de Louvain.
	Houdaille, Jacques (1988). Les signatures au mariage. 1670-1739. Population (French Edition), 43 (1), 208-12.	1716-1784 French army soldiers data set created by John Komlos and cooperators.
	Houdaille, Jacques (1977). Les signatures au mariage de 1740 à 1829. Population (French Edition), 32 (1), 65-90.	1851, 1856, 1861 and 1866 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.

Germany	Hinrichs, Ernst (1982). Zum Alphabetisierungsstand in Norddeutschland um 1800. Erhebungen zur Signierfähigkeit in zwölf oldenburgischen ländlichen Gemeinden. In: Ernst Hinrichs and Günter Wiegelmann (eds.): Sozialer und kultureller Wandel in der ländlichen Welt des 18. Jahrhunderts, 21-42. Wolfenbüttel: Herzog August Bibliothek.	1523 conscription list for Balingen, see City archive of Balingen, Musterungsliste 1523, A 28a M21. Special thanks go to Lisbeth Zahawi.
	Hofmeister, A., Prass, R. and N. Winnige (1999). Elementary Education, Schools, and the Demands of Everyday Life, Northwest Germany ca. 1800. Central European History 31, 329-84.	1675 soul register (Status Animarum) of Strückhausen processed by Jürgen Rode, available at www.juergen-rode.de.
	Norden, Wilhelm (1985). Die Alphabetisierung der oldenburgischen Küstenmarsch im 17. und 18.Jahrhundert. In: Ernst Hinrichs and Wilhelm Norden (eds.) mit einem Beitrag von Brigitte Menssen u. Anna-Margarete Taube, Regionalgeschichte. Probleme und Beispiele. Hildesheim: Lax.	1704 census data (Mecklenburg-Schwerin) provided by the Arbeitskreis Volkszahl-Register (AKVZ, www.akvz.de).
		1718-1832 death register Berlin (Garnisonskirchen-district) collected from the Landesarchiv Berlin (signatory A Pr.Br.Rep. 005 A). Special thanks go to Anne Braenzel for collecting the data in the context of her diploma thesis.
		1749 Soul register (Status Animarum) of Emsdetten from the record office of the diocese Muenster (Bistumsarchiv Muenster, signatory GV Hs. 150, 399) provided by Westfälische Gesellschaft für Genealogie und Familienforschung, available from http://www.genealogy.net. Special thanks go to M.Thérèse Haemers-Van Roey Neerpelt.
		1749 Soul register (status animarum) of Greffen from the record office of the diocese Muenster (Bistumsarchiv Muenster). Special thanks go to

		M.Thérèse Haemers-Van Roey Neerpelt.
		1749 Soul register (Status Animarum) of Walstedde from the record office of the diocese Muenster (Bistumsarchiv Muenster, signatory Hs 152, 780-793), processed by Helmut Winterscheid, available from www.genealogienetz.de.
		1749 soul register (registrum parochiae) of Buldern from the record office of the diocese Muenster (Bistumsarchiv Muenster, signatory: GV-Hs. 152, fol. 1-223), processed by Hans Joachim Große Gorgemann, available from www.genealogienetz.de.
		1749 soul register (registrum parochiae) of Rheine und Elte from the record office of the diocese Muenster (Bistumsarchiv Muenster), processed by Edgar Winter, available from www.rheineahnen.de.
		1749/50 soul register (catalogus accolarum/incolarum) of Haltern from the record office of the diocese Muenster (Bistumsarchiv Muenster), processed by Dagmar Strauss and Bodo Stratmann, available at www.westfalengen.de.
		1750 soul register of Lippborg, City archive of Lippborg, Catalogus Familiarum ... parochiae Libborgensis de dato 20. Martii 1750. Special thanks go to M.Thérèse Haemers-Van Roey Neerpelt.
		1769 census data (Duchy of Holstein) provided by the Arbeitskreis Volkszahl-Register (AKVZ, www.akvz.de).
		1788 soul register Tuebingen-Bebenhausen collected from Evangelisches Landesarchiv Stuttgart-Moehringen. Special thanks to Anne Braenzel for

		processing the data in the context of her diploma thesis.
		1791-1850 age data from the health insurance and funeral expenses fund Berlin (Neu-Cölln district), collected from Geheimes Staatsarchiv Preussischer Kulturbesitz. Special thanks to Anne Braenzel for processing the data in the context of her diploma thesis.
		1799 “Einwohnerliste Esch-Auweiler-Pesch (Koeln)“ from the North Rhine-Westphalia public record office in Duesseldorf (NRW Hauptstaatsarchiv Düsseldorf, Roerdepartement 1722 / Heft 3), processed by Heinz Nellessen, available from www.vfg-pulheim.de.
		1799 “Stadt Pulheim: Frz. Einwohnerlisten“ from the Hessian State Archive in Darmstadt (Hessisches Staatsarchiv Darmstadt), processed by Josef Wißkirchen, available from www.vfg-pulheim.de.
		1803 census data (Duchy of Holstein) provided by the Arbeitskreis Volkszahl-Register (AKVZ, www.akvz.de).
		1910 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
Hungary	Tóth, István György (2000). Literacy and written culture in early modern Central Europe. Budapest : CEU Press.	1706 and 1708 conscription lists from the Hungarian National Archive, Budapest, processed by Rita Benyus.
		1771 and 1791 soul register from the Hungarian National Archive, Budapest, processed by Rita Benyus.

		1784/5 census data from the Hungarian National Archive, Budapest, provided by Rita Benyus.
		1869 census data for the comitates of Borsod, Csik, Györ, Pest and Hrnvasarhely from the Hungarian National Archive, Budapest. Special thanks to Tamás Faragó and to Rita Benyus for processing the data.
Ireland	Johnson, Paul (1994). 'Health and welfare of women in the United Kingdom, 1785-1920' (with Stephen Nicholas), in R. Floud and R. Steckel (eds), Health and Welfare during Industrialization (NBER/University of Chicago Press, 1997), pp. 201-49.	1831-35 New South Wales (NSW) convict data provided by Deborah Oxley.
Italy	Brambilla, Elena (1991). Alfabetismo e Società nelle Marche in Età Napoleonica. In: A. Batoli Langeli and X. Toscani (eds.), Istruzione, Alfabetismo, Scrittura. Saggi di Storia dell'Alfabetizziazone in Italia (Sec. XV-XXIX). Milano: Franco Angeli.	1427 Veneto and Tuscany tax register,, see the Catasto data set by Herlihy und Klapisch-Zuber (1978) (including 1425 and 1502 age data for Verona and 1458 and 1480 samples for Florence).
	Grendler, Paul F. (1989). Schooling in Renaissance Italy: Literacy and Learning 1200-1600. Baltimore: The John Hopkins University Press.	1489 Pozzuoli and 1561 Sorrento census data,, see Julius Beloch (1937), Bevölkerungsgeschichte Italiens, I, S. 29-31, published by Gaetano de Sanctis.
	Marchesini, Daniele (1993). Sposi e Scolari. Sottoscrizioni Matrimoniali e Alfabetismo tra Sette e Ottocento. Quaderni Storici 18, 601-23.	1802 Napoleonic census of Bergamo, collected from the State Archiv Bergamo. Special thanks to Nadia Vernice for collecting the data.
		1804/05 Napoleonic census of Milan, see Zangheri, Renato (1963), La popolazione italiana in età napoleonica, Bollettino del Museo del Risorgimento, year VIII, pp. 3-220, I-CIII.
		1804/05 death register Milan, see Zangheri, Renato (1963), La popolazione italiana in età napoleonica, Bollettino del Museo del Risorgimento, year VIII,

		pp. 3-220, I-CIII.
		1871 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
		1880 census of Tyrol, see Österreichische Statistik 1882, Wien.
Netherlands	Kuijpers, Erika (1997). Lezen en schrijven. Onderzoek naar het alfabetiseringsniveau in zeventiende- eeuws Amsterdam. Tijdschrift voor Sociale Geschiedenis 23, 490-522.	1514 Dutch government reports,, see Jan Luiten van Zanden and Tine de Moor (2006), *Vrouwen en de geboorte van het kapitalisme in West-Europa*, Amsterdam: Uitgeverij Boom.
	Woude, A.M. van der (1980). De alfabetisiering. Algemene Geschiedenis der Nederlanden 7, 257-264.	1796 *Inwonerslijst* for Eijsden/Limburg available at www.prevos.net/heugem/huishoudens.htm. Special thanks go to Mrs Haemers-Van Roey.
	Zanden, Jan Luiten van (2004). De timmerman, de boekdrukker en het ontstaan van de Europese kenniseconomie over de prijs en het aanbod van kennis voor de Industrielle Revolutie, Utrecht : Utrecht University Press.	1796 *Inwonerslijst* for Houthem 1796, processed by Fons Heijnens, available from www.houthem.info. Special thanks go to Mrs Haemers-Van Roey.
		1849, 1859, 1869, 1879, 1889, 1899 and 1909 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
Norway		1801, 1865 and 1900 census data available from http://digitalarkivet.uib.no.
Poland		1801-1850 Marriage register Radzionkau, see Zbigniew Kwasny (1996), Ksiega slubów parafii radzionkowskiej z lat 1801 - 1850 / wydal i wstepem opatr. Zbigniew Kwasny. - Wroclaw : Wydawn. Uniw. Wroclawskiego.
		1811-1850 Death register Radzionkau,, see Zbigniew Kwasny (1998), Ksiega pogrzebów parafii

		radzionkowskiej z lat 1811 - 1850 / wydal Zbigniew Kwasny. - Wroclaw : Wydawn. Uniwersytetu Wroclawskiego.
		1835 German Mennonites census data,, see Johann Epp (2004), Die Volkszaehlung im Molotschnaer Mennonitengebiet 1835, Bielefeld.
Portugal	Magalhães, Justino (1994). Ler e escrever no mundo rural do Antigo Regime: um contributo para a história da alfabetização e da escolarização em Portugal. Braga: Universidade do Minho, Instituto de Educação.	
Russia	Mironov, Boris N. (1991). The Development of Literacy in Russia and the USSR from the Tenth to the Twentieth Centuries**.** History of Education Quarterly 31(2), 229-252.	1715/1720 Tula and 1710/1717 Vyatka census data,, see Daniel Kaiser and Peyton Engel (1993), Time- and Age-Awareness in Early Modern Russia, Comparative Studies in Society and History, vol.35, no.4, pp.824-839.
		1897 Census of the Russian Empire, see Central statictical bureau of the Ministry of Internal Affairs (1905), The First Total Census of Russian Empire, Editor N. A. Troynitsky, Moskau. Special thanks for processing the data go to Andrej Volodin.
Scotland	Houston, Robert A. (1985). Scottish literacy and the Scottish identity**:** illiteracy and society in Scotland and Northern England 1600-1800. Cambridge: Cambridge University Press.	1911 and 1921 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
Serbia		1850 census data for Serbia from the Serbia National Archive. Special thanks go to Heinz Gruber for making the data available.
Spain	Bennassar, B. (1983). Les resistances mentales. In : J.-P. Amalric et al., Aux origines du retard économique de 1'Espagne XVIè-XIXè Siècles. Paris : Editions du CNRS.	
	Cerdá Díaz, Julio (1986). Libros y lectura en la Lorca del siglo	

	xvii. Murcia: Caja Murcia, Obra Cultural.	
	Frago, Antonio (1990). The History of Literacy in Spain: Evolution, Traits, and Questions. History of Education Quarterly 30 (4), 573-599	
	Gelabert Gonzáles, J.E. (1982). Lectura y escritura en una ciudad provinciana del siglo xvi: Santiago de Compostela. Bulletin hispanique 84, 264-90.	
	Larquie, C. (1981). L'alphabétisation à Madrid en 1650. Revue d'histoire moderne et contemporaine xxviii, 132-57.	
	Nalle, Sara (1989). Literacy and Culture in Early Modern Castile. Past and Present 125, 65-96.	
	Rodriguez, M.-C. and Bennassar, B. (1978). Signatures et niveau culture1 des temoins et accuses dans les proces d'inquisition du ressort du Tribunal de Tolede (1525-1817) et du ressort du Tribunal de Cordoue (1595-1632). Caravelle 31, 17-46.	
Sweden	Markussen, Ingrid (1990). The Development of Writing Ability in the Nordic. Countries in the Eighteenth and Nineteenth Centuries. Scandinavian Journal of History 15 (1), 37-63.	1880, 1890, 1900 and 1910 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
Switzerland	Wartburg-Ambühl, Marie-Louise von (1981). Alphabetisierung und Lektüre: Untersunung am Beispiel einer ländlichen Region im 17. u. 18. Jahrhundert. Bern: Lang.	1560-1800 death register Geneva, see Roger Mols S.J. (1956), Introduction a la demographie Historique des villes d'Europe du XIVe au XVIIIe siècle. Tome Troisieme, Annexes,Université de Louvain.
		1860, 1870, 1880, 1888, 1900, 1910 and 1920 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.
United Kingdom	Chartier, Roger (1985-7). Les Pratiques d'Écrit. In : Philippe Ariès and Georges Duby (eds.), Histoire de La Vie Privée. Paris : Seuil.	1624-1637 Ship passenger lists departing from England, provided by The OLIVE TREE Genealogy

		(http://olivetreegenealogy.com).
	Clark, Gregory (2007), A Farewell to Alms: A Brief Economic History of the World (Princeton Economic History of the Western World), Princeton: Princeton University Press.	1769 soul register for Corfe Castle collected by Aravinda Guntupalli from the Population database at Cambridge (thanks to Richard Smith).
	Houston, Robert A. (1985). Scottish literacy and the Scottish identity**:** illiteracy and society in Scotland and Northern England 1600-1800. Cambridge: Cambridge University Press.	1782 soul register for Cardington, see David Baker (1973), The inhabitants of cradington 1782, Historical Record Society, vol.52. Special thanks go to Aravinda Guntupalli and Richard Smith for providing the data.
	Zanden, Jan Luiten van (2004). De timmerman, de boekdrukker en het ontstaan van de Europese kenniseconomie over de prijs en het aanbod van kennis voor de Industrielle Revolutie, Utrecht : Utrecht University Press.	1851 and 1881 census data available from the UK Data Archive at http://www.data-archive.ac.uk/.
		1911 and 1921 census data, see Franz Rothenbacher (2002), The European population 1850-1945, Basingstoke: Palgrave Macmillan.

3 New Evidence and New Methods to Measure Human Capital Inequality Before and During the Industrial Revolution: France and the U.S. In the 17th to 19th Centuries

Abstract
We explore pre- and early industrial inequality of numeracy using the age heaping method, in combination with anthropometric strategies. For France, we can measure the differential numeracy between the upper and lower segments of a sample population for 27 regions from the late seventeenth century, and study the subsequent impact of inequality on welfare growth in those regions – in interaction with political variables such as proximity to central government. The Northern U.S. had a relatively egalitarian distribution of numeracy, before industrialization started there on a large scale. With the onset of industrial development and immigration, inequality strongly increased.

This chapter is based on a working paper written jointly with Joerg Baten (University of Tuebingen). The working paper has been revised for the *Economic History Review* and currently under review. The concept for the paper was developed jointly, the analyses and writing were equally shared.

I. Introduction

Recent models of the Industrial Revolution argued that a relatively modest initial inequality in England might have been an influential factor in creating a modern market for consumer goods (and, perhaps, the motivation for an industrious revolution).[40] It is, however, very difficult to measure inequalities before and during the Industrial Revolution period. This study measures human capital inequality by employing a set of methods that developed around the phenomenon of *age heaping*, i.e. the tendency of poorly educated people to round their age erroneously or because of missing number discipline.[41] For example, they answer more often '40', if they are in fact 39 or 41, compared with better educated people.

In a related study, we found that the relationship between illiteracy and age heaping for Less Developed Countries (LDCs) after 1950 is relatively close.[42] The age heaping and illiteracy for not less than 270,000 individuals that were organized by 416 regions, ranging from Latin America to Oceania, produced a correlation coefficient as high as 0.63.[43] A number of other studies supported a close correlation between age heaping and other human capital measures for earlier periods. Data from the U.S. census manuscripts showed a very consistent and robust relationship, as well as a panel of 17 European countries for the late Medieval and Early Modern period (see below section II for a more detailed report).

The crucial advantage of those age heaping methods is that data are widely available for the early modern period, because many people were asked for their age in a more or less standardized way, when entering the military voluntarily, when they married etc. Even women

---

[40] Voigtlaender Voth, 'Why England?'.
[41] Mokyr, *Why Ireland starved.*
[42] A'Hearn, Baten and Crayen, 'Age heaping'.

who were accused of witchcraft in court were asked for their age, so one could even analyze the human capital of 'witches'.[44] In addition, age accuracy reflects basic numerical skills even more than literacy skills, which could be important as a precondition for technical, commercial and craftsmen activities in the production process. Even if the ability to report an exact age is certainly not a sufficient base for those activities, the lack of basic numeracy would prevent it. We apply those methods to two countries for which very large data sets are available. We assess the quality of the data carefully, by scrutinizing the institutional framework, and selection processes, as far as this can be reconstructed.

One interesting sample is the data set of the French army that was originally collected by John Komlos and his French collaborators. This data set is large enough to reconstruct numeracy even in relatively small regions of France since the late seventeenth century. We present below maps of French numeracy for several periods that give hints about regional inequality of human capital formation. From the degree of heaping and some anecdotal evidence we can be sure that soldiers were actually directly asked for their age (no birth certificates were demanded, and no comparisons with other sources were done by the registration officers). One advantage of this data set is the inclusion of anthropometric variables, so that we can assess the relationship between net nutritional status and numeracy. Another informative source comes from the census records of the United States 1850-1910, which includes some 650,000 cases. We concentrate on France and the U.S., because comprehensive data sets for the intended study are available for those two core countries.

Moreover, we use different units to assess inequality: We measure inequality between

---

[43] In an unpublished study, the authors compared the level of age heaping also to numerical skills as measured in the PISA survey (Programme for International Student Assessment). The data yielded a correlation coefficient of 0.80.

[44] on the relative age heaping of women see De Moor and Van Zanden, *Vrouwen.*

taller and shorter individuals (reflecting their nutritional status and social stratification), and differences between occupations of middle/upper versus lower social status. The results are multi-facetted, as the history of inequality has always been. One particularly important result is that height was a good predictor of numeracy in France: The taller half of the height distribution displayed a much lower age heaping tendency, and hence higher numeracy. Moreover, the size of the difference varied by region.

We use two different dependent variables for welfare growth, namely height and value added growth. How do those measures differ conceptually? The strength of GDP per capita is, of course, its comprehensive account of purchasing power and its comparability over time if given in standardized monetary units (such as 1990 Geary-Khamis $). It is clearly the most important measure for economic well-being. One of the disadvantages of GDP as a welfare measure is its bias against subsistence farming and production within the household. In general, non-traded goods and goods produced and consumed within households are often underreported. Moreover, other forms of informal markets, such as black markets, can often not be captured. Finally, the data requirements for GDP estimates are very large. Strong assumptions have to be made in many cases. Finally, GDP concentrates on purchasing power and does not include other 'biological' living standard components such as health, longevity, and the quality of nutrition. Anthropometric techniques were developed to provide an additional welfare indicator that covers biological aspects and includes not only subsistence farmers, but also local rulers, craftsmen, and unskilled day-laborers. Although the quality of nutrition is partly determined by income, heights also reflect other living standard components, such as the disease environment, hygienic behavior, and non-market nutrition factors such as proximity to the production of perishable proteins (e.g. milk or offals, which could not be transported and traded over longer distances before the mid-twentieth century).[45] In a seminal study, Margo and Steckel found a remarkable deviation of height and income for the mid-nineteenth century United States.[46] We will use this concept as welfare measure for French regions below. For the U.S. case, we will assess the value-added growth in industry in order to measure the productive potential of the individual states.

The present paper is organized as follows. We first review the literature on human capital inequality and its relationship with welfare growth. We then suggest in the second section the new measurement strategy *age heaping* that can be applied to numeracy inequality for the period before and during the industrial revolution. After discussing our data in more detail in section III, we apply the age heaping technique in section IV to 27 French provinces during the late seventeenth century, producing an estimate for human capital inequality for the first time. We also apply a modified version of this measure to 25 U.S. states in the early and later nineteenth century. Section V tests whether there is a negative impact of inequality in human capital on welfare increase in French regions between the seventeenth and nineteenth centuries. We base this analysis on a large data set of French soldiers born between the 1650s and 1760s. As a proxy for welfare growth, we use height growth between the late seventeenth century and the late nineteenth century. Finally, we test if inequality in numeracy in the United States, measured for birth cohorts before 1840, had a significant impact on economic growth between 1850 and 1880. Our last section concludes.

II. Views of the literature on inequality and growth

The new growth theory, pioneered by Romer, Lucas, Barro and Barro and Sala-i-Martin has aimed at explaining long-term growth within their models, rather than by some exogenously

---

[45] See Baten, 'Kartographische Residuenanalyse'; Baten and Murray, 'Nineteenth century Bavaria'; Komlos, 'The antebellum puzzle'; Moradi and Baten, 'Sub-Saharan Africa'.

growing variables like unexplained technological progress.[47] Moreover, new growth economics has strongly emphasized the role of human capital formation and its persistence in nations over time. During the last decade, many scholars extended this endogenous growth theory and addressed distributional aspects of income, suggesting that inequality might reduce growth.

This idea stands in contrast to the earlier study of the inequality phenomenon, which tended to view inequality as the variable to be explained. Most prominently, Simon Kuznets analyzed the effect that economic growth had on inequality of income when he formulated the well-known hypothesis that postulates an 'inverted U' relationship between income and inequality.[48] According to that hypothesis, the degree of inequality would first increase and then decrease with economic growth. During the 1990s, a fairly large body of literature has sought to test Kuznets' hypothesis of an inverted U relationship but the results have been ambiguous.[49] Deininger and Squire pointed out that most of the empirical work on this topic used cross-country instead of longitudinal data to test an inter-temporal relationship.[50] According to them, the inverted U shape tends to vanish if world region dummies are included.

Williamson and Williamson and Lindert argued that the inverted U relationship holds for the historical development during the nineteenth century.[51] Van Zanden could even measure inequality cycles for the pre-industrial period.[52] Both mainly supported the notion that during growth periods, some parts of the population developed special skills and hence earned higher incomes, whereas the remaining 'traditional' sector followed only after some lag. Oshima

---

[46] Margo and Steckel, 'The antebellum period'.

[47] Romer, 'Increasing Returns'; Lucas, 'Economic development'; Barro and Sala-i-Martin, *Economic Growth*; Barro, 'Inequality and growth'.

[48] Kuznets, 'Growth and inequality'.

[49] see for example Bourguignon and Morrison, 'Income distribution and foreign trade'; Bourguignon, 'Human resources'; Milanovic, 'Transition economies'; Polak and Williamson, 'Poverty, policy, and industrialization'.

[50] Deininger and Squire, 'Inequality and growth'.

[51] Williamson, *British Capitalism;* Williamson and Lindert, *American inequality* (but see the critical review of Feinstein, 'The Williamson curve').

[52] Van Zanden, 'Kuznets Curve'.

explained the phenomenon rather by considering physical capital: technology in the late nineteenth century, such as the steam engine, was characterized by large indivisibilities in production: Large factories might have been needed to achieve sufficient economies of scale.[53] These prevented all, but the richest part of the population from accumulating capital, and required a large group of unskilled workers, thus facilitating industrialization only at the cost of growing inequality over time. By contrast, it is argued, that the link between growth and inequality does not turn out as significant in recent times, as current technology is much more divisible and improved capital markets enable a much larger share of the population to make profitable investments. Barro pointed out that recent empirical work might have failed to confirm the Kuznets curve as, according to these theories, inequality would depend on the time passed since a new technological innovation was introduced into the economy, and not on per capita GDP.[54]

Beside this literature on the Kuznetsian determinants of inequality, a new line of research studied the effects of the distribution of income, i.e. the opposite direction of causality. This was triggered by the contrast of relatively low levels of inequality and high growth rates in Asian countries, at a time of high inequality and much lower growth rates in Latin American economies.[55] This research can be roughly grouped according to the specific channels through which inequality might affect growth. One idea is that by lowering the income of the median voter relative to the national average, greater inequality increases the pressure for redistribution, implying a higher tax rate. This, in turn, reduces incentives for productive investment, by inducing economic distortions.[56] Barro argues that even if no redistribution takes place, higher inequality implies higher costs for the rich to prevent redistribution.[57] The lobbying activities

53 Oshima, 'Technological transformation'.
54 Barro, 'Inequality and growth'.
55 Lindert and Williamson, 'Globalization'.
56 Alesina and Rodrik, 'Distributive politics'; Persson, and Tabellini, 'Is inequality harmful?'.
57 Barro, 'Inequality and growth'.

would consume resources and promote official corruption and tend to hamper economic performance. A second branch of theories is based on the idea that sociopolitical conflict reduces the security of property rights, thereby discouraging accumulation of capital. Perotti found that sociopolitical instability is enhanced by higher inequality, which in turn hampers economic development.[58] Besides this, as Barro argued, crime and riots deter investment and reduce the productivity of an economy.[59] In the recent past (such as the 1960s to today), a third channel might be imperfect credit markets, which prevent asset-poor people from making economically profitable investments in physical and especially human capital.[60] For the historical period, however, Gerschenkron had argued that a rise in inequality tends to raise investment, as the rich could save more and hence invest in physical capital.[61]

Only very recently, research has focussed on inequality not only of wealth or income but also explicitly on inequality of human capital. Loury studied an overlapping two-generation model: individuals have varying abilities, and those are not perfectly correlated with parental investment capabilities into children's human capital.[62] Hence, in his model, income inequality affects inequality of human capital through decreasing returns to human capital investment at the individual level – after all, richer people are not able to increase their human capital as much with higher income. Redistribution would lead to higher human capital investment among poor people, and increased aggregate human capital investment. Checchi analyzed the relationship between educational achievements and inequality of incomes using education Gini coefficients for 94 countries between 1960 and 1995.[63] The results of his fixed-effects-panel regressions suggest that increased access to education reduces income inequality if, and only if, the initial

[58] Perotti, 'Income distribution and democracy'.
[59] Barro, 'Inequality and growth'.
[60] Ibid., Deininger and Squire, 'Inequality and growth'; Galor and Zeira, 'Income distribution and macroeconomics'.
[61] Gerschenkron, *Economic backwardness.*
[62] Loury, 'Intergenerational transfers'.

level of educational attainment is sufficiently low and if average educational attainment is raised sufficiently rapidly.

Thomas, Wang, and Fan focused on the effect of development on the inequality of education.[64] Providing a unique data base of educational Gini indices covering 85 countries for the second half of the twentieth century, they found strong evidence that an education Kuznets curve between countries over time exists. Lopez, Wang and Thomas provided a theoretical background to the detrimental effects of inequality of human capital on economic growth. Since talents are bound to individuals and not tradable like most other goods, they assume that the marginal product of human capital across individuals is not generally equalized, if inequality is substantial.[65] Hence, aggregate production depends not only on the total level of human capital but also on its distribution. They found empirical evidence that more unequal education tends to have a negative impact on per capita income in most countries. The results also indicate that an equal distribution of human capital does not enhance growth by itself, but that it needs an adequate policy environment, which protects property rights.

In conclusion, there are newer theoretical approaches that directly focus on human capital inequality as a negative determinant of growth. We will put those to the test below. However, given that inequality of income and human capital might be closely correlated, we have to take interpretations into account, which see income inequality behind human capital inequality. Moreover, the reverse causality (Kuznets' view that growth causes inequality) needs to be discussed below. New evidence is clearly needed, and the seventeenth to nineteenth century histories can potentially shed light on this important debate, if we find ways to measure inequality of human capital (and income).

---

[63] Checchi, 'Access to education'.
[64] Thomas, Wang and Fan, 'Education inequality'.
[65] Lopez, Thomas and Wang, 'Education puzzle'.

III. A new method to measure human capital inequality

We will first discuss the general potential -- and the problems -- of the age-heaping method in this section, before discussing our measures for inequality of human capital below. The most widely used measure for age heaping is the Whipple Index, which adds the ages which are reported as multiples of five, and divides the total number of age statements. Given that in most populations, the ages ending in 5 and 0 should account for 20% of all age statements (as there are ten digits on which a number can end). If the share is, say, 40%, then clearly age heaping is present. Many authors emphasized that a person who cannot report her exact age, but rather reports a rounded age, normally has difficulties with numbers in general.[66] It is important, however, to counter-check whether census-takers or recruitment officers did ask at all for the age (and did not 'correct' the reported ages afterwards). This was the case for the samples studied here, as the level of age heaping would not be possible if a correction would have been done.[67]

How close is the relationship between age heaping and other human capital indicators such as literacy and schooling? A'Hearn, Baten, and Crayen used the large U.S. census sample to perform a very detailed analysis of this relationship.[68] They subdivided by race, gender, high and low educational status and other criteria. In each case, they obtained a statistically significant relationship. The highest correlation is among samples which reflect lower educational status, whereas the adjusted R-squares are lower in situations when almost everybody knows her exact age and knows how to read and write. Nevertheless, in all sub samples the correlation between

---

[66] The same effect occurs of an enumerator made a guess at the age from the appearance and entered a rounded age when a person could not state his or her own age exactly.

[67] Even if the precise birthday (often related to a saint's day or a holiday) is known to the individual, it might well be that the exact amount of years since birth means little to an individual although the annual event is celebrated again and again.

[68] A'Hearn, Baten and Crayen, 'Age heaping'.

age heaping and literacy was significant. Remarkable is also the fact that the coefficients are relatively stable between samples, i.e. a unit change in age heaping results in similar changes in literacy across the various tests. Those results are not only valid for the Northern U.S.: In any country studied so far which had substantial age-heaping, the correlation was both statistically and economically significant.[69]

In order to assess the robustness of those U.S. census results and the similar conclusions which could be drawn from late twentieth century Less Developed Countries, as mentioned in the introduction to this study, A'Hearn, Baten, and Crayen also assessed age heaping and literacy in 17 different European countries between the Middles Ages and the early nineteenth century.[70] Again, they found a positive correlation between age heaping and literacy, although the relationship was somewhat weaker than for the nineteenth or twentieth century data. It is likely that the unavoidable measurement error when using early modern data induced the lower statistical significance. One source of error was the heterogeneous data coverage since the regional units for literacy and age heaping did never perfectly match.

The possibly widest geographical sample studied so far has been created by Crayen and Baten, who could include up to 70 countries for which both, age heaping and schooling data (as well as other explanatory variables) were available.[71] They found in a series of cross-sections between the 1880s and 1940s that primary schooling and age heaping were closely correlated, with R-squares between 0.55 and 0.76 (including other control variables, see below). Again, the coefficients were relatively stable over time, with ten percent additional schooling reducing age heaping between 2.0 and 3.7 percent (the latter value for the 1880s is based on 22 cases only and can be considered as an outlier, given that the other six decades had all values between 2.0 and

69 On the regions of Argentina, see for example Manzel and Baten, 'Colonial and post-colonial Latin America'.
70 A'Hearn, Baten and Crayen, 'Quantifying quantitative literacy'.
71 Crayen and Baten, 'Global trends in numeracy'.

2.7). This large sample also allowed testing of various other potential determinants of age heaping. Sceptics of the method might wonder whether the degree of bureaucracy and government interaction with citizens is likely to influence the knowledge of one's exact age, independently of personal education. For example, mighty administrative authorities such as the Roman Empire might have performed many censuses to count potential soldiers and tax-payers, which might have led to better knowledge of one's age without extra education (which might not be necessarily true, just to give a potential example). To assess this effect, Crayen and Baten used the number of censuses performed for each individual country up to the period under study as explanatory variable for their age heaping measure. Except for countries with a very long history of census taking, all variations of this variable turned out insignificant, which would suggest that such a bureaucracy effect was rather weak. Only after six or more enumerations had been done in a country, there was some kind of 'training effect' giving rise to better knowledge of one's age and thus to lower age heaping values. It cannot be ruled out, however, that already a smaller number of censuses conduces to a more age-aware population but that this effect is captured by the schooling variable because authorities that have a preference for population statistics also tend to invest in schooling, and this was the variable with the most explanatory power. Crayen and Baten also tested whether the general standard of living had an influence on age heaping tendencies (using height as well as GDP per capita as welfare indicators) and found a varying influence: in some decades, there was a statistically significant correlation, in others there was none. Finally, age heaping values in East Asia were significantly lower than in other countries, suggesting that probably due to their inclination for calendars and astrology the East Asian culture increased age awareness or facilitated the calculation of the exact age by the enumerator without a similarly high level of general schooling.[72]

[72] Ibid.

In conclusion, the correlation between age heaping and other human capital indicators are quite well established, and the ‘bureaucratic’factor is not invalidating this relationship. A limitation of the age heaping method exists for very high human capital levels (and correspondingly low age heaping levels), implying that for most rich countries after the late nineteenth century the method cannot be used. Though not studied as yet, the authors acknowledge that regional factors could also impact on the correlation, for example, if harbours generate trade and eventually high numeracy. For example, A’Hearn, Baten, and Crayen find that Dutch immigrants in the U.S. were more numerate than literate, which might be an indication of a better acquaintanceship with numbers originating from their lives in harbour settlements where relatively many trade activities took place.[73] Another caveat relates to other forms of heaping (apart from the heaping on multiples of five), such as heaping on multiples of two and twelve. Heaping of multiples of twelve might be present in some Medieval samples, although the evidence for this is quite limited.[74] Heaping on multiples of two is quite widespread among children and teenagers and to a lesser extent among young adults in their twenties. At higher ages, this heaping pattern is mostly negligible, but interestingly somewhat stronger among populations who are numerate enough not to round on multiples of five. A potential measure for all sorts of heaping is the Bachi Index, which captures all deviations of the reported from the smoothed age distribution. Unfortunately, the Bachi index has several shortcomings compared to the Whipple Index. Most notably, it is statistically scale dependent. That is samples with identical heaping pattern but different sample size score dissimilar Bachi Index values since deviations caused by random sampling variation are not cancelled out.[75]

Please note that the Whipple Index reports a lower bound estimate of age misreporting.

---

[73] A’Hearn, Baten and Crayen, ‘Age heaping’.
[74] Thomas, ‘The prothero lecture’.

The Index does not take into account, ages that are untruthfully stated as ages other than multiples of five – if somebody reports '29'while being truly 28, this goes unnoticed. In a similar vein, age misstatements that neutralize each other are not captured. That is, if a 30 year old claims to be 35, this is counterbalanced by a 35 year old reporting himself 5 years younger.

Duncan-Jones recommends calculating age heaping measures for the population aged 23-62 only, and we followed him, excluding those below 23 and above 62 since a number of possible distortions affect those specific age groups, leading to age reporting behaviour, different to the one featured by the adult group in between.[76] Many young males and females married in their early twenties or late teens, they had to register as voters, military conscripts etc. At such occasions, they were sometimes subject to minimum age requirements, a condition which gave rise to increased age awareness. Moreover, individuals physically grow during this age group, which makes it easier to determine their age with a relatively high accuracy.[77] Another reason for excluding the very young is brought forth by the fact that it is plausible that parents and other relatives returned the ages for children and for a good share of young teenagers. All these factors tend to deflate age heaping levels for children and young adults, compared with the age reporting of the same individuals at higher ages. Empirically we observe indeed that rounding on multiples of five tends to be dominated by rounding on multiples of two until the age of 20. In fact, this phenomenon continues into the twenties, but less strong and less systematic and decreasing with age. The peak at age 30 is usually the strongest in the 23-32 age span.

The age heaping pattern of very old individuals is subject to upward as well as downward bias for the following reasons. Firstly, older people have a propensity to overstate their age, where old age is considered a sort of distinction. Secondly, selective mortality might affect the

---

[75] Its scale independence, efficiency, precision and reliability suggest that the Whipple Index should be the preferred age heaping measure. A'Hearn, Baten and Crayen, 'Age heaping' give a detailed discussion.

[76] Duncan-Jones, *The Roman Economy*.

educational composition of very high age groups, while a steep gradient in the right end of the age distribution due to high mortality is misinterpreted by the Whipple Index as digit avoidance. Finally, younger household members are likely to have reported the age of the elderly if those were considered as not being able to do so anymore. To be on the safe side and to avoid the risk of measurement error from this source, we exclude the very old, as well as the very young in the present study.[78] Moreover, for military samples featuring extremely skewed age distributions centred at late teen years and early twenties, the age span should be restricted to at least 23 plus to prevent us from interpreting military-specific age distributions as digit preference.

There remains some uncertainty about whether age heaping in the sources contains information about the numeracy of the responding individual, or rather about the diligence of the reporting personnel who wrote down the statements. A potential bias always exists if more than one person is involved in the creation of a historical source. For example, if literacy is measured by analysing the share of signatures in marriage contracts, there might have been priests who were more or less interested in obtaining real signatures, as opposed to just crosses or other symbols. We find it reinforcing that we observe generally much more age heaping (and less numeracy) among the lower social strata, and among the half of the sample population which had lower anthropometric values. Moreover, the regional differences of age-heaping are similar to regional differences in illiteracy.

We conclude that the age heaping method is now arguably a well-established indicator for numeracy of groups, but the problem remains how upper and lower group members can be distinguished from each other for historical populations for which we typically have no individual income data. Occupations have been often used to classify upper versus lower income group

[77] A 17-year-old might round off to 16 or 18, but not to 15 or 20.
[78] In principle, it might be possible to analyse the even heaping of teenagers and heaping patterns of the elderly, but given the newness of the age heaping method we opt for the most robust heaping of /adults aged 23-69 only.

individuals, and we will apply this to U.S. census data below. Of course, occupations such as 'day-laborer'or 'agricultural worker'typically yielded a low income, whereas professionals, noblemen, factory owners, and skilled craftsmen had higher incomes. However, in some occupational systems there can be large variation of income even within a given occupation. Probably, the most important case is the 'farmer", whose income can range from the one of a large land-owner to a small peasant (and especially if occupation is self-reported, even the smallest cottager sometimes becomes a farmer). Moreover, there was much structural change since the late nineteenth century, when workers in large textile firms, for example, advanced to relatively high incomes, whereas farmers typically lost in relative terms.

In this study, we propose an alternative, similarly rough proxy to distinguish social groups, based on human stature: We contrast those above and below mean height, and aggregate human capital characteristics by sample half. Almost all anthropometric studies that considered occupational or income groupings found that the well-off strata of society were taller, as long as regional differences are held constant (such as proximity to protein production). A second very interesting aspect to this strategy is that tall individuals are much less likely than short individuals to have suffered from infant protein deficiency syndrome (IPDS) that reduces learning abilities to a certain extent. The syndrome was wide-spread during the seventeenth to nineteenth centuries, when malnutrition was so severe that most populations were severely stunted (with adult males being shorter than 170 cm on average. The following review is based on Baten, Crayen, and Voth.[79] Support for this claim comes from biologists and psychologists who have conducted experiments on the influence of protein malnutrition in childhood and the intellectual ability later in life. While the ethical backgrounds of those experiments are debatable, the results cannot be ignored. For example, Lucas finds that children who had received less nutrient-rich diets showed

---

[79] Baten, Crayen and Voth, 'Poor, hungry and stupid'.

markedly lower neurodevelopment during the first two years of life, compared to a control group.[80] Even as late as age 7.5 the IQ scores are significantly lower. A randomized experiment in Guatemala suggests that protein supplements can produce marked improvements in cognitive ability.[81] Especially complex numerical abilities could be affected. Similar findings are brought forward by Paxson and Schady.[82] They use a sample of over 3000 preschool age children from Ecuador to identify determinants of children's language ability. The authors find that household socio-economic characteristics have a significant effect, and that its importance increases with the child's age. Most relevant for the present paper is the report by Grantham-McGregor since it points to a link between heights and cognitive ability even on the individual level.[83] Her study on stunted children showed that nutritional supplements can produce important gains in intellectual development. Several other studies indicate that the persistent exposure to undernourishment and poverty produce a cumulative effect on cognitive ability.[84] The longer a child's nutritional and educational needs go unmet, the greater the overall cognitive deficits. This applies not only to severe but also to mild undernourishment. While Magnusson, Rasmussen, and Gyllensten reason that genetic influences cannot fully explain the correlation between heights and cognitive ability from observing stature and intelligence of Swedish siblings. One caveat to the proposed anthropometric method is clearly that genetic height variation remains on individual level.[85] Nonetheless, we are confident that most individual variation can be averaged out by means of sufficiently large sample sizes.[86]

Like in all inequality studies, we have to consider which social strata are actually covered

[80] Lucas, 'Programming by early nutrition'.
[81] Brown and Pollitt, 'Malnutrition and intellectual development'.
[82] Paxson and Schady, 'Cognitive development'.
[83] Grantham-McGregor, 'Growth retardation and cognition'.
[84] e.g., Gorman, 'Cognitive development in children'; Strupp and Levitsky, 'Early malnutrition'.
[85] Magnusson, Rasmussen and Gyllensten, 'Swedish men'.
[86] Please note that we calculate the inequality within regions, hence possible genetic differences between regions should not play a decisive role.

by a given sample: In many types of source, lower classes are over-represented; hence inequality might look smaller, compared with samples of broader coverage. There are different degrees of concentration on the lowest income ranks, within the lower class. It is very important to discuss the representativeness of each type of sample.

In sum, we apply the age heaping method (our non-numeracy indicator) to the taller half and the shorter half of sample populations, as well as to the upper and lower occupational strata. As a first test, we use very different height and age heaping samples and calculate inequality scores for them (Table 1). Firstly, we can compare the numeracy of those from higher versus lower income occupations in the U.S. (based on U.S. census data). We excluded immigrants as their numeracy might have been determined in the countries of origin before they migrated. In general, age heaping was modest among the Northern United States population that was born between the 1800s and 1870s. Inequality was very low in the Northern U.S., whereas it was much higher in the South. Secondly, we have a very large sample of French army soldiers that was created by Komlos and co-operators, described in detail in Komlos' study.[87] Judging from his description of recruitment practices, this sample has a lower class bias of medium strength. Analysing more than 20,000 recruits, we found that short soldiers of the Southwest had very strong heaping (low numeracy), whereas it was low among tall Frenchmen in the North.[88] In addition, inequality in the province of Ile de France (Paris) was higher than in any of the other samples.

One important result from these samples is that the half of the population which was taller or had higher-status occupations (and came probably from more advantaged family backgrounds) had always lower or equal heaping values, whereas numeracy tends to be lower among the shorter half of the population, and those with lower-status occupations. Moreover, it is interesting

---

[87] Komlos, 'Early modern France'.

that great inequality of numerical human capital was visible in the Southern U.S. and Southwest of France, which developed poorly in the later nineteenth century industrialization waves. In contrast, Northeastern France and the Northern United States was characterised by only modest inequality. According to recent theoretical models, as described in more detail in section I, low inequality might impact favourably on economic growth. If the low inequality – growth relationship postulated in section I held, we would expect the earliest industrial development in the Northern U.S. and the Northeast of France – and that is were it took place. Paris is somewhat of an outlier to this relationship. Perhaps its positive development in spite of high inequality was stimulated by the proximity of the central government. In the following, we will take a much more disaggregated view of France and the U.S., as this relationship would otherwise be based on too few cases.

IV. The Data

The French soldiers were recruited between the 1690s and 1780s.[89] We draw on his description in the following. Most of the soldiers enlisted under the reign of Louis XV and XVI, and were preserved in the French military archive in Chateau de Vincennes (Komlos lists 80 archival signatures).[90] The information reported in the sources includes age, height (in French *pied*, *pouce*, and *ligne*), place of birth, the company into which the soldier enlisted, and for a smaller share the father's and own previous occupation. The anthropometric measurement quality appears relatively high, as 42 percent of heights were not reported as a (rounded) integer. With regard to place of birth, Komlos reports an overrepresentation of urban places, and argues that recruits

[88] We excluded those younger than 23 years and those of unknown province.
[89] Ibid.
[90] Ibid.

tended to report the nearest urban locality.[91] Given that we study the provincial averages here, those inaccuracies can be neglected. We dropped some 1,000 foreigners enlisted in the army since we study birth cohorts born in French provinces only.

How representative was the French army of the underlying French population? Unfortunately, only fourteen percent of the archival entries contained the father's occupation, and eleven percent of the soldier's own occupation before recruitment. Among the recorded soldiers´ occupations, 'craftsmen'was the most frequent category (44%). It is likely that some of those listed under 'textiles'(13%), 'food processing'(7%), and 'middle class'(7%) also belonged to this group (the distribution of father's occupations was similar). Given that the largest part of the French population was agricultural, the proportion of occupations belonging to 'agriculture' (7%) and 'worker/laborer' (jointly 17%, referring mostly to agricultural but also to industrial labor) was probably less than the corresponding share in the French society. However, second sons of farmers or of agricultural laborers were less likely than craftsmen recruits to report an occupation when enlisted and the true share of those with rural backgrounds is underestimated on the basis of the available occupational information. It is thus difficult to obtain a reliable evaluation of the representativeness of the data based on occupational structure. To obtain a rough impression we compare the average age heaping values for both the military sample and other sources of the French population. From Paris death registers we know the level of age heaping in the families of Parisians who died in 1740 was 123 for the 23-32 year-olds and 151 for the 33-42 year-olds.[92] Age accuracy in the death registers is thus slightly worse than among Parisian soldiers aged 23-42 born during the corresponding birth decades (90 percent of Parisian soldiers were aged 23-42) who featured a Whipple Index of 122. In other words, in terms of numeracy the soldier sample

[91] One possible reason for this could have been to facilitate the enlistment procedure for the recruitment officer for whom the names of larger places were more familiar and thus easier to spell correctly.

[92] A'Hearn, Baten and Crayen, 'Quantifying quantitative literacy', based on Mols, *Démographie Historique.*

might be modestly biased towards more numerate persons, at least in the case of Paris.[93] We conclude that the existing evidence does not speak for a negative selectivity of soldiers, as one might have expected. Given the positive relationship between height and numeracy, the minimum height requirement might have had a positive impact on the selection process. Are there sufficient numbers of cases for the regional units under study? The highest numbers of cases are available for Guyenne and Ile-de-France (which included Paris). Guyenne comprised a large area, including the populous city of Bordeaux, so the high number of soldiers indicates certain representativeness.

We also studied the very large U.S. census database provided by the Integrated Public Use Microdata Series (IPUMS) of the Minnesota Population Centre (MPC).[94] The data is composed entirely of individual persons and household records from population censuses. Records for nearly 650,000 men and women with data on race, age, literacy, and birthplace were extracted from 1% samples of the decennial population censuses 1850 - 1870 and 1900. The instructions for the census enumerators concerning the variables relevant for the present study were broadly consistent for the 1850, 1860 and 1870 enumerations. The 1900 census questionnaire included an additional question on year of birth in order to increase the accuracy in age reporting. The attempts seem to have been in vain, however, since age accuracy had not increased markedly in the 1900 enumeration, which led to the renewed omission of the very same question in the 1910 census questionnaire.[95] Since the U.S. data is based on a census (and it was relatively skillfully done), it can be assumed to be representative of the underlying population.

---

[93] Note that age accuracy might be reduced by the fact that age is not stated by the individual her- or himself but by a family member or other person, as it is obviously the case for death registers.

[94] The IPUMS data are available online from the Minnesota Population Center at the University of Minnesota. Steven Ruggles, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004.

[95] Bailey and Parmelee, 'Age returns'.

V. Regional patterns of inequality in France and the U.S.

Is the differential numeracy between the taller and shorter half of the population even observable on a low regional level? We examine this issue on the basis of French provinces in the late seventeenth and early eighteenth century, and for the nineteenth century US States. Especially French regional human capital formation in the early modern period has been a white area on the map so far, hence we report the data in somewhat greater detail.

For the French regional level, we have to make sure that our regional units contain sufficient observations to calculate differential numeracy values for the lower and upper half. Some French provinces of the *ancien régime* were rather small and provided only a small number of soldiers; hence we aggregated them with neighbouring provinces. We reduced the number of French regions from 36 to 27 to obtain more similar unit sizes. Since the merged provinces were either very small provinces adjacent to a large neighboring province or two small provinces next to each other, we consider the loss of information as small. Aggregation had to be made not only spatially but also over time, i.e. birth decades. The resulting 50-year birth cohorts comprise about 200-1300 soldiers each, which give reasonably good estimation results with the age heaping method.[96]

From Figure 1, it can be inferred how human capital, measured in terms of numeracy, was distributed across French Provinces in the seventeenth century. The darker the shaded areas, the lower the level of numeracy (i.e. higher Whipple indices of 'non-numeracy'). We find the lowest level of numeracy in a belt stretching from Brittany southeast to Auvergne. The low level of numeracy there stands in sharp contrast to the high levels in the provinces Picardie/Hainaut, Anjou, Orléans, Bourbonnais, and the Rhone valley.

In order to assess not only the level of numeracy, but also its inequality, we apply the technique described in section II. We thus group the birth cohorts in tall and short soldiers, measuring their relative height on provincial level. That is, the recruit is considered being tall if his body height is above the mean height in the province where he was born. We then divided the resulting 'non-numeracy' of the shorter half by the one of the upper half, hence a value of 1.5 indicates that the shorter half displays a heaping behaviour that is 50 per cent stronger than in the taller half, whereas a value of 1.0 or even below indicates little difference (=low numeracy inequality, see Figure 2). A low value might be expected in regions that required few skills for those that were active, for example, in cattle herding, and hence had non-market access to protein (in mountainous parts of Franche-Comté or the Basque mountains, for example). The highest levels of inequality in the late seventeenth century were prevalent in North-Central France between Ile de France, Saint-Onge, Bourbonnais, and the Rhone valley, i.e. the large regions south of Paris and along the large rivers that include much of the grain and wine belt that produced surpluses for the urban development of Paris and the big cities (Figure 2). In contrast, we find low inequality around Franche-Comté in Eastern France, and in the southwestern area of Gascony/Bearn/Foix. It is quite interesting that these are all peripheral mountain regions, where income and education might not be correlated as strongly as in the rest of France. Low inequality is also detected in Limousin and Marche where both tall and short soldiers had exceptionally low levels of numeracy. These are actually two of the poorest regions of France today. Moreover, this area was characterized by a particularly adverse environment (swamps etc.), so the height-education relationship might have been distorted by heterogeneous disease environments.

What might determine the size of numeracy inequality? Are more dynamic regions more equal? It seems as if in the late seventeenth Century France, higher inequality of numeracy was more often apparent in urbanized, more densely populated regions which tended to industrialize

[96] The Appendix reports regression results using 25-year birth cohorts, i.e. 1660-1684, 1685-1709, 1710-134 and

more strongly in the subsequent period. This would not support Lopez, Thomas and Wang's hypothesis, which predicts that inequality, hampers economic development. In contrast, the evidence is consistent with a Kuznetsian view that the direction of causality was reverse: The more urbanized and industrially developed parts of France generated higher inequality. Sanderson argued that there was a decline of literacy between the 1780s and 1800s especially among the industrial workers of Lancashire.[97] He argues that during this period, investment in schooling was not sufficient in the rapidly growing cities of this most industrial country, and the demand for workers' literacy skills was also not given. At the same time, he emphasizes that clerks, entrepreneurs, and the richer parts of society probably kept or expanded their education, which would be an alternative explanation of initial Kuznetsian patterns in the rapidly growing industrial cores. From an alternative perspective, one could also argue that those regions were relatively market-integrated and concentrated on wheat growing. The high degree of market integration might have led to a stronger division between the upper and lower income groups. In summary, on the basis of the underlying data, it does appear as if more dynamic regions in France are initially characterized by higher inequality. As industrial development was strongest in the North and Northeast, we cannot confirm Lopez, Thomas and Wang's hypothesis for early industrial France that high inequality of human capital adversely affects economic growth.

Turning to the U.S. case, we base the following part on the large U.S. census database provided by the Integrated Public Use Microdata Series (IPUMS). Heights are not given in this data set, but we can distinguish inequality of numeracy by occupational groups. All occupations have been classified by the IPUMS team according to the median total income (in hundreds of 1950 dollars) of all persons with that particular occupation in 1950, which is certainly not

---

1735-1759. The effects discussed here proved to be robust to this alternative birth cohort specification.

[97] Sanderson, 'Literacy and social mobility'.

unproblematic for our purposes, but it can provide a first guidance. We omitted those individuals with occupations that were classified as zero income. The zero income group was very heterogeneous, since many of these people were in the labor force, but, because they did not report an occupation, they received an income score of zero. This group also includes some persons of typically high education (such as retired businessmen as well as gentlemen and ladies with inherited wealth) and some persons with typically quite low education (e.g. unemployed people, Native Americans living on reservations). We also omitted all women in this step, as a very large share of them were housewives or were at least recorded in the census as having no occupation (although they might have been active in the family business). We also excluded all immigrants from our analysis, as a large share of them migrated when they were already young adults. Therefore, migrants can be considered as having received their education in greater part in their country of origin unless they migrated with their family when they were young.[98]

In order to group the population into two groups of different social status, we distinguished the upper from the lower part of the occupational hierarchy. We opted for a 'top 40 vs. bottom 60' distinction, assigning a lower social standing to the large (14% of the underlying population) and indivisible occupational group with an income score of 20, which consists to 90 per cent of labourers. Given the gradient in numeracy and literacy from this group to the adjacent class with higher income scores, this cut-off point appears reasonable to the authors (see Appendix 2). The upper class of occupations is lead by physicians and other academic jobs, which had an occupational income score of 8,000 Dollars (in 1950). In the lower segments of those 'top 40' per cent of the male population, the occupational groups featuring an income score of 22 and 21 mainly comprise skilled workers such as bookkeepers and secretaries. As a caveat we have to note that there were also some groups that had a relatively high income around 1950,

[98] Family migration was wide-spread until the mid-19th century, but in the later 19th century the single (or

but their social and economic status in the nineteenth century might have been lower (workers in textile factories, for example). While all farmers were classified by the IPUMS research team as being in the ‘bottom 60%’ segment, we have to keep in mind that some farmers might have been relatively wealthy and skilled. Having excluded all states where either the upper or the lower segment was based on fewer than 100 observations, we computed the Whipple Index for both social groups and the ratio of both index values (lower to upper class), which we used to measure inequality. The ratio thus gives the relationship between the lower and upper part of the occupational scores, and the higher this ratio, the larger inequality of numeracy.[99]

We hypothesize that the highest inequality of numeracy may be found in the southern plantation belt of the U.S., where a relatively small group of whites benefited strongly from the export-oriented plantation system, whereas the majority of smaller farmers and other occupations might have had much lower income and financial means to invest in children’s education. We actually observe this pattern for those cohorts born before 1840 (i.e. before large-scale industrialization in the U.S. started, see Figure 3). States like Alabama, Louisiana, Georgia, and Mississippi clearly had the highest inequality of numeracy, produced by a particularly low numeracy among individuals who belong to the lower 60 per cent segment of the occupational hierarchy. Please note that slaves are not included in this statistic, as they were coded in the ‘zero income’ occupational group.

It is intriguing that inequality was very low in the Northeast. These differences would support the interpretation of Lopez, Thomas and Wang that lower inequality of human capital stimulates economic growth, as the strongest economic growth took place in the Northeast. We

---

couple) migrant became quite frequent.

[99] The question of endogeneity naturally arises: did they achieve a good position in the upper half of the occupational hierarchy, because they had good numeracy? Or was social mobility still relatively low, and hence most sons took over the occupations of their parents? Probably both patterns existed. However, for our purposes, the

can roughly summarize the situation before 1840: numeracy was low and unequally distributed in the South. In contrast, the North, and especially the Northeast, had higher numeracy and a relatively egalitarian distribution. For the post-1840 birth cohorts, the economic environment in the U.S. changed dramatically. Large parts of the country heavily industrialized and attracted workers from other parts of the world. Kuznets hypothesized that inequality rises in the first stages of industrial development, and we would expect this effect in the newly emerging centres of industrial expansion. Moreover, the arrival of more and more immigrants at the East coast (which tended to be less skilled than previous immigration waves) might have increased inequality in the industrial centers, as lower skilled Native Americans had to compete with the newly arriving immigrants. We observe that the inequality of numeracy in industrial states such as Connecticut and New Jersey moved into the top ten of American inequality in this period (Figure 4). The higher inequality in north eastern states such as Connecticut and New Jersey can be attributed to the fact that numeracy levels of the less skilled occupational groups worsened or did not improve as fast as the numeracy levels of the better skilled 40 per cent of the population. In contrast, inequality of numeracy declined significantly in most of the southern states, after the plantation system did not have an inequality effect anymore for the cohorts born after 1840. [100] This is even more remarkable as now the freed African Americans are mostly included in the statistics, and we would have expected that it took them some time to acquire numeracy, hence their presence should have increased inequality. But we actually find that the reduction in inequality in the South is not only due to improved numeracy on the part of the lower class but also to declining numeracy among the middle and upper classes in Virginia and South Carolina.

---

question is not as relevant, since we simply want to measure the distance of numeracy between the more and the less successful half of the population within as many states as possible.

[100] In the southern states of the U.S., inequality of numeracy increased markedly in Louisiana and Maryland only.

These states experienced a reduction in their average numeracy of 16 and 11 per cent, respectively.

In summary, we conclude for the early period: inequality of numeracy was in fact low for the Northeastern U.S. where industrialization was strongest later-on. For the birth cohorts after 1840 we find that industrialization in the U.S. led to increased inequality in the centers of industrial development, which supports Kuznets' hypothesis. Moreover, it can be observed, how the aftermath of the Civil War and the end of the Southern plantation system had actually egalitarian effects. There was even some deskilling of the Southern middle- and upper class.

VI. Is there an influence of numeracy inequality on welfare growth?

In the recent world, the inequality of human capital has a strong negative influence on welfare growth, as the studies cited above demonstrated.[101] However, growth in today's world might be much more driven by human capital intensive processes that require schooling and numeracy of a large share of the society, compared with the pre-twentieth century period. Gerschenkron argued that inequality of income might actually have been more conducive to nineteenth century growth.[102] Economic development during the high time of the industrial revolution was strongly focused on physical capital formation: it mattered much more whether the investments could be undertaken during the nineteenth century, whereas in the late twentieth century, the skills of the workforce are likely to have been a stronger determinant. To the extent that income inequality is correlated with human capital inequality, we have opposing views for different centuries, but rigorous empirical tests are scarce for the later period, and non-existent for the early period. Hence it remains an open issue which theoretical concept applies to the real world situation in the

---

[101] Lopez, Thomas and Wang, 'Education puzzle'.

[102] Gerschenkron, *Economic backwardness.*

long run.

In the following section, we test the effect of human capital inequality on growth, using the inequality of numeracy measures that we presented in the previous section for seventeenth and eighteenth century French regions. Is Gerschenkron's positive or Lopez, Thomas and Wang's negative view of inequality visible in the history of French regions? Or possibly, opposing forces were at work, leaving the empirical picture inconclusive, because both mechanisms were important.

Which other variables should be taken into account? Thinking of French regional development, the focus on centralism immediately comes to mind. France has a reputation for centralizing many important institutions and gathering decision-makers and the most talented people in Paris, since at least the Napoleonic period. Only recently, during the last few decades of the twentieth century, the opposite movement has been put in motion, reallocating some institutions back to the periphery. During the nineteenth century, Paris was a strong attractor of French elites, and of taxes: public goods were centered disproportionately on Paris, partially financed by the provincial tax-payers. Paris was gradually transformed in the nineteenth century from an overcrowded and unhealthy Moloch into the cultural capital of the world. It is not a rare phenomenon that the capital region is favored by spending on public goods: In today's LDCs (Less Developed Countries), for example, recent studies argue that the government spends much more on the capital region in order to satisfy its nearby citizens, as they pose a direct threat in times of crisis and unrest. For the French case, this would imply that centralization might have been strongly reinforced by the events of the French revolution, which demonstrated the peril dissatisfied nearby citizens compared to the ruling classes.

In sum, we need to control for *centralité*, and its potential interactions with the inequality of numeracy, the variable of main interest in this study. We measure *centralité* by the distance from Paris, assigning nine different categories, from Ile de France coded as zero and the Provence coded as eight (Corse, for which we reserved category 9, is excluded due to an insufficient number of observations). We interact the distance variable with the inequality index values as we measured them for the late seventeenth and eighteenth century's French provinces.[103] Since standard indicators of economic development such as industrial value added, real wages, income per capita, or life expectancies are unavailable for seventeenth century French regions, we use the change in height as a proxy of the biological standard of living as dependent variable. Besides its availability, this proxy indicator has some other, more conceptual advantages for our purposes. One advantage compared with income is that the biological standard of living is an outcome indicator that takes into account the health effects of public goods: it has been frequently argued that centers of economic development were unhealthy places generating an overall much lower standard of living than unadjusted incomes would suggest. Height in contrast also reflects the amount of public goods that might improve hygienic and health conditions, such as waste removal or isolation of infectious disease carriers.

The increase of human stature between the seventeenth and the nineteenth century was quite different across French regions (Figure 5). In the central regions, between Anjou, Ile de France (incl. Paris), Bourgogne, and Poitou, heights increased substantially. In contrast, the periphery did not improve much compared with the catastrophic seventeenth century minimum of height: Languedoc/Roussillon, Provence/Venaisson, Lorraine, and Alsace were clearly unsuccessful compared to the average French development. The Lyonnais developed also quite badly.

While this pattern does not surprise on the background of the French *centralité* development, we are particularly interested in the potential additional effects of the interaction

[103] To obtain sufficient observation numbers, we opted for the broad birth cohorts 1660-1709 and 1710-1759.

between the French *centralité* and the inequality of numeracy, approximated by our age heaping method. The adverse provinciality effect could have been much more detrimental in an unequal region, compared with more egalitarian ones, if we follow Lopez, Thomas and Wang's view of human capital inequality effects.[104] As can be deduced from Table 2, this interaction term had a significant negative effect as large as the French *centralité* for the whole period of the seventeenth to nineteenth centuries as well as the later sub period of the eighteenth to nineteenth centuries. Adding a square term of distance from government, we can conclude that the detrimental effect of the French *centralité* on the periphery was non-linear and decreasing with distance. The regressions yield robust results after the inclusion of other control variables such as initial height level, initial level of numeracy, and initial inequality of numeracy.

We measured the explanatory variable in column 4 with a separate inequality estimate for each eighteenth century province (column 4-7, whereas columns 1-3 referred to our seventeenth century estimates). Given the similarity of results, we are confident that those results are robust in spite of the obvious measurement error that should bias significance downwards (see also notes to Table 2). Please note that in the French case, only the interaction of inequality with distance from the government was significant, whereas inequality of numeracy never had a significant influence (col. 3). This can be interpreted on the basis of Gerschenkronian effects which were counteracting, or even reverse causality -- Kuznetsian forces might have increased inequality. We find that the level of numeracy does not matter for the biological welfare, while we can observe some convergence in height. These results confirm our hypothesis that centralism in France was crucial to regional development. This could be explained by selective migration to the capital as well as by the central allocation of state funds.

For the U.S. case, we also test the theories brought forward by Lopez, Thomas and Wang

[104] Ibid.

vs. Gerschenkron in a regression framework, using our inequality measures for the United States from section IV of our paper as an explanatory variable.[105] More specific, we examine if inequality of numeracy had a statistically significant effect on subsequent economic growth in the U.S.[106] To address this question, we use manufacturing samples from the manuscripts of the decennial federal census for 1850, 1860, 1870 and 1880 collected by Atack, Bateman and Weiss.[107] The authors drew random samples of each state and territory that included between 6,000 and 10,000 firms in each census and cover 22 of the 25 states that we used in our descriptive analysis above. [108]

As a proxy for economic performance, we calculated the value-added (total value of output - total value of raw materials) of the individual firms in the manufacturing census samples and aggregated them on state level. Since input and output values were reported in nominal Dollar amounts, we adjusted for changes in prices using Atack, Bateman and Margo's price deflators.[109] We used the change over time (relative to 1860) of the state-specific deflated value added as dependent variable, in the following referred to as real growth. From Figure 6 it can be inferred that the estimates indicate negative development for Massachusetts, Maine, Mississippi and Alabama. Moreover, we see that the South was characterized by rather high inequality and low growth while the North experienced rather high growth rates and low inequality.[110] For early industrializing states such as Massachusetts there might have been less room for additional growth – the classical convergence effect.

---

[105] Ibid.;Gerschenkron, *Economic backwardness.*

[106] Note that we considered inequality before 1840 and economic growth for the period of 1850-1880 only, so that endogeneity is not a problem.

[107] See Atack and Bateman, 'U.S. industrial development' for a description of the data set.

[108] For Rhode Island, all census years are missing, while for Georgia and Louisiana, only 1880 is available.

[109] Atack, Bateman and Margo, 'American Manufacturing'.

Table 3 reports the coefficients of our empirical analysis. Inequality of numeracy (measured for the birth cohorts 1800s-1830s) have a significant and negative impact on economic growth during 1860-1880 in a regression framework that also includes the deflated value added in 1860 to control for convergence effects. Embedded in a multivariate regression with additional variables, the variables remain significant. Moreover, the level of age heaping (non-numeracy) also had a negative influence – in other words, both the level and equality of numeracy mattered positively for subsequent industrial growth. We also included a dummy variable that takes the value 1 for the 11 southern states and 0 otherwise.[111] The dummy variable was designed to capture effects caused by the different industrial structure between North and South as well as the consequences of the civil war. The differences between the U.S. South and the North are certainly manifold, and were very much so during the nineteenth century. Not only did the South suffer severe destruction of its infrastructure and population during the Civil War in the early 1860s, but its general economic structure was also very different from the North. Rather surprisingly, the dummy variable for the South turns out statistically insignificant. All in all, the model can explain as much as 55 per cent of the variance of value-added growth.

We can thus conclude that our empirical analysis of nineteenth century United States shows mild support for a negative impact of inequality on economic development, as predicted by Lopez, Thomas and Wang.[112] Ideally, our results will be confirmed on the basis of a more comprehensive data base including additional variables such as policy environment and economic structure.

---

[110] We grouped the data on the basis of the Census divisions used by the US Census Bureau. Accordingly, we created the group 'North' (including the divisions East North Central, Middle Atlantic and New England) and 'South' (West South Central, East South Central and South Atlantic).

[111] The southern states include Alabama, Delaware, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee and Virginia.

[112] Lopez, Thomas and Wang, 'Education puzzle'.

VII. Conclusion

Age heaping provides an exciting new way to measure the numeracy of populations before the Industrial Revolution, if a number of institutional factors can be taken into consideration. Moreover, if there are proxies available for the lower versus middle/upper class origin of individuals in a sample (such as occupational group or anthropometric percentiles), we can even study the inequality of numeracy for pre-industrial economies. We have assessed this methodology for France and the U.S. during the seventeenth to nineteenth centuries and produced maps of numeracy inequality for France in the late seventeenth century, and the U.S. in the nineteenth century. Next we studied whether recent theories, which suggest a negative influence of human capital inequality on subsequent welfare growth, can be confirmed for this period. We find that this was indeed the case for the U.S., as the Northeast had very low inequality and strong growth afterwards. Moreover, we run a regression using the relative change over time of value added as a proxy for economic growth. We find that inequality of numeracy exerts a significant negative relationship on subsequent economic development in the United States. In France, the relationship was less clear. Taking also political factors into account, the proximity of a region to the central government in Paris turned out to be an influential variable for nineteenth century France. Interestingly, this factor in interaction with numeracy inequality also had explanatory power: Peripheral regions with high inequality experienced especially low welfare growth.

The U.S. case allowed to trace the path of inequality into the industrial period. We find that the industrial (and immigration) centers in the Northeast actually had strong surges in inequality in the spirit of Kuznet’s inverse U hypothesis, whereas the opposite development took place in the post-bellum South. Some of the middle and upper class Southerners even had declining numeracy. Kuznetsian tendencies might also underlie the French inequality development, as the North-Central region of France tended to be an early developing region with higher educational inequality.

References

A'Hearn, B., Baten, J. and Crayen, D., 'Quantifying quantitative literacy: Age heaping and the history of human capital', Universitat Pompeu Fabra Economics, Working Paper Series, 996 (2006).

A'Hearn, B., Baten, J. and Crayen, D., 'Quantifying quantitative literacy: Age heaping and the history of human capital', revised version, Working Paper Franklin and Marshall College/ University of Tuebingen (2008).

Alesina, A. and Rodrik, D., 'Distributive politics and economic growth', *Quarterly Journal of Economics*, 108 (1994), pp. 465–90.

Atack, J. and Bateman, F., '19th century U.S. industrial development through the eyes of the census of manufactures', *Historical Methods*, 32 (1999), pp. 177-88.

Atack, J., Bateman, F. and Margo, R., 'Capital Deepening in American Manufacturing, 1850-1880', *NBER Working Paper Series*, w9923 (2004), pp. 1-36.

Bailey, W. B. and Parmelee, J.H., 'The age returns of the twelfth census (1900)', *Publications of the American Statistical Association,* 12 (1910), pp. 110-23.

Barro, R. and Sala-i-Martin, X., *Economic Growth* (New York, 1995).

Barro, R., 'Inequality and growth in a panel of countries', *Journal of Economic Growth,* 5 (2000), pp. 5–32.

Baten, J., 'Kartographische Residuenanalyse am Beispiel der regionaloekonomischen Lebensstandardforschung ueber Baden, Wuerttemberg und Frankreich', in D. Ebeling, ed., *Historisch-thematische Kartographie: Konzepte - Methoden - Anwendungen* (Bielefeld, 1999), pp. 98-109.

Baten, J., Crayen, D. and Voth, H.-J., 'Poor, hungry and stupid: numeracy and the impact of high food prices in industrializing Britain, 1790-1840', Working Paper Universitat Pompeu Fabra /University of Tuebingen (2008).

Baten, J. and Murray, J., 'Heights of men and women in nineteenth century Bavaria: Economic, nutritional, and disease influences', *Explorations in Economic History,* 37 (1995), pp. 351-69.

Baten, J., 'Height and real wages: An international comparison', *Jahrbuch fuer Wirtschaftsgeschichte,* 2000-1 (2000), pp. 61-76.

Bourguignon, F., 'Growth, distribution, and human resources', in G. Ranis, ed., *En route to modern growth: essays in honor of Carlos Diaz-Alejandro* (Washington, DC, 1994), pp. 43–69.

Bourguignon, F. and Morrison, C., 'Income distribution, development and foreign trade: A cross-sectional analysis', *European Economic Review*, 34 (1990), pp. 1113–32.

Brown, J. and Pollitt, E., 'Malnutrition, poverty and intellectual development', *Scientific American,* 274 (1996), pp. 38-43.

Checchi, D., 'Inequality in incomes and access to education. A cross-country analysis (1960-95)', *Labour*, 17 (2003), pp. 153-201.

Crayen, D. and Baten, J., 'Global trends in numeracy 1820-1949 and its implications for long-run growth', Working paper University of Tuebingen (2008).

Deininger, K. and Squire, L., 'New ways of looking at old issues: inequality and growth', *Journal of Development Economics*, 57 (1998), pp. 259-87.

De Moor, T. and Van Zanden, J.L., *Vrouwen en de geboorte van het kapitalisme in West-Europa* (Amsterdam, 2006).

Duncan-Jones, R., *Structure and Scale in the Roman Economy* (Cambridge, 1990).

Feinstein, C. , 'The rise and fall of the Williamson curve: Review article', *Journal of Economic*

History, 48 (1988), pp. 699-729.

Galor, O. and Zeira, J., 'Income distribution and macroeconomics', *Review of Economic Studies,* 60 (1993), pp. 35–52.

Gerschenkron, A., *Economic backwardness in historical perspective* (Cambridge, MA, 1962).

Gorman, K., 'Malnutrition and cognitive development in children: evidence from experimental/quasi-experimental studies among the mild-to-moderately malnourished', *Journal of Nutrition*, 125 (1995), pp. 2239S-44S.

Grantham-McGregor, S. ,'Linear growth retardation and cognition', *Lancet*, 359 (2002), p. 542.

Keynes, J., 'The general theory of employment, interest and money', in D. Moggridge, ed., *Collected writings of John Maynard Keynes*, *Vol. 7: The general theory* (1936).

Komlos, J., 'Anomalies in economic history: Reflections on the antebellum puzzle', *Journal of Economic History*, 56 (1996), pp. 202-14.

Komlos, J., 'An anthropometric history of early modern France', *European Review of Economic History*, 7 (2003), pp. 159–89.

Kuznets, S., 'Economic growth and income inequality', *American Economic Review*, 45 (1955), pp. 1–28.

Lindert, P., and Williamson, J., 'Does globalization make the world more unequal?' *NBER Working Paper Series,* w8228 (2001).

Lopez, R., Thomas,V. and Wang, Y., 'Addressing the education puzzle', *World Bank Policy Research Working Paper Series*, 2031 (1998).

Loury, G.C., 'Intergenerational transfers and the distribution of earnings', *Econometrica*, 49 (1981), pp. 843-67.

Lucas, R.E., 'On the mechanics of economic development', *Journal of Monetary Economics*, 42 (1988).

Lucas, A., ‘Programming by early nutrition: An experimental approach’, *Journal of Nutrition*, 128 (1998), pp. 401S-6S.

Magnusson, P., Rasmussen, F. and Gyllensten, U., ‘Height at age 18 years is a strong predictor of attained education later in life: Cohort study of over 950,000 Swedish men’, *International Journal of Epidemiology,* 35 (2006), pp. 658-63.

Manzel, K. and Baten, J., ‘The development of numeracy in colonial and post-colonial Latin America (1640-1949)’, Working paper University of Tuebingen (2008).

Margo, R.A. and Steckel, R.H., ‘Heights of native-born whites during the antebellum period’, *Journal of Economic History,* 43 (1983), pp. 167-74.

Milanovic, B., ‘Poverty, inequality and social policy in transition economies, transition economics division’, *World Bank Research Paper*, 9 (1995).

Mokyr, J., *Why Ireland Starved* (1983).

Mols, R., *Introduction a la démographie Historique des villes d'Europe du XIVe au XVIIIe siècle* (Louvain, 1956).

Moradi, A. and Baten, J., ‘Inequality in Sub-Saharan Africa 1950-80: New estimates and new results’, *World Development*, 33 (2005), pp. 1233-65.

Paxson, C. and Schady, N.,‘Cognitive development among young children in Ecuador: The roles of wealth, health and parenting’, *Journal of human resources*, 42 (2007), pp. 49-84.

Perotti, R., ‘Growth, income distribution and democracy’, *Journal of Economic Growth*, I (1996), pp. 149-87.

Persson, T. and Tabellini, G., ‘Is inequality harmful for growth*?. American Economic Review,* 84 (1991), pp. 600–21.

Polak, B. and Williamson, J., ‘Poverty, policy, and industrialization: Lessons from the distant past’, *PRD Working Paper Series*, 645 (1990).

Oshima, H., 'The impact of technological transformation on historical trends in income distribution of Asia and the West', *Developing Economies*, 32 (1994), pp. 237–55.

Romer, P., 'Increasing returns and long-run growth', *Journal of Political Economy*, 94 (1986), pp. 1002-37.

Sanderson, M., 'Literacy and social mobility in the Industrial Revolution in England', *Past and Present*, 56 (1972), pp. 75-104.

Strupp, B. and Levitsky, D., 'The enduring cognitive effects of early malnutrition: A theoretical reappraisal', *Journal of Nutrition*, 125 (1995), pp. 2221S-32S.

Thomas, K., 'Numeracy in early modern England: The prothero lecture', *Transactions of the Royal Historical Society*, 5th Ser., 37 (1987), pp. 103-32.

Thomas,V., Wang, Y. and Fan, X., 'Measuring education inequality', *World Bank Policy Research Working Paper Series*, 2525 (2001).

Van Zanden, J.L., 'Tracing the Beginning of the Kuznets Curve: Western Europe during the early modern period', *Economic History Review*, XLVIII (1995), pp. 643-64.

Voigtlaender, N. and Voth, H.-J., 'Why England? Demographic Factors, Structural Change and Physical Capital Accumulation during the Industrial Revolution', *Journal of Economic Growth*, 11 (2006), pp. 319-61.

Williamson, J.G., *Did British Capitalism Breed Inequality?* (Boston, 1985).

Williamson, J.G. and Lindert, P.H., *American inequality: a macroeconomic history* (New York, 1980).

Tables

Table 1: Difference in non-numeracy by tall versus short individuals in various samples

Country/Region	Institution	Birth Decades	Whipple Index by status Low	High	Ratio
US (North)	US census	1800s-1830s	127	125	1.01
US (South)	US census	1840s-1870s	180	138	1.37
US (North)	US census	1830s-1870s	126	116	1.08
US (South)	US census	1830s-1870s	165	129	1.38
Paris	French Army	1650s-1760s	141	102	1.38
France (Northeast)	French Army	1650s-1760s	125	117	1.07
France (Southwest)	French Army	1650s-1760s	142	125	1.14
France (Total)	French Army	1650s-1760s	135	123	1.10

Note: a high Whipples index stands for low numeracy, and higher values of the 'ratio'indicate higher numeracy advantages of the taller half of the population, compared with the shorter half.

Table 2: Determinants of height change in France: distance, distance interacted with inequality, and other variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable:	17^{th}-19^{th} C	17^{th}-19^{th} C	17^{th}-19^{th} C	18^{th}-19^{th} C	18^{th}-19^{th} C	18^{th}-19^{th} C	18^{th}-19^{th} C
Dis * ineq of numeracy 1660-1709		-0.15 (0.00)					
Dis * ineq of numeracy 1710-1759					-0.16 (0.00)		-0.11 (0.01)
Distance to Government	-0.16 (0.00)		-0.41 (0.00)	-0.17 (0.00)		-0.38 (0.01)	
(Distance to Government)^2			0.03 (0.04)			0.04 (0.06)	
Level of Age Heaping 1660-1709			0.01 (0.23)				
Inequality of Age Heaping 1660-1709			0.43 (0.25)				
Level of Height 17^{th} C			-0.19 (0.00)				
Level of Age Heaping 1710-1759						-0.01 (0.11)	-0.00 (0.31)
Inequality of Age Heaping 1710-1759						-0.52 (0.38)	-0.07 (0.91)
Level of Height 18^{th} C						-0.22 (0.00)	-0.21 (0.00)
Constant	2.77 (0.00)	2.74 (0.00)	328.40 (0.00)	1.54 (0.00)	1.51 (0.00)	373.1 (0.00)	356.60 (0.00)
Observations	83	83	83	83	83	83	83
R-squared	0.10	0.10	0.43	0.09	0.10	0.52	0.50

Robust p values in parentheses. Note: We calculated the dependent variable as the height change between adjusted Komlos & Kim estimates (adjusted with the method proposed by Komlos 2003, p. 64) on the basis of the provinces, and heights in French departments from the Annuaire statistique (cited from Baten 1999). This means that sometimes two or three department heights are compared with the joint province height estimate, introducing some randomly distributed measurement error. Due to space constraints, we excluded the insignificant results for the sub period 17^{th}-

18th centuries.

Table 3: Inequality as a determinant of Economic Growth in the 19th Century U.S.?

Dependent Var.: Real Growth 1860-1880	(1)	(2)
Inequality of Numeracy before 1840	-1.76	-0.77
	(0.00)	(0.05)
Value Added 1860	-0.82	-0.79
	(0.00)	(0.00)
Level of Age Heaping before 1840		-0.61
		(0.05)
Dummy for Southern states		-0.24
		(0.27)
Constant	2.56	2.39
	(0.00)	(0.00)
Observations	22	22
R-squared	0.39	0.55

Robust P-values in parentheses. Notes: For demonstration purpose, Value Added 1860 and Level of Age Heaping before 1840 are multiplied by 100.

Figures

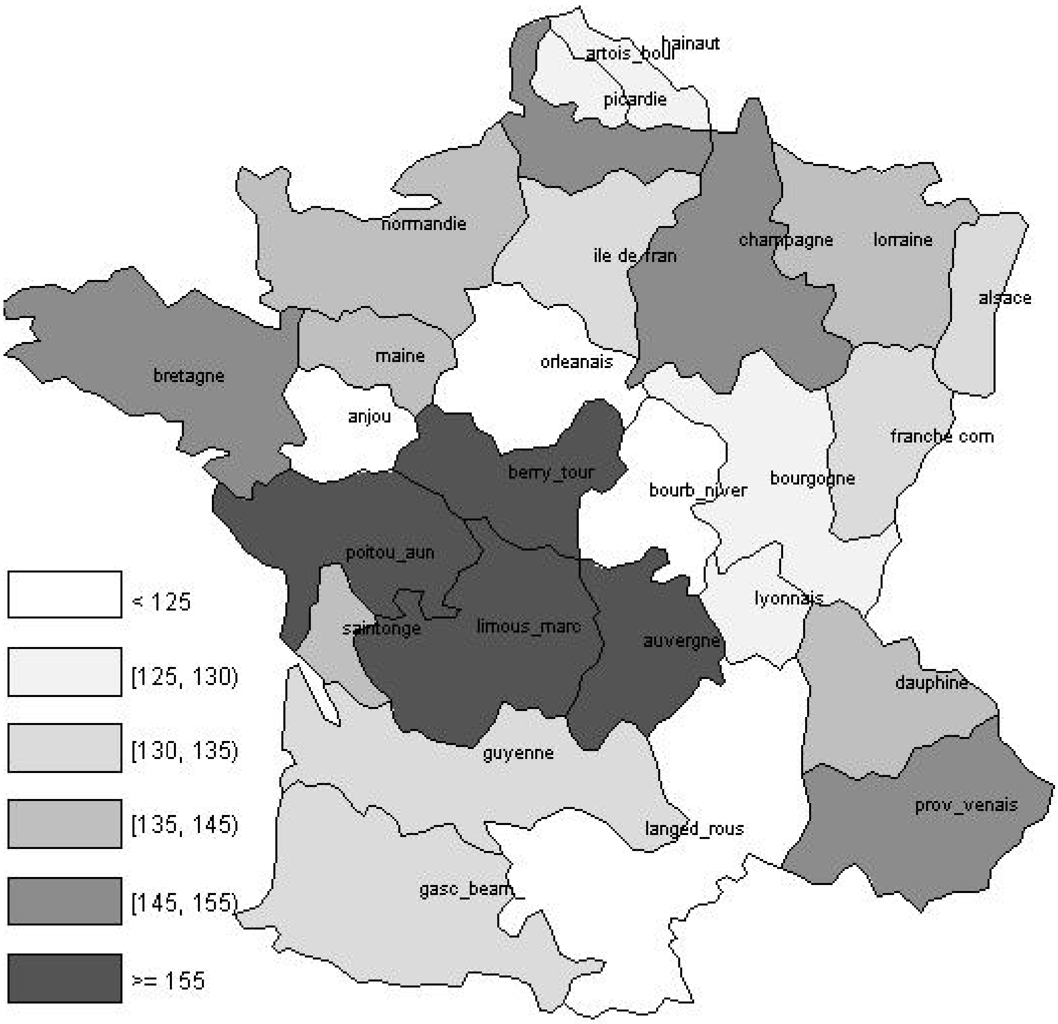


Figure 1: Non-numeracy in seventeenth century France (Whipple Index values of 1660-1709).

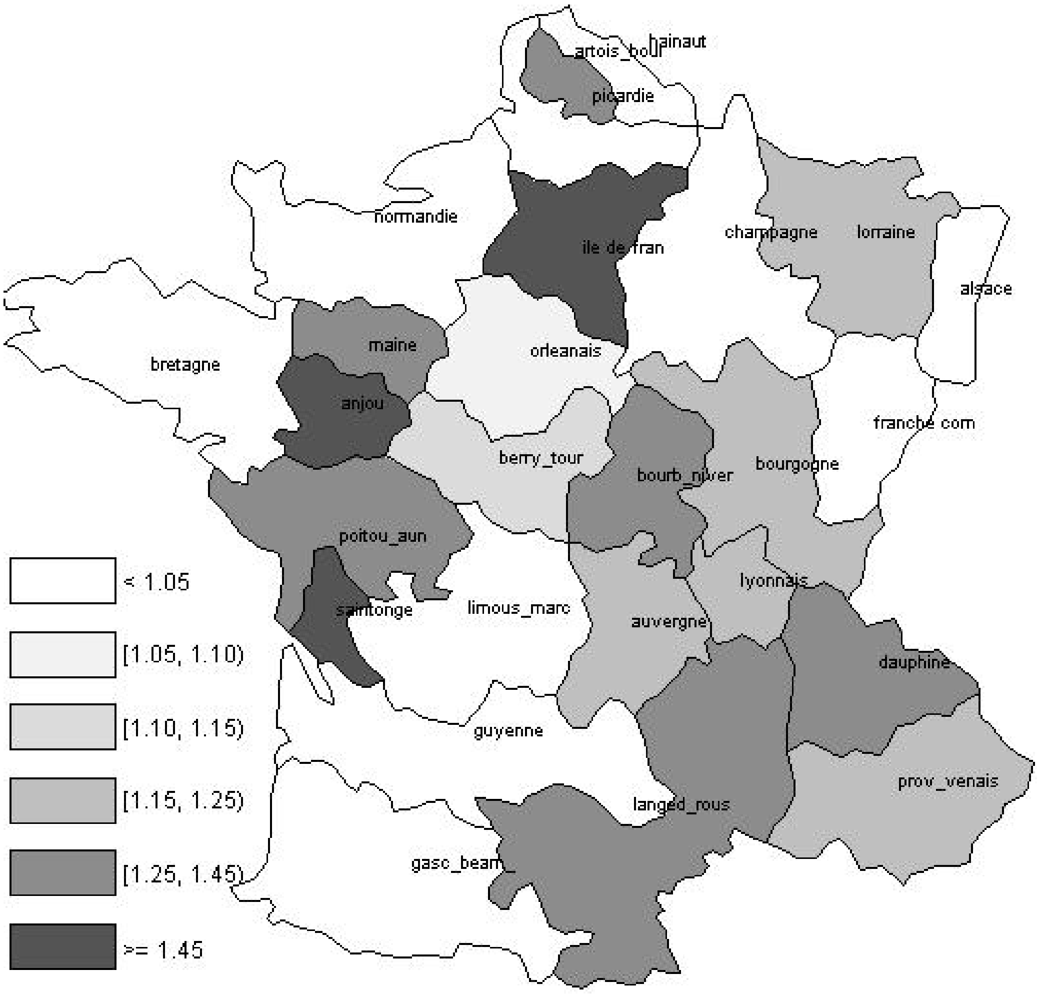


Figure 2: Inequality of Numeracy in seventeenth century France (Inequality of Whipple Index values of 1660-1709).

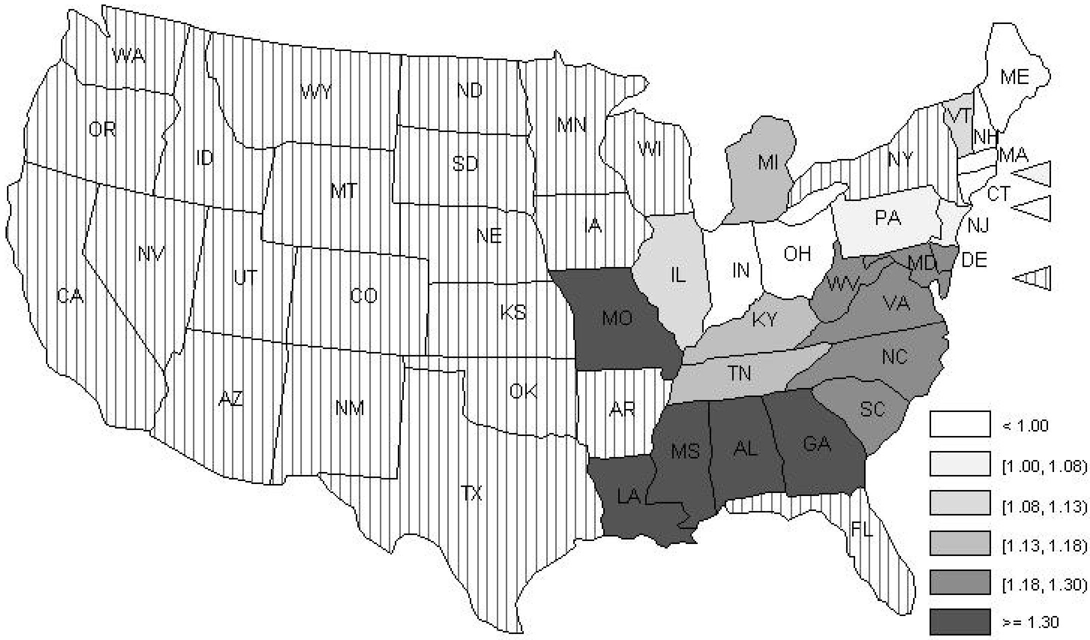


Figure 3: Inequality of Numeracy in the U.S. before 1840

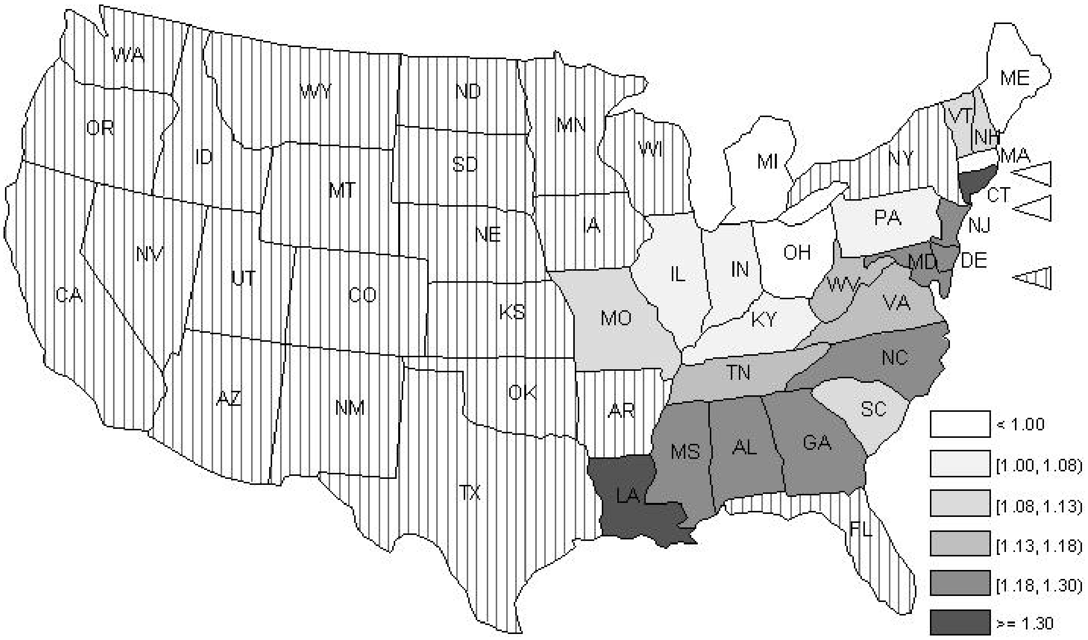


Figure 4: Inequality of Numeracy in the U.S. after 1840

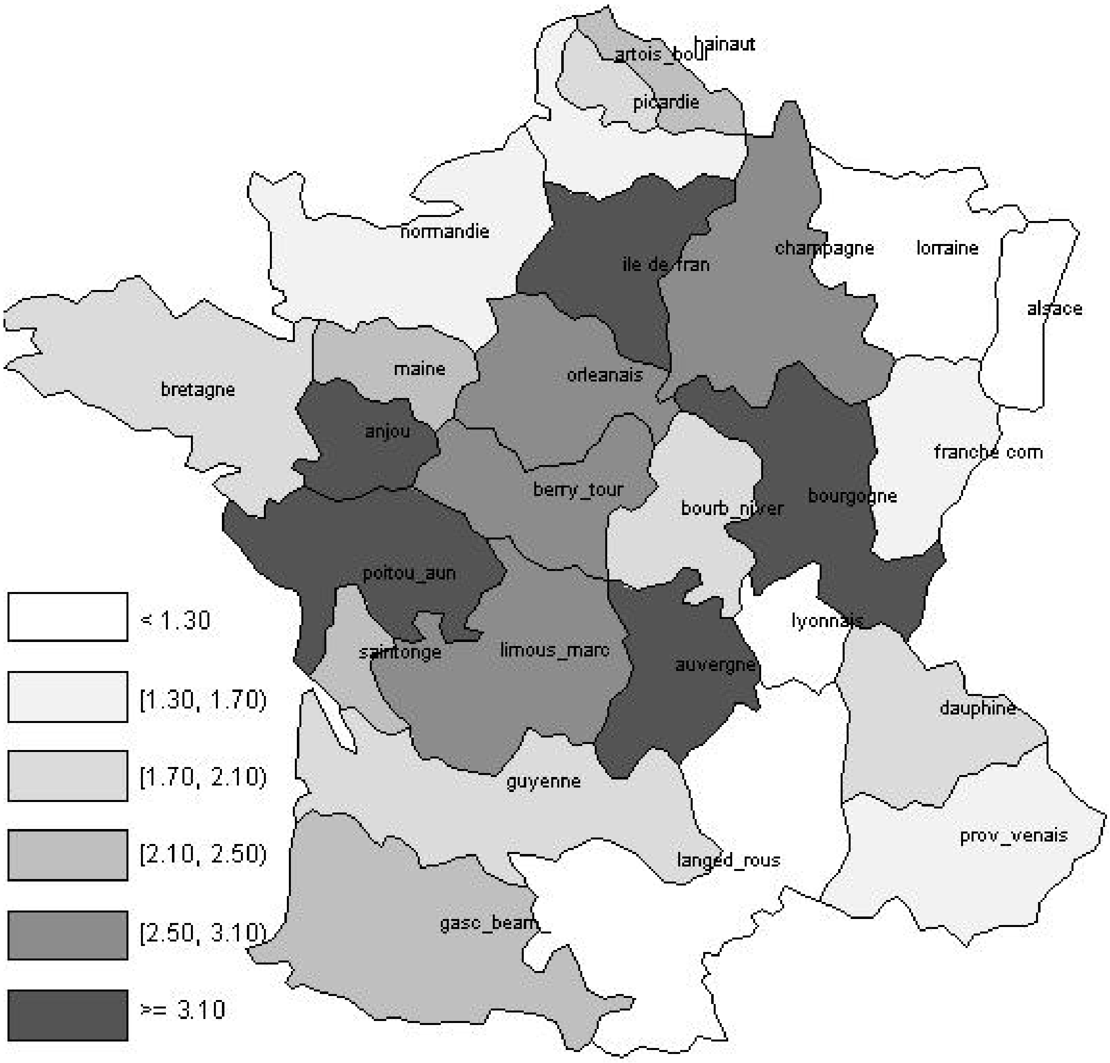


Figure 5: Change in height (in cm), seventeenth to nineteenth century.

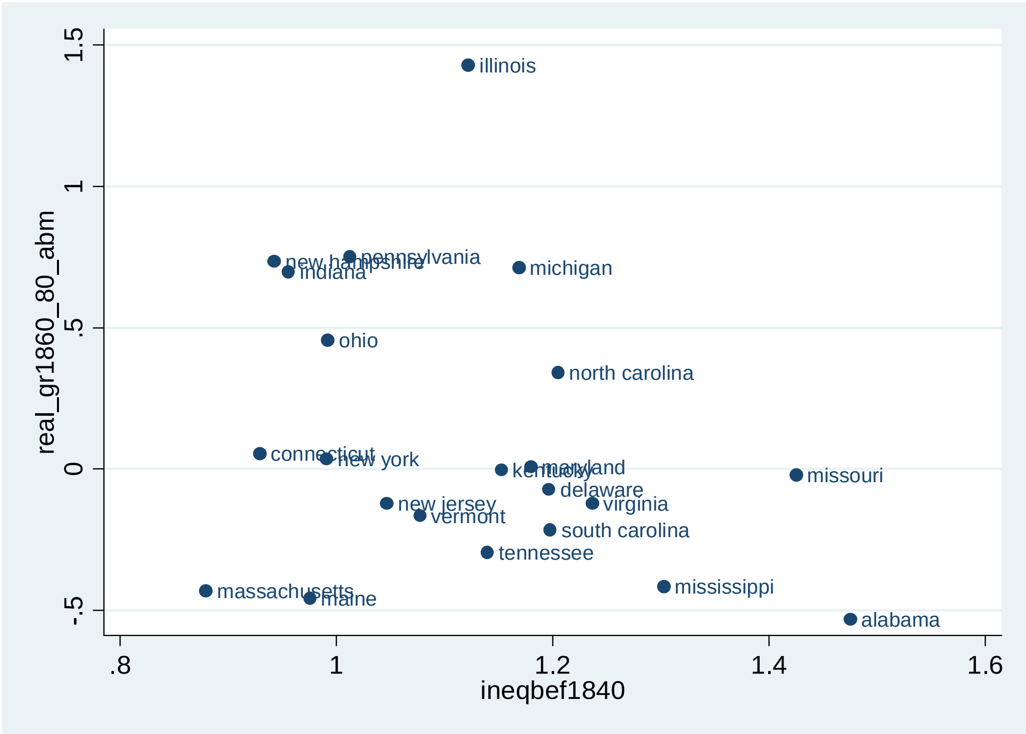


Figure 6: Inequality's impact on economic growth in nineteenth Century United States.

Appendix

Appendix 1: Sensitivity of findings to the length of birth cohorts (not be included in the publication)

Table A1-1: Determinants of height change: Distance interacted with inequality, and other variables, using 25-year birth cohorts

	(1)	(2)	(3)	(4)
Dependent Variable:	17th-19th C	17th-19th C	18th-19th C	18th-19th C
Dis * ineq of numeracy 1660-1684	-0.16			
	(0.00)			
Dis * ineq of numeracy 1685-1709		-0.14		
		(0.00)		
Dis * ineq of numeracy 1710-1734			-0.11	
			(0.04)	
Dis * ineq of numeracy 1735-59				-0.0437
				(0.22)
Constant	2.80	2.72	1.34	1.17
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	83	83	83	79
R-squared	0.12	0.10	0.05	0.02

Robust p values in parentheses. Notes: See Table 2.

Appendix 2: Occupational structures in nineteenth century United States (not be included in the publication).

Table A2-1: Distribution of occupational scores among US men born 1800-1879

income score	Freq.	Percent
3	35	0.01
4	25	0.01
6	848	0.32
7	168	0.06
8	6	0
9	26,805	10
11	1,381	0.52
12	956	0.36
13	883	0.33
14	89,944	33.56
15	500	0.19
16	1,318	0.49
17	1,190	0.44
18	934	0.35
19	1,251	0.47
20	36,686	13.69
21	29	0.01
22	1,601	0.6
23	16,150	6.03
24	22,979	8.57
25	11,307	4.22
26	4,273	1.59
27	2,316	0.86
28	2,631	0.98
29	4,836	1.8
30	2,259	0.84
31	266	0.1
32	5,629	2.1
33	1,126	0.42
34	1,238	0.46
35	726	0.27
36	3,755	1.4
37	403	0.15
38	1,034	0.39
39	245	0.09
40	585	0.22
41	82	0.03
42	16,568	6.18
43	366	0.14
46	615	0.23
47	62	0.02
48	18	0.01
54	85	0.03
62	1,502	0.56
63	263	0.1
80	2,116	0.79
Total	267,996	100

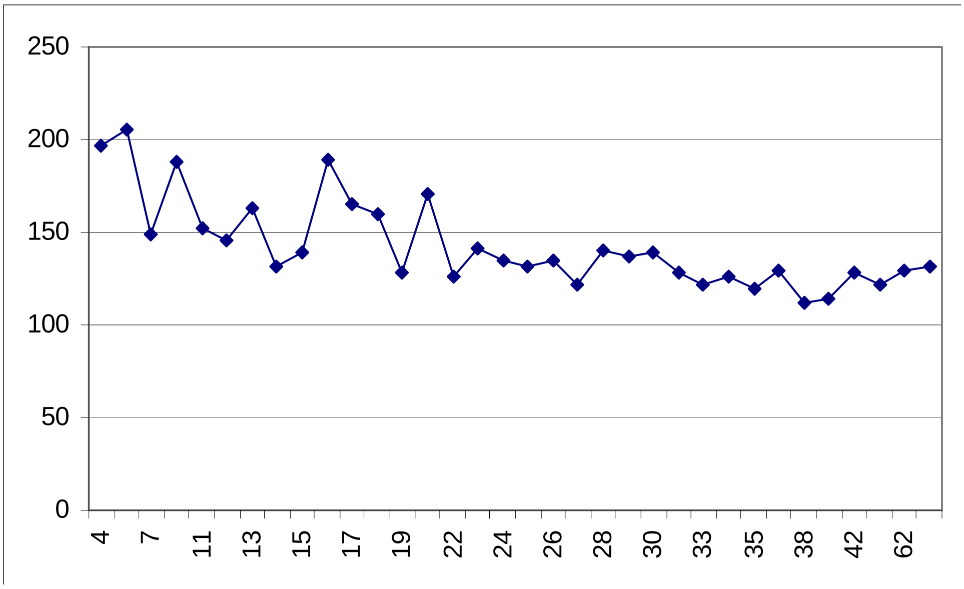


Figure A2-1: Age Heaping (measured in Whipple Index values) by occupational income score.

Appendix 3: Height and human capital outcomes during the 1950s-1970s (not be included in the publication)

How plausible is it to use the human capital values of the tallest 50 per cent versus the shorter 50 per cent? We applied the method to the 'Demographic and Health Surveys", which recorded for a large number of women in 38 less developed countries (406,171 individual cases born 1950-79, aged 23-52) height, literacy and years of schooling. We calculated the ratio of literacy for the tallest 50% to the shortest 50% in each country and decade of birth. This ratio was almost always above 1, indicating that the tallest 50% were always more literate (Figure A3-1, horizontal axis). Doing the same exercise for years of schooling, again the tallest 50% had almost always the higher schooling values (Figure A3-1, vertical axis). Moreover, the correlation between the literacy ratio and the school year ratio was very close (correlation coefficient 0.91), suggesting that the relationship is consistent. Possible explanations for this relationship range from a better family background over positive parental height discrimination to infant protein deficiency syndrome, or a combination of those factors. Since the main objective of this little exercise is to produce evidence to support or reject the strategy to proxy social standing within a given population by division into taller and shorter half, a detailed study of the underlying causal mechanisms is not conducted here. We consider the results based on DHS data to support our strategy.

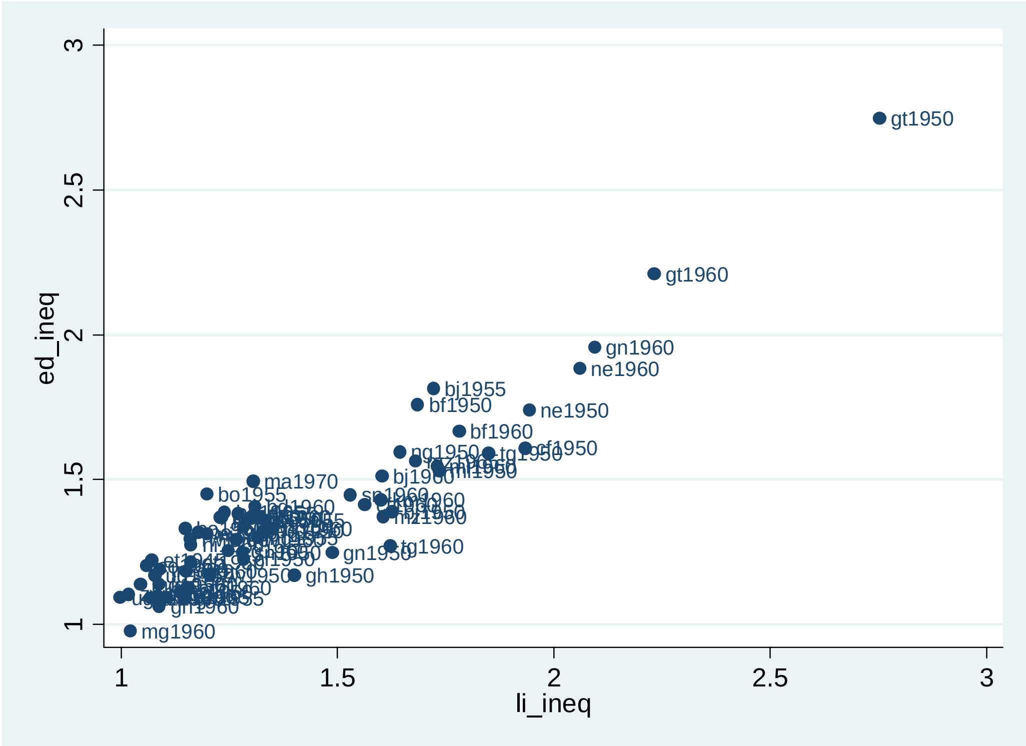


Figure A3-1: The literacy ratio and the years of schooling ration in linear form (tallest vs. shortest 50%)

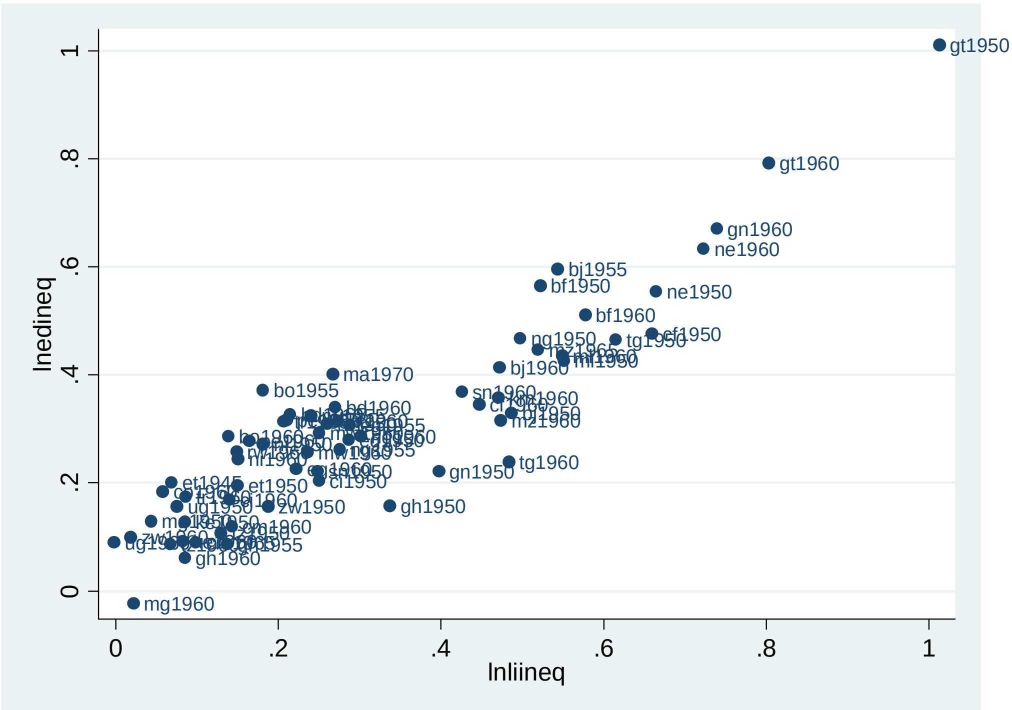


Figure A3-2: The literacy ratio and the years of schooling ration in logarithmic form (tallest vs. shortest 50%)

Is the Whipple index differential by tall and short correlated with the literacy differential? We regressed the literacy ratio as defined above on the Whipple index ratio, taking only those cases into account which had a substantial age heaping (i.e. Whipple indices of 105 or higher). There is clearly a positive relationship between the two, as the R-square is around 0.23. Moreover, the size of the coefficient is reasonably large. Hence, although the DHS surveys might produce Whipple indexes with quite a bit of measurement error, the relationship between literacy inequality and numeracy inequality is again quite substantial.

Table A3-1: Regression of Literacy inequality (tallest vs. shortest 50%) on numeracy inequality (same definition), DHS-countries, birth cohorts 1950s-1970s

Dependent Var.: Inequality of Literacy (log)	(1)
Inequality of Numeracy (log)	-2,07
	(0.00)
Constant	0.29
	(0.00)
Observations	49
R-squared	0.23

Notes: only observations based on 250 and more cases included.

4 POOR, HUNGRY AND STUPID: NUMERACY AND THE IMPACT OF HIGH FOOD PRICES IN INDUSTRIALIZING BRITAIN, 1780-1850

Abstract: This paper argues that low levels of nutrition impaired cognitive development in industrializing England, 1780-1850. Age heaping is used as an indicator of numeracy, as derived from census data. We show that numeracy declined markedly for those born during the war years, especially when wheat was dear. Also, individuals born in periods or countries characterized by higher age heaping sorted into occupations with limited intellectual requirements which also paid lower wages. In addition, we document that England's nascent welfare state mitigated the adverse effects of high food prices on cognitive skill. Where the Old Poor Law provided for generous relief payments, the adverse impact of high prices for foodstuffs was mitigated.

This chapter is based on a working paper written jointly with Hans-Joachim Voth (Universitat Pompeu Fabra) and Joerg Baten (University of Tuebingen). The concept for the paper was developed jointly, the analyses and writing were equally shared.

I Introduction

Nutrition in the past was often poor. Countries whose populations today rank amongst the tallest in the world two centuries ago had low average heights. Dutchmen and Norwegians males in the 18th century had a mean height of only 165 cm, below the 10th percentile of the US population today, and below the levels attained by the British, Irish and French (Fogel 1994). Social differences in heights were also marked. Boys from the London slums were up to 9 inches (23 cm) shorter than those attending the Royal Military Academy at Sandhurst (Floud et al. 1990).[113] Stature is known to be a good indicator of cumulative net nutrient intake during an individual's growing years. While short-term deficits can normally be compensated – a phenomenon known as catch-up growth – sustained shortfalls affect terminal heights. Because Europeans' genetic composition has changed little in the last two centuries, historic heights reflect how severe chronic malnutrition must have been in the more distant past.

Using a newly-constructed dataset on numeracy in England, 1770-1850, we argue that scarce nutrition harmed not just physical development, but cognitive ability as well. Our conclusions are in line with recent research showing that nutritional status is highly correlated with laobr market success. In modern studies, wages are more strongly correlated with intelligence than with education (Murnane, Willett and Levy 1995).[114] Recent work has also shown that heights are strongly correlated with IQ measures, and have considerable explanatory power for wage outcomes (Persico, Postlewaite and Silverman 2004; Case and Paxson 2006).

Our evidence has an important implication: output in pre-modern times may have been low because many workers were stupid. In his controversial *Farewell to Alms,* Clark (2007)

---

[113] The same is true of children from single-parent families (Horrell, Humphries and Voth 2001). The average difference reported is from recalculations of the Sandhurst and Marine Society data (Komlos 2005).

[114] Komlos (1989) argues that nutrition mattered at the opposite end of the skill spectrum as well. He observed that many innovators of the Industrial Revolution in the UK were born during the good times of the 1730s, when food prices were low.

argues that the poverty of both the European past and many modern-day societies is driven by cultural norms constraining output. He concludes that neither differences in education nor in infrastructure nor in machinery can account for the low output of Europeans in the early modern period, or of textile workers in India. Clark's finding that most of the differences in per capita income in the past were driven by TFP, not factor inputs, echoes the conclusions of Hall and Jones (1998). In contrast, our results suggest that, instead of 'socially induced lethargy' (Clark 2007), output may have been constrained by low numeracy and limited intellectual ability in general.

The indicator of cognitive skill we use is numeracy. Numeracy has high predictive power for wages and employment.[115] Unfortunately, data on IQ or math test scores are not available for the more distant past. Instead, we focus on age heaping. Self-reported age data often show a tendency for people to 'round off' to the nearest multiple of 5 or 10 (Mokyr 1983, Myers 1976). Roman tombstones, for example, show high rates of age heaping (Duncan-Jones 1990). Age heaping is widely regarded as a good indicator of numerical skill. Gradual changes in heaping over longer periods can reflect a number of factors, such as schooling, the importance of administrative procedures relying on age, and evolving cultural norms. These factors are unlikely to explain abrupt changes over short periods. Short, sharp shocks to numeracy are more likely to reflect environmental factors. In historical data, numeracy, as measured by more accurate age reporting, is correlated with stature. We argue that numeracy in the past can be captured by using age heaping, and – in part – reflects the influence of a nutritional component. The paper concludes that negative shocks to the availability of nutrition during the Napoleonic Wars resulted in much higher age heaping in industrializing Britain.

[115] Rivera-Batiz 1992.

Using a new dataset derived from the 1851 and 1881 censuses, we demonstrate that the Napoleonic Wars saw a sharp decline in cognitive ability amongst Britons born in these decades. The decline in numeracy also had significant effects on labor market outcomes, such as occupational choice and wages. We first compile measures of age heaping for those born from the 1780s onwards, and demonstrate that numeracy fell precipitously in England between 1790 and the early 1800s as grain prices surged. During the Revolutionary and Napoleonic Wars, the price of wheat more than doubled; the number of erroneously reported ages at multiples of five doubled, from 4 to 8 percent. Cross-sectional evidence supports our hypothesis. We merge our new dataset on numeracy with information on the generosity of poor relief under the so-called Old Poor Law from Boyer (1990). As the nutritional crisis unfolded, English regions that offered substantive support to their poor saw much smaller declines in numeracy. Age heaping still increased systematically in those parishes that were relatively generous to the poor, but much less so than in those areas with limited support payments. Britain was one of the first states to offer income supplements for able-bodied adults in need of support (outside a workhouse). Moreover, the system was comparatively generous. Its cost was high, consuming as much as two percent of GDP (Mokyr 1993). At a time when an agricultural laborer could expect to earn 22-35 shillings a year, relief expenditures per recipient ranged from 7 to 19 shillings (Boyer 1986). Generosity was determined at the parish level, by the overseers of the poor. Funding was also raised locally, through property taxes. Economic factors explain part of the differences in generosity. Some regions had much greater incentives to retain a large number of able-bodied poor than others (for example, to retain a workforce of sufficient size to cope with the harvest). We control for these factors separately. Overall, we argue that generosity of poor relief largely reflected idiosyncratic differences in attitudes towards supporting the poor.

Related literature also includes the extensive work on nutrition and cognitive development. There is growing evidence for nutrition affecting intelligence directly. We discuss this literature in more detail in Section II. Our results also relate to recent anthropometric research that has sought to measure nutrition in the past, mainly based on heights (Steckel 1995, Komlos 1994, Fogel 1994). Other related research includes work on human capital formation in industrializing Britain (Mitch 1998, Schofield 1973). Finally, our findings have a bearing on research into the origins of accelerating growth after 1850. One class of unified growth models (Galor and Weil 2000, Sunde and Cervellati 2005) has aimed to join human-capital based interpretations with models of fertility choice, arguing that more investment in the skill of the workforce was crucial for the transition to self-sustaining growth. While we do not examine these arguments directly, we document how nutrition constrained a key dimension of pre-modern human capital – numeracy.

Section II reviews the literature on the link between IQ, malnutrition, and labor market performance. Section III describes our preferred measure of numeracy based on age heaping, and Section IV discusses the datasets we use in more detail. Our results are presented in section V. We show evidence from difference-in-difference estimation that nutritional availability in industrializing Britain influenced numerical ability. We also document that Englishmen born in the hungry decades of the 1790s and 1800s sorted into jobs with lower skill requirements – especially those from areas with limited poor law support. Section VI concludes.

II Nutrition, cognitive ability and occupational outcomes

In this section, we review the literature linking nutrition and cognitive ability. There is strong experimental evidence that protein availability in childhood influences intellectual ability. In one study of preterm infants, the protein content of the diet was varied on a random basis (Lucas 1998). Children receiving less nutrient-rich diets showed markedly lower neurodevelopment (lower mental development scores and psychomotor scores) at the 18 month follow up than the control group. These effects could still be detected as late as at age 7.5, when IQ scores were significantly lower. Since intelligence tests at age 7-8 predict adult cognitive ability, it is likely that infant nutrition can harm cognitive development in a major way. Other randomized trials of stunted children similarly show that nutritional supplements can produce important gains in intellectual development (Grantham-McGregor 2002). Studies on mammals suggest that both pre- and post-natal nutrition have a strong impact on brain development (Winick and Rosso 1975). Low birth weight in humans in particular predicts lower cognitive scores (Richards et al. 2002). Malnutrition between 1 and 16 months also is a strong predictor of poor cognitive outcomes (Lloyd-Still 1976).

While sensitivity is great in utero and in early childhood, nutrition during the second decade of life also appears to have strong effects. Recent studies found a clear cumulative effect of persistent exposure to undernutrition and poverty. The longer a child's nutritional, emotional and educational needs are not filled, the greater his or her cognitive deficits (CHP 1998, Paxson and Schady 2005). In this study, we follow the literature and consider the impact of food prices during the 10 year period following the birth decade to capture the most important and cumulative nutritional effects during childhood.

A positive correlation of heights and cognitive scores also points in this direction. The heights of individuals are in part determined by parental genes. The same is probably true of intelligence. In populations, however, the gene pool stays approximately constant over time. Changes in average heights primarily reflect the influence of environmental factors up to young adult ages (Steckel 1995). Intelligence is likely to be affected in the same way. Richards et al. (2002) used data on IQ scores and height at various ages for a large British post-war sample, and find that the variables are strongly and positively correlated. In particular, maximum height gain during early childhood and the timing of the adolescent growth spurt predict cognitive ability. There is also some evidence that rising IQ scores in developed countries may partly reflect improving nutrition, and not better education (Hiscock 2007; Lynn and Vanhanen 2002). A randomized experiment in Guatemala suggests that protein supplements can produce marked improvements in cognitive ability (Brown and Pollitt 1996). Genetic factors play a role, but do not dominate. While results vary, studies of Scandinavian twins reason that genetic influences cannot explain the correlation between heights and cognitive ability (Magnusson, Rasmussen, and Gyllensten 2006).[116] Earlier studies of malnourished children and their (better fed) siblings also suggested that nutrition, in addition to parental IQ scores, is a prime determinant of cognitive performance (Craviato and deLicardie 1975).

There also appears to be little 'catch-up' in cognitive scores, except in the case of very brief shocks. Different studies have tracked the effects of a disadvantageous early environment into late middle age and beyond retirement. Abbott et al. (1998) conclude that men in their 70s show lower cognitive ability if they were shorter. Richards et al (2002) and Richards and Wadsworth (2004) conclude that the negative effects of a deprived childhood can be found in IQ scores measured at all ages up to 53. Paxson and Schady (2005) find that, in a large sample from

---

[116] Sunder et al.(2005) argue that height and intelligence may be jointly determined by parental genes, and argue that

Ecuador, test scores for shorter children are significantly lower than for taller ones. Alderman et al. (2006) show that in rural Zimbabwe, civil war and weather shocks caused childhood malnutrition. Long-term consequences include reduced stature, lower grades in school, as well as lower attendance overall. The one exception appears to be the Dutch hunger winter in 1944-45, where retreating German forces left part of the population starving for 5-6 months. Children exposed to the famine in utero or immediately afterwards showed no systematic sign of lower cognitive development in later life, despite reductions in birth weight (Stein and Susser 1976). This suggests that for shocks to have long-term effects, they have to last for an extended period.

Cognitive ability also has a clear effect on labor market outcomes. Zax and Rees (2002) show that intelligent members of the workforce earn substantially more. Heckman (1995) also finds IQ to be one important predictor of wages. Based on controlled experiments in today's Third World, Behrman (2006) argues that the correlation between height and wages may largely reflect nutrition's impact on cognitive development, and not strength or resilience to disease. This is in line with the argument in Case and Paxson (2006). In related work, Lynn and Vanhanen (2002) offer a broad-ranging survey and argue that intelligence is a prime determinant of the wealth of nations. Jones and Schneider (2006, 2007) shows that, in a simple Ramsey framework, present-day cross-country differences in IQ can account for a large proportion of the gap in productivity.

III Numeracy

Using age heaping as an indicator of numeracy is not new. Bachi (1951) and Myers (1976) showed that across countries and within them, richer, more educated populations were less prone to show age heaping. Historical applications include the work of Herlihy and Klapisch-Zuber (1978) on fourteenth century Florence, Mokyr (1983) on selectivity bias among Irish emigrants,

---

this accounts for approximately 30% of the observed co-movement.

and Duncan-Jones (1990) on the Roman Empire. Over the very long run, numeracy as proxied by age heaping varies strongly with income, and is highly correlated with literacy (Clark 2007, A'Hearn, Baten and Crayen 2006).

The most commonly used indicator of age heaping is the Whipple index.[117] It calculates the number of self-reported ages that are multiples of 5, relative to the number expected with a uniform distribution of ages:

$$W = 100 \frac{\sum (n_{25} + n_{30} + n_{35} ... n_{60})}{\frac{1}{5}\sum_{i=23}^{62} n_i} \qquad (1)$$

The range of ages has to be chosen so as to include the same number of ages for each terminal digit (in this case, 23 to 62). There is substantial evidence that the Whipple index dominates competing estimators like the Bachi measure, in particular in terms of accuracy at low levels of heaping (A'Hearn, Baten and Crayen 2006). The index ranges from 0 to 500. Accordingly, a Whipple Index of 0 (500) implies no (only) ages ending in multiples of 5. At 100, it would imply that exactly 20% of the population report ages ending in multiples of 5 .

It could be argued that the ability to recall one's age correctly is indicative of schooling, the bureaucratization of life, and changing cultural norms rather than of cognitive development. Where it varies considerably over short periods, it is unlikely to be a result of cultural norms and administrative procedures. Since the use of age and birthdays to identify individuals and the prevalence of schooling have generally been on the rise of the last two centuries, there is an asymmetry in how we should interpret short-term fluctuations. Increases could be driven by, say, the introduction of compulsory schooling (in the later 19th century in most European countries).

Where numeracy falls sharply, on the other hand – and without a collapse of the school system at the same time – other factors should be at work. This is the logic behind our examination of the sharp falls in numeracy in England around 1800.

IV Data

We combine two main datasets – the Poor Law data collected by Boyer (1990), and the 2% and 5% samples from 1851 and 1881 census, respectively. These are then supplemented by additional information on grain prices and on historical weather patterns.

Boyer compiled information on the generosity of outdoor relief under the Old Poor Law. His data is based on a survey by the Poor Law Commission, conducted in the summer of 1832. Motivated by growing concern about the surging cost of poor law provision, it sent out a questionnaire, called the Rural Queries, to the approximately 11,000 parishes of England. They received answers from ca. 10% of them. Of these, Boyer used a sample of 329 parishes in 21 counties in Southern England.[118] The returns include information on average relief expenditure, summer and winter unemployment, the existence of allotments, the percentage of land used for grain production, and the presence of cottage industry, as well as the annual income of agricultural laborers. We do not have access to all necessary data at the parish level. The data at the county level is publicly available. We therefore aggregated the age data from parishes at county level. This had the additional benefit that it allows for a more accurate calculation of Whipple indices.

The age heaping data is derived from self-reported data on ages as they emerge from the 1851 and 1881 censuses. The collection methodology in each case was similar. The aim was to collect information on all individuals who spent the night of 30 March 1851 in a particular home.

---

[117] For an overview of different indicators, cf. A'Hearn, Baten, and Crayen (2006).
[118] A total of 735 returns came from Southern parishes. Boyer selected the most complete ones.

Information on the age of household members was self-reported. The Census samples were taken and coded from the original returns by Michael Anderson (1987). Occupations were initially classified according to the scheme devised by Armstrong (1972). We later adapt this scheme to more modern standards.[119] In addition to the reported age, we use information on gender, the county of birth, and occupational information.

That age reporting was not fully accurate in the censuses has been known for some time.[120] The General Report for the 1891 census argued that 'a very large proportion of persons, not improbably the greater number of adults, do not know their precise ages and only report it approximately'.[121] Thomson (1980) traced individuals' self-reported ages across the 1861, 1871, and 1881 censuses. He found that for both men and women, the correct age (found by adding 10 or 20 to the earlier reported age) was only given by 38-64 percent of respondents aged 60 and over. Up to 30 percent gave answers that were wrong by more than two years.[122] Various other studies checked age recording between two or several 19th century British censuses and found that between 52 and 65% of all men and women reported their ages consistently, while 2-11 percent were inconsistent by more than two years (Anderson 1972, Higgs 1989, Robin 1995, Tillot 1972). Adjustments, on the whole, were as likely to be up as down, suggesting that genuine mistakes – and not a desire to appear older or younger – were to blame. As late as 1951, only 94 percent of men and 64 percent of women reported their ages correctly.[123]

We use the snapshot data from the 1851 and 1881 censuses to compile information on age heaping by birth decade. Our earliest birth decade is the 1780s; the latest, 1850. In order to include a sensible number of multiples of five, we use the period 1779-1788 for the 1780s, 1789-

---

119 They are available at http://www.data-archive.ac.uk/.
120 Apart from the heading of the appropriate column in the household schedule which said 'Age [last birthday]', no general instruction was given to households how to report their age.
121 1891 Census, p. 27.
122 We find markedly lower rates of age heaping.
123 1951 Census, p. 36.

1798 for the 1790s, etc. Whipple indices range from 94.4 (indicating underreporting of ages ending in a multiple of five) to 176.5, with an average of 118. These scores indicate that the Englishmen in our sample are not from a population with particularly low literacy, by historical standards. On average, about 4-8 percent of respondents misreported their age.

Age heaping in the 1851 and 1881 census varied considerably between the different counties in our sample. Figure 1 shows the differences in age heaping between the average Whipple by individual county, over the entire sample period 1780-1850. The range is substantial. While the least numerate county (Wiltshire) shows Whipple scores nearly 125, the most numerate one (Buckinghamshire) scores a much better 109. Most counties fall into the range 120-125, indicating that multiples of 5 are overreported, with 25 percent higher than if all respondents had remembered their ages perfectly. One crucial question concerns age-specific changes in the respondents' ability to remember their age correctly.[124] If age alone leads to a deterioration of numeracy, we should find that, say, the 60-year-olds in the 1881 sample have higher Whipple scores than the 30-year-olds in the 1851 sample. Table 1 shows Whipple scores for the same birth cohorts from the two census samples. While age heaping is never identical, there appears to be no systematic pattern that would undermine belief in the usefulness of the indicator or the samples chosen. In particular, three out of four differences between the two datasets indicate lower age heaping in the later sample, when respondents were older. If anything, greater age should have made it more difficult for people to recall their ages. We conclude that here is no simple mapping from age to age heaping, and that the use of the 1881 census for our purposes is unproblematic.

The grain price data w as collected by Liam Brunt and Edmund Cannon from contemporary prize gazettes.[125] Acts of Parliament ordered the collection of grain price data

[124] In the main empirical section, we will use a fixed-effects approach to bypass some of the underlying difficulties.
[125] The authors kindly made their data available to us as county-year averages. The source is described in more detail in Brunt and Cannon (2005).

during the period 1770 to 1863.[126] In most years, between 140 and 290 towns reported prices. While information on a number of different grains was recorded, we focus on the price of wheat. It was the main staple of eighteenth and nineteenth century British diets. As Figure 2 illustrates, wheat flour alone accounted for 27% of working class expenditure on food.[127] Bread – largely baked from wheat as well – took up another 20% to the food budget. Together with oatmeal, grain-based food accounted for 60% of the food budget, or 40% of the consumer basket overall. To gauge the importance of wheat in particular, and grain more generally, we also have to add part of the 10% spent on drink. The largest share of this would have been consumed in the form of beer, derived in large part from wheat and barley.

During our sample period, average wheat prices in our sample rose sharply in 1795/96, 1800/1801 and in the late 1810s. At their peak, they were more than twice as high as they had been in the 1780s. It was the pressure produced by these food prices that inspired Frederic Eden and David Davies to undertake their surveys of working class families on which the budget data cited earlier is based.[128]

Economic historians do not agree on the factors responsible for the price surges apparent in Figure 3. Britain had become a net food importer in the 1760s. In years with poor harvests, her import needs were substantial. Mancur Olson (1963) described 'food…as the weakest link in Britain's chain of defense'. The Revolutionary and Napoleonic Wars limited the extent to which domestic misharvests could be mitigated by grain from abroad. Insurance rates for shipping to the Baltic were high throughout the war (possibly three times their post-war level), as both sides used privateering to destroy the merchant fleet of their adversary (Jacks 2007; Mokyr and Savin 1976).

---

[126] 10 George III, 31 George III, 1 and 2 George IV, 9 George IV, and 5 Victoria

[127] The figure is from Voth 2003, and is based on data from Feinstein 1998.

[128] Data in Horrell (1996) show a strongly negative own-price elasticity of demand for bread, and a positive (but smaller) one for flour. This suggests that price increases led to sharp falls in bread consumption, and that some substitution to home-baked bread took place.

In a bid to hurt the island's trade, the Berlin decree of 1806 instituted the Continental System, denying European ports to British ships. Neutral shipping was also severely curtailed. The system was at its peak in the years 1807-12. While the French supplied Britain with grain in 1810, they did so while charging export licensing fees that more than quadrupled the price of grain at source (Jacks 2007).[129] In some years, when price differences were greatest, imports continued to flow into Britain, even from direct military adversaries. However, transactions costs inevitable rose, limiting the extent to which domestic weather shocks could be arbitraged away. Bread riots in 1795, 1800, and 1812 reflect the extent to which the precarious food situation put pressure on vulnerable groups in English society.

Table 2 contains the data descriptives for our key variables. Since our unit of analysis is birth decade, county, and gender, half of our sample by definition is female. Grain prices fluctuate over time and across localities (Figure 3). Relief payments vary widely between parishes (Figure 4). Grain-growing areas make up 12 percent of our sample.

[129] According to some estimates, the UK imported around 15% of its total food in 1810 (Jacks 2007).

V Empirical Results

In this section, we show that across a wide range of samples, from different time periods, countries, and social groups, the well-nourished show greater numeracy. We then document that numeracy fell precipitously in England as grain prices during the Napoleonic Wars surged. Declines in numeracy were particularly pronounced in counties where (i) grain prices were particularly high (ii) income support for the poor was less generous. We show that the exogenous component of grain price changes, as driven by weather shocks, was an important determinant of numeracy. Finally, we examine issues of endogeneity, present evidence on the effect that nutrition-related shocks had for labor market outcomes, and discuss some caveats.

The first question concerns the link between nutrition, cognitive ability and numeracy. While the influence of nutrition on cognitive ability appears well-established, the connection between age-heaping and cognitive facilities is less clear-cut. A large number of factors unrelated to cognitive ability – such as schooling, changing cultural norms, and bureaucratization – might influence the extent of age-awareness. In the appendix, we show that in modern data, greater heaping is associated with lower cognitive scores. To address this issue, we turn to heights as an indicator of cumulative nutritional status since childhood. As is well-known, well-nourished individuals stand a better chance to reach their genetic potential in terms of height. In Table 3, we present data from the US, France, Ireland, and the UK, from the 1660s to the 1840s. The samples are divided into 'tall' and 'short', according to whether they are above or below the median. We then calculate Whipple indices for both groups. Throughout, the tall are less likely to misreport their age. In some cases – such as the data from Wandsworth prison– the difference is small. In other cases, such as the Irish prisoners sent to Australia, and French Army recruits from Paris, the differences are marked, with Whipple indices that are 20-40 percent higher for the shorter group than for the taller one. Since the samples are drawn from relatively homogenous backgrounds, this strengthens the *prima facie* case in favor of a link between nutrition and our indicator of cognitive ability, age heaping. In his analysis of nineteenth century Bavaria conscripts, Schuster (2005) finds that the shortest also had unusual high rates of incidence for exceptionally low intelligence. If we believe the existing evidence from modern-day sample documenting the nexus between nutrition and cognition, it also strengthens the belief in the usefulness of age heaping as a measure of numeracy.

Did years of high prices affect numeracy? Figure 5 plots the Whipple indices over time. After the outbreak of the Napoleonic wars, Whipple indices rise sharply in both generous and less generous counties. However, counties with limited relief show a markedly higher peak. Their Whipple scores stay above those for the generous counties until the 1850s. While not conclusive proof that the poor in parishes with low income support suffered nutritional insults and harm to the cognitive development of their offspring, the pattern is consistent with such an interpretation.

We next examine these patterns more systematically. Table 4 shows OLS regressions, using the Whipple index as the dependent variable and national wheat prices as well as relief generosity as explanatory ones. Grain prices consistently drive up age heaping in our sample. On average, a one standard deviation increase in national wheat prices pushed up the Whipple index by 5.2 points (eq. 1 and 2). Counties with generous relief lowered their Whipple scores by 4 points (eq. 2). Equation 5 employs a continuous transformation of the poor relief variable to test if numeracy declined consistently in those parishes where relief payments were smaller. Instead of the simple dichotomous variable that codes counties as generous or not, we define [R_{max}-R_i], where R_{max} is the maximum relief payment per capita, and R_i is the relief payment in county i. It expresses the shortfall of relief payments relative to the most generous county (Sussex) in our sample. We find that lack of poor relief consistently and strongly predicts higher Whipple scores, and that the use of this continuous measure does not undermine the size and significance of the grain price variable.

Do higher grain prices lower cognitive ability to the same extent in generous and in 'mean' counties? Equations 3 and 4 split the sample into two, according to whether parishes are above or below the median for poor relief. We find that in those parishes where support is limited, grain prices have a particularly strong effect on Whipple scores (eq. 4), with a one standard deviation increase pushing up age heaping scores by 5.8 points. In those counties where support payment are generous, however, the effect of higher grain prices is much smaller. Figure 6 illustrates this graphically. We plot Whipple scores by county and birth decade against grain prices (during the birth decade). Observations from counties with above-average poor relief payments are marked 1; less generous counties are marked 0. While there is a lot of variation that our setup cannot account for, the regression line for the more generous counties much flatter than the one for the more restrictive counties.

The evidence in Table 4 suffers from one important drawback – possible bias from unobserved heterogeneity. To address the issue, we use difference-in-difference estimation:

$$W_{i,t} = a_i + \beta G_{i,t} + \gamma X'_{i,t} + \varepsilon \qquad (2)$$

where $W_{i,t}$ is the Whipple index for county i at time t, *a* is a county-specific intercept, $G_{i,t}$ is the grain price in county i at time t, and X' is a vector of controls. Alternatively, we use a dichotomous relief payment indicator, based on whether parishes are above or below the median for poor relief. The results are presented in Table 5. The estimated coefficient on G is identical to the OLS results, with a rise of 2.7 Whipple points for every 10 additional points of grain prices (eq. 1), and also the results for the relief dummy variable are very similar. In counties that are relatively generous to their poor residents, Whipple indices are on average 4.35 points higher (eq.

3). The magnitude of this effect more than doubles after controlling for grain prices. When we include other explanatory variables such as cottage industry, population density and the share of grain growing area, the negative effect of grain prices on numeracy even increases slightly (eq. 5).

Endogeneity

There could be some endogeneity in our setup, with higher grain prices causing the workforce to be less well-fed and energetic. This, in turn, could cause a reduction in a county's grain output. To sidestep potential endogeneity issues, we use an instrumental variable approach. In eq. 2 and 7, we predict the main explanatory variable with the ratio of annual spring rain to its long term average. More rain in the spring was bad for crops, raising prices – the first stage regression has a t-statistic of 16 on the spring ratio, and an R^2 of 0.5. In equations 6 and 7, in a setup similar to Rajan and Zingales (1998), we replace G with $G*[R_{max}-R_i]$. The idea is to examine if counties offering less poor relief suffering higher grain prices had systematically higher Whipple scores. The coefficient on the interaction term is highly significant. It is also large: a one standard deviation increase raises age heaping scores by almost 11 points. If we include year dummies in the fixed effects estimation (not reported in Table 5), the results are nearly identical.

The Skill Requirements of Occupations

Did the declining numeracy levels matter for occupational outcomes? We use occupational information given in the census data to classify individuals according to the coding from the Dictionary of Occupational Titles (England and Kilbourne 1988). Their study offers scores for the skills required for a wide range of jobs. Our main interest is in the cognitive skill requirement of jobs performed by individuals in our sample who suffered from high grain prices. In nineteenth century Britain, class and parental income were major determinants of access to higher education. Families that could send children to university were unlikely to suffer from the dear food prices during the Napoleonic wars. We therefore exclude the professions requiring the highest skills (code 1-199 in the England-Kilbourne scheme – basically all professionals such as architects, medical doctors, civil engineers, etc.). For the rest, we find that those born in years and counties with lower average numeracy caused by high grain prices or low relief payments selected into occupations that required less intelligence, compared to their peers born in years and counties of better numeracy levels.

Table 6 reports the fixed effects regression results. Given that the Kilbourne/England scheme is structured like a school grading system, that is higher scores mean lower intelligence requirements, our results suggest that the Whipple Index indeed lowered the average the intelligence requirements of occupations (eq. 1). Since we include time and county fixed effects, general structural change is accounted for. When we include additional control variables, the coefficient of the Whipple Index proves robust but loses some of its significance (eq. 2). To correct for potential endogeneity, we use national wheat price and the high relief dummy variable as instruments for the Whipple Index (eq. 3 and 4). The results are nearly identical to the previous regressions, implying an increase in the intelligence scores by 15% of the standard deviation for every 1-standard deviation rise in the Whipple Index. These effects are relatively small, partly reflecting the highly aggregated nature of our data.

Economic Impact

So far, we have documented that being born in a low-poor relief parish increased age-heaping. It thereby also raised the likelihood of sorting into jobs with low skill requirements. The economic impact remains to be examined. To do so, we use the census information about occupation to impute average earnings by occupational group. We attribute average earnings provided by Long (2006) and Williamson (1980, 1982) to each individual in our dataset, and then use our set of explanatory variables to assess the impact of declining numeracy levels on labor market outcomes several decades hence.

Table 7 gives the regression results. Higher whipple scores go hand in hand with lower earnings. This is true in both the OLS and IV estimates. We use the national grain price index as an instrument, as well as the variation in spring ratio. Both show similar results. The largest estimate suggests that for every standard deviation increase in the Whipple index (12.3 points), average earnings at the county level declined by up to 2.3 pounds sterling, or 4 percent relative to the sample average (25% of 1 standard deviation). Since there is likely to be attrition bias arising from the aggregation, these effects are substantial.

Caveats

Our analysis assumes that wheat prices are a good proxy for the general price of food, but alternative sources of calories were clearly available. Those suffering from high grain prices could have substituted away from relatively dear sources of calories, thus mitigating the impact of dear wheat. A more comprehensive measure of the price of food should also capture that cheaper substitutes such as potatoes. In crisis years, their price also rose dramatically. While wheat prices increased drastically between 1798 and 1800 by 73%, substitution possibilities were limited. Rye prices also increased by 55%, and potato prices reacted even more sharply, rising by 78 percent. The magnitude of price changes was similar in 1812. When wheat prices increased by 34%, potato prices shot up by 81 percent (compared to the non-crisis level in 1806). In general, the correlation between wheat and potato prices in the difficult period between 1793 and 1817 was 0.57. In short, while many Englishmen clearly tried to avoid hunger in its most extreme form, by switching from wheat bread to potatoes, this was not a simple solution for the hungry masses. Rapid price increases for all staples caused a deterioration of diets overall during crisis years. Crucially, little or no money would have been left to purchase food rich in proteins, such as meat, fish, eggs and milk. Since the effect of nutrition on cognitive development probably depends on protein availability (Lucas 1998), this must have biased downwards the chances for infants to develop their full potential.

The decline in numeracy is concentrated during the Revolutionary and Napoleonic wars. Britain fought a war that required unprecedented military, fiscal, and economic mobilization (Brewer 1990). Alternative mechanisms could have caused increased age heaping. For example, wartime dislocation brought about by the absence of fathers may have led to family instability. Passing on information about the age of children could have been disrupted by large-scale mobilization. We think this is unlikely, for a number of reasons. First, since Britain was still fighting the American War of Independence until 1783, and the Fourth Anglo-Dutch War until 1784, establishment size of the armed forces was not that much smaller in the baseline period of the 1780s compared to the 1790s and 1800s. The actual date range for the decade is 1779-1788, comprising five war years (1779-83). Second, the single best indicator for family instability – illegitimacy rates – showed only a small uptick, increasing from 4.6 percent in 1750-74, 5.9 percent in 1775-1799 and to 6.2 percent in 1800-24 (Wrigley et al. 1997). Even if all of the additional 33,000 illegitimate births were caused by the wars, this would pale in comparison with the total rise in misreporting.[130] Third, the army and Royal Navy did not satisfy their demand for manpower by recruiting bachelors who would otherwise have gone on to found stable families. As George Chalmers (1812) put it, in Britain, 'the sword had not been put into *useful* hands.'

[130] Average whipple scores increased by 15 points in our sample between the two periods. This corresponds to 8.1 percent of Englishmen born in the period, and surviving to 1851 or 1881, misreporting their age – an additional 3.7

Press gangs routinely rounded up vagrants and other unproductive elements. Impressment was limited to 'such able-bodied men as had not any lawful calling or employment'.[131] Also, many men in the armed forces – officers and privates alike – joined in their teens. Average age at marriage in England in 1800 was 25 for men (Wrigley et al. 1997). This means that probably less than half of the men in the armed forces would have been of an age when they would have married onshore – and many did regardless of their profession. Finally, the numbers for total enlistment include foreigners recruited into the British army. The British army in 1813, for example, consisted of 203,000 British troops and 53,000 foreign ones (Smith 1998; Hall 1992). Since one out of five British soldiers were not from the British isles, any negative effects on family stability that there might have been was probably mitigated.

Our results establish a prior that the availability of adequate nutrition was important for numeracy, and that Poor Law provisions helped the most vulnerable parts of the country fight the effects of high grain prices. Nonetheless, we cannot rule out that other factors – to the extent that they are correlated with the generosity of poor relief – were responsible for our results in the cross section. There is no direct evidence on the prosperity of individual counties. The validity of our results hence rests on the plausibility of the mechanism we describe, with no possibility of controlling for other variables that might also have provided alternative 'safety cushions' for the poor.

Access to schooling may have suffered during the Napoleonic wars. Educating children is an investment. If general economic conditions deteriorated during the war, a decline in schooling – rather than a decline in nutritional standards – could be responsible for the lower numeracy attained during these decades. However, information on trends in basic literacy – as proxied by

---

percent compared with those born 1779-1789. In a population of 21 million, this implies 777,000 extra cases of misremembered information on age.

[131] Cit. acc. to Brewer 1990, p. 49-50.

the ability to sign one's name – do not support this alternative interpretation. Schofield (1973) found that illiteracy rates for men and women thus measured were broadly stable or gradually declining between 1750 and 1840. The general view is that the acquisition of basic skills in England took place outside day schools before the 1870s (Mitch 1992). There is no evidence of a sudden fall in signature rates during the Napoleonic wars. Nicholas and Nicholas (1992) examine convict data, and find that illiteracy by the end of the wars was lower than it was at its outbreak.[132] To the extent that the ability to sign one's name is a more basic skill than remembering one's age, it could be argued that our findings suggest that only the performance of more complex tasks suffered. Alternatively, it could be argued that schooling continued unabated, but that the ability to acquire more advanced skills was limited in many cases. There is no evidence to suggest that a sudden collapse in school attendance could be responsible for our results.

VI Conclusions

This paper has argued that cognitive shortcomings in the past were partly driven by inadequate nutrition. To demonstrate the importance of this new explanation for low living standards in the past, we exploit a quasi-natural experiment. When industrializing Britain went to war with France in the 1790s, grain supplies from the continent were cut off in many years. In other years, the costs of imports was unusually high. Prices for wheat and other staples surged. Market integration within Britain declined as privateers preyed on coastal shipping. We examine the impact of these exogenous shocks to food availability, and show that it lowered average numeracy throughout the country. Subjects born in the hungry decades of the 1790s and 1800s were much less likely to remember their age correctly, or to perform the calculation necessary to

[132] Subsequently, they document an increase. Their data may suffer from greater problems of representativeness and small sample bias than Schofield's.

derive it without errors. The detrimental effect of high food prices was particularly pronounced in those areas that did little to help the poor. In areas where relief was generous, higher grain prices caused only mild reductions in numeracy. We show that in addition to causing lower numeracy, the careers of those of those affected by high grain prices and low support payments suffered. In particular, and in line with our broader argument, individuals from counties hit by particularly high prices, or without much income support, had on average lower cognitive ability. As a consequence, they selected into occupations that were on average less demanding in terms of cognitive skills. They also earned less than their peers. This suggests that the 'first welfare state' offered an effective way to improve living conditions for the poorer groups of society.

In his Nobel address, Robert Fogel (1994) sought to determine the contribution of better nutrition to higher productivity over the last 200 years. Focusing on the increase in life expectancy, and the greater resilience and strength of humans today, Fogel concluded that 20-30 percent of total output growth could be attributed to improved food intake. In his work, one factor does not feature prominently – greater cognitive ability. One of the potential implications of research is that cognitive ability may have been a key factor limiting output in the past. While we offer no final proof, the findings presented in this paper suggest that the transition to self-sustaining growth in industrializing Europe could owe a great deal to improved nutrition and higher cognitive ability. There are also possible implications for the more recent past. Flynn (1984) has shown that cognitive scores underlying IQ tests have been rising for several decades in the 20th century. Between 1930 and 1900, average cognitive ability scores rose by the equivalent of 0.6 IQ points per year (Hiscock 2007). The benefits of higher cognitive scores in the labor market today are well-known (Case and Paxson 2006). If cognitive ability was in part affected by poor food intake, as our results suggest, we will have to take seriously the hypothesis that life was 'nasty, brutish, and short', in the words of Hobbes, because earlier populations' cognitive development was nutritionally constrained. Our results offer an alternative to Greg Clark's (2007) recent suggestion that it was cultural norms in the European past – and around the globe today – that severely constrained output. The cause of lower productivity may well have been 'within' economic agents, in the sense of Clark, but it may not be culture and socialization that harmed output, but low cognition as a result of poor nutrition.

References

[1] Abbott, R., White, L., Ross, G., Petrovitch, H., Masaki, K., Snowdon, D. and J. Curb (1998). Height as a Marker of Childhood Development and Late-life Cognitive Function: The Honolulu-Asia Aging Study. Pediatrics 102(3): 602-609.

[2] A'Hearn, B, Baten, J. and D. Crayen (2006). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital. UPF working paper.

[3] Alderman, H., J. Hoddinott, and B. Kinsey (2006). Long Term Consequences of Early Childhood Malnutrition. Oxford Economic Papers 58: 450-474.

[4] Anderson, M. (1972). The study of family structure. in: E. A. Wrigley, ed., Nineteenth-century Society: Essays in the Use of Quantitative Methods for the Study of Social History, Cambridge: CUP, 47-81.

[5] Anderson, M. (1987), The 1851 Census: A National Sample of the Enumerators Returns, Chadwyck-Healey.

[6] Armstrong, W. (1972). The use of information about occupations, in: E. A. Wrigley, ed., Nineteenth-century Society: Essays in the Use of Quantitative Methods for the Study of Social History, Cambridge: CUP, 191-310.

[7] Ashton, T. (1959). Economic Fluctuations in England 1700-1800. London: Oxford University Press.

[8] Bachi, R. (1951). The Tendency to Round Off Age Returns: Measurement and Correction. Bulletin of the International Statistical Institute 33: 195-221.

[9] Behrman, J., H. Alderman and J. Hoddinott (2004) Copenhagen Consensus – Challenges and Opportunities: Hunger and Malnutrition.

[10] Behrman J. (2006). Consequences of Height in Developing Countries. Presentation at the Conference on Height, Health, and Living Standards, Princeton, June.

[11] Boyer, G. (1990). An Economic History of the English Poor Law, 1750-1850. Cambridge: CUP.

[12] Boyer, G. (1986). The Old Poor Law and the Agricultural Labor Market in Southern England: An Empirical Analysis. Journal of Economic History 46 (1): 113-135.

[13] Brewer, J. (1990). The Sinews of Power: War, Money and the English State, 1688-1783. Cambridge, MA: Harvard University Press.

[14] Brown, J. and E. Pollitt (1996). Malnutrition, Poverty and Intellectual Development.

Scientific American 274(2): 38-43.

[15] Brunt, L. and E. Cannon. (2005). Asset Prices in the Industrial Revolution: Evidence from the Grain Market. University of Bristol, Department of Economics working paper.

[16] Case, A, and C. Paxson (2006). Stature and status: Height, ability, and labor market outcomes, Princeton University working paper.

[17] Census (1891), General Report, British Parliamentary Papers 1893-4, vol. CVI.

[18] Census (1958), General Report on the 1951 Census.

[19] Center on Hunger and Poverty/CHP (1998). Statement on the Link Between Nutrition and Cognitive Development in Children, Brandeis University Press.

[20] Chalmers, George (1812), An Historical View of the Domestic Economy of Great Britain, Edinburg: Constable.

[21] Clark, G. (2003). The Great Escape: The Industrial Revolution in Theory and in History. UC Davis mimeo.

[22] Clark, G. (2007). Farewell to Alms. Princeton: PUP.

[23] Cravioto, J. and E. DeLecardi (1975), Neurointegrative Development and Intelligence in Children Rehabilitated from Severe Malnutrition. in: J. Prescott, M. Read and D. Coursin, Brain Function and Malnutrition, New York: Wiley.

[24] Duncan-Jones, R. (1990). Structure and Scale in the Roman Economy. Cambridge: CUP.

[25] England, P. and B. Kilbourne (1988). Occupational Measures from the Dictionary of Occupatinal Titles for 1980 Census- Detailed Occupations. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

[26] Ericson, A. and B. Kallen. (1998). Very Low Birthweight Boys at Age 19. Archives of Disease in Childhood – Fetal and Neonatal Edition 78 (May): F171-4.

[27] Feinstein, C. (1998). Pessimism Perpetuated: Real Wages and the Standard of Living in Britain during and after the Industrial Revolution. Journal of Economic History 58: 625–658.

[28] Floud, R., K. Wachter, and A. Gregory (1990) Height, Health, and History: Nutritional Status in the United Kingdom, 1750–1980. Cambridge: CUP.

[29] Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. Psychological Bulletin, 95: 29-51.

[30] Fogel, R.W. (1994). Economic Growth, Population Theory, and Physiology: The

Bearing of Long-Term Processes on the Making of Economic Policy, American Economic Review, 84, (3), 369-395.

[31] Galor, O. and D.N. Weil (1999). Population, Technology, and Growth. American Economic Review, 90, 806-828.

[32] Grantham-McGregor, S. (2002). Linear growth retardation and cognition. Lancet 359: 542.

[33] Hall, R. and C. Jones (1998). Why Do Some Countries Produce So Much More Output Per Worker Than Others? Quarterly Journal of Economics 114(1): 83-116.

[34] Heckman, J. (1995), Review: Lessons from the Bell Curve. Journal of Political Economy 103: 1091-1120.

[35] Herlihy, D., and D. Klapisch-Zuber(1978). Les Toscans et leurs familles: une étude du catasto florentin di 1427. Editions de l'Ecole des Hautes Etudes en Sciences Sociales; Presses de la Fondation Nationale des Sciences Politiques, Paris. Chap. XIII.

[36] Higgs, E. (1989). Making Sense of the Census-The Manuscript Returns for England and Wales, 1801-1901. Public Record Office Handbook No23. London: Her Majesty's Stationary Office.

[37] Hill, Ch. (1992). British Strategy in the Napoleonic War 1803-1815. Manchester: Manchester University Press.

[38] Hiscock, M. (2007). The Flynn Effect and its Relevance to Neuropsychology, Journal of Clinical and Experimental Neuropsychology, 29: 514-529.

[39] Horrell, S., J. Horrell and H.-J. Voth (2001). Destined for Deprivation? Intergenerational Poverty Traps in Eighteenth-Century Britain. Explorations in Economic History 38.

[40] Horrell, S. (1996). Home Demand and British Industrialization. Journal of Economic History, 56: 561–604.

[41] Jacks, D. (2007). Foreign Wars, Domestic Markets: England, 1793-1815, unpublished manuscript, Simon Fraser University.

[42] Jones, G., and W. J. Schneider (2007). IQ and Productivity, mimeo.

[43] Jones, G. and W. J. Schneider (2007). Intelligence, Human Capital, and Economic Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. Journal of Economic Growth, 11: 71-93.

[44] Komlos, J. (1989), Nutrition and Economic Development in the Eighteenth-Century Habsburg Monarchy: An Anthropometric History. Princeton, NJ: Princeton University

Press.

[45] Komlos, J., ed., Stature, Living Standards, and Economic Development: Essays in Anthropometric History. Chicago: Chicago University Press.

[46] Komlos, J. (2005). On English Pygmies and Giants: the Physical Stature of English Youth in the late-18th and early-19th Centuries, Munich Working Paper 06.

[47] Lloyd-Still, J. (1976). Clinical Studies on the Effects of Malnutrition During Infancy and Subsequent Physical and Intellectual Development, in: J. Lloyd-Still, Malnutrition and Intellectual Development, Lancaster: MTP.

[48] Long, J. (2006). The socioeconomic return to primary schooling in Victorian England. Journal of Economic History 66: 1026-1053.

[49] Lucas, A. (1998). Programming by Early Nutrition: An Experimental Approach. The Journal of Nutrition. 128 (2): 401S-406S.

[50] Lynn, R. (1989). A nutrition theory of the secular increases in intelligence, positive correlation between height, head size and IQ. British Journal of Educational Psychology 59: 372-77.

[51] Lynn, R. and T. Vanhanen (2002). IQ and the Wealth of Nations. New York: Praeger.

[52] Magnusson, P., F. Rasmussen, and U. Gyllensten. (2006). Height at Age 18 Years is a Strong Predictor of Attained Education Later in Life: Cohort Study of Over 950,000 Swedish Men. International Journal of Epidemiology 35 (January): 658-63.

[53] Mitch, D. (1998). The Role of Education and Skill in the British Industrial Revolution. in: Joel Mokyr, ed., The British Industrial Revolution: An Economic Perspective. Boulder: Westview Press, pp. 241-279.

[54] Mitch, D. (1992). The Rise of Popular Literacy in Victorian England: The Influence of Private Choice and Public Policy. Philadelphia: University of Pennsylvania Press.

[55] Mokyr, J. (1993). Editor's Introduction, in: Joel Mokyr, ed., The British Industrial Revolution: An Economic Perspective, Boulder: Westview Press, 1-131.

[56] Mokyr, J. (1983). Why Ireland Starved. London: George Allen & Unwin.

[57] Mokyr, J. and N.E. Savin (1976). Stagflation in Historical Perspective. Research in Economic History, 1: 198-259.

[58] Murnane, R.J., Willet, J.B., and F. Levy (1995). The growing importance of cognitive skills in wage determination. Review of Economics and Statistics 77(2): 251-266.

[59] Myers, R. (1976). An Instance of Reverse Heaping of Ages. Demography 13 (4): 577-

580.

[60] Nicholas, S. and J. Nicholas (1992). Male Literacy, 'Deskilling', and the Industrial Revolution, Journal of Interdisciplinary History, 23 (1): 1-18.

[61] Olson, M. (1963). The Economics of Wartime Shortage. Durham: Duke University Press.

[62] Paxson, C. and N. Schady (2005). Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health and Parenting. Princeton working paper.

[63] Persico, N., A. Postlewaite, and D. Silverman (2004). The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height. Journal of Political Economy 112: 1019-1053.

[64] Rajan, R. and L. Zingales (1998). Financial Dependence and Growth. American Economic Review 88: 559-586.

[65] Richards, M. and M. Wadsworth (2004). Long-term effects of early adversity on cognitive function. Arch DisChildhood, 89, 922-27.

[66] Richards, M., Hardy, R., Kuh, D. and M. Wadsworth (2002). Birthweight, postnatal growth and cognitive function in a national UK birth cohort. International Journal of Epidemiology 31 (2): 342-48.

[67] Rivera-Batiz, F. (1992). Quantitative literacy and the likelihood of employment among young adults in the Unites States. Journal of Human Resources 27: 313-328.

[68] Robin, J. (1995). From childhood to middle age: cohort analysis in Colyton, 1851-1891. Cambridge Group for the History of Population and Social Structure, 64, [23] p., Cambridge Group for the History of Population and Social Structure Working Paper Series No. 1.

[69] Schofield, R. (1973). Dimensions of illiteracy, 1750–1850. Explorations in Economic History 10: 437–454.

[70] Schuster, K. (2005). Wirtschaftliche Entwicklung, Sozialstruktur und biologischer Lebensstandard in München und dem südlichen Bayern im 19. Jahrhundert. St Katharinen: Scripta Mercaturae Verlag.

[71] Smith, A. (1976). An Inquiry into the Nature and Causes of the Wealth of Nations. Oxford: Clarendon Press.

[72] Smith, D. (1998). The Greenhill Napoleonic Wars Data Book. London: Greenhill Press.

[73] Steckel, R. (1995). Stature and the Standard of Living, Journal of Economic Literature

33 (4): 1904-40.

[74] Stein, Z. and M. Susser (1976). Prenatal Nutrition and Mental Competence, in: J. Lloyd-Still, Malnutrition and Intellectual Development, Lancaster: MTP.

[75] Sunde, U. and M. Cervellati (2005). Human Capital Formation, Life Expectancy and the Process of Economic Development. American Economic Review 95: 1653-1672.

[76] Sunder, J., Tambs, K., Harris, J., Magnus, P., and T. Torjussen (2005). Resolving the Genetic and Environmental Sources of the Correlation between Height and Intelligence: A Study of Nearly 2600 Norwegian Male Twin Pairs. Twin Research and Human Genetics 8: 307-11.

[77] Thomson, D. (1980). Age Reporting by the Elderly and the Nineteenth Century Census, Local Population Studies 25:13-25.

[78] Tillot, P.M. (1972). Sources of inaccuracy in the 1851 and 1861 censuses, in: E. A. Wrigley, ed., Nineteenth-century Society: Essays in the Use of Quantitative Methods for the Study of Social History, Cambridge: CUP, 82-133.

[79] Voth, H.-J. (2003). Living Standards and Urban Disamenities, in: R. Floud, P. Johnson (eds.), Cambridge Economic History of Britain, Cambridge: Cambridge University Press.

[80] Williamson, J.G. (1980). Earnings Inequality in Nineteenth-Century Britain. Journal of Economic History 40:457-75.

[81] Williamson, J.G. (1982). The Structure of Pay in Britain, 1710-1911. Research in Economic History 7: 1-54.

[82] Winick, M. and P. Rosso (1975). Malnutrition and General Nervous System Development, in: J. Prescott, M. Read and D. Coursin, Brain Function and Malnutrition, New York: Wiley.

[83] Wrigley, E.A., Davies, R., Oeppen, J. and R. Schofield (1997). English Population History from Family Reconstitutions, 1580-1837, Cambridge: Cambridge University Press.

[84] Zax, J. and D. Rees (2002). IQ, Academic Performance, Environment, and Earnings. Review of Economics and Statistics, 84(4): 600-616.

Tables

Table 1: Whipple scores by birth decade, census date, and gender

	birth decade	census 1851	census 1881	difference
male	1810	120.54	112.93	-7.61
	1820	112.21	116.60	4.39
female	1810	122.53	114.92	-7.61
	1820	117.14	114.86	-2.28

Table 2: Descriptive Statistics

Variable	N	Mean	Std. Dev	Min	Max
relief	290	16.30	4.77	9.61	26.04
female	305	0.50	0.50	0.00	1.00
regional wheat price	250	99.67	22.24	60.79	147.97
national wheat price	263	99.57	21.67	67.29	134.40
grain growing area	290	12.27	4.11	4.48	19.09
Whipple Index	305	118.34	12.30	94.43	176.47
birth decade	305	1817.90	21.99	1780	1850

Table 3: Stature and Whipple Ratios

Country/Region	Birth Decade	Average Height		Ratio Height	Whipple Index		Ratio Whipple
		Short	Tall		Short	Tall	
England	1800-1840	62.66	67.11	0.93	133	129	1.03
Ireland	1790-1810	63.65	67.70	0.94	160	131	1.22
US	1800-1830	65.75	69.81	0.94	124	114	1.09
France -Paris	1660-1760	61.80	63.98	0.97	141	102	1.38
France -northeast	1660-1760	61.64	64.17	0.96	125	117	1.07
France -southwest	1660-1760	61.43	63.98	0.96	142	125	1.14
France-total	1660-1760	61.52	64.08	0.96	135	123	1.10

Table 4: Regression Analysis: Whipple Scores and Grain Prices (Whipple Index as dependent variable)

Eq.	(1)	(2)	(3)	(4)	(5)
			Relief payments		
Sample Coefficient	all	all	< median	≥ median	all
National wheat price (t+1)	0.27***	0.27***	0.30***	0.23***	0.27***
	(0.04)	(0.04)	(0.05)	(0.06)	(0.04)
Female	-1.09	-1.14	-2.04*	-0.12	-1.14
	(1.17)	(1.18)	(1.09)	(2.20)	(1.18)
Reliefhigh		-4.06**			
		(1.43)			
Relieflack (R_{max}-R_i)					0.35**
					(0.14)
Constant	93.24***	95.63***	93.49***	94.68***	90.05***
	(3.22)	(3.40)	(4.12)	(5.29)	(3.45)
Observations	210	210	109	101	210
R-squared	0.19	0.23	0.25	0.13	0.22
Impact of 1 Std. Dev.					
National wheat price (t+1)	5.23	5.23	5.81	4.46	5.23
Relieflack					1.70

Note: Robust standard errors (clustered at county level) in parentheses, *** p<0.01, ** p<0.05, * p<0.1
a)defined as [R_{max}-R_i], where R_{max} is the maximum relief payment per capita, and R_i is the relief payment in county i.

Table 5: Fixed Effects Panel Estimation (Whipple Index as dependent variable)

Eq. Coefficient	(1) OLS	(2) IV	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) IV
National wheat price (t+1)	0.27***	0.14***		0.27***	0.29***		
	(0.04)	(0.04)		(0.04)	(0.03)		
female	-1.17	-1.21	0.01	-1.17	-1.38	-1.17	-1.16
	(1.24)	(1.28)	(1.00)	(1.24)	(1.28)	(1.24)	(1.11)
Pop density					-5.51***		
					(0.07)		
Grain area					-0.76***		
					(0.00)		
Cottage industry					-19.12***		
					(0.08)		
reliefhigh			-4.35***	-9.28***			
			(0.03)	(0.05)			
Wprice(t+1) *[R_{max}-R_i]						0.02***	0.03***
						(0.00)	(0.03)
Instrument		Spring ratio (t+1)					Spring ratio (t+1)
Constant	81.91***	93.35***	112.6***	91.18***	112.40***	85.76***	81.31***
	(3.57)	(6.28)	(0.51)	(3.54)	(3.19)	(3.60)	(5.12)
Observations	210	210	305	210	198	210	210
R-squared	0.35	0.09	0.11	0.35	0.38	0.34	0.33

Note: Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Intelligence Equations (Kilbourne/England intelligence scores as dependent variable)

Eq.	(1)	(2)	(3)	(4)
Estimation	OLS	OLS	IV	IV
Coefficient				
whipple	0.002*	0.002[1)]	0.003***	0.003***
	(0.001)	(0.001)	(0.000)	(0.000)
female	0.319***	0.307***	0.270***	0.270***
	(0.016)	(0.019)	(0.022)	(0.022)
Cottage industry		0.020**		
		(0.009)		
Pop density		0.013***		
		(0.003)		
Grain area		0.005***		
		(0.001)		
Instrument			National wheat price	National wheat price Reliefhigh
Fixed Effects				
County	Yes	Yes	Yes	Yes
Time	Yes	Yes	Yes	Yes
Constant	2.874***	2.755***	2.675***	2.675***
	(0.089)	(0.110)	(0.028)	(0.028)
Observations	304	273	209	209
R-squared	0.86	0.83	0.79	0.79

Note: Robust standard errors in parentheses , *** p<0.01, ** p<0.05, * p<0.1. One standard deviation of the Whipple Index increases the Kilbourne/England intelligence score by about 0.02 units or 15% of 1 Std.
[1)] p-value: 0.109.

Table 7: Earnings Equations (average earnings in Pounds Sterling per year as dependent variable)

Eq.	(1)	(2)	(3)	(4)
Estimation	OLS	OLS	IV	IV
Coefficient				
whipple	-0.18***	-0.16**	-0.23***	-0.22***
	(0.06)	(0.07)	(0.01)	(0.01)
female	-18.80***	-14.52***	-12.68***	-16.26***
	(3.71)	(1.91)	(1.55)	(3.66)
Cottage industry		0.42		
		(0.81)		
Pop density		-1.87***		
		(0.20)		
Grain area		-0.48***		
		(0.05)		
Instrument			National wheat price (t+1)	Spring ratio (t+1)
Fixed Effects				
County	Yes	Yes	Yes	Yes
Time	No	No	Yes	Yes
Constant	83.84***	86.69***	79.85***	80.50***
	(7.06)	(9.79)	(1.82)	(1.86)
Observations	304	273	209	220
R-squared	0.74	0.68	0.59	0.64

Note: Robust standard errors in parentheses , *** p<0.01, ** p<0.05, * p<0.1. One standard deviation of the Whipple Index reduces average earnings by about 2.26 Pounds Sterling or 25% of 1 Std.

Figures

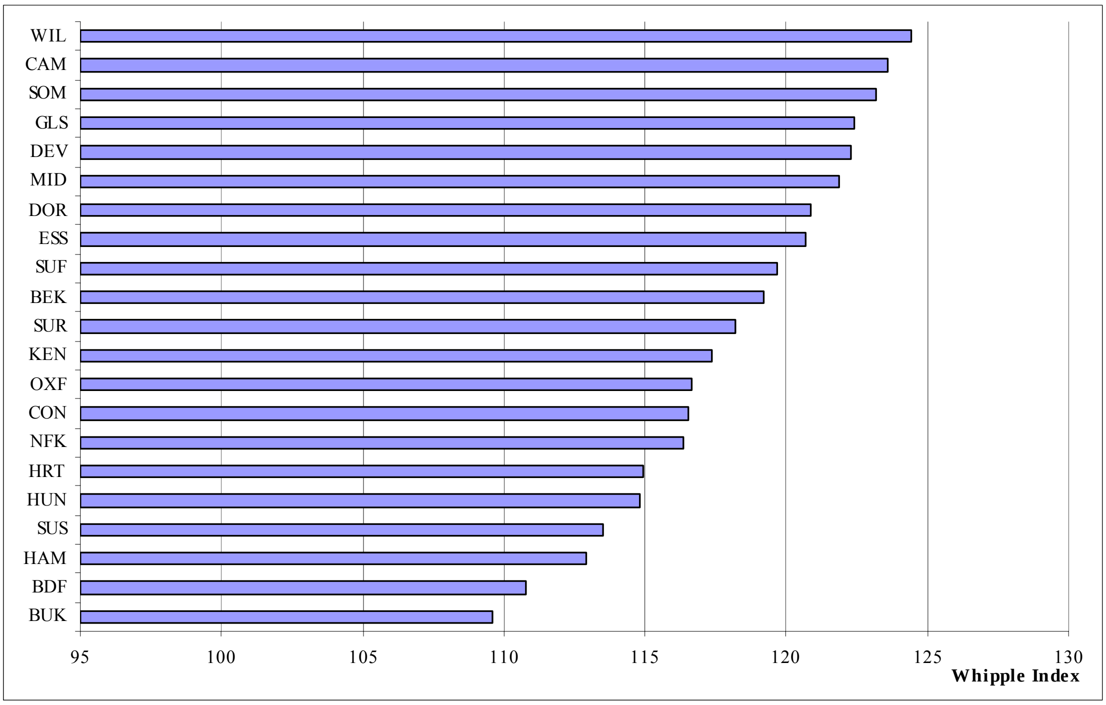


Figure 1: Whipple Index by County

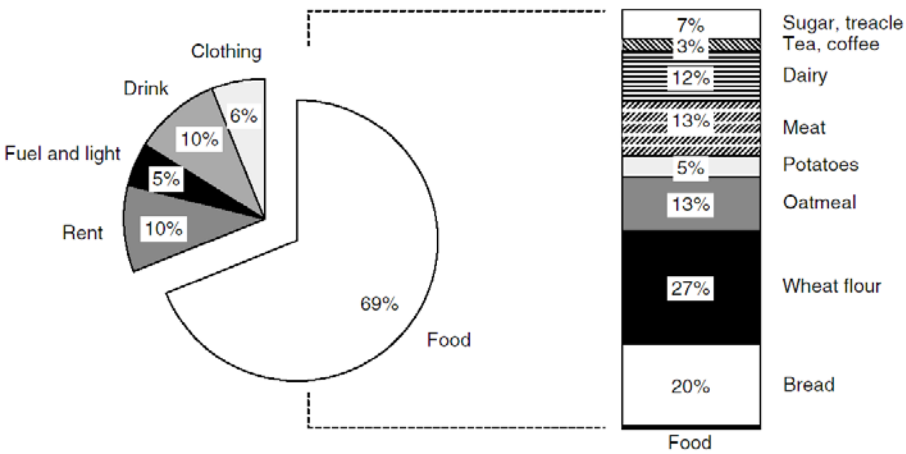


Figure 2: Composition of Working Class Expenditure, 1788-92

Source: Voth (2003)

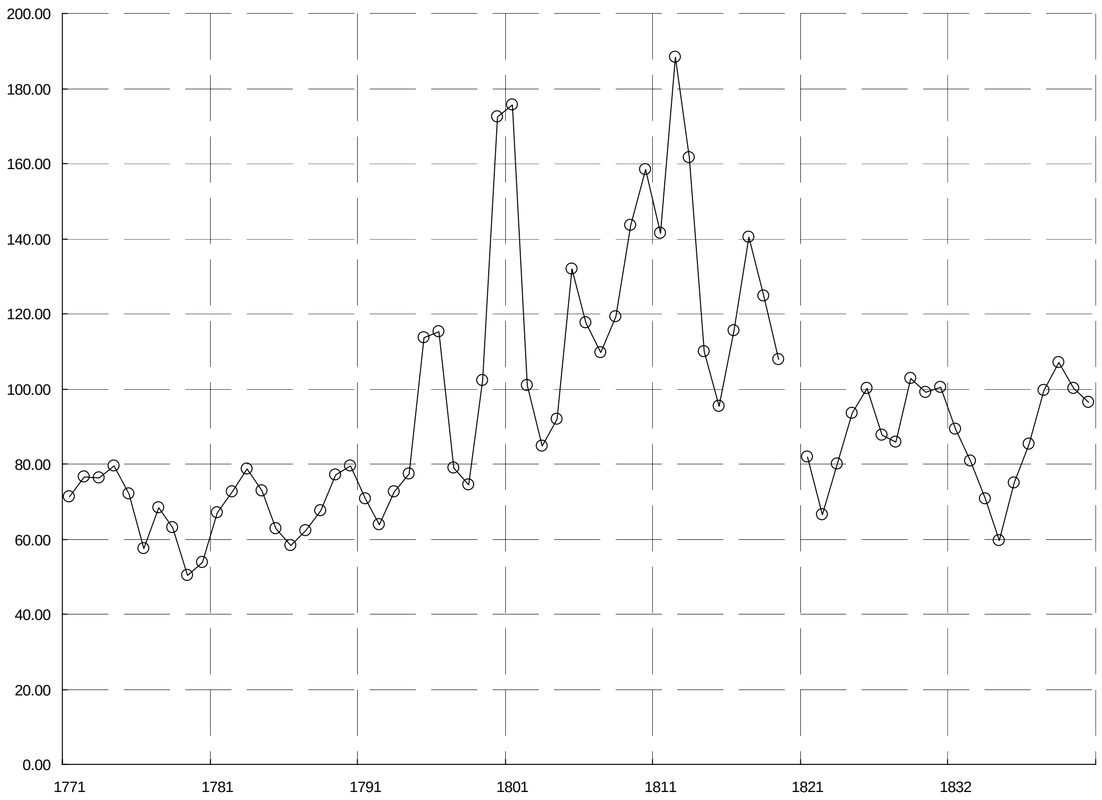


Figure 3: Grain prices in England (1800=100)

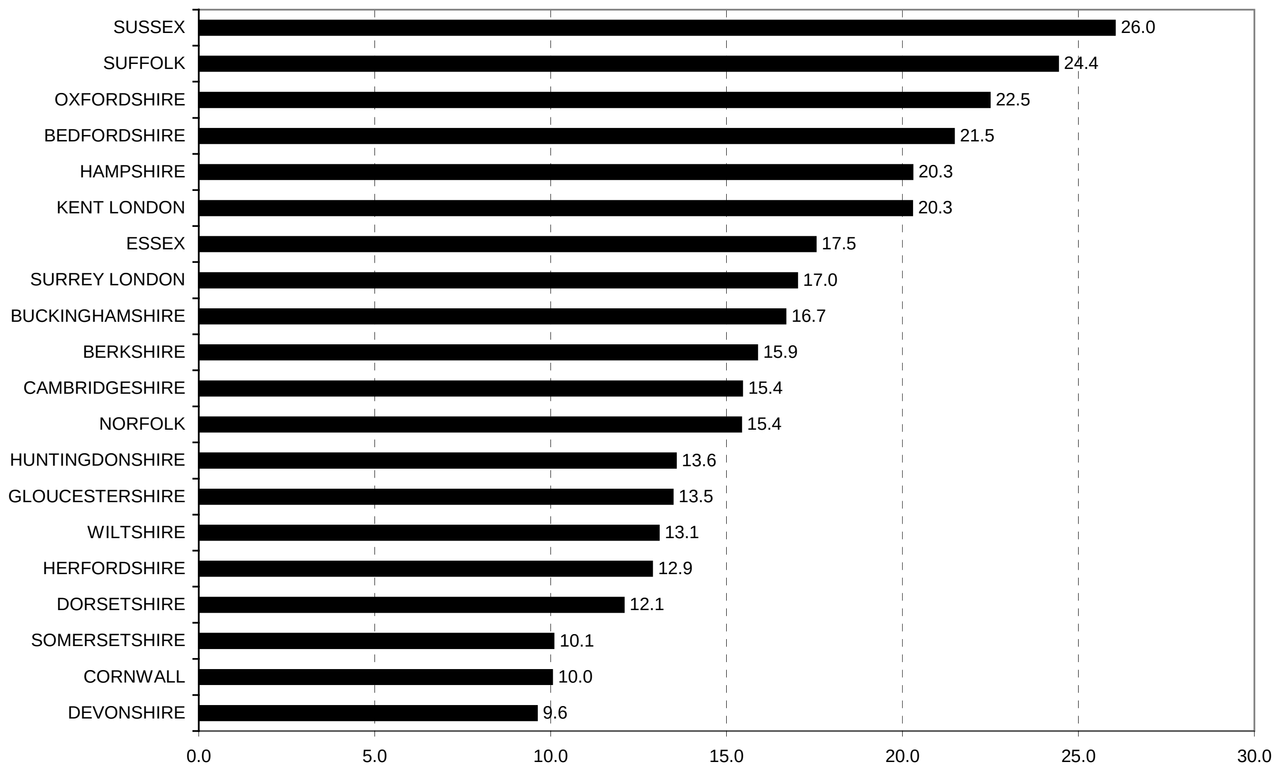


Figure 4: Poor relief per capita, in shilling

Source: Boyer 1990.

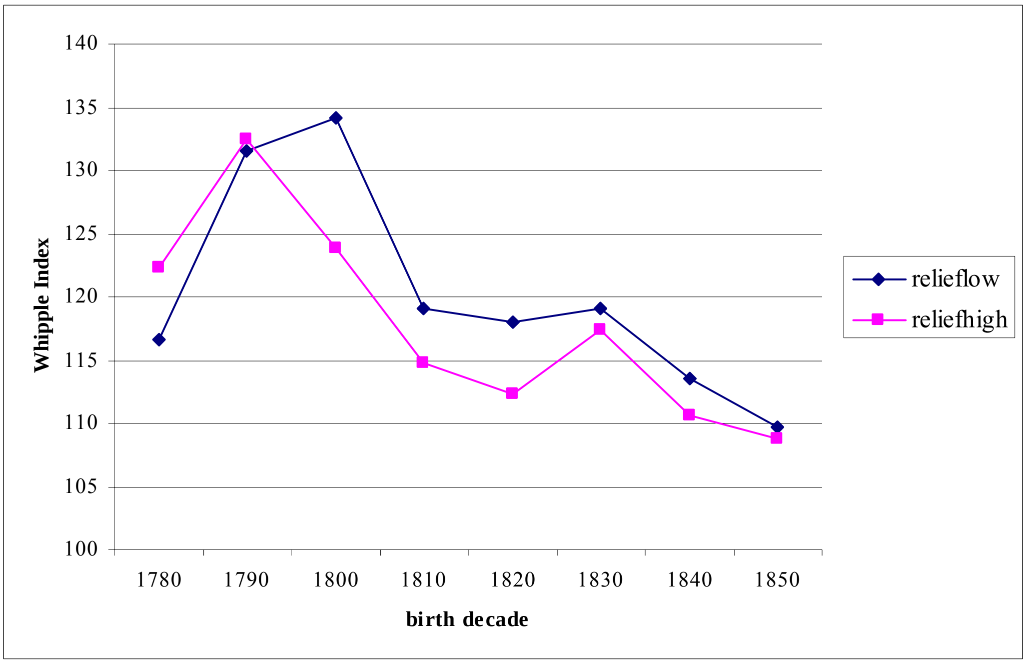


Figure 5: Whipple Indices over Time, by Generosity of Poor Relief

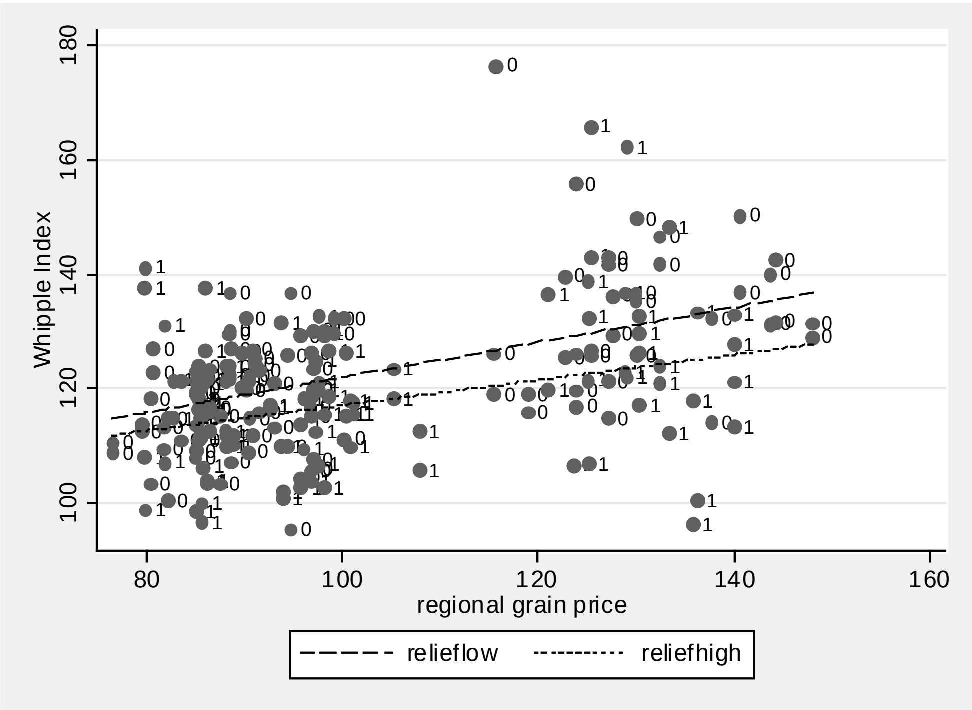


Figure 6: Whipple scores and grain prices

Appendix

One crucial assumption underlying our work is that heaping of reported ages is a good proxy for cognitive ability. To test this assumption, we used the 1993 Assets and Health Dynamics Among the Oldest Old (AHEAD) dataset from the Health and Retirement Survey. This contains information on the individual, health, cognitive, and income characteristics of the American elderly. In total, 8,222 returns from individuals are available.

There are numerous measures of cognitive ability in the AHEAD survey. One indicator involves asking individuals to count backwards from 100 in steps of 7. This is used as an overall indicator of numerical ability. There are also questions involving word recall (number of words remembered correctly, of a list of 10) and the current day, month, and year. We use all of these as measures of cognitive ability.

Age in the AHEAD survey is not self-reported. However, individuals are questioned about the age of death of their parents. This is not exactly the same as the information we would ideally like to use (self-reported age), but it is similar to measures of age-heaping used in the literature (such as the age given on Roman tombstones, etc.). We find that in a sample of 7,317 respondents reporting the age of their father, 2,073 provide an answer that is a multiple of five. In a truly random sample, we should expect no more than 1,463 responses naming a multiple of five. This corresponds to a Whipple index of 142.

If age heaping are indicators of cognitive (and numerical) ability, we should expect that the HRS measures of are correlated with reported ages of death of the father at multiples of five. As Table A1 shows, those *not* reporting the age of their father as a multiple of five were more likely to remember a large number of words correctly, to know the current year, and to perform a series of subtraction exercises without error.

Table A1: Heaping in reported ages and cognitive ability

Dependent variable	Number of words remembered correctly	Current year correctly remembered	Number of subtractions correct
Panel A:			
Age of father's death NOT reported as multiple of 5	0.025 (1.96)	0.10 (1.79)	0.044 (2.5)
Estimator	Poisson	Probit	Poisson
(Pseudo)-R^2	0.0001	0.0011	0.0003
N	7163	7363	7382
Panel B:			
Age of neither parent's death reported as multiple of 5	0.055 (2.71)	0.34 (4.3)	0.11 (3.4)
Estimator	Poisson	Probit	Poisson
(Pseudo)-R^2	0.03	0.006	0.0005
N			

The vast majority of those reporting the age of their father's death as a multiple of five will remember the right age. To demonstrate the power of age-heaping, we also constructed a second variable that takes the value of 0 if only one or none of the ages of the parent's death are reported as a multiple of five, and 1 otherwise. In our sample, if age of death was random, only 329 individuals should be reporting both parents having died at an age that is a multiple of five. In actual fact, 655 reported that this was the case, suggesting that the rate of mistakes is particularly high. Reporting age of death twice as a multiple of five is the explanatory variable in Panel B of Table A1. As expected, the size and significance of the coefficient rises markedly compared to Panel A, where we only used the age of the father's death. These results establish a strong prior that age-heaping in historical data is related to cognitive ability.

5 GLOBAL TRENDS IN NUMERACY 1820-1940 AND ITS IMPLICATIONS FOR LONG-RUN GROWTH

Abstract

This study is the first to explore long-run trends of numeracy for the 1820-1949 period in 165 countries, and its contribution to growth. Estimates of the long-run numeracy development of most countries in Asia, the Middle East, Africa, America, and Europe are presented, using age heaping techniques. Assessing the determinants of numeracy, we find school enrolment as well as Chinese instruments of number learning to have been particularly important. We also study the contribution of numeracy as measured by the age heaping strategy for long-run economic growth. In a variety of specifications, numeracy mattered quite strongly for growth patterns around the globe.

Keywords: Human Capital, Age Heaping, Growth, Industrial Revolution, Numeracy
JEL codes: I21, N01, N30, 015

This chapter is based on a working paper written jointly with Joerg Baten (University of Tuebingen). It is currently under review for *Explorations in Economic History.* The concept for the paper was developed jointly, the analyses and writing were equally shared.

I Introduction

Human capital is at the heart of modern economic growth studies (Lucas 1988, Romer 1989, Mankiw, Romer and Weil 1992 and Jones 1996 among others). Growth economics has strongly emphasized the role of human capital formation and its persistence in nations over time. Also Unified Growth Theory has underlined the role of human capital in the transition from the Malthusian stagnation to the contemporary era of modern economic growth (Galor and Weil 2000, Galor and Moav 2002, Galor 2005). Becker, Murphy and Tamura (1990) placed investments in human capital at the centre of their study, assuming that the return on growth-enhancing investments in human capital rises rather than declines as the stock of human capital increases. Their framework is characterized by multiple steady states that differ in regard to schooling decisions and opportunity costs of child care. Initial human capital stock and other major exogenous shocks play an important role in the determination of fertility, growth rates and the wealth of countries. Their predictions being inconsistent with broad historical evidence, more recent growth theories have accepted the empirical argument of Mitch (1991) and Mokyr (1983) that some countries' growth paths do not fit into this demographic-educational pattern: for example, Britain might have experienced stagnating literacy between the mid-18th and mid-19th century, while France experienced an early fertility decline already around 1800 without becoming a driving force of growth in 19th-century Europe. While education and a slowing down of population growth cannot explain the first Industrial Revolution satisfactorily, many long-run growth economists believe that these factors played a key role in the later transition to a regime of sustained economic growth (Boucekkine, de la Croix and Licandro 2003, Glaeser et al. 2004, Cervellati and Sunde 2003 among others). Economic factors eventually altered the parental quality-quantity decision and stimulated human capital investments, reinforcing technological progress and economic growth.

Given that human capital accumulation is a crucial factor in long-run economic growth theory, efforts have been made to strengthen the available empirical evidence. O'Rourke and Williamson (1997), for example, were able to include schooling in European convergence regressions for 16 countries for the 1870-1913 period, concluding that globalization forces were in fact a much more important influence on comparative development.[133] When going further back in time to the early nineteenth century and beyond, schooling data dry up even for Europe, and literacy must generally be inferred from a proxy - the ability to sign one's name on marriage registers and legal documents (Reis 2005). Leaving Europe, it becomes increasingly difficult to find systematic, comparable data. Crafts (1997) reports enrolment and literacy rates for 17 advanced economies since 1870 while Lindert (2004) provides school enrolment rates and teacher-student ratios on some fifty countries, substantially improving the Mitchell (2003a-c) data set. Benavot and Riddle (1988) have compiled additional schooling data for LDCs for the 1870-1940 time period. Morrisson and Murtin (2007) have revised and extended Mitchell's (2003a-c) data set, using national census data to obtain an educational attainment data set for 1870 to 2000. Since census data is scarce for the developing world prior to the end of the 19th century, the authors had to assume enrolment rates of 1 or 0.1% for the LDC world in 1820.

Although these studies represent a clear improvement of our knowledge, about half of the existing 165 countries with populations above 500,000 are not yet documented for the late 19th century. In consequence, we think that existing samples of human capital data are biased towards today's richer spectrum of countries, as those were the first to introduce schooling statistics. Studies on human capital development in the poorer half of the world would provide important insights into overall human development.

This study is a first attempt to achieve almost global coverage since the late 19th century. Moreover, quite a number of additional countries can be included for the 1820-1890

---

[133] Tortella (1994), using literacy data, offers a different interpretation, at least for southwestern Europe.

period. Another aim of this article is to broaden the human capital literature by constructing a numeracy index. Why can numeracy be seen as a historical measure of human capital? We would argue that number knowledge and number discipline are even more crucial for economic growth than the ability to sign one's name in a marriage register. Numeracy is highly complementary to technological abilities, and it is a precondition for the modern commercial economy. For Weber, Sombart, and Schumpeter, numeracy was at the very heart of modern rational capitalism. They traced the latter's roots back to the invention of double-entry bookkeeping in late-medieval Italy. Carruthers and Espeland (1991) have described in some detail the process of abstraction and organization inherent in compiling a ledger, which made possible the development of concepts like capital, depreciation, and the rate of profit. It is no accident that the introduction of Arabic numerals to Europe and the earliest accounts of mathematical education stem from the same time and place. Hence, in this paper, we go beyond traditional literacy and enrolment indicators by presenting proxies for numeracy based on age heaping.

What is age heaping? Demographers normally use age data to describe a population's age structure and to forecast population growth. In contrast, the idea behind this study is to use irregularities in the reporting of ages to estimate the level of education in an economy. Such irregularities appear in the form of heaped data, i.e. the age distribution does not run smoothly but exhibits sharp jumps and clustering at certain ages. This phenomenon is attributed to age heaping, a term which describes the ignorance of one's own age or the tendency to round ages. Age heaping is a well-known phenomenon among demographers and applied statisticians. However, while they perceive age heaping mainly as a data problem because it leads to biased vital rates on the one hand, and the degree of heaping as a measure of data quality only on the other hand, we use it as a proxy for non-numeracy.

Already a half-century ago, influential study by Bachi (1951) and Myers (1954) investigated age heaping and its correlation with education levels within and across countries.

Thereby, Bachi (1951) was able to analyze the degree of age heaping among Jewish immigrants to Israel in 1950 and among Muslims in Mandated Palestine in 1946, finding, amongst other things, that the increasing spread of education resulted in a better knowledge of age. Myers (1976) found a correlation at the individual level between age awareness and income. Another innovative example is the study by Herlihy and Klapisch-Zuber (1978) who used successive Florentine tax enumerations and found distinct heaping on multiples of five for adults, which declined substantially in the period from 1371 to 1470. Furthermore, they showed that age heaping was more prevalent among both women in rural areas and small towns, and among the poor. Duncan-Jones (1990) employed this technique to study age data from Roman tombstones. Mokyr (1983) was the pioneer who established the age heaping measure as an explicit numeracy indicator in economic history. He employed the degree of age heaping to assess the labor-quality effect of emigration on the Irish home economy during the first half of the nineteenth century, as emigrants from pre-famine Ireland were less sophisticated than those who stayed behind. Thomas (1987) considered the slight but discernible improvement in the accuracy of age reporting as evidence that numerical skills in England had improved between the 16^{th} and 18^{th} century. Budd and Guinnane (1991) studied Irish age-misreporting in linked samples from the 1901 and 1911 censuses. They found considerable heaping on multiples of five in the 1901 census, which was also more frequent among the illiterate, poor, and aged. More recent research was conducted by Long (2005, 2006) who analysed age data from the 1851 and 1881 British population censuses, identifying urban migrants in Victorian Britain as being educated beyond average. By exploiting repeated observations, he was able to show that individual age discrepancies (which are also a measure of missing age awareness) had a significant negative impact on socio-economic status and wages. De Moor and van Zanden (2006) studied the relative numeracy of women during the Middle Ages, and Clark (2007) has recently reviewed the evidence.

From the literature cited above, we can conclude that demographic data exhibited significant age heaping at least until the turn of the 20th century, and that the degree of heaping varied across individuals or groups in a way that makes age heaping a plausible measure of human capital. The correlation of age heaping and the prevalence of illiteracy among the population was explored in more detail by A'Hearn, Baten and Crayen (2006) who found in their analysis of 52 countries or 415 separate regions that the level of age heaping is indeed correlated with illiteracy. The authors also concluded that the probability to report a rounded age increases significantly with regional and personal illiteracy.

In the following, we will first discuss important methodological aspects in the second section. The third section takes a look at the (non-interpolated) country level data especially for the Islamic world and the industrialized countries. The next section will attempt a first estimate for eight world regions, using quite a bit of interpolation, in order to describe the global development of numeracy during the period 1820-1949. Section V will discuss the potential determinants of numeracy. Finally, section VI tests the implications for economic growth, and the seventh part concludes.

II Methodological aspects

2.1 The age heaping technique

The age heaping technique can be applied to many sources of age data such as census returns, military enrolment lists, legal or hospital records, and tax data. However, care must be taken as to who asked questions regarding age and how, and whether self-reported ages were counterchecked with birth registers. Counterchecked age information does normally not reflect any age heaping besides minor, random fluctuations and can hence not be used. If the enumerator asked for both age and birth year, the resulting data became sometimes unusable for our purposes due to mixed-age and birth-year heaping. In consequence, we used only census data, which lacked such problems in order to maximize the representativeness of our numeracy proxies.

The Whipple Index as a measure of age heaping was designed to capture heaping on ages ending in 0 or 5. Applied to an age range divisible by 10 (i.e. in which every digit occurs with the same frequency), it sums the frequencies of all ages ending in a multiple of five and expresses the result relative to one fifth of the sample size. The resulting ratio is multiplied by 100, yielding an index, which ranges from 0 to 500. Accordingly, a Whipple Index of 0 (500) implies no (only) ages ending in multiples of 5. Generally it is assumed that a Whipple Index of 100 reflects the "true" age distribution with exactly 20% of all ages ending in a multiple of five-digit. The Whipple Index is linear, and a 50% increase in the share of ages ending in multiples of five translates into an increase of the Whipple Index by 50%.[134]

2.2 Age-group boundaries

As mentioned above, one technical requirement for the calculation of the Whipple Index is an age range of 10 successive single years. Since less people are alive at age 69 than at age 60, the total share of people whose age ends in 0 should be naturally larger than the total share of people reporting an age ending in 9. Thus, age heaping is likely to be overestimated if we calculate the Whipple Index over an age range such as 20-29, 30-39, etc. In order to mitigate this effect and to spread the final digits of 0 and 5 more evenly across the age ranges, we calculated the Whipple Index for fixed age ranges starting with the final digit 3 and ending with the final digit 2, such as 43-52, for instance. The age-group-specific Whipple Indexes were used as proxies for the numeracy levels of the decades in which most individuals were born. For example, individuals belonging to the age-group 23-32 enumerated in the 1881 census were born between 1849 and 58. The corresponding Whipple Index was used to proxy the numeracy level for the birth decades of the 1850s. We also considered the question of

[134] An even more intuitive linear transformation of the Whipple Index (WI) is the ABCC Index which reports a society's share of individuals who probably know their true age (named after A'Hearn, Baten and Crayen as well

different ages having different age-specific Whipple indices, as described in the appendix below.

III A first glance at country level data

For the industrialized countries, the coverage of our data set is very good (Figure 1). Most countries conducted several censuses during the 19th and 20th centuries, and our cohort analysis covers the majority of the Western world (incl. Japan). However, given that many of these countries had already experienced the peak of their decline before 1820, the observations cluster together strongly between 100 and 120 for the early-to-mid 19th century, and between 100 and 110 for later years. There are two extremes and a few notable exceptions. The extreme cases are Greece and Cyprus that started out from a very high age heaping level in the late 19th century and then improved rapidly. The numeracy retardation of those countries might have been caused by the backward educational institutions of the Ottoman Empire which may still have impacted on numeracy some time after Greece became independent in 1829, and Cyprus was ceded to Britain in 1878; or possibly the Greeks did not improve their institutions considerably.

Notable exceptions are the U.S., Canada, and Italy, which initially had quite strong age heaping, and Spain, Portugal, and Ireland, which experienced an adverse development during the mid-to-late 19th century (on the age heaping increase in Spain during the mid-19th century famine period, see Manzel 2007). For the U.S., age heaping was mainly a Southern phenomenon (Crayen and Baten 2006). Even the UK had substantial age heaping during the early 19th century, whereas Scandinavia and central Europe had relatively good numeracy values during this period.

A strong contrast to the industrialized countries existed in the Middle East and North Africa, but even here, interesting differences are observable which might be even more

---

as Greg Clark, who first gave us this comment for our 2006 paper). The formula is ABCC=1-(WI/100-1)/4. In informative thon the absolute level of age heaping (Figure 2). However, we must note here that our sources vary in quality over time. We have a number of censuses from the 1930s-1950s period; hence the oldest groups that we could analyze for this region were born in the 1880s or 1890s. Two sources date from earlier periods. One of them is the Egyptian census of 1907 that lists individuals born as early as the 1830s. Another source is a census of the Turkish province of Kars, which was under temporary government by the Russian Czar in 1878-1918 so that we are able to obtain information about Kars from the Russian census of 1897. In this period, mainly Turks, Armenians, Kurds, Azerbaijanis, Greeks, and a minority of Russians (7%) inhabited the Kars region. Yet how representative might Kars have been for the territory, which is now Turkey? Of course, literacy and numeracy were presumably much lower in the countryside than in the metropolis Istanbul. But otherwise, when judging from height data, for example, this region seems fairly representative of Turkey's rural and small-town landscape. If we compare the final years of the two early series with the Egyptian and Turkish series starting with the 1880s, they are not too far apart. Those two early series are both very high, although Egypt had much worse values than the Turkish Kars region.

In general, the countries in this world region had high age heaping values that remained so during the early 20th century. But there are also exceptions. For example, Algeria (abbreviated as 'dz'), Tunisia and Jordan had much better (i.e. lower) values than the other countries. It might be the case that the French educational policy had some positive effects on the former two countries, even if it was probably insufficient in general. Moreover, the French settlements in both countries might have caused spill-over effects to the Arabic and Berber majorities of those countries. Iran performed fourth-best in the region. Iraq did also quite well initially, but was then overtaken by countries such as Turkey, Morocco, Bahrain (abbreviated as 'bn'), and Kuwait later on. Afghanistan and Egypt performed worst in the early 20th century.

---

this version of the paper, we still report the Whipple Index, as it is the UN standard measure.

To sum up our evidence for the Middle East and North Africa, age heaping was quite high and improvement slow during the 19th century. During the early 20th century, numeracy improved strongly but was still far from the European (or East Asian) level in the mid-20th century. This is quite an astonishing result, given that Arabic numerals were a huge innovation in Europe in the later Middle Ages and early modern period. In this era, the Islamic world must thus have been far advanced in terms of numeracy compared with the Europeans, also because other important innovations also originated in the Middle East. Unfortunately, we do not yet have age heaping information on the early modern or late medieval period in the Islamic world. Yet any attempt to associate the lagging development of the Middle Eastern region in the 19th and early 20th century with an "Islamic" mentality must be clearly dismissed, since neither the Christian populations of the nearby Caucasus and Balkan regions (Georgians, Armenians, Serbs, Greeks) nor most Hindu regions were doing much better in terms of numeracy. Rather, it was probably the adverse institutional and educational infrastructure of this world region that led to low numeracy levels, just as there were positive exceptions such as the Algerian, Jordanian and Tunisian cases.

IV World region estimates

We also present a very preliminary estimate of numeracy trends for the regions of the world (Figure 3). Our basic strategy consisted of collecting as many census data as possible and calculating the Whipple Index for the age-groups 23-32, 33-42 and so on, until 63-72. For the industrialized countries and East Asia, we obtained many of the necessary country-birth decade observations, whereas for the Middle East and North Africa, documentation was much more sparse, especially before the 1880s. For East and Southeast Asia, we were able to produce estimates for the period from the 1840s onwards, while values for sub-Saharan Africa and Latin America could only be traced back to the 1880s and 1890s. The remaining gaps were interpolated, mostly using a benchmark for a given country and then applying the growth rates of the most similar neighboring country (or countries) for which data were available. The idea here is that trends in neighboring countries were often quite similar, although they might have differed in levels (the interpolation decisions are documented in an appendix available from the authors). The final step was to calculate population-weighted averages for the eight world regions.

East Asia (clearly dominated by China, as Japan was put in the group of the industrialized countries) is the only region for which we partly relied on sample information. Census data did not become available for the birth cohorts before the 1890s. For earlier periods, we worked with the estimates of Baten, Ma, Morgan and Wang (2007) who used data on Chinese migrants to the U.S., Australia, and Indonesia, as well as on soldiers in Beijing and legal records to estimate a Chinese age heaping trend. Especially the close correspondence between all those series led Baten et al. to the conclusion that age heaping levels were substantial during the early and mid-19th century, culminating in the civil war and famine period of the 1840s and 1850s. During the 1870s and 1880s, in contrast, Chinese age heaping vanished. The authors took great care in assessing cultural explanations of Chinese age reporting preferences, such as assessing tiger year preferences, 4 avoidance, 8 preferences etc. As a result, Chinese age heaping was not found to have been fundamentally different from rounding patterns in other parts of the world, although the Chinese cultural interest in the calendar and astrology had a positive effect on number abilities and number discipline.
For Southeast Asia (SE Asia), the estimates are halfway representative from the 1890s onwards. For earlier times, we relied mostly on Myanmar (the British province of Burma).
For Latin America and the Caribbean (LAC) as well as sub-Saharan Africa, we restricted our preliminary estimates to the last four to six decades, as we did not have representative data for the preceding period.

What we can infer from these world regional trends is that the Middle East/North Africa and South Asia had the highest age heaping (low numeracy) during most of the 19th

century, whereas a substantial improvement took off in the 20th century. On the other extreme, the industrialized countries displayed noticeable age heaping until the 1880s, although it was quite moderate and disappeared around 1880. Eastern Europe/Central Asia[135] (EEU) was similar to the industrialized countries after 1900, and had only slightly higher values in the two decades before. Between those five extreme world regions of high and low numeracy in the late 19th century, we can identify three world regions of medium numeracy: South East Asia did considerably better than South Asia (but not as good as East Asia), and Latin America's numeracy developed even more favorably. Sub-Saharan Africa performed fifth best (or third worst) around 1900.

The global distribution of numeracy is shown in Figure 4. We do not have estimates for the territories of present day Libya, Israel, West Sahara, Uzbekistan, Turkmenistan, Tadzhikistan and Laos (and countries with less than 500,000 inhabitants), but the remainder of the world is roughly covered. In only 23 cases were we forced to interpolate by using the values of similar neighbouring countries, whereas we had direct data for 73 countries in 1890. For another 49 countries, we had direct benchmark values for the following decades and could estimate growth rates by using the growth rates of similar countries. As a result, we find that the least numerate populations of this period lived on the stretch between Egypt, Sudan, southern Arabia, India and what today is Bangladesh. The remaining northern Islamic region between Persia and Turkey was doing better, with the same applying to Islamic northwestern Africa (except Morocco).

Within other world regions, there are also interesting differences. Africa had much better values for its southern part, which reached as far northward as Angola. The Sahel zone did slightly better than the highly populated Gold and Ivory Coast in this period, and the least numerate Africans were located between Nigeria and the Congo. Similarly, Latin America

[135] The aim was to estimate numeracy for the formerly socialist countries (reference year: 1980). We were able to include Siberia and Kazakhstan but could not yet create estimates for the other Central Asian and Caucasian countries. Omitting Kazakhstan changed the estimates only minutely.

had a strong North-South gradient, but what is even more interesting is that Brazil and Venezuela performed better than the countries located between Bolivia and Mexico.

Finally, within Europe, the expected West-East differential holds, but this does not apply to the far West and Southwest (Ireland, Spain, Portugal). The relatively poor Scandinavians were the "Impoverished Sophisticates" frequently described in the literature. Similar descriptions could apply to the Chinese and Southeast Asians (except for Indonesia and the Philippines).

Summing up, we were able to describe global trends in numeracy for eight world regions, two of which started out with very low numeracy levels (the Middle East/North Africa and South Asia), while three world regions had quite good values (the industrialized countries, Eastern Europe after 1880, and East Asia). Among the three medium regions, the case of Africa is particularly remarkable, as it performed better than South Asia and the Middle East in terms of numeracy. Given its educational standing today, it was unexpected that Africa would have started out from a relatively promising level. Another noteworthy case is the deterioration of numeracy in East Asia during the mid-19th century civil war catastrophes (albeit still on a relatively modest scale).

V Determinants of age heaping

Age heaping mainly originates either from a respondent's ignorance of his or her exact age, or missing number discipline. Nowadays, we can assume that most people living in industrialised countries know their exact age or their year of birth, and otherwise that they can check this information in registers, passports etc. if necessary. By contrast, people living in developing countries will more often have only a vague idea about their age or the year when they were born (this even applies to societies where birth registration has become well established). In such countries, age-awareness can still be quite low (Kaiser and Engel 1993). As discussed above, one likely determinant of age heaping is the degree of schooling a person received. In school, children are taught numbers, and the formal schooling system is likely to improve their structural thinking skills in general, which in turn might enhance their number knowledge and discipline. A second potential determinant of age heaping is the degree to which an individual interacts with the state or religious and other administrative authorities, and the degree to which that authority is involved with market transactions. Another numeracy-enhancing factor besides the state bureaucracy, market demand or schooling success is the frequent use of calendars or astrological elements as in the Chinese culture, for example.

Finally, in many historical societies and the poorest economies today, the infant protein malnutrition syndrome (IPMS) plays a role since it limits an adult's cognitive abilities. This factor is especially important when we observe increases of age heaping although there is no deterioration in the political, economic or educational sphere that would sufficiently explain this development, as was the case in China or Ireland in the mid-nineteenth century, or in England around 1800 (see Manzel 2007 on the Spanish regions; Baten, Crayen and Voth 2007 on the English regions; Baten, Ma, Morgan and Wang 2007 on China). In principle, an IPMS reduction could also explain the decline of age heaping during the 19^{th} and 20^{th} centuries, when protein access improved. However, the IPMS phenomenon is easier to identify in times of rising or stagnating age heaping. Thus, age awareness yields valuable information both about individuals and the society they inhabit.

To explore the correlation of age heaping with schooling, protein malnutrition, and state development, we included primary school enrolment data from the famous Lindert data set, supplemented by Benavot and Riddle (1988) in our model specification. As our age heaping data were organised by birth decades whereas primary school enrolment takes place at approximately age 10, we lagged the Whipple Index by one decade. To approximate state development, we used the data on state antiquity by Bockstette, Chanda and Putterman (2002) who argued that state antiquity proxies the strength of the state and the quality of its institutions. Although time-invariant, this variable also provides valuable information on the history and importance of state-level institutions that might have generated a demand for the knowledge of one's age even if the prevailing human capital level was not too high. As an alternative (time-variant) measure for state development and for the demand for the knowledge of one's age in society, we computed the number of censuses taken in each country since 1600 up to and including the birth decade.[136] Following Domschke and Goyer's *Handbook of National Population Censuses* the highest score for 1880 was assigned to Greece, where as much as 16 censuses had been conducted until 1889. In our model, we use three dummy variables denoting '1 or 2', '3 to 5' and '6 or more' past censuses. Fourthly, we included a (time-invariant) dummy variable for those nations, which were influenced by the Chinese calendar/astrology/culture.[137] Finally, we used the recent global height estimates by Baten (2006) to control for infant protein malnutrition (also lagged by one decade). In order to measure non-constant marginal effects, we included square and square root terms but found the log specification to perform best. Unfortunately, the inclusion of these variables led to a decline in the number of cases. Moreover, we included only those age heaping data points, which were not interpolated here – using not even the benchmark method described above.

As a result, all specifications yield that schooling is by far the strongest determinant of age heaping patterns (Table 1). In the birth decade-specific OLS regression, schooling has a very robust coefficient and very large t-values. It is important to note that the coefficients remain similar if the underlying samples are large enough. In contrast, height has a variable influence, which is evident only for the decades for which we have the largest number of observations (1920s-1940s, plus the 1900s). The state history variable which proxies the

---

[136] Censuses were counted only if they covered the vast majority of the population (i.e., colonial censuses enumerating the white population only were not included) and if the censustakers asked for the age.
[137] The dummy variable takes the value 1 for China, Hong Kong, Taiwan, North and South Korea, Japan, Singapore, and Vietnam.

authority of the state is never significant in this series of cross-sections, and East Asia has the expected reducing effect: in China and its neighbouring countries, we observe less schooling but more age awareness. We also tested several panel regression models and found our results confirmed (Table 2). Notably, schooling is closely correlated with age heaping in the fixed-effects regression, which controls for unobserved heterogeneity between countries as well. The schooling investment variable is statistically and economically significant: a change of one standard deviation of the schooling variable (1.193) in the random-effects specification, multiplied with the coefficient, accounts for 44% of the standard deviation of the dependent variable (0.472). In contrast, the same procedure for the height and state history variables leads to the conclusion that a rise of those variables by one standard deviation leads to a less than 10% increase of the standard deviation of the dependent variable. In the case of the height variable, we also have to admit the possibility of endogeneity (numeracy might improve the welfare level), but this should bias the significance levels upwards. Moreover, height is also closely correlated with GDP, so that its effect might proxy a general developmental effect rather than an infant protein malnutrition effect. Once the dummy variables indicating the number of past censuses are integrated, the height variable loses significance. From those considerations, we conclude that the importance of the height variable might not be given in global cross-sections and panels, whereas the importance of malnutrition in time series of age heaping was shown in related studies (Manzel 2007; Baten, Crayen, and Voth 2007; Baten, Ma, Morgan and Wang 2007). At the same time, countries where census taking has become standard, i.e. countries with 6 or more past censuses, have significantly lower age heaping levels. The effect is not very large though: an one standard deviation increase in the census variable improves the logarithm of the Whipple Index by 3% of its standard deviation only.

In order to provide a schooling estimate based on age heaping, we also performed the regression with schooling as the dependent variable (see Table 3). The results apply to all world regions except East Asia. Again, the coefficients are very stable over time, suggesting that one additional percentage point in age heaping approximates 1.8-2.0 percent less schooling. Given that this is time-invariant for the period between the late 19th century and the mid-20th century, we would suggest using a value of 1.9 to estimate schooling throughout the 19th century.

VI Implications for empirical growth economics

What are the implications of our new estimates for empirical growth economics? Can age heaping based on numeracy explain growth capabilities in a large number of countries? In the following we argue that this is indeed the case. We employ all available GDP growth figures between 1820 and 1870, as well as between 1870 and 1913, and combine them with our numeracy estimates of the respective periods. We can identify a set of 62 GDP growth figures for those two crucial periods in the 19th century that can also be documented with initial numeracy. The sample is quite comprehensive, as it does not only include today's rich countries, but also Less Developed and Middle Income Countries such as Myanmar, Egypt, Malaysia, and Armenia (see the notes to Table 4). We can also account for later oil exporters such as Iraq and Iran, and some growth successes of the 20th century (Korea, for example). In a set of regressions, we follow the standard procedure developed by Barro (1991, 1999) and many others, who regress growth rates on a set of "growth capabilities" measured in levels. The level of human capital is such a "growth capability", since theory suggests that after controlling for the initial welfare level (which might also proxy a country's capital stock, as Barro (1991) has argued), only countries with high human capital can achieve successive welfare growth.

The numeracy estimates given in the Whipple Index have considerable explanatory power: its coefficient is consistently negative, as expected, and highly significant. In a general regression (column 3 in Table 4), we control for the initial level of GDP, an East Asia dummy (due to the Chinese calendar usage, numeracy in East Asia might be lower than would be expected from the low age heaping levels) and state antiquity. A standard deviation of the logarithm of the Whipple index of those 62 cases which we could include in column (1), (3) and (5) amounts to 0.53. Hence, the difference in annual growth rates between average cases and those that feature age heaping levels one standard deviation higher is 0.37 percentage points (column 3), which is clearly economically meaningful, as the standard deviation of the dependent variable is only 0.71 percentage points. We also regressed GDP growth on Whipple indices alone, and found a significant impact, albeit a smaller coefficient (column 1). The East Asia dummy is in fact negatively significant, which might be caused either by the measurement effect described above, or by the disappointing growth in East Asia for political and other reasons during this period. Initial GDP is not significant in the first specification, but it turns significant once we include a dummy variable for the period 1870-1913, as this pooled cross-section might be characterised by heterogeneous intercepts (but its insignificance suggests it is not, column 5). Hence there is apparently some conditional convergence observable, once we control for initial human capital and other variables. State antiquity has a negative effect here (quite the opposite to its effect in the 1960-90 period). This might suggest that strong government institutions are not inevitably favorable. In times of massive industrial change, a strong state dominated by vested interest groups and conservative feudalists might have been a hindrance.

Unfortunately, we are not able to provide a matching test for schooling effects on growth, as schooling is available for only 25 cases in the 1870-1913 period. For this period, schooling is consistently insignificant at the 5% level. But we cannot interpret this as evidence that numeracy explains more than schooling. For a direct comparison, the sample (for which we have both numeracy and schooling values) comprises 16 cases only, the reduced sample size leading to insignificance of all potential explanatory factors. However, we can conclude that the age heaping strategy allows capturing the human capital impact on 1820-1913 growth successes and failures, whereas our knowledge about schooling is insufficient to do so. Measuring numeracy is clearly an important activity in order to understand long-run economic growth.

VII Conclusion

In this paper, we presented age heaping as an indicator of human capital, which of course has its limitations. However, this is the nature of all human capital indicators. The limitations of signature ability as a measure of functional literacy are obvious, but can also be raised with respect to the self-reported “ability to read.” School enrolment as an input measure is conceptually problematic, as we do not know about the quality of schooling and the concept and educational methods of the teacher. We think that the comparison of different human capital indicators is the most promising way to establish a reliable database that can inform growth economics, economic policy recommendations, and many other fields of research.

South Asia and the Middle East had relatively low numeracy levels in the 19th century (as opposed to the Western world and East Asia). Also, South East Asia as exemplified by Burma, for instance, had better values than South Asia. China stands out as a country with very low age heaping levels in the late nineteenth century - comparable to Western industrialised countries. This would have suggested good prospects for the future economic development of China, had it not been for the civil war and other political obstacles to economic and social development.

It is remarkable that the northern Islamic countries stretched between Iran and Turkey, and the northwestern part (Algeria, Tunisia etc.) of the Islamic world performed much better than the Southeast. Within Europe, the Greeks, Cypriotes, Irish, Portuguese and Spanish stood out as having had relatively low numeracy during the 19th century. Especially for the birth cohorts after the Great Famine, there was even a temporary deterioration before Ireland and Spain converged back to Western European levels. Other parts of Europe experienced their numeracy revolution already in the 17th and 18th centuries. The United States (or rather its Southern part) took a position of lagging numeracy in the first decades of the 19th century, but developed rapidly in the second half of the century. For sub-Saharan Africa and Latin America, the data sources do not allow estimates for the time before the 1880s and 1890s. Preliminary outcomes suggest that Africa belonged to a middle group between Southeast Asia and the Middle East.

We ran explorative regressions on the determinants of age heaping. We found schooling to be the most important correlate. Protein malnutrition might have played a role, but the results were much less consistent, and endogeneity could not be ruled out. While the age and authority of the state bureaucracy did not seem to play a significant role, our results indicated that more generally, state demand for frequent age reporting improved people's numeracy or number discipline, as we can see from the fact that countries with a long tradition of census taking had slightly lower age heaping levels.

Finally, we assessed the contribution of numeracy as measured by the age heaping strategy for long-run economic growth. In a variety of specifications, numeracy mattered strongly. We can conclude that age heaping allows capturing the human capital impact on 1820-1913 growth successes and failures, whereas our knowledge about schooling is insufficient to do so. Measuring numeracy is clearly an important activity in order to understand long-run economic growth around the globe. Overall, the age heaping technique allows a more nuanced view on human capital formation in most of the world's regions during the 19th and 20th century.

References

A'Hearn, B., Baten, J., Crayen, D. 2006. Quantifying quantitative literacy: Age heaping and the history of human capital. Universidad Pompeu Fabra Economic Working Paper No. 996.

Bachi, R. 1951. The tendency to round off age returns: measurement and correction. Bulletin of the International Statistical Institute 33, 195-221.

Barro, R. 1991. Economic growth in a cross section of countries. Quarterly Journal of Economics CVI , 407-443.

Barro, R. 1999. Determinants of democracy. Journal of Political Economy 107 (S6), S158-29.

Baten, J. (2006). Global Height Trends in Industrial and Developing Countries, 1810-1984: An Overview. Working Paper Tuebingen.

Baten, J., Crayen, D., Voth, J. 2007. Hungry, Stupid and Poor: Cognitive Ability, Income Support, and Economic Growth in Industrializing Britain. Working Paper, University of Tuebingen/Universidad Pompeu Fabra.

Baten, J., Ma, D., Morgan, S., Wang, Q. 2007. New Findings on Historical Living Standards and Human Capital in China: an anthropometric and quantitative approach. Working Paper.

Baten, J., van Zanden, J.L. 2006. The impact of human capital on early modern welfare growth, 1450-1913: book production as a human capital indicator. UPF Working Paper.

Becker, G.S., Murphy, K.M., Tamura, R. 1990. Human capital, fertility, and economic growth. Journal of Political Economy 98, 12-37.

Benavot, A., Riddle, P. 1988. The expansion of primary education, 1870-1940: Trends and Issues. Sociology of Education 66 (3), 191-120.

Bockstette, V., Chanda, A., Putterman, L. 2002. States and markets: The advantage of an early start. Journal of Economic Growth 7, 347-369.

Boucekkine, R., de la Croix, D., Licandro, O. 2003. Early Mortality Declines at the Dawn of Modern Growth. Scandinavian Journal of Economics, 105, 401-418.

Budd, J.W., Guinnane, T. 1991. Intentional age misreporting, age heaping and the 1908 old age pensions act in Ireland. Population Studies 45 (3), 497-518.

Carruthers, B.G., Espeland, W. 1991. Accounting for rationality: double-entry bookkeeping and the emergence of economic rationality. American Journal of Sociology 97 (1), 31-96.

Cervellati, M., Sunde , U. 2005. Human Capital Formation, Life Expectancy and the Process of Development. American Economic Review 95(5), 1653-1672.

Clark, G. 2007. A Farewell to Alms: A Brief Economic History of the World. Princeton University Press.

Crafts, N. 1997. The Human Development Index and changes in standards of living: Some historical comparisons. Éuropean Review of Economic History I, 299-322.

Crayen, D. and Baten, J. 2006. Numeracy, Inequality, Age Heaping, and Economic Growth: New Estimation Strategies for Western Europe and the U.S. (17th - 19th centuries). Working Paper Univ. Tuebingen.

De Moor, M., van Zanden, J.L. 2006. Vrouwen en de geboorte van het kapitalisme. Boom, Amsterdam.

Dillon, L. 2007. The Shady Side of Fifty: Age and Old Age in Late Victorian Canada and the United States. McGill-Queen's University Press, Kingston.

Domsche, E., Goyer, D.S. 1986. The handbook of national population censuses: Africa and Asia. Greenwood, New York.

Domsche, E., Goyer, D.S. 1986. The handbook of national population censuses: Europe. Greenwood, New York.

Domsche, E., Goyer, D.S. 1986. The handbook of national population censuses: Latin America and the Caribbean, North America and Oceania. Greenwood, New York.

Duncan-Jones, R. 1990. Structure and Scale in the Roman Economy. Cambridge University Press, Cambridge.

Ewbank, D.C. 1981. Age Misreporting and Age-selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis. National Academic Press, Washington, D.C.

Galor, O. 2005. From Stagnation to Growth: Unified Growth Theory, in: Ph. Aghion and S. Durlauf (eds.) Handbook of Economic Growth, North-Holland.

Galor, O., and O. Moav 2002. Natural Selection and the Origin of Economic Growth, Quarterly Journal of Economics 117, 1133-1192.

Galor, O., and D.N. Weil 2000. Population, Technology and Growth: From the Malthusian Regime to the Demographic Transition, American Economic Review 110, 806-828.

Glaeser, Glaeser, E.L., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer 2004. Do Institutions Cause Growth? SSRN Working Paper id556370.

Herlihy, D., Klapisch-Zuber, D. 1978. Les Toscans et leurs familles: une étude du catasto florentin di 1427. Editions de l'Ecole des Hautes Etudes en Sciences Sociales / Presses de la Fondation Nationale des Sciences Politiques, Paris. Chap. XIII.

Jones, C. 1996. Human capital, ideas and economic growth. Mimeo, Stanford University.

Kaiser, D., Engel, P. 1993. Time and age-awareness in early modern Russia. Comparative Studies in Society and History 35 (4), 824-839.

Lindert, P. 2004. Growing Public. Cambridge University Press, Cambridge.

Long, J. 2005. Rural-urban migration and socioeconomic mobility in Victorian Britain. Journal of Economic History 65 (1), 1-35.

Long, J. 2006. The socioeconomic return to primary schooling in Victorian England. Journal of Economic History 66 (4), 1026-1053.

Lucas, R.E. 1988. On the Mechanics of Economic Development. Journal of Monetary Economics, 3-42.

Mankiw, N.G., Romer, D., Weil, D.N. 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics 107, 407-37.

Manzel, K. 2007. Long run development of human capital and the impact of the "hungry forties" in Spain. University of Tuebingen Working Paper.

Mitch, D. 1991. The Rise of Popular Literacy in Victorian England. University of Pennsylvania Press, Philadelphia.

Mitchell, B. 2003a. International Historical Statistics: the Americas, 1750-1993. M.Stockton Press, New York.

Mitchell, B. 2003b. International Historical Statistics: Europe, 1750-1993. M.Stockton Press, New York.

Mitchell, B. 2003c. International Historical Statistics: Africa, Asia and Oceania, 1750-1993. M.Stockton Press, New York.

Morrisson, C., Murtin, F. 2007. Education Inequalities and the Kuznets curves: a global perspective since 1870. Paris School of Economics Working Paper No. 2007-12.

Mokyr, J. 1983. Why Ireland Starved. George Allen & Unwin, London. Chap. 8.

Myers, R. 1976. An instance of reverse heaping of ages. Demography 13 (4), 577-580.

Myers, R. 1954. Accuracy of age reporting in the 1950 United States census. Journal of the American Statistical Association XLIX, 826-831.

Narasimhan, R. L., Retherford, R.D., Mishra, V., Arnold, F., Roy, T.K. 1997. Comparison of fertility estimates from India's Sample Registration System and Family National Health Survey. National Family Health Survey Subject Reports 4. IPPS, Mumbai.

O'Grada, C. 2006. Dublin Jewish demography a century ago. UCD Centre for Economic Research Working Paper Series WP 06/01.

O'Rourke, K., Williamson, J.G. 1997. Around the European periphery 1870-1913: globalization, schooling and growth. European Review of Economic History 1, 153-190.

Reis, J. 2005. Economic growth, human capital formation and consumption in western Europe before 1800, in: Allen, R.C., Bengtsson, T., Dribe, M. (eds.), Living Standards in the Past. Oxford University Press, Oxford, 195-225.

Retherford, R.D., Mishra, V. 2001. An evaluation of recent estimates of fertility trends in India. National Family Health Survey Subject Reports 19. IPPS, Mumbai.

Romer, P. 1989. Human Capital and Growth: Theory and Evidence. NBER-Working Paper 3173.

Thomas, K. 1987. Numeracy in Early Modern England: The Prothero Lecture. Transactions of the Royal Historical Society, 5th Ser., 37.

Tortella, G. 1994. Patterns of economic retardation and recovery in south-western Europe in the nineteenth and twentieth centuries. Economic History Review 47 (1), 1-21.

Tables

Table 1: Determinants of age heaping, measured by the logarithm of the Whipple Index, by birth decade

Coefficient	1880	1890	1900	1910	1920	1930	1940
Schooling	-0.37	-0.26	-0.26	-0.27	-0.24	-0.20	-0.22
	(0.05)	(0.06)	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)
Height	0.52	-1.68	-2.73	-1.38	-2.09	-1.70	-1.44
	(1.17)	(1.38)	(1.19)	(1.09)	(0.83)	(0.63)	(0.59)
State Antiquity	-0.35	-0.20	-0.13	-0.06	-0.10	-0.17	-0.13
	(0.13)	(0.28)	(0.15)	(0.13)	(0.10)	(0.10)	(0.08)
East Asia				-0.91	-0.63	-0.39	-0.33
				(0.3)	(0.17)	(0.11)	(0.07)
Cons.	6.44	9.34	11.03	8.86	9.82	9.00	8.63
	(1.95)	(2.18)	(1.85)	(1.68)	(1.27)	(0.98)	(0.92)
Obs	22	31	45	60	69	70	64
R squared	0.76	0.57	0.64	0.59	0.63	0.55	0.61

Notes: Robust standard errors in parentheses. The dark and light, grey-shaded areas denote significance at the 1% and 5% level, respectively. Dependent variable: Whipple Index of age heaping (logarithm), lagged by 10 years. Independent variables: schooling: Primary school enrolment in logs. height: Height lagged by 10 years. State Antiquity: State history and authority. East Asia: Dummy for East Asia. Sources: see text.

Table 2: Panel regression results: determinants of age heaping, measured by the logarithm of the Whipple Index

Coefficient	RE (general)	RE (specific)	FE (general)	FE (specific)
Schooling	-0.17	-0.16	-0.14	-0.15
	(0.02)	(0.02)	(0.01)	(0.01)
Height	-0.64		-0.12	
	(0.24)		(0.37)	
State Antiquity	-0.18			
	(0.09)			
Censuses (1 or 2)			-0.01	
			(0.02)	
Censuses (3 to 5)			-0.01	
			(0.02)	
Censuses (6 or more)			-0.05	
			(0.02)	
East Asia	-0.44			
	(0.2)			
Cons.	7.11	5.85	5.92	5.76
	(0.38)	(0.10)	(0.60)	(0.05)
Obs	402	477	471	477
Number of countries	73	91	90	91
R squared (overall)	0.58	0.49	0.49	0.49

Notes and sources: Robust standard errors in parentheses. The dark and light, grey-shaded areas denote significance at the 1% and 5% level, respectively. Dependent variable: Whipple Index of age heaping (logarithm), lagged by 10 years. Independent variables: schooling: Primary school enrolment in logs. height: Height lagged by 10 years. State Antiquity: State history and authority. Censuses: Dummy variables indicating the number of censuses taken up to and including the specific birth decade. East Asia: Dummy for East Asia. Sources: see text.

Table 3: Age heaping as a determinant of student enrolment ratios, by birth decade (East Asia omitted)

Coefficient	1880	1890	1900	1910	1920	1930	1940
Whipple	-2.02	-1.78	-2.01	-1.94	-2.07	-2.06	-2.14
	(0.30)	(0.35)	(0.30)	(0.19)	(0.26)	(0.29)	(0.25)
Cons.	15.79	14.37	15.53	15.18	15.74	15.74	16.18
	(1.45)	(1.72)	(1.44)	(0.93)	(1.28)	(1.39)	(1.19)
Obs	23	35	51	71	82	84	77
R squared	0.72	0.50	0.58	0.62	0.53	0.51	0.56

Notes: Robust standard errors in parentheses. The dark shaded areas denote significance at the 1% level. Dependent variable: schooling (primary school enrolment in logs). Independent variable: Whipple Index (logarithm), lagged by 10 years.

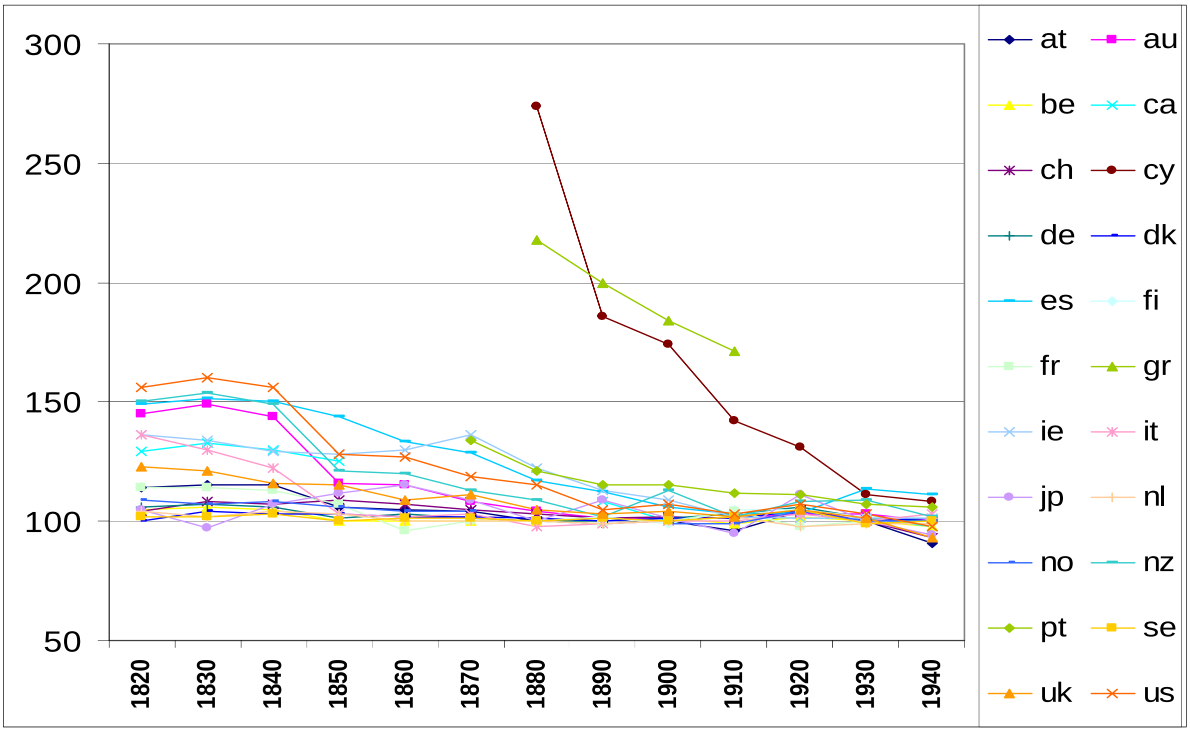
Table 4: Growth Regressions with Age heaping (Whipple Index) and other determinants (pooled cross-sections: 1820-1870 and 1870-1913)

	(1)	(2)	(3)	(4)	(5)
Whipple Index	-0.32		-0.70		-0.75
	(0.08)		(0.21)		(0.21)
Schooling		0.10		0.43	
		(0.06)		(0.22)	
Initial GDPpc			-0.0003	-0.0004	-0.0004
			(0.0002)	(0.0002)	(0.0002)
East Asia			-1.06		-1.12
			(0.28)		(0.27)
Period 1870-1913					0.23
					(0.13)
State Antiquity			-1.62	-0.33	-1.67
			(0.64)	(0.45)	(0.64)
Constant	2.57	0.68	6.18	-44.90	6.47
	(0.46)	(0.42)	(1.74)	(1.07)	(1.75)
Observations	62	25	55	22	55
R-squared	0.05	0.05	0.47	0.18	0.49

Notes: : Robust standard errors in parentheses. The dark and light, grey-shaded areas denote significance at the 1% and 5% level, respectively. Dependent variable is the average growth rate (geometric mean of $GDPC_{1820}$ and $GDPC_{1870}$ or of $GDPC_{1870}$ and $GDPC_{1913}$). The first three independent variables are in logs and refer to the initial year of the dependent variable (the Whipple in 1820 for the growth rates 1820-70 etc.). Countries included in column (1) are (for both periods, unless otherwise mentioned: AM (1820), AT, AU, BE, CA (1820), CH, CN, DE, DK, EG, ES, FR, GR, IN (1870) IQ, IR, IT, JP, KP, KR, MA, MM (1870), MY (1870), NL, NO, NZ, PL (1870), PS, PT, SE, TR, TW, UK, US.

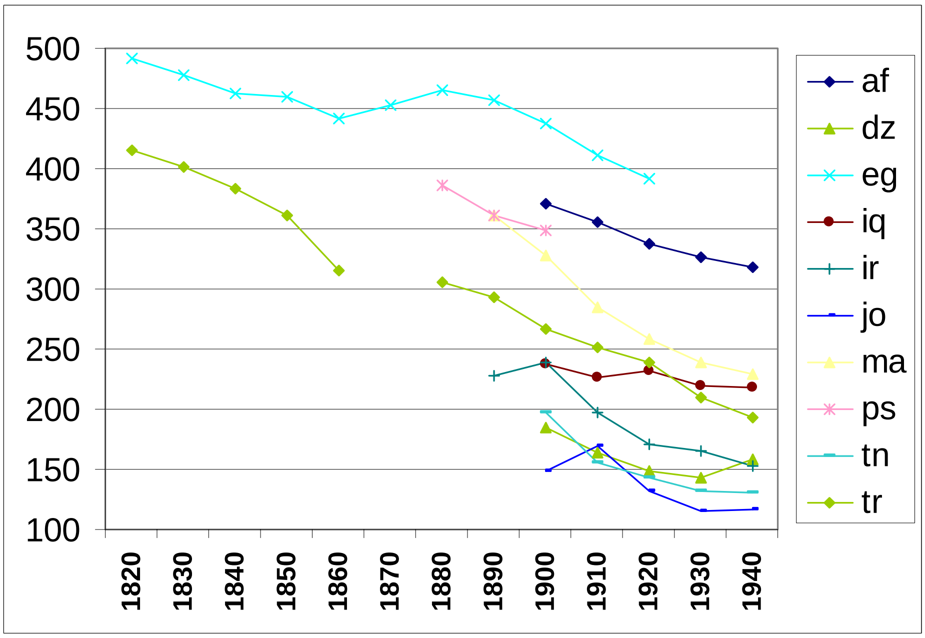
Figures

Figure 1: Age heaping levels in industrialized countries



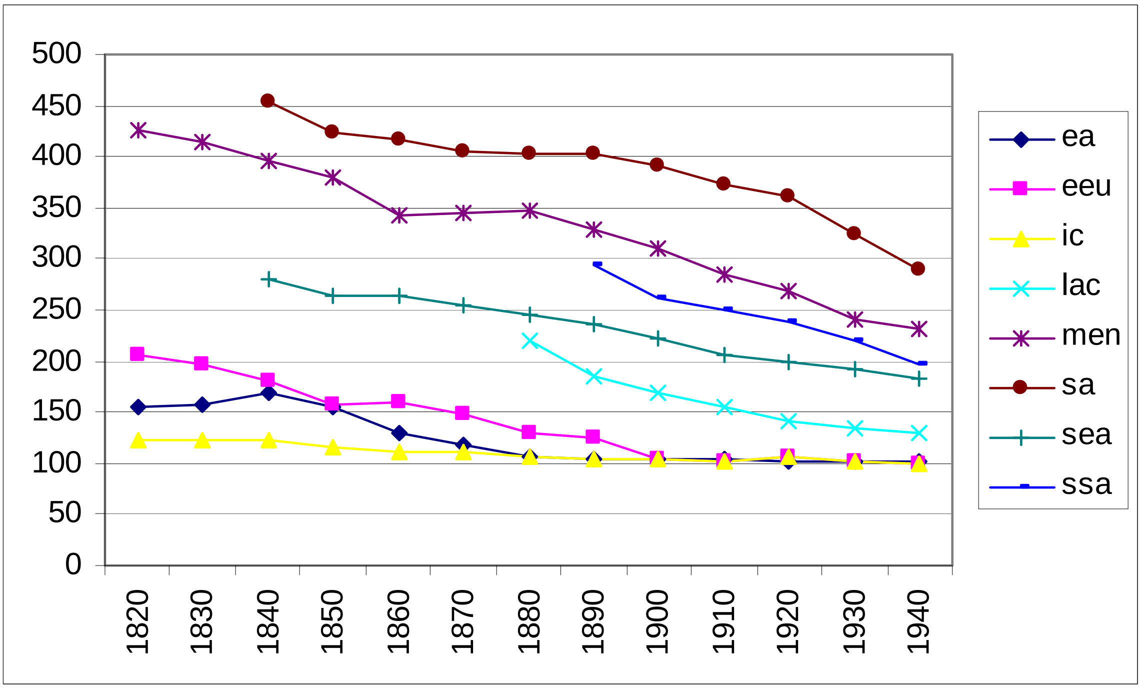
Sources: see text.

Figure 2: Age heaping levels in the Middle East and North Africa



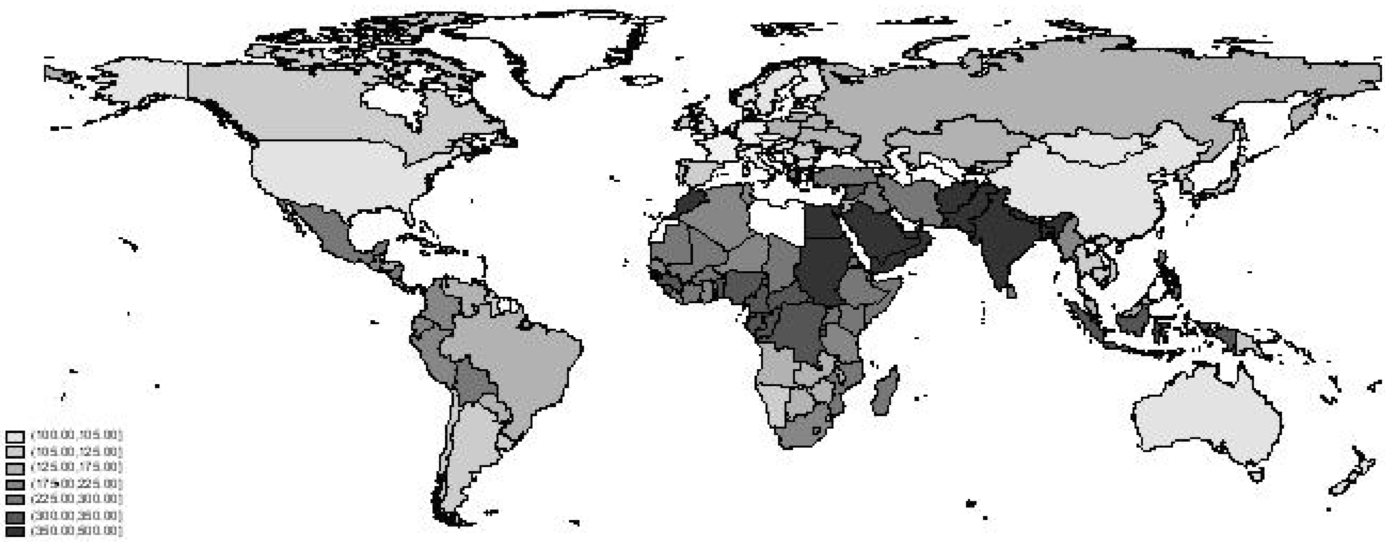
Sources: see text.

Figure 3: Age heaping in the different world regions



Sources: see text.

Figure 4: Worldwide overview of age heaping levels for the birth decade 1890



Sources: see text.

Appendix

Appendix 1: Age adjustment: society-induced age awareness and survivor bias issues

A'Hearn, Baten and Crayen (2006) found evidence for significantly lower heaping levels among young people, i.e. for the age-group 23-32, compared to higher ages. The probability of reporting a heaped age can vary with age itself (even when controlling for cohort effects) for several reasons. In the first place, one could imagine that older individuals are more likely to forget their age.[138] Secondly, a society may provide different incentives to different groups for distorting their true age. For example, women who are still single in their thirties or even later in life might feel pressured to report a low age in order to increase their "marriageability," a distortion which could then be associated with heaping, although heaping could also take the form of excess frequencies of numbers like 29 rather than multiples of five.[139] Thirdly, young people may be more aware of their exact age since they are in the process of passing through important stages of their societal life, acquiring increasingly more rights, but also more responsibilities. Minimum age requirements for marriage, military enlistment, and participation in elections are only a few examples of the importance of age in the second and third decades of life. This point is illustrated by Keith Thomas (1987) who states for early modern England that "[…]ages were reported with precision for people under twenty, because differences in age could be of considerable social and administrative importance for young persons. But when people reached adulthood their exact numerical age had much less social meaning; and it was much more vaguely reported". In accordance with this argument, Dillon (2007) concluded from Canadian and U.S. census data taken between 1870-1901 that being above 24 increased the propensity to round off one's age significantly. Using separate logistic regression for each census, she found the age-groups 35 to 54 and 35

[138] See the discussions in Kaiser and Engle (1993) and Ewbank (1981).
[139] See Retherford and Mishra (2001) and Narasimhan et al. (1997).

to 64 (for the 1870/1 and 1900/01 censuses respectively) to have the highest probability to report a rounded age.

Among the higher age-groups, two different effects might play a role. On the one hand, people might tend to pay less attention to their age or forget it, as described above. But there might also be a small positive selection effect with regard to older people, as more educated people might have a lower mortality risk. Although we omitted the age-groups with high mortality rates (such as children and adolescents below 23 as well as old people above 72), there is the possibility of a survivor bias among the older age-groups. The dimension of the bias will be determined by the degree of selectivity and the inequality of numeracy in the population. However, even under the assumption of extreme selectivity and a mortality rate difference of about 10% between age-groups, the Whipple Index will hardly be underestimated by more than a few percentage points for any reasonable level of numeracy inequality. Due to successive census taking, we were most often able to estimate Whipple Indexes for mixed age-groups, a strategy which will reduce the survivor bias further. Egypt and Turkey are examples where we did not have multiple age-groups for certain birth decades. Nevertheless, the data did not suggest any survivor bias in these cases: the Whipple Index for Turkey stagnates in spite of the shift from the youngest group in the earlier census (birth decade 1860) to the oldest age-group in the later census (birth decade 1880), while the Whipple Index even increases for the Egyptian time period 1870 to 1880, which marks the line between the two succeeding censuses. As a second reason for a potential old-age-effect, the literature reports the observation that beyond a certain age, people begin to be proud of their age. This pride might either lead to more mental investment into recalling one's exact age, or to bragging about an extremely old age ("I'm already 100 years old" was a frequent statement in Soviet Central Asia, which biased life expectancy estimates upwards). Although individuals of 73 years and older were excluded from the sample, it is not clear whether the positive or the negative effect prevails.

In sum, all of the above boils down to the following question: is there a systematic deviation of one particular age-group's Whipple Index from the other Whipple Indexes measured for different age-groups, holding birth cohorts constant? That is, do people – with a certain level of education/numeracy given - heap more or less in different stages of their life due to changing number discipline and demand for the knowledge of their age? For many countries, we have several numeracy estimates for the same birth decade because of successive census taking. These estimates are based on the self-reported ages of varying age-groups, depending on the census year and the birth decade. If heaping behaviour changes with age, the "true" numeracy of the whole population in a given country could thus be under- or overestimated.

We would expect that the degree to which individuals change their heaping behaviour over time, and in particular the degree to which they round less in the age-group 23-32 depends on the overall numeracy level at a given time. Hence, we formed groups of different heaping levels and then ran regressions separately for each group. We distinguished the following heaping categories: Whipple Indexes of 105-124, 125-149, 150-174, 175-224, 225-299, 300-349, and 350 and above (see Table A.1 for the distribution).

Table A.1: Sample sizes for the age effect analysis

heaping category	sample size
105	420
125	306
150	183
175	302
225	184
300	148
350	51

Sources: see text.

Birth decade dummies were included to account for cohort effects. Over the long run, heaping patterns changed considerably. Russia started out with high heaping values - around 250 in some regions - which declined over the century to almost no heaping. We therefore grouped the countries not only by their overall heaping level across the entire data set, but generated a category which grouped them by half-century and country (or region in the case of large countries).

We then ran level-specific regressions on age-group and birth decade dummies. The 33 to 42-year-olds served as a reference category, as did the birth decade 1920 (chosen based on the highest number of observations). Following the reasoning above, we particularly tested the hypothesis that younger people (aged 23-32) are more aware of their age due to their proximity to birth, marriage, and military service. We might also observe higher heaping values for the oldest age-groups due to lower memory abilities and declining number discipline, or lower heaping values due to positive selection and pride of one's old age.

In general, those aged 23-32 did heap systematically less (Table A.2).[140] The result is consistent for all levels of age heaping. For those aged 23-32, marriage and similar events took place or had taken place not too long before, resulting in a better knowledge of their age or better number discipline and less heaping for this age-group than the 33-42 age-group in the same birth decade. The results for the 53-62- and 63-72-year-olds living in relatively numerate societies were not consistent. Two level-specific coefficients of these age-groups indicated that 53-62-year-olds and 63-72-year-olds heaped in fact slightly less than the reference category (represented by the constant).[141] Among the less numerate societies with Whipple Indexes above 300, the negative effect for older age-groups seems to have been stronger (resulting in large coefficients).

In conclusion, to create a data set which is comparable to estimates from other age-groups on the basis of the 23-32 age-group, we need an adjustment. For this aim, we used the coefficients of the 23-32-year-olds - which reflected the only systematic age effect - and included them in a second regression framework in which we sought to model the relationship

[140] Below 105, age heaping can be regarded as randomly fluctuating around 100. Hence, we omitted this group in our age effect analysis.

between the heaping level and the magnitude of the age effect in a continuous rather than a discrete and group-specific way (see Figure A.1).

Table A.2: Coefficients of the age-group specific dummy variables

heaping category	age-group effect				
	23-32	43-52	53-62	63-72	constant
105	-5.19	2.23	-2.38	-5.65	112.37
	(1.39)	(1.88)	(1.96)	(1.65)	(1.73)
125	-8.74	2.53	-10.3	-7.65	138.48
	(3.41)	(3.80)	(4.87)	(7.44)	(5.73)
150	-25.42	1.6	-3.18	-7.37	166.31
	(12.41)	(10.67)	(9.20)	(9.18)	(10.66)
175	-23.03	4.81	-1.16	16.42	194.62
	(8.99)	(8.23)	(8.49)	(13.76)	(8.67)
225	-42.39	-3.98	10.39	5.92	271.89
	(15.78)	(17.35)	(19.08)	(14.14)	(27.53)
300	-43.22	19.7	31.3	-0.27	318.1
	(15.15)	(16.11)	(18.47)	(27.02)	(17.70)
350	-70.91	30.77	34.64	115.98	417
	(19.22)	(16.46)	(18.26)	(16.4)	(19.22)

Note: Robust standard errors in parentheses. The dark and light, grey-shaded areas denote significance at the 1% and 5% level, respectively. Dependent variable: Whipple Index by heaping group. The age-group coefficients reported above were obtained through regressions including birth decade dummy variables. Sources: see text.

Table A.3: Age adjustment for the age-group 23-32: OLS-regression of the 23-32 age-group effect on the general heaping level in 8 heaping categories

Coefficient	
Age Heaping level	-0.21
Constant	-3.03
N	8
Adj. R^2	0.95

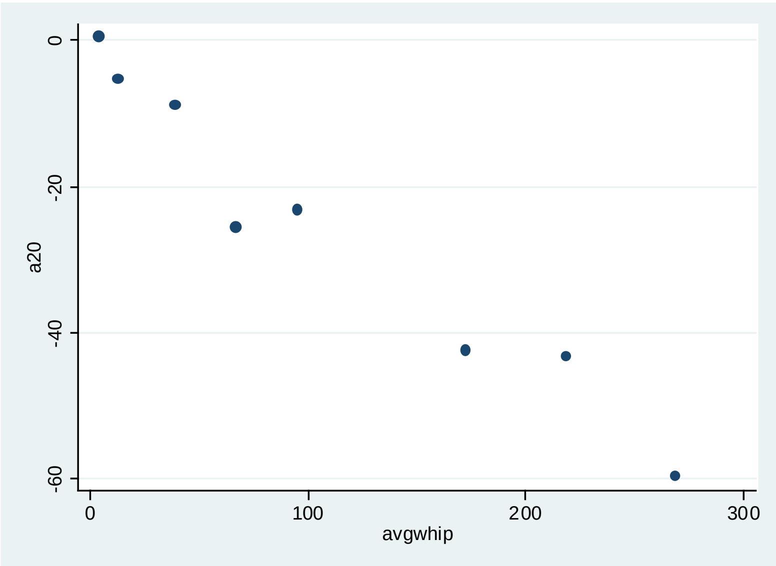
Note: the independent variable is the level-specific regression constant minus 100, which refers to age heaping in the age-group 33-42. The dependent variable comprises the coefficients for the age-group dummy for 23-32-year-olds from the level-specific regressions.

We measured the heaping level by the level-specific regression constant minus the 'zero-heaping'-Whipple Index of 100. We found that the larger the heaping level, the stronger the age effect for the age-group of the 23-32-year-olds, which suggests the following formula to correct for age-induced heaping biases: add 0.2 Whipple units for every Whipple unit above 100 for the age-group of the 33-42 year-olds (Table A.3). For example, if the level of age

[141] Here, we would not suggest a systematic adjustment.

heaping in a given country and period can be described by a Whipple Index of 150 for those aged 33-42, the Whipple Index for 23-32-year-olds should be adjusted upward by 50*0.2=10 units, leading to an age-adjusted Whipple Index of 160.[142] We counterchecked this adjustment for regions for which we had overlapping birth cohorts, and it yielded plausible results.

Figure A.1: Heaping level and age effect among 23-32-year-olds



Note: avgwhip is the level-specific regression constant minus 100, which refers to age heaping in the age-group 33-42. Sources: see text.

142 Alternatively, if the Whipple Index is 150 for the 23-32-year-olds, then the "unbiased" level of the older age-groups is around 160. Hence, the adjustment should be 60*0.2=12, resulting in a Whipple Index of 162 for those aged 23-32. We applied this adjustment to the data above.

6 LIFE EXPECTANCY IN MYSORE STATE DURING THE EARLY 20TH CENTURY: THE DUAL PURPOSE OF SINGLE YEAR AGE DISTRIBUTIONS

Abstract

This paper contributes to an academic dispute in Indian historical demography. While it has been accepted as a fact that there is a regional divide in demographic parameters such as infant mortality and fertility between northern and southern provinces of post-WW2 India, the question whether this north-south divide can be traced back to the colonial period and also be extended to other demographic parameters such as life expectancy levels as Saito, Takahama and Kaneko (2005) suggested has created need for further evidence. This paper employs adjustment procedures for age heaping, migration and underreporting of events of death. It presents new estimates of life expectancy for Mysore State during the early 20th century that indicate relatively favorable living conditions similar to Madras although the region was relatively poor. Mysore province is hence an interesting case study for research on high life expectancy levels in comparably poor countries as initiated by Riley (2005).

I. Introduction

The research on living standards for the last two centuries is receiving renewed interest in economic history and historical demography. Besides level and trends of real wages and GDP per capita, the biological components of the standard of living as well as mortality and life expectancy are central research fields. Given that the history of mortality and life expectancy lies at the core of the development of living standards, whether gauged economically or culturally, new findings can provide input to discussions in many fields of research.

One hotly contended subject-matter is the spatial inequality and the development of living standards in India. It has been accepted as a fact that there is a regional divide in demographic parameters such as infant mortality and fertility between northern and southern provinces of post-WW2 India. Some researchers have argued that this north-south divide can be traced back to the colonial period and also be extended to other demographic parameters such as life expectancy levels (Saito, Takahama and Kaneko 2005). However, the underlying evidence is sparse and as yet inconclusive since previous studies differ in the interpretation of their results and their assessment of the quality of the data used. The main aim of my study is to contribute new evidence to this academic dispute in historical demography. Moreover, I will address the paradox of poor countries featuring relatively high life expectancy levels (Caldwell 1986, Riley 2005), prominently discussed by Sen (1999) who found that China, Costa Rica and the Indian State of Kerala have achieved large increases in life expectancy despite low per capita income, in contrast to, for example, the low life expectancy of black American New York residents.

Life expectancy at birth is a classic indicator used in the research of living standards but it is quite difficult to calculate for populations in the past. As outlined in Crayen and Baten (2008a, 2008b), A'Hearn, Baten and Crayen (2006, 2008) and Baten, Voth and Crayen (2008), the ongoing work on age heaping has unearthed from archives numerous age distributions of populations as well as death registers. Beside their usefulness for the age heaping research, these data also open up new scope for research in other historical disciplines. To exemplify the dual purpose of the age data the present paper undertakes a specific demographic case study using Indian census samples and vital registration data and show how the obtained data can be deployed in historical demography for the estimation of life expectancies. My results suggest that the 19th century Mysore State in South India was characterized by levels of life expectancy higher than expected given its economic standing, providing supportive evidence for the thesis of a north-south divide of demographic regimes in colonial India.

Section II reviews the ongoing debate in the literature about the north-south gradient of life expectancy in India, and other related literature. Section III describes the underlying dataset and general methodology, while section IV identifies techniques to overcome possible biases. The estimation results are presented and discussed in section V. Section VI concludes.

II. Related Literature

Historical demography of India remains largely unexplored territory. Yet, especially in the developing world, few countries can match India's wealth of historical demographic data. The recent past has seen a growing interest in the field by scholars in economic history, historical demography and applied statisticians. Also, technical advances have created new possibilities for analyzing old data sets and have thereby pushed the frontier of knowledge creation. Better knowledge regarding India's population history not only has an intrinsic value but can also gain valuable insights for social and economic historians of India, and economists in general.

This paper aims to make a contribution to the academic dispute prevailing in India's historical demography. In the relevant literature, it has been accepted as a fact that there is a regional divide in demographic parameters between northern and southern provinces of post-

WW2 India. Whereas the north features high mortality and fertility rates, life expectation at birth and infant mortality tends to be lower in South India. The causes of the phenomenon are debated. One prominent explanation draws on the labour-market structure. Bardhan (1974) was among the first to link the economic value of women to the survival chances of girls. He suggested that the labour-intensive production of rice typical of southern India depended on female labour, leading to a higher social standing of the female population in general. In contrast, the mechanized cultivation of wheat in North India excluded women from production. In a similar vein, Rosenzweig and Schultz (1982) argued that intrahousehold resource allocation was motivated by the anticipated future contribution to household income. That is, in regions with higher female labour force participation rates, and where the potential contribution of female children to household income was relatively high, more household resources were allocated to girls. Another strain of literature focused on the on culture and kinship systems as determinants of the spatial structure in Indian demography. Sopher (1980) assessed the geographic differences in the context of traditional cultural regions and provided a detailed map of the cultural and linguistic variety on the Indian subcontinent. Dyson and Moore (1983) found a significant correlation between sex differential in infant mortality and marriage and kinship patterns as well as female autonomy. The authors argue that the higher women's decision making control, the better the survival chances of their children, and particularly of their daughters.

Saito, Takahama and Kaneko (2005) tried to trace back these demographic contrasts into the past as proposed by Dyson (1997) in relation to infant and child mortality. They identified 17 districts in Madras Province and 8 districts for Punjab (see Figure 1) for which they found the quality of historic pre-independence registration statistics to be sufficient to conduct detailed demographic analysis. The authors showed that even under British colonial rule both fertility and mortality rates were higher in Punjab than in Madras province, although economic conditions were generally more favorable in Punjab (see Collins 1999, Williamson 1998 on regional real wage levels).

Disagreement prevails over the hypothesis of Dyson and Moore (1983) for post-WW2 India and whether it can also be applied to the pre-WW2 period. They plot the boundary of the demographic divide in terms of post-war infant mortality as stretching from the Chota Nagpur Plateau in Eastern India to the Satpura Hills in Central India.[143] Dyson (1989b) estimated life expectancy at birth in pre-independence Berar (Berar was part of the Central Provinces, see Figure 1) to be higher than the all-Indian average. Due to its moderate levels of fertility and comparatively favorable mortality the author found his hypothesis confirmed that the central province of Berar, lying south of the Satpura Hills, belonged to the demographic scheme of south India. In contrast, Saito, Takahama and Kaneko (2005) assigned Berar to the demographic regime of northern India since Berar's life expectation and fertility regime was closer to Punjab's than to the one of Madras.

Table 1 gives an overview of regional and national life expectancy estimates for India. Several observations can be made. Firstly, Madras province consistently features very high life expectancy values while all other provinces so far analyzed are characterized by much lower levels. Secondly, while many studies come to the conclusion that life expectancy for both male and female declined from 1901-1910 to 1911-1920, it seems to have increased in Punjab and Bombay Presidency. Thirdly, Berar's level was close to Punjab's for the later period after levels in Punjab caught up substantially. While Berar's life expectancy values were above all-Indian averages in both periods, living conditions in Punjab used to be relatively unfavorable in the first period. Fourthly, Bombay, which lies further south than Berar, had comparably low life expectancy levels for both periods reviewed here. And lastly, while Clark (1989) also ascribed Madras the highest life expectancy values among the provinces for the beginning of the twentieth century, her estimates for the later period imply a substantial decrease in the number of expected years to live and rank Madras close to Punjab. It thus seems that a clear "north-south" divide of Indian demography is as yet not sufficiently supported by evidence. The present study seeks to provide new evidence for regional life expectancy levels in India to advance the research in this field.

Demographic patterns in India vary drastically not only over space but also over time: Dyson (1989a) maintains that demographic history in India can be divided into four broad phases. The first spans from the second half of the 18th to the beginning of the nineteenth centuries. Characterized by wars and social conflicts, population growth at that time was tiny if not negative. During the subsequent recovery period of the nineteenth century population growth picked up, but to an unknown magnitude. This phase has so far only been subject of regional analysis for Uttar Pradesh province and Madras presidency, leaving wide scope for further research. Dyson (1989a) estimated that the third phase began at around 1891, after which he found the population growth rate to slow down again due to an increase in epidemics and famines. The last and fourth phase presumably started in the 1920s, indicated by an increased control over disease and famines and economic development bringing about a substantial increase of population growth rates.

The comparatively slow increase in population numbers during the time assigned to the third phase by Dyson (1989a) has attracted academic interest and initiated a lively discussion if mortality really worsened during this period. This long-standing debate feeds from the often ideologically charged controversy whether the British rule entailed 'immiserization or progress' for India or not.[144] For instance, Davis (1951), Visaria and Visaria (1983) and Mari Bhat (1989) found increasing mortality trends for all-India in that period, a trend that Indian economic historians mostly ascribed to the high land revenue during British rule that forced many cultivators to grow on-food crops for the market, thereby increasing vulnerability to crop failure and the famine death tolls. McAlpin (1983), however,

[143] See also Libbee (1980) and Sopher (1980).

who studied famines in the Bombay Presidency 1860-1920 came to the conclusion that the colonial government merely enforced a socio-economic transformations towards a market economy, which contributed to improved standards of living, better nutrition, and a fall in mortality.

Analysis of Indian historical data can also contribute to an increased understanding of demographic characteristics and challenges of the newly industrializing countries today: Since the late 1960s several publications have taken a critical look at the idea that life expectancy gains are made exclusively possible by gains in economic development and national income (e.g., Kosa, Antonovsky and Zola 1969, Morris 1979). Caldwell (1986) emphasized the role of social justice or equality, meaning the equitable distribution of goods and services between different strata of society as well as between men and women. In the more recent past, life expectancy has further risen in most countries, including regions without equal advancements in per-capita-income. These gains have also occurred among countries not at all committed to equality such as Oman, which has been an autocracy for most of its modern history, or Jordan and Syria. At the same time, countries being then part of the Soviet Union have lost life expectancy in spite of relatively egalitarian income distributions and access to health facilities. Research on this paradox has received fresh stimulus by Riley (2005) who argued that all attempts to understand how some poor countries managed to achieve high life expectancy often need to embrace a historical perspective since health-related behaviour develops in a path-dependent ways over longer time spans. Whereas hitherto scholars mainly focused on policies being followed during the 1970s and 1980s to explain simultaneous gains in survivorship, Riley (2005) attracted notice to past circumstances to explain present phenomena. Though the analysis of determinants of life expectancy levels in South India is beyond the scope of the present paper, my study will add data to the described line of research

---

[144] See Clark (1989) for a more comprehensive overview.

by the analysis of spatial patterns in India and by the identification of potential future case studies.

III. Data and Methods

Indian historical demographic data stems from different sources of varying quality. Under the influence of British tradesmen and authorities India had introduced vital registration, i.e. the continuous registration of births and deaths on local level, as early as the second half of the nineteenth century. The earliest reported data give merely total numbers of deaths and births for the whole region, not disaggregated by district, age or sex. While the annual series of the local vital registration enables researchers to explicitly look at short-term fluctuations such as mortality crises following famines and epidemics, for most provinces the data quality of these statistics remains questionable. Dyson (1989b) reasoned that local arrangements for recording births and death appear to have been comparatively advanced at least in Berar, Punjab and Central Provinces. For most districts and provinces, however, the resulting birth and death rates are very defective as can be shown by comparison with census data (Thygarajaiyar 1923).

For work on the more recent historical demography of India the census is the single most valuable source. I analyze the age and sex distributions and the survivorship of the Indian population using single year age distributions given in the subsidiary tables of the 1901, 1911 and 1921 census reports. The data collection procedure was consistent between enumerations. Instructions given for recording age were the same in 1921 as in 1911 and 1901.[145] The census enumerators trusted that there was fair preparedness to answer questions among the population, given that there was ample notice that census inquiries would be made. However, we see extreme age heaping in the reported age data since many Indians did not

[145] Age at last birthday was to be entered. In the case of children under one year, 'infant' was to be written in the age column. If no statement of the interviewed person was possible, the enumerators were urged to ask further questions, for example the relations to refer to some well-known event of local importance, or to make a guess at the age from the appearance (Thygarajaiyar 1923).

know their exact age. Because of the irregularities in the single year age distributions census officials generally published age data in broad age groups only, making estimates of demographic parameters difficult.

This study is the first to use (adjusted) single year age distributions to derive life expectancy estimates for early twentieth century India. It employs representative samples of 100,000 men and women displayed in the subsidiary tables of the census reports of 1901, 1911 and 1921 originally created to illustrate the phenomenon of age heaping.[146] As a consequence, the life expectancy estimation procedure described in the present paper stands out from most other studies by the representativeness and the completeness of the data and by the thorough smoothing of the empirical age distributions, being possible due to its high degree of age data disaggregation.[147]

Estimation of life expectancy using successive census age distributions

Rooted in the theoretical model of Bennett and Horiuchi (1981), the method proposed by Preston and Bennett (1983) draws on the information contained in the change in size of successive age cohorts of population from one census to the next. It compares age groups to their equivalents across census years, and computes intercensal growth rates for each age group: In a closed population with two accurate censuses t years apart, the population aged $x+t$ at the second census represents the survivors of the population aged x at the first census, so that the intercensal survivorship probability from age x to age $x+t$ can be calculated. Traditional mortality estimates can then be obtained from the sequence of survivorship probabilities for successive initial ages x using the sum of the intercensal growth rates to each age, and the mid-decade population of each age (UN, 1983). However, the requirements that the population be closed to migration and the enumerations accurate are rarely met. The census-based method, like most other demographic techniques, is thus subject to biases

[146] The relative age distribution was transformed into absolute figures using sex-specific population totals.

induced by varying degrees of data quality, age misreporting, and migration. Moreover, changing data collection methods and geographical coverage of the enumeration needs to be assessed. These issues will be discussed in this section further below and in section IV.

It lies in the nature of this method that life expectancy for the youngest and the oldest age groups cannot be derived, since they lack the counterparts in the previous or following census enumeration. To compensate for the lack of cohort data, I resort to age-and sex-specific death rates collected from the Indian vital registration system to derive survivorship ratios as suggested by Preston, Heuveline and Guillot (2000). As explained above, total figures derived from the local registration offices are considered as inaccurate, and mortality levels are likely to be downward biased. Since I report no single estimates but band widths of life expectancy, a broad picture of the age-specific mortality pattern by sex can still be deduced as long as the underreporting of events of death does not systematically vary with age. The age- and sex-specific death rates are summarized in period-specific life tables from which I can calculate the probability of surviving to each age or, in other words, the probability of dying before the next birthday.[148] These life tables are then used to complete life expectancy estimates for the very young and old ages.

Regional selection

My choice of Mysore State for the present study stems from four main considerations. First, although there was considerable district swapping, boundary changing, and territorial acquisition among the Indian presidencies, Mysore State experienced no significant boundary changes for the period 1901-1921 reviewed here, which would complicate analysis. The second reason for selecting Mysore bases in its location in South India. The plateau of Mysore is situated like a wedge between two mountain ranges, the Eastern and Western Ghats, and the

[147] For his estimates of life expectancy for all-India, Mari Bhat (1989) also devised a complex method to adjust for age misreporting.

[148] For the period life table construction, I follow Preston, Heuveline and Guillot (2001, pp.48-51).

Nilgiris where the Ghats converge. The general elevation rises from about 2000 feet above sea level along the north and south frontiers to about 3000 feet at the central watershed which separates the basin of the Krishna to the north from that of the Cauvery to the south. Due to its elevation above sea, the climate of Mysore is temperate, in spite of its situation within the tropics. As a consequence, health conditions are considered generally favorable during the whole year. Mysore State covered an area of about 29,475 square miles, at that time being on par with the size of Scotland or of Bavaria. What makes its case so interesting is the fact that the province is surrounded on all sides by the Madras Presidency except on the north and north-west where it's neighbours are the districts of Dharwar and North Kanara respectively of the Bombay Presidency and towards the south-west where Coorg adjoins it. In other words, it lies between Bombay Presidency that can be assigned to the "North Indian Demographic Regime", given the life expectancy estimates by Clark (1989), and Madras Presidency, which Saito, Takahama and Kaneko (2005) instanced of the "South Indian Demographic Regime".[149] Thirdly, featuring 224 inhabitants per square mile, Mysore ranked 10th among the 22 states in terms of population density. Similarly, it ranked 11th in terms of population size. Given that changes in population growth rates, proportions urban and exposure to diseases and famine prevalent in Mysore State mirrored broad trends that characterize demographic patterns of most regions in India, its socio-economic environment can be assumed reasonably representative for India (particularly South India) and thus the results more generally valid. Last but not least, availability of both single year age data and death records allowed for reasonably reliable life expectancy estimates.

I chose the State as unit of analysis (as opposed to districts or natural divisions) because migration data was available at this level of aggregation, and because inner-state migration between districts was probably more common than inter-state migration, increasing

[149] For an overview of the different regional and national studies on early twentieth century Indian life expectancy see Table 1.

the magnitude of the migration-induced bias of district-level analysis. Nevertheless, for the mortality estimation methodology used here, prior estimation of intercensal migration is crucial.

IV. Possible biases and adjustment procedures

Section III above listed several potential sources of error in the data we use. As explained in the same section, the problems of changing standards and geographical coverage of the enumeration do not apply for Mysore State 1901-1921 and will thus not be dealt with any further. The issue of over- or underenumeration of the population, in form of changing degrees of enumeration completeness from one census to the other, can submerge the effect of mortality and thus give rise to misleading results (UN, 1983). An advantage for our analysis (but less positive for the responsible census officers at that time), is the fact that there is no indication of significant improvement in the age returns between the Mysore State census enumerations of 1901, 1911 and 1921. One possible explanation is the fact that British authorities started census taking in India already in the 1870s, so that procedures had been standardized already in 1901 (Thygarajaiyar 1923). Given the fact that instructions and training of the enumerators did not change until 1931, when the age question was slightly changed, it appears a reasonable assumption that data quality more or less stagnated during the period under study. In contrast, age misreporting, leading to extreme heaping of ages, was prevalent and constitute a severe data problem. Likewise, our data needs to be adjusted for migration, since positive (negative) net migration inflates (reduces) cohort sizes, in a way that varies over time and by age group, and thus induces upward (downward) distortions in survivorship ratios. I will also discuss the problem of underreporting of events of death in the vital registration system underlying the mortality estimates for very young and old ages.

Age misreporting

To most people living in the distant past, the exact number of years they have lived was of little or no practical importance and therefore, a considerable amount of ignorance prevailed on the subject. Age misreporting can distort the results severely since a strong preference for certain ages when declaring age will introduce considerable variability into the estimated survivorship ratios. As people are systematically assigned to the wrong age cohort (by declaring a wrong age), cohort sizes are erroneous and can gravely bias life expectancy estimates. To overcome this problem, we use Sprague's formula for osculatory graduation that smoothes the single year age distribution (for a detailed description of the method, please refer to Appendix A1).

Whereas general age heaping can be adjusted for, we lack the information to correct for possible asymmetries in age misreporting, that is, if more people make themselves older than younger, particularly the age exaggeration at old ages that Mari Bhat (1995) identified as a potential source of error. He found that South Indian adult males have a greater propensity than women to overstate their age. This is why results indicating higher male than female life expectation must be considered with caution. Nevertheless, it is unlikely that the bias in male life expectancy and sex ratios at high ages induced by age overstatement varies across regions and therefore should impact this study only mildly, if at all. Preston, Elo and Steward (1998) argued that among the very old more individuals inappropriately move upwards into an age interval than who move downwards out of it because the age distribution peters out rapidly in this segment. Thus, the steeper the age distribution at old ages, the more likely numbers of old people are inflated. We try to mitigate this effect by summarizing all age above 60 years. In conclusion, while this study addresses the most important issue of symmetric age heaping, the results might be slightly upward biased, since mortality estimates at old ages, especially for males, might be biased downwards by asymmetric age misstatements.

Migration

There are no direct statistics of migrations of Indian people in the past, and they have to be inferred only from birth places of the enumerated population. There are three sets of inputs into the migration estimation procedure to adjust the census age distribution before mortality rates are calculated: the total number of foreign population (i.e., not born in Mysore) given in the census, and the number of people born in Mysore and enumerated in other Indian states, which Zachariah (1964) used for his estimates of sex-specific migration figures.[150] Furthermore, the census report of 1921 provides a rough age distribution of immigrants to the Bangalore Gold Fields in Mysore, on the basis of which I will construct age-specific migration adjustment factors.

Since the economic and social gains from movement are generally highest at young and middle adult ages, and men are generally more prone to migrate than women (UN 2004), the sex ratios in Table 2 offer a first indication that Mysore received more immigrants than it lost emigrants. Interestingly, sex ratios for the young adult ages below 30 are close to one, suggesting that either migration did not affect these ages as much as older adult ages or that in these age groups as much women as men migrated. For the age groups above 30, while there are indications that life expectancy for the female sex was more favorably than for men for the middle-aged population (see section V), working-age men still outnumbered women in the original data of Mysore collected for the present study. As Figure 2 illustrates, this presumption is corroborated by census figures underlying Zachariah's (1964) calculations that yield more immigrants than emigrants for Mysore, as well as more men than women immigrating. His estimates figure approximately the same amount of men and women born in Mysore but enumerated in a different province (depending on the census year 45,000-60,000

[150] As it was a custom of the Hindu people for the females to go to their mother's house, or a sister's house for confinement, particularly for the first two confinements, there must be a large number of people who list a place of birth different from where they are permanently domiciled although they cannot be counted as regular immigrants. Intermarriages between families in different provinces are rare, however, and not systematic into one direction.

each) whereas close to 190,000 men migrated to Mysore State, as opposed to 140,000 women. That is, the positive net-migration at a volume of about 170,000-220,000 (or roughly 3-4% of the population) pushes up the sex ratio in Mysore State. The roughly constant migration numbers in Figure 2 also suggests that the overall bias of migration was rather modest, unless migration patterns varied extremely by age group.

In the absence of statistics to show the age-distribution of migrants, it is not possible to frame any accurate measure of the disturbing influence of migration on the age distribution. An approximation of the migration effect can still be obtained, however, thanks to the census officials of the 1911 enumeration who published figures on immigrants in Bangalore City and the Kolar Gold Fields as can be seen in Table 3.[151] Thereby more insights can be gained about what part of the age distribution is affected most by immigration and in what direction it influences our estimates. Moreover, I can use the percentages displayed in Table 3 in combination with Zachariah's (1964) estimates to derive adjustment factors for migration-inflated cohort sizes. Apparently the bulk of foreign-born belongs to the age group 15-40, when the economic gains from immigration are largest. Hence, where a large number of people left an area for working outside, the chances are that the majority of them were of working years. Given that the numbers of immigrants rose slightly over from 1901 to 1911 more than the overall population, it can be assumed that the numbers of the middle-age group given in the following census are inflated and life expectancy estimates therefore biased upwards. For the following analysis, the relative age distribution of immigrants shown in Table 3 was multiplied by the net-immigration numbers to obtain the figures to be subtracted from the totals of the broad age ranges 0-14, 15-39 and 40 + years.[152]

Within these broad age ranges, subtraction was done proportionally to the relative share of the more differentiated age cohorts. Figure 3 shows that the adjusted population

---

[151] The term 'immigrant' refers to people enumerated in Mysore but born in a different Indian state.
[152] This was done by sex and by census year.

numbers are marginally lower than the originals for the age groups up to 40, after which the migration effects vanishes. Estimation procedures conducted by the author that did not take account of migration effects produced life expectancy estimates that were 2-3 years different from the estimates based on age distributions that were adjusted for this effect.

Underreporting of events of death

Since the census based method by Preston and Bennett (1983) cannot produce life expectancy estimates for the very young and very old cohorts, mortality estimates must be added in order to extrapolate survivorship ratios for the open age intervals. For this purpose, demographic studies often make use of standardized life tables that summarize the demographic history of more or less homogeneous groups of states, for which data is available. As such, Coale and Demeny (1983) have reviewed the relation between mortality levels and mortality age patterns in many populations and have fit a line to the international and intertemporal data. The popular Model West tables, for instance, are mainly based on the mortality experience of Western European populations, but also on the experience of Taiwan and Japan.

There are two reasons why the present analysis draws on deficient mortality data from the Indian vital registration system instead of relying on commonly used Coale-Demeny Model West life tables. Firstly, this study follows Mari Bhat (1989) who estimated trends in mortality and fertility for all-India using modern census data. He argued that India differed from most demographic models in the way that it is characterized by relatively heavy adult mortality during the early decades of the 20th century, and not by a regime with extremely high mortality at young ages as commonly found in more recent demographic data of developing countries.

Secondly, Clark (1989) suggested that the age pattern of Indian mortality has been different in bad years, compared to good years. In times of crisis, such as the two famines in 1896/97 and 1899-1901, she argued, very young children who were protected by breast-feeding, were not as much affected as older children and adults. This implies that the increase in adult mortality must have been much more pronounced than the rise in infant mortality. By using Indian vital registration data, we can capture India-specific mortality crises.

The reported birth and death-rates of the Indian vital registration records are extremely low. Thyagarajaiyar (1923) assessed the extent to which omissions occur by comparing the survivors of those born 1911-1921 according to the vital registration records with the age cohort 0-10 in the 1921 census tables. He came to the conclusion that underreporting happened very largely as the census reported almost half a million more survivors than the total number of births given in the vital registration records during that decade would suggest, a magnitude that cannot nearly be explained by immigration. A similar exercise for vital records 1921-1930 suggested that in the more backward rural parts of Mysore only 50 per cent of deaths and an even smaller proportion of births were reported (Iyengar 1931). Hence, the data quality of vital registration numbers is questionable, and the mortality pattern based on reported death and birth numbers must be interpreted with caution. In the following paragraphs, I will first discuss mortality trends and possible biases before switching to mortality levels.

From Figure 4 it can be inferred that according to the vital registration system, average mortality increased from 1901-1910 to the following decade by about 10 per cent for men and 16 per cent for women. In both decades, mortality was higher among men than among women. To test if the increase in the reported death rate was merely due to improvements in the registration system, i.e. if it was a common upward trend (i.e., a continuous increase in death rates), we examine annual death rates for selected years, for which we were able to gather death figures.

As illustrated in Figure 5, the higher death rate visible for the decade 1911-1920 is not a consequence of continuous improvement in death reporting but it was produced solely by extremely high mortality in 1918, the year of a devastating influenza crisis. In fact, mortality rates tripled in 1918 after a first small increase in reported deaths in 1917. As death figures plunged again to 1913 levels in the following year, the rise in mortality is almost certainly driven by non-statistical factors only. We can thus be reasonably sure that the change in mortality during the early twentieth century was caused by the 1918 influenza and not by improving data quality of the Mysore vital registration system. The degree of underreporting and its impact on mortality levels, however, is likely to be substantial. Since we lack more detailed information to derive more sophisticated controls, the present paper reports an upper and a lower bound of life expectancy instead of giving a single "best" estimate.

As upper bound of life expectancy, we will report the estimate based on the original (and deficient) death registration data, which most probably underestimates mortality and therefore amplifies life expectancy values. As a lower bound, life expectancy is estimated based on the pessimistic assumption that 50% of the events of death are not reported (see Iyengar 1931). For both females and males, the upper and lower bound estimates differ by more than 6 years, reflecting the sensitivity of demographic estimates to low data quality.

V. Results

A summary of my new estimates of life expectancy at birth is given in Table 4.[153] It is impossible to quantify the findings accurately, since the upper and lower bound estimates differ greatly, by between 6.3 and 7.3 years. But several conclusion can still be drawn from the table. For both the male and the female population, living conditions seem to have worsened during the period under study. The Spanish influenza crisis of 1918 that spread through India from the Bombay area had severe effects on India's demography although it lasted a couple of months only. Killing about 7 million people in India alone (about 3% of the population), with roughly four times more people suffering from the disease at some stage, it had devastating direct and indirect effects on Indian death rates. While mortality rose

[153] For intermediate steps involved in the Preston and Bennett (1983) and Preston et al. (2001) method, please refer to Appendix A2.

substantially in North and West India, its impact was milder in eastern and southern India. Mills (1986) found that life expectation at birth in the Bombay presidency had fallen as low as six years. Against this background, the drop in life expectancy in Mysore by 2.5-3.3 years for women seems modest, while the average male suffered graver deteriorations of their living conditions.

The findings indicate that, compared to other Indian provinces, Mysore enjoyed a relatively high level of live expectancy. With the “true figures” of life expectancy probably lying somewhere in the middle of the band width, life expectancy in Mysore exceeded the levels estimated for Madras Presidency by Clark (1989) and comes close to the Madras figures proposed by Saito, Takahama and Kaneko (2005). That is, in both periods Mysore scores life expectancy values better than the Indian average. Notwithstanding, male life expectancy might have fallen to lower levels than in Punjab during 1911-1920.

Taking a closer look at the age pattern of life expectancy (approximating the remaining years to live at a given age), more observations can be made. Figure 6 and Figure 7 show how life expectancy changed with age for each sex in both periods.[154] For all cases, we see the distinctive increase in life expectancy after early childhood, during which the susceptibility to diseases is typically high (and resistance particularly low). All curves show a reduction in life expectancy for the teenage years, followed by a slight improvement during young adulthood and a adjacent continuous fall in remaining years to live for the average Mysore resident.

A male person living in Mysore during the early twentieth century who has survived to the age of 55 can thus expect to live another 15 years. The influenza-induced rise in mortality in the period of 1911-1920 impacted on young male age groups most strongly while men over 50 were more or less unaffected and had about the same life expectancy as the decade before. The picture is different for the female population that did not suffer as severely as their

[154] For illustrative purposes, the lower bound estimates are shown only. Except for the very first age groups, the age pattern is practically identical to the upper bound estimates.

brothers and fathers: The deterioration of life expectancy among the women of Mysore pertained to all age groups but young girls.

Figure 8 plots female mortality rates by age group before, during and after the Influenza crisis. In the year of the pandemic, death rates for young adults (starting in the late teenage years) up to adults in their forties spiked, as well as for infants and the very old. This W-shaped mortality age profile (death rates having a mode in the middle - age group) is unusal for epidemics, which often have a U-shaped profile, but typical for the Great Influenza mortality crisis (Nicholls 2006). Phillips and Killingray (2003) provide findings from laboratory research deciphering the genetic code of the Influenza virus, explaining why the Great Influenza was generally particularly lethal to young adults aged 20-40. The age-mortality picture is similar for the male population, however, the rise in mortality for young male adults is not as sharp. Since the decadal mortality rate is used for the life expectancy estimates, the influence of the Great Influenza on young adult men is partly averaged out. Another explanation for the age pattern in influenza-induced mortality can be found in the same paper. Phillips and Killingray (2003) suggest that regions and individuals that are linked to a reasonably advanced transportation system were more exposed to the first mild strain of the virus so that the immune system were better prepared when the pandemic stroke with full force. Though this might be one possible explanation to why life expectancy did not decrease as much for relatively mobile middle-aged men, it cannot explain why women were less affected than their male peers.

Noymer and Garenne (2000) are two of the few demographers studying the sex differential in mortality during the pandemic. They convincingly argued that in the United States, the 1918 influenza had a strong and long-lasting effect on differential mortality by sex, caused by the higher incidence of tuberculosis in men than in women: The mechanism they identified is a selection effect, whereby those with tuberculosis in 1918 were more likely than others to die of influenza. This outcome led to sex differentials in mortality since tuberculosis morbidity was disproportionately male. Moreover, the authors reason that the narrowing of mortality differentials by sex in the United States after the Great Influenza was induced by the deaths of fragile individuals during the pandemic, being to a larger share men than women, who otherwise might have died from tuberculosis. Herring and Sattenspiel (1998) examine individual-level travel patterns and trade in Canada to study the spread and subsequent individual-level mortality during the Influenza pandemic. They highlight the correlations between the social organization of everyday life and mortality at the level of families and filiations in communities struck by the 1918 pandemic, which could also help to explain the different degree of affliction of male and female population in Mysore.

There are some noticeable features with respect to gender differences in life expectancy. This inequality can be observed most easily in the ratio of male to female life expectancy (Figure 9). For the first period under study, the graph reveals a ratio that is constantly above 1, implying higher life expectancy values for men than for women at all ages. A systematically unequal intra-household allocation of health-care and nutrition at the expense of female children has long since been identified as the underlying cause of this development. Although this practice is and was more widespread in the northern part of India than in the South (Drèze and Sen 1995, Clark 1989), its effects are still visible in the southern state of Mysore. Figure 9 also shows that for young to middle-aged adult women around 1901 the conditions were peculiarly harsh. These are typically the age groups, in which young women start entering married life and start bearing children, entailing female mortality increases relative to male mortality rates in the absence of adequate medical facilities. Women who have passed the age of bearing children, which is associated with relatively high female mortality rates, also stood life better causing the gender inequality of life expectancy to fall again. In contrast, the inequality pattern emerging for the later period displays ratios below 1 only, indicating that the average man passing through life constantly had lower life expectancy than the average woman. As discussed in the section above, this probably largely reflect the mortality disadvantage that males suffered during the Great Influenza pandemic in 1918.

VI. Conclusions

The paper has demonstrated the dual purpose of single year age data collected for the age heaping research, contributing to academic disputes in other fields of historic research. This study combines census, complemented where necessary by vital registration data, to derive life expectancy estimates for the Indian state of Mysore. The estimation procedure involves techniques to correct for age heaping, migration, and underreporting of cases of death. It presents tentative evidence that life expectancy levels in Mysore during the early twentieth century were relatively high compared to other Indian regions. Life expectancy in Mysore fell for both the female and male population from 1901-10 to 1911-20. This worsening of living conditions, however, can most likely be attributed to the devastating effects of the Influenza pandemic in 1918. The estimates also indicate that Mysore men suffered worse from the Great Influenza than women, a phenomenon that has been observed for other countries, too (Noymer and Garenne 2000).

This study thus adds support for the thesis by Saito, Takahama and Kaneko (2005) that relatively favorable demographic conditions in South India were not just a phenomenon of the Madras presidency alone. This is especially intriguing, given that the average real wages were relatively low in Madras, Mysore and the Central Provinces (including Berar) compared to Punjab and Bombay in the early twentieth century (Collins 1999, Williamson 1998). Additional evidence is presented by Williamson (1998) who found that around 1900, real wages in the northern and the southern regions of India were comparatively low, contrasted by higher wages in the east and the west of India. Mysore province is hence an interesting case study for research on high life expectancy levels in comparably poor countries (Riley 2005).

Similar to Clark (1989) who calculated demographic measures for early twentieth century Madras, Uttar Pradesh and Bombay Presidency, I find that sex differentials in life expectancy (whatever their origins) were not persistent. Due to the short period under study, these results are tentative only but they suggest that the social status of women is probably less important for India's historic demographic trends than the economic and epidemiological context.

References

A'Hearn, Brian, Joerg Baten and Dorothee Crayen (2006). Quantifying quantitative literacy: Age heaping and the history of human capital. Universitat Pompeu Fabra Economics Working Paper Series 996.

A'Hearn, Brian, Joerg Baten and Dorothee Crayen (2008). Quantifying quantitative literacy: Age heaping and the history of human capital (revised version). Working Paper Franklin and Marshall College/ University of Tuebingen.

Banthia, Jayant Kumar (2004). India administrative atlas 1872-2001 : a historical perspective of evolution of districts and states in India. New Delhi: The Government Press.

Bardhan, Pranab (1974). On Life and Death Questions. Economic and Political Weekly, 10: 32-34.

Baten, Joerg, Dorothee Crayen and Hans-Joachim Voth (2008. Poor, hungry and stupid: numeracy and the impact of high food prices in industrializing Britain, 1790-1840. Working Paper Universitat Pompeu Fabra /University of Tuebingen.

Bennett, N. and S. Horiuchi (1981). Estimating the completeness of death registration in a closed population. Population Studies 47, pp. 207-21.

Caldwell, J C. (1986). Route to Low Mortality in Poor Countries. Population and Development Review, 12(2).

Clark, Alice (1989). Mortality, fertility, and the status of women in India, 1881-1931. In Dyson (1989a).

Coale, Ansley and Paul Demeny (1983), Regional Model Life Tables and Stable Population, Princeton: Princeton University Press.

Collins, W. *(*1999). Labor mobility, market integration, and wage convergence in late 19th century India. Explorations in Economic History 36 3, pp. 246-277.

Crayen, Dorothee and Joerg Baten (2008a). Global trends in numeracy 1820-1949 and its implications for long-run growth. Working paper University of Tuebingen.

Crayen, Dorothee and Joerg Baten (2008b). New evidence and new methods to measure human capital inequality before and during the Industrial Revolution: France and the U.S. in the 17th to 19th centuries (revised version). Working paper University of Tuebingen.

Davis, K. (1951). The population of India and Pakistan. New Jersey: Princeton University Press.

Drèze Jean and Amartya Sen (1995). India: Economic Development and Social Opportunity. Oxford University Press.

Dyson, Tim and Mick Moore (1983). On Kinship Structure, female Autonomy, and demographic behaviour in India. Population and Development Review, 9:1.

Dyson, Tim ed. (1989a). India’s Historical Demography: Studies in Famine, Disease and Society. London: Curzon Press.

Dyson, Tim (1989b). The Historical Demography of Berar 1881-1980. In Dyson (1989a).

Dyson, Tim (1997). Infant and Child mortality in the Indian subcontinent. In A. Bideau, B. Desjardins and H. Brignoli, (eds.): Infant and Child Mortality in the Past. Oxford: Oxford University Press.

Herring, Ann and Lisa Sattenspiel (1998). Structured epidemic models and the spread of influenza in the Norway House District of Manitoba, Canada. Human Biology 70, pp.91-115.

Iyengar, Venkatesa (1931). Census of India, 1931 Volume XXV MYSORE – Report,. Bangalore: The Government Press.

Keyfitz, Nathan (1977). Applied mathematical demography. New York: Wiley.

Klein, Ira (1973). Death in India, 1871-1921. The Journal of Asian Studies, 32:4, pp. 639-59.

Kosa, John, Aaron Antonovsky and Irving Zola (1969). Poverty and Health. Cambridge Mass.: Harvard University Press.

Libbee, M.J. (1980). Territorial endogamy and the spatial structure of marriage in rural India, in D.E. Sopher (ed.), An Exploration of India: geographic perspectives on society and Culture. Cornell University Press: Ithaca.

Mari Bhat, P.N. (1989). Mortality and fertility in India, 1881-1961: a reassessment. In Dyson (1989a).

Mari Bhat, P.N. (1995). Age misreporting and its impact on adult mortality estimates in South Asia’, Demography India 24, pp. 59-80.

Mayer, Peter (1999). India's Falling Sex Ratios. Population and Development Review, 25:2, pp. 323-43.

McAlpin, Michelle (1983). Subject to famine: Food Crisis and Economic Change in Western India 1860-1920. Princeton: Princeton University Press.

Mills, I.D. (1986). The 1918-1919 Influenza Pandemic-the Indian Experience, Indian Economic and Social History Review 23, pp.1-40.

Morris, David Morris (1979). Measuring the condition of the world's poor: the physical quality of life index; Physical quality of life index. Pergamon Policies Studies, no. 42. New York: Pergamon Press.

Nicholls, Henry (2006). Pandemic Influenza: The Inside Story. PLoS Biology 4(2), p. e50.

Noymer, Andrew and Michel Garenne (2000), The 1918 Influenza Epidemic’s Effects on Sex Differentials in Mortality in the United States, Population and Development Review, 26, pp. 565-581.

Phillips, Howard and David Killingray (2003). The Spanish Influenza Pandemic of 1918-19: New Perspectives. London and New York: Routledge.

Preston, Samuel and Neil Bennett (1983). A census-based method for estimating adult mortality. Population Studies 37, pp. 91-104.

Preston, Samuel, Irma Elo and Quincy Stewart (1998). Effects of Age Misreporting on Mortality Estimates at Older Ages. PARC Working Paper Series WPS 98-01.

Preston, Samuel, Patrick Heuveline, and Michel Guillot (2000). Demography. Measuring and Modelling Population Processes. Oxford: Blackwell.

Rice, G.W. (1990). Australia and New Zealand in the 1918-19 influenza pandemic, Occasional Papers in Medical History Australia 4, pp. 67-74.

Riley, James (2005). Poverty and Life Expectancy. The Jamaica Paradox. Bloomington: Indiana University.

Rosenzweig, Mark and Paul Schultz (1982). Market opportunities, genetic endowments, and intrafamily resource distribution: child survival in rural India. American Economic Review, 72:4.

Saito, Osamu, Mihoko Takahama and Ryuichi Kaneko (2005). Contrasts in Vital rates: Madras and Punjab in the Colonial Period. Hitosubashi University Research Unit: Discussion paper series No.68.

Sen, Amartya (1999). Development as freedom. New York: Knopf.
Shryock, Henry and Jacob Siegel (1973). The methods and materials of demography. New York: Academic Press.

Sopher, David (1980). The geographic patterning of culture in India, in D.E. Sopher (ed.), An Exploration of India: Geographical perspectives on Society and Culture, Cornell University Press: Ithaca.

Thygarajaiyar, V.R. (1923). Census of India, 1921 Vol-XXIII MYSORE - Report. Bangalore: The Government Press.

United Nations (UN, 1983). Manual X - Indirect techniques for demographic estimation. Population Studies 81. New York: United Nations Publication.

United Nations (UN, 2004). World Economic and Social Survey 2004: International migration trends. New York: United Nations Publication.

Visaria, Leela and Pravin Visaria (1982) : Population (1757-1947). In: Dharma Kumar, Maghnad Desai (eds.): The Cambridge economic history of India. Vol 2: c. 1757 - c. 1970. Cambridge: Cambridge University Press.

Williamson, Jeffrey (1998). Real Wages and Relative Factor Prices in the Third World 1820-1940: Asia, Harvard Institute of Economic Research Discussion Paper No. 1844, Cambridge (Mass.): Harvard University.

Zachariah, K.C. (1964). A Historical Study of Internal Migration in the Indian Sub-Continent, 1901-1931. Bombay: Asia Publishing House.

Tables

Table 1: Estimates of life expectancy in India

Study	Region	Sex	1901-1910	1911-1920
Saito, Takahama and Kaneko (2005)	Madras - 5 distr.	M	36.9	36.9
		F	37.7	37.3
	Madras - 12 distr.	M	31.5	30.8
		F	32.7	31.6
	Punjab	M	19. 2	24. 4
		F	16.3	22.4
Dyson (1989b)	Berar	M	26.4	24.0
		F	27.6	26.2
Clark (1989)	Madras	M	29.0	25.1
		F	26.7	22.5
	Uttar Pradesh	M	18.1	17.2
		F	19.6	17.6
	Bombay Presidency	M	20.6	21.1
		F	19.0	20.1
Mayer (1999)	India	M	22.5	19.2
		F	23.4	21.3
Das Gupta (1971)	India	M	22.9	23.0
		F	24.1	23.3
Mari Bhat (1989)	India	M	25.3	21.8
		F	25.5	22.0
Hardy, cited by Ira Klein (1971)	India	M,F	22.9	20.1

Table 2: Sex ratios (males to females) in Mysore by age group, 1901-1921

age	1901	1911	1921
0-10	0.92	0.94	0.94
10-19	0.97	1.09	1.05
20-29	1.03	0.99	0.89
30-39	1.11	1.07	1.04
40-49	1.25	1.20	1.06
50-59	1.26	1.13	1.07
60 and over	1.01	0.99	1.19

Source: Own calculations based on Census of India, 1901-1921, MYSORE-Reports: Subsidiary tables.

Table 3: Relative age distribution of locals vs. immigrants in Bangalore and Gold Fields

Area	0-14	15-39	40 +	sum
Whole state	37.4	39.2	23.4	100
Bangalore and Gold Fields, excluding immigrants	43.2	36.5	20.3	100
Bangalore and Gold Fields, including immigrants	33.2	47.4	19.4	100
Immigrants in Bangalore and Gold Fields	23.0	58.5	18.5	100

Source: Census of India, 1911- Report.

Table 4: Estimates of life expectancy at birth for Mysore 1901-1910 and 1911-1920

	1901-1910		1911-1920	
	upper bound	lower bound	upper bound	lower bound
female	35.6	29.2	33.1	25.9
male	37.4	30.1	29.6	22.3

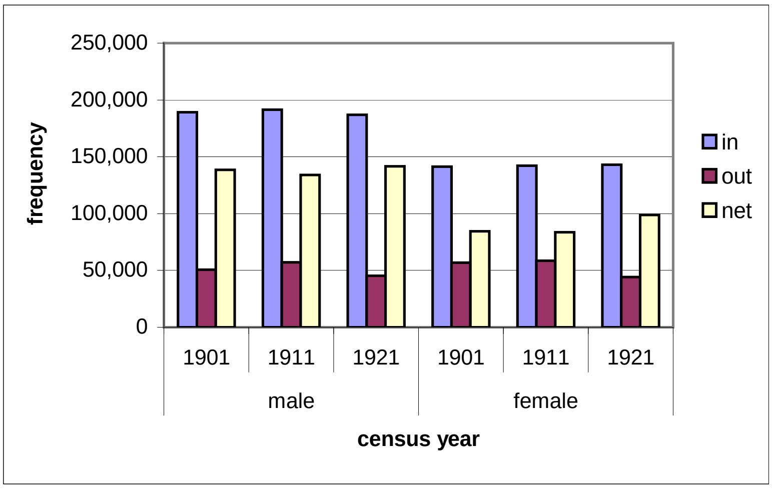
Sources: Own calculations based on subsidiary tables in the census reports of 1901, 1911 and 1921.

Figures

Figure 1: Indian Provinces under British Rule

Source: Based on Banthia (2004). The red line represents the approximate location of the Satpura hill range, the red circle the approximate location of the Chota Nagpura plateau.

Figure 2: Migration to and from Mysore

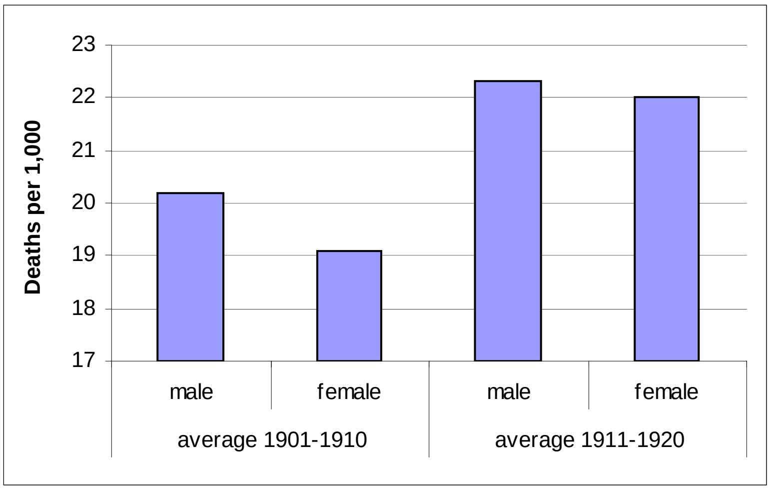


Source: Zachariah (1964).

Figure 3: Mysore male age distribution in 1911 adjusted and not adjusted for migration.

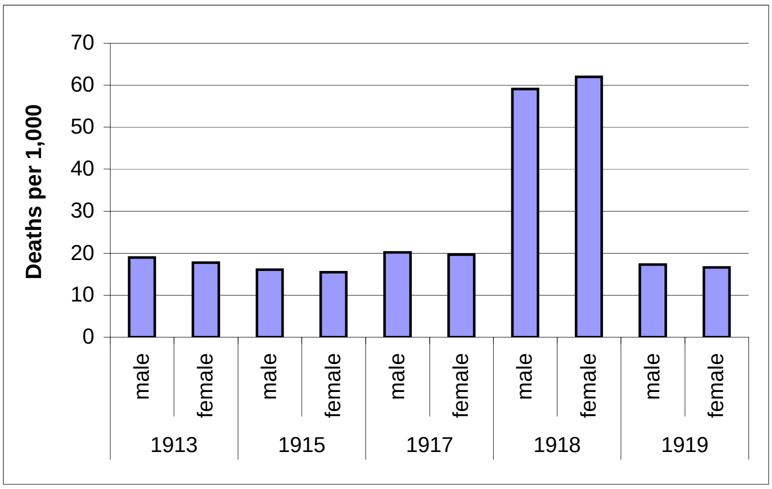
Source: Own calculations based on Census of India, 1911, MYSORE-Reports: Subsidiary tables.

Figure 4: Death rate in Mysore 1901-1910 and 1911-1920



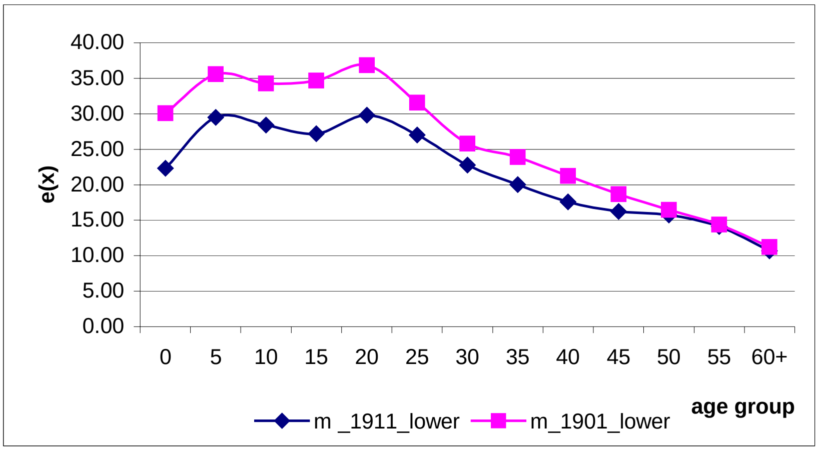
Source: Vital registration data as reported in Census of India, 1911- Report and Census of India, 1921- Report, Subsidiary tables.

Figure 5: Death rate in Mysore for selected years



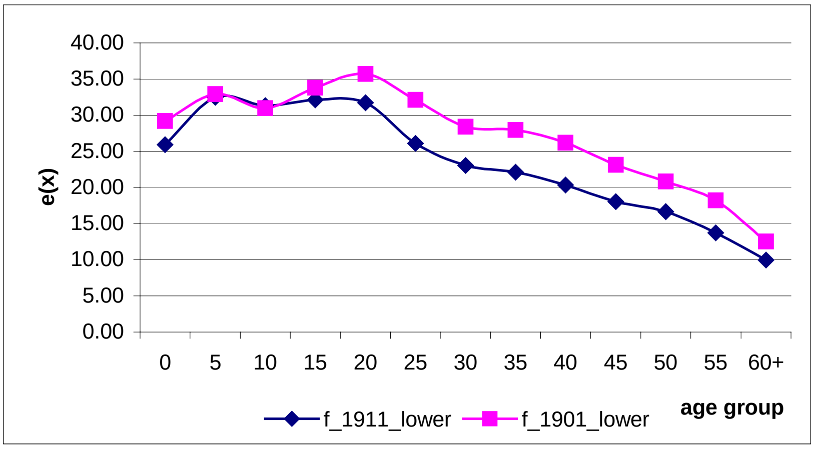
Source: Vital registration data as reported in Census of India, 1921- Report, Subsidiary tables.

Figure 6: Lower bound estimates of male life expectancy at birth for Mysore 1901-1910 and 1911-1920



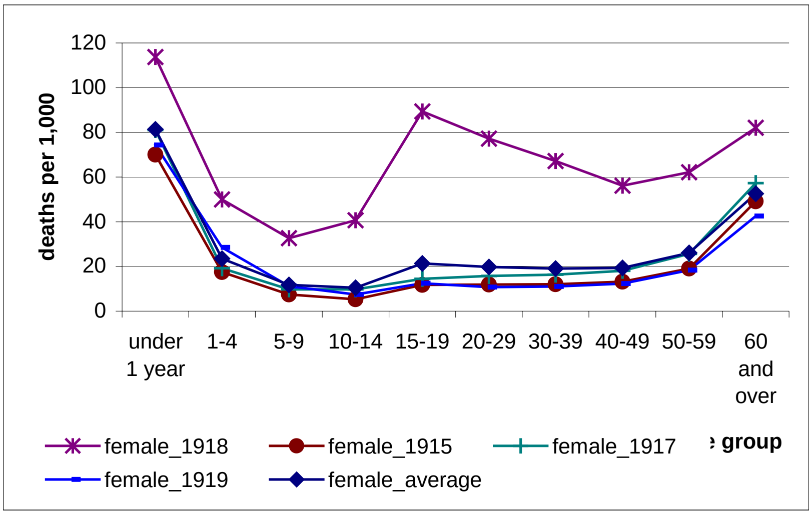
Source: Own calculations based on Census of India, 1901-1921, MYSORE-Reports: Subsidiary tables.

Figure 7: Lower bound estimates of female life expectancy at birth for Mysore 1901-1910 and 1911-1920



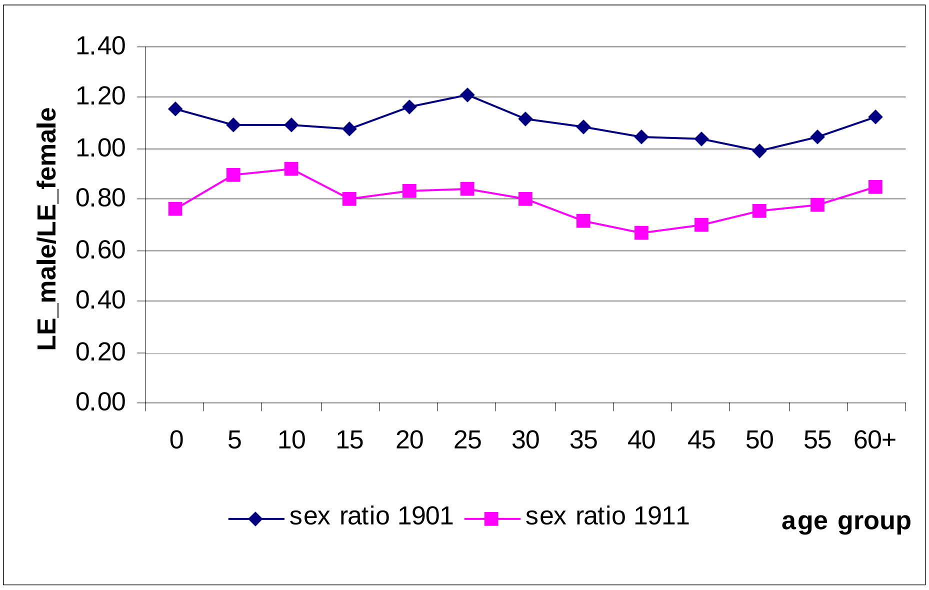
Source: Own calculations based on Census of India, 1901-1921, MYSORE-Reports: Subsidiary tables.

Figure 8: Mortality crisis in Mysore in 1918



Source: Vital registration data as reported in Census of India, 1921- Report: Subsidiary tables.

Figure 9: Gender Inequality in Mysore State 1901 and 1911: Ratio of male to female life expectancy



Source: Own calculations based on Census of India, 1901-1921, MYSORE-Reports: Subsidiary tables.

Appendix

A1. Osculatory interpolation

In order to calculate unbiased estimates of demographic parameters it is necessary to de-heap the reported age distribution to arrive at the "true" age data. In general, the data given in single years of age is combined into data for 5-year age groups, which are then interpolated to obtain smoothed estimates of data by single years of age. Polynomial interpolation is a method of graduation where the age group is assumed to conform to an equation of the general type $y=a+bx+cx^2+dx^3+\ldots+nx^n$. For each age group one polynomial is fitted. A problem that might arise when fitting these polynomial interpolation formulas to adjust rough data, is that, at points where two interpolation curves meet, there might be sudden breaks in the values of the first-order differences (Shryock and Siegel 1973). In order to effect a smooth junction of the interpolations made for one range of data with the interpolations made for the next range, osculatory interpolation can be used. This method of graduation combines two overlapping polynomials into one equation. Interpolation is only conducted in the range of data where the two polynomials overlap. The second of the two polynomials in the first range, i.e. the polynomial that begins later and ends later than the other, then becomes the first polynomial in the second range and so on. This set of arcs for successive ranges is continuous as the polynomials have common tangents or other specific conditions in common when they overlap (Keyfitz 1977). These specific conditions are usually accomplished by forcing the first or the first two derivatives to be equal. Among the osculatory interpolation equations which are used the most often is Sprague's fifth-difference equation. Sprague's formula for interpolation can be expressed in linear compound form, that is, in terms of coefficients or multipliers that are applied to the given data. The Sprague formula interpolates over five age groups, and additionally provides special multipliers for the second age group from the beginning of the distribution and for the next-to-last age group from the end of the distribution and thereby provides a very accurate smoothing (Shryock and Siegel 1973).

A2. Intermediate steps based on the Preston, Heuveline and Guillot (2000) and the Preston and Bennett (1983) method

Results are shown for the lower bound estimates. The upper bounds estimate differ by the death rate nmx (fourth column in a.1-a.4).

a. Life table construction as described by Preston, Heuveline and Guillot (2000) to obtain mortality estimates for the very young and old cohorts.

a.1 Females 1901-1910:

Age x	nNx	nDx	nmx	nax	nqx	npx	lx	ndx	nLx	Tx
0	47,727	6,548	0.137	0.413	0.127	0.873	100,000	12,698	92,550	2,604,149
1-4	283,853	9,878	0.035	1.265	0.127	0.873	87,302	11,096	318,857	2,511,599
5-9	408,729	8,338	0.020	2.500	0.097	0.903	76,206	7,396	362,540	2,192,742
10-14	335,792	6,917	0.021	2.456	0.098	0.902	68,810	6,735	326,918	1,830,202
15-19	263,225	9,897	0.038	2.639	0.173	0.827	62,076	10,719	285,071	1,503,284
20-29	432,748	14,627	0.034	2.515	0.270	0.730	51,357	13,854	409,874	1,218,212
30-39	353,782	10,826	0.031	2.525	0.249	0.751	37,503	9,340	305,214	808,338
40-49	240,893	8,624	0.036	2.568	0.283	0.717	28,164	7,964	222,447	503,124
50-59	148,544	7,160	0.048	2.588	0.355	0.645	20,200	7,174	148,830	280,677
60 and over	117,484	11,607	0.099	3.488	1.000	0.000	13,026	13,026	131,847	131,847

a.2 Males 1901-1910:

Age x	nNx	nDx	nmx	nax	nqx	npx	lx	ndx	nLx	Tx
0	45,591	7,313	0.160	0.476	0.148	0.852	100,000	14,795	92,240	2,496,339
1-4	252,871	9,508	0.038	1.199	0.136	0.864	85,205	11,594	308,348	2,404,099
5-9	371,230	7,722	0.021	2.500	0.099	0.901	73,611	7,277	349,861	2,095,751
10-14	327,948	6,100	0.019	2.456	0.089	0.911	66,334	5,890	316,684	1,745,889
15-19	242,532	7,616	0.031	2.639	0.146	0.854	60,443	8,835	281,358	1,429,206
20-29	432,696	13,933	0.032	2.515	0.259	0.741	51,609	13,391	415,858	1,147,848
30-39	418,756	14,573	0.035	2.525	0.276	0.724	38,218	10,554	303,283	731,989
40-49	294,818	13,031	0.044	2.568	0.333	0.667	27,664	9,204	208,231	428,706
50-59	179,768	10,570	0.059	2.588	0.410	0.590	18,460	7,560	128,567	220,475
60 and over	123,271	14,620	0.119	3.488	1.000	0.000	10,900	10,900	91,908	91,908

a.3 Females 1911-1920:

Age x	nNx	nDx	nmx	nax	nqx	npx	lx	ndx	nLx	Tx
0	64,841	6,108	0.163	0.481	0.150	0.850	100,000	14,996	92,224	2,262,860
1-4	283,112	6,963	0.047	1.193	0.166	0.834	85,004	14,118	300,389	2,170,637
5-9	365,142	4,387	0.024	2.500	0.111	0.889	70,886	7,899	334,684	1,870,247
10-14	323,982	3,400	0.021	2.456	0.100	0.900	62,988	6,278	298,966	1,535,563
15-19	238,155	4,590	0.043	2.639	0.194	0.806	56,709	11,022	257,524	1,236,597
20-29	477,530	8,144	0.040	2.515	0.305	0.695	45,687	13,956	352,415	979,073
30-39	355,484	8,314	0.038	2.525	0.296	0.704	31,732	9,390	247,118	626,658
40-49	285,485	7,429	0.039	2.568	0.300	0.700	22,341	6,701	173,605	379,540
50-59	188,056	6,218	0.052	2.588	0.375	0.625	15,640	5,870	112,890	205,935
60 and over	179,754	10,226	0.105	3.488	1.000	0.000	9,770	9,770	93,045	93,045

a.4 Males 1911-1920:

Age x	nNx	nDx	nmx	nax	nqx	npx	lx	ndx	nLx	Tx
0	59,927	11,854	0.198	0.576	0.182	0.818	100,000	18,249	92,260	2,209,129
1-4	255,939	13,514	0.053	1.094	0.183	0.817	81,751	14,969	283,504	2,116,869
5-9	348,887	8,513	0.024	2.500	0.115	0.885	66,782	7,679	314,712	1,833,365
10-14	343,681	6,599	0.019	2.456	0.092	0.908	59,103	5,410	281,752	1,518,653
15-19	256,778	8,576	0.033	2.639	0.155	0.845	53,693	8,311	248,843	1,236,901
20-29	461,144	15,218	0.033	2.515	0.265	0.735	45,382	12,010	363,927	988,058
30-39	398,381	15,537	0.039	2.525	0.302	0.698	33,372	10,077	258,391	624,130
40-49	332,290	14,355	0.043	2.568	0.327	0.673	23,295	7,618	176,333	365,740
50-59	204,352	12,016	0.059	2.588	0.410	0.590	15,677	6,420	109,188	189,407
60 and over	171,232	19,760	0.115	3.488	1.000	0.000	9,257	9,257	80,219	80,219

b. Method proposed by Preston and Bennett (1983) to estimate life expectancy

b.1 Females 1901-1910

Age	rx	pop	R(x)	5L*x	l*x	T*x	ex0
0-4	0.005	339,766		392,613	108,798	2,889,631	29.20
5-9	-0.011	386,935	-0.028	376,211	77,123	2,497,018	32.90
10-14	-0.004	329,887	-0.065	309,084	68,530	2,120,807	30.95
15-19	-0.010	250,690	-0.099	227,058	53,614	1,811,723	33.79
20-24	0.014	236,796	-0.089	216,626	44,368	1,584,664	35.72
25-29	0.005	218,343	-0.041	209,616	42,624	1,368,038	32.10
30-34	-0.003	205,456	-0.035	198,446	40,806	1,158,422	28.39
35-39	0.005	149,177	-0.029	144,892	34,334	959,977	27.96
40-44	0.019	161,674	0.030	166,617	31,151	815,085	26.17
45-49	0.014	101,515	0.112	113,585	28,020	648,468	23.14
50-54	0.023	116,722	0.206	143,407	25,699	534,883	20.81
55-59	0.024	51,578	0.325	71,366	21,477	391,476	18.23
60 and over	0.044	72,611	0.495	119,068	67,148	840,397	12.52

b.2 Males 1901-1910

Age	rx	pop	R(x)	5L*x	l*x	T*x	ex0
0-4	0.006	307,164		375,764	115,720	3,074,867	30.06
5-9	-0.006	360,059	-0.015	354,531	77,280	2,699,102	35.56
10-14	0.005	335,814	-0.019	329,407	68,394	2,344,572	34.28
15-19	0.006	249,655	0.007	251,319	58,073	2,015,165	34.70
20-24	0.018	212,923	0.067	227,668	47,899	1,763,845	36.82
25-29	-0.005	233,997	0.102	259,002	48,667	1,536,177	31.57
30-34	-0.006	219,236	0.075	236,222	49,522	1,277,176	25.79
35-39	-0.004	189,332	0.050	199,084	43,531	1,040,954	23.91
40-44	0.012	183,937	0.071	197,525	39,661	841,870	21.23
45-49	0.012	129,617	0.131	147,731	34,526	644,345	18.66
50-54	0.013	127,109	0.192	154,034	30,177	496,614	16.46
55-59	0.013	64,951	0.256	83,912	23,795	342,580	14.40
60 and over	0.030	74,958	0.364	107,924	56,936	638,695	11.22

b.3 Females 1911-1920

Age	rx	Pop	R(x)	5L*x	l*x	T*x	ex0
0-4	0.013	371,464		411,406	123,001	3,503,294	25.93
5-9	0.024	414,289	0.060	439,794	93,734	3,091,888	32.46
10-14	0.014	348,340	0.155	406,576	84,637	2,652,094	31.33
15-19	0.002	240,530	0.195	292,216	69,879	2,245,519	32.13
20-24	0.006	261,070	0.214	323,497	61,571	1,953,302	31.72
25-29	0.009	234,268	0.251	301,055	62,455	1,629,805	26.10
30-34	0.005	207,717	0.285	276,241	57,730	1,328,751	23.02
35-39	-0.004	149,994	0.288	199,991	47,623	1,052,510	22.10
40-44	-0.008	169,671	0.257	219,355	41,935	852,519	20.33
45-49	-0.006	105,710	0.221	131,917	35,127	633,163	18.02
50-54	0.007	134,968	0.225	168,961	30,088	501,247	16.66
55-59	-0.001	57,629	0.240	73,285	24,225	332,286	13.72
60 and over	0.014	94,894	0.274	124,764	57,356	572,732	9.99

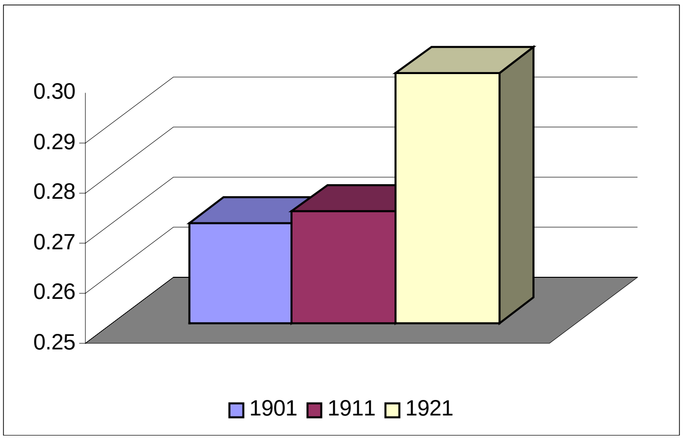
b.4 Males 1911-1920

Age	rx	pop	R(x)	5L*x	l*x	T*x	ex0
0-4	0.014	336,831		400,588	126,851	3,204,300	22.33
5-9	0.023	389,439	0.057	412,097	93,376	2,803,712	29.49
10-14	0.013	366,894	0.146	424,648	84,145	2,391,614	28.42
15-19	-0.004	252,050	0.169	298,320	72,297	1,966,967	27.21
20-24	-0.006	226,318	0.145	261,717	56,004	1,668,647	29.80
25-29	-0.001	227,716	0.130	259,221	52,094	1,406,930	27.01
30-34	0.002	214,399	0.132	244,691	50,391	1,147,709	22.78
35-39	-0.004	181,843	0.125	206,149	45,084	903,018	20.03
40-44	-0.019	178,103	0.067	190,351	39,650	696,869	17.58
45-49	-0.020	125,060	-0.030	121,323	31,167	506,518	16.25
50-54	-0.002	134,126	-0.084	123,379	24,470	385,195	15.74
55-59	-0.003	68,083	-0.095	61,891	18,527	261,816	14.13
60 and over	0.016	93,963	-0.062	88,289	44,340	473,473	10.68

A3. Fertility in Mysore

Another interesting observation that can be made from the newly constructed data set is that fertility levels soared right after the pandemic. Modern studies most often refer to Total Fertility Rates, the estimation procedure of which requires age-specific female population figures as well as number of births by the age of mother. Given the lack of this kind of detailed data, I employ a crude measure of fertility, the child/women ratio (CWR), i.e. the ratio of children aged 0-4 to the female population aged 15-44. Figure 10 illustrates the sharp increase in the CWR in 1921, which could be either a

Figure 10: Child/Women Ratio in Mysore 1901, 1911 and 1921.



Source: Own calculations based on Census of India, 1901-1921, MYSORE-Reports: Subsidiary tables.

statistical relict of the Influenza crisis, during which a relatively large share of middle-aged women died, reducing the denominator of the ratio. Or, the risen CWR could actually mirror higher birth rates, a phenomenon not rare for periods following crises such as wars and epidemics.

7 SUMMARY AND DIRECTIONS FOR FUTURE RESEARCH

Age heaping has the potential to broaden and deepen the empirical foundation for research on human capital and its role in economic development. As signature rates, despite their limitations, can proxy for basic literacy, so age heaping can be used as an index of basic numeracy. Both measures offer a partial view of human capital, and both reflect not only individual but also broader social capabilities. Age heaping is thus both a complement and a substitute for literacy-based measures: It complements literacy by revealing different cognitive abilities – abilities that are equally important in determining individual labor market outcomes today, and which arose in tandem with the market economy historically. It is also complementary in the sense of yielding insights not only about individuals but also about the societies in which they live. Age heaping can be a substitute or proxy for (il)literacy, as the two are well correlated between and within countries, in aggregate and individual data, in contemporary and historical contexts.

Though tenuous at very low levels of illiteracy and age heaping, this correlation is remarkably robust over the full range of observed variation in contemporary and historical data. The type of age heaping that predominates in historical data (on multiples of five) can be accurately measured using a computationally simple but reliable and efficient indicator: the Whipple Index. Alternatively, a simple transformation of the Whipple Index yields an easily interpretable numeracy index that provides an estimate of the share of individuals that correctly report their age.

The primary contribution of age-heaping based numeracy estimates to the study of human capital, however, is to extend coverage to times and places where signature based literacy estimates are not available. While our findings must remain tentative at this early stage, they are certainly intriguing. In the early Middle Ages, numeracy may have remained at or below levels observed in antiquity, a thousand years earlier. In Western Europe and in the German speaking lands of Central Europe, a slow but steady improvement appears to have set in by late medieval times. This period appears to have been the point of maximum divergence from Eastern Europe, where numeracy rates were only beginning to rise from very low levels. The development of numerical cognitive abilities and administrative capabilities thus had deep roots – long predating the spread of mass schooling – and displayed persistent variation across countries. The early and substantial rise of numeracy in the West lends strong support to accounts that assign human capital the role of cause, rather than effect, of the industrial revolution. The next stages in the development and implementation of the method include a more comprehensive survey of potential data sources and expansion of the database; an investigation of whether age-heaping or literacy is more closely correlated with measures of economic development; and the extension of the study to contexts in which literacy measures are completely absent or inherently more problematic, as with the ideographic writing systems of East Asia.

Moreover, age heaping can provide new input for research fields other than the long-run development of human capital, such as inequality studies. Employing age heaping inequality measures in a state-level analysis in the third chapter, we found that low levels of inequality in the U.S. were followed by strong subsequent economic growth, while high degrees of inequality of numeracy exerted a significant negative relationship on subsequent economic development in the United States. At the same time, industrialization increased inequality levels in the following periods, as suggested by Kuznet's inverse U hypothesis. In nineteenth century France, inequality of numeracy appears not to have been a crucial driver of welfare growth, in contrast to the proximity of a region to the central government in Paris, which turned out to have a significant and positive effect.

We also employed age heaping as a proxy for cognitive abilities. Exploiting the quasi-natural experiment of exogenous shocks to food availability in industrializing Britain in the 1790s, we assessed the importance of cognitive shortcomings driven by inadequate nutrition for low living standards in the past. In our empirical analysis, we used various instrumental variables to account for endogeneity issues and found detrimental consequences of high wheat prices for county-level numeracy values. This effect was particularly strong in counties with low poor relief payments. Our research also suggested that the decline in numeracy had significant effects on labour market outcome: Individuals born during the crisis years sorted into occupations with limited intellectual requirements that also paid lower wages. One of the implications of our research is that cognitive ability may have been a key factor limiting output in the past. Future research based on the findings presented in this chapter could elaborate further whether the transition to self-sustaining growth in industrializing Europe could owe a great deal to improved nutrition and higher cognitive ability.

The next chapter presented new evidence on human capital development in the 19th and early 20th century not only for Europe but for all world regions. Across as well as within culturally homogeneous regions we find substantial variation in basic numerical skills. Within Europe, for instance, Southern Europe and Ireland stood out as having had relatively low numeracy during the 19th century. Especially for the birth cohorts after the Great Famine, there was even a temporary deterioration before Ireland and Spain converged back to Western European levels. Comparing world regions, South Asia and the Middle East had relatively high age heaping levels in the 19th century (as opposed to the Western world and East Asia). China was characterized by very low age heaping levels in the late nineteenth century - comparable to Western industrialized countries.

We also assessed the impact of culture and bureaucracy on the degree of age rounding. While the Chinese calendar seems to have led to a very accurate knowledge of one's age, bureaucracy does not seem to have mattered much. State demand for frequent age reporting improved people's numeracy or number discipline, but only in the long-run, as indicated by our result that countries with a long tradition of census taking had slightly lower age heaping levels. In a variety of long-run growth regressions, numeracy had a positive and significant effect on subsequent growth. Due to the unavailability of schooling data for a comparable set of countries for the period under study, we could unfortunately not compare the importance of both human capital proxies, an exercise which is left for further research on modern developing countries.

The sixth chapter of my thesis demonstrated the dual purpose of single year age data collected for the age heaping research, contributing to an academic dispute in historical demography. I discussed adjustment procedures for age heaping, migration and underenumeration, before I presented new life expectancy estimates for the Indian state of Mysore during the early twentieth century. My results suggested that life expectancy levels in Mysore were relatively high compared to other Indian regions, although mortality rose tremendously as a consequence of the Influenza pandemic in 1918. My findings thus provide support for the thesis that relatively favorable demographic conditions prevailed in comparably poor South India.