Essays in Economic Growth

Ph.D. Thesis - Nicolai Kaarsen

University of Copenhagen
September 2011

Contents

Summary

The dissertation is comprised of three self-contained papers all in the field of growth economics. In the first two papers, I investigate what role the quality of human capital plays in accounting for cross-country income differences. In the third paper, I seek to understand why the Neolithic Revolution occurred earlier in some places than in others.

The grand question in growth economics is why some countries are rich and other countries are poor. The development accounting literature, a subfield of growth accounting, seeks to answer this question by quantifying how important differences in factors of input, such as the education of the labour force and physical capital stocks, are to differences in income. The two first papers belong to that literature. Traditionally, the development accounting literature uses average years of schooling to quantify human capital. However, the cognitive skills obtained from one year of schooling could be larger in, say, the U.S. than in Zimbabwe. If this is the case, differences in human capital could be larger than previously thought. Consequently, differences in human capital could account for a larger fraction of income differences.

In the first two paper, I construct indices for the quality of human capital defined as the cognitive skills obtained from one year of schooling. I then use these estimates to find out whether accounting for differences in the quality of human capital improves our understanding of income differences across the world.

The first paper constructs a cross-country measure of the quality of human capital using a novel approach based on international test scores data. The first main finding is that there are large differences in the quality of human capital - one year of schooling in the U.S. is equivalent to three or more years of schooling in a number of low-income countries. I incorporate the estimated series for the quality of human capital in an accounting framework calibrated using evidence on Mincerian returns. This leads to the second important finding, which is that the fraction of income differences explained by the model rise substantially when one includes the quality of human capital; the increase is around 25 percentage points.

The second paper estimates the quality of human capital for 34 developing countries. Whereas existing studies have derived quality indices based on either student test scores or immigrant earnings in the U.S. labour market, the second paper takes a novel approach by estimating human capital quality from national representative household surveys on years of schooling and literacy. The approach has the significant advantage that quality indices

fully reflecting the current labour force can be obtained; in contrast, student test scores only reliably speak to younger individuals in a population, and the human capital quality of people who emigrate to the U.S. may not be representative of the source population. A development accounting analysis of cross-country income differences reveals that the inclusion of human capital quality doubles the contribution from human capital and lowers the contribution from TFP by ten percentage points.

The third paper belongs in a literature which seeks to identify the historical roots of differences in income per capita. Development accounting is a useful tool to understand the question of why some countries are rich and other countries are poor. However, it leaves out some important questions. First of all, as shown in the two first papers, a large fraction of income differences remain unexplained. Secondly, it does not answer the question of why there are differences in the input factors such as physical capital and education. According to some scholars we need to turn to the historical record to understand fully the long-run processes generating differences in factors of production and productivity.

A prominent theory in this literature claims that differences in the transition from hunter-gathering to agriculture (the Neolithic Transition) had important and long-lasting consequences for economic development. A number of scholars present evidence that number of years passed since the Neolithic Transition are positively correlated with contemporary as well as historical measures of development. This finding, however, leads to another question which is why there are differences in the timing of the Neolithic Transition in the first place. This is the question I seek to answer in the third paper.

In that paper, I first document that in most cases the transition was caused by diffusion of crops from four different centers of cereal-based agriculture. Hence, the timing of the Neolithic Transition was determined by the speed of the spread of these crops. I then proceed to test two prominent yet, hitherto untested theories that explains the speed of diffusion. The first is that latitudinal distance is more inhibiting to diffusion of crops than is longitudinal distance. The second is that seasonal variation in precipitation is a barrier to diffusion of crops. I find support for both hypotheses.

Danish Summary

Afhandlingen består af tre selvstændige papirer. I de første to papirer analyseres betydningen af kvalitetsforskelle i humankapital for indkomstforskelle på tværs af lande. I det tredje papir undersøges hvorfor den neolitiske transition fandt sted tidligere i nogle lande end i andre.

I det første papir konstrueres et indeks for kvaliteten af humankapital på tværs af lande baseret på data fra internationale tests af skoleelever. Den første hovedkonklusion er, at der er store forskelle i kvaliteten af humankapital - et års skolegang i USA svarer til tre eller flere års skolegang i en række lavindkomstlande. Den anden hovedkonklusion er at disse forskelle i humankapital har stor betydning for forskelle i indkomst på tværs af verdens lande. Modellens evne til at forklare indkomstforskelle stiger betragteligt når kvaliteten af humankapital inkluderes.

Det andet papir har også til formål at konstruere et indeks for kvaliteten af humankapital. Bidraget i dette papir er at benytte data, som er repræsentativt for befolkningen i det pågældende land til at estimere kvaliteten af humankapital. Tidligere studier har benyttet sig af to forskellige tilgange til at estimere et mål for humankapitalskvalitet. Ved den første tilgang benyttes data for internationale tests af skolelever. Ved den anden tilgang benyttes løndata fra immigranter i USA til at forklare humankapitalskvaliteten i det pågældende fødeland. Problemet med disse tilgange er, at hverken skoleelever eller emigranter nødvendigvis udgør et repræsentativt udsnit af arbejdsstyrken i det land, hvor kvaliteten af humankapital ønskes estimeret. I dette papir benyttes et repræsentativt mikrodatasæt til at estimere kvaliteten af humankapital for 34 udviklingslande. Ligesom i det første papir er hovedkonklusionen at modellens forklaringskraft øges når kvaliteten af humankapital inkluderes.

I det sidste papir analyseres årsagerne til globale forskelle i timingen af den neolitiske transition. I de fleste tilfælde skyldes transitionen spredningen af afgrøder fra andre områder. Kun i få tilfælde er landbrug opstået uafhængigt. Således bliver hastigheden hvormed afgrøder spredes fra et sted til et andet en vigtig forklarende faktor for forskelle i timingen af den neolitiske transition. I papiret testes to fremtrædende teorier vedrørende spredningshastigheden af landbrug. Den første teori tilsiger at afgrøder spredes hurtigere langs længdegrader end langs breddegrader. Den anden teori tilsiger at spredningshastigheden mindskes ved store forskelle i sæsonvariation i nedbør. Konklusionen er, at begge teorier understøttes af data.

Acknowledgements

I am grateful to a number of people who have supported me throughout my time as a Ph.D. student at the Department of Economics, University of Copenhagen. First of all, I would like to thank my advisor Carl-Johan Dalgaard. The countless discussions I have enjoyed with Carl-Johan over the years have been an everlasting source of inspiration and encouragement. His criticism, suggestions and comments have been invaluable.

I would also like to thank my fellow students, the faculty members, and the staff at the Department of Economics for making my years as a student more fun and enjoyable. Because of these people, the Department of Economics has been an inspiring and motivating workplace.

In the spring of 2009, I had the pleasure of visiting Brown University. My stay at Brown was extremely rewarding professionally as well as personally. I would like to thank Oded Galor for giving me the opportunity to visit Brown, and for taking the time out to discuss my research. I also owe a thanks to my advisor Carl-Johan for establishing the contact to Oded Galor. My visit abroad had not been possible without the financial support of Augustinus Fonden, Knud Højgaards Fond and Oticon Fonden, for which I am very grateful.

Finally, I thank my family and friends for supporting me throughout all of the years and bearing with me in times of distress and hardship. A special thanks goes out to my girlfriend Sys for her support and understanding.

Cross-Country Differences in the Quality of Human Capital*

Nicolai Kaarsen†

September 2011

Abstract

This paper constructs a cross-country measure of the quality of human capital using a novel approach based on international test scores data. The first main finding is that there are large differences in the quality of human capital - one year of schooling in the U.S. is equivalent to three or more years of schooling in a number of low-income countries. I incorporate the estimated series for the quality of human capital in an accounting framework calibrated using evidence on Mincerian returns. This leads to the second important finding, which is that the fraction of income differences explained by the model rise substantially when one includes the quality of human capital; the increase is around 25 percentage points.

*I would like to thank Carl-Johan Dalgaard, Oded Galor, Michael Svarer, David Weil, Asger Moll Wingender and seminar participants at Brown University, University of Copenhagen and the 25th Annual Congress of the European Economic Association for helpful comments and suggestions.

†Contact: University of Copenhagen, Øster Farimagsgade 5B, Building 26, DK-1353 Copenhagen K, Denmark. Nicolai.kaarsen@econ.ku.dk.

1 Introduction

How important is human capital in determining income per capita? The literature on development accounting seems to agree that human capital is an important determinant of income, but that the lion's share of the gap in income between poor and rich countries is not attributable to differences in human capital or physical capital endowments[1]. Instead, the main cause of the world income differences lies in differences in a residual productivity term which is unexplained. However, recent work by Manuelli and Seshadri (2006), Hanushek and Woessmann (2009) and Schoellman (2011) suggest that the role of human capital may be underappreciated. The central charge is that the literature hitherto has ignored differences in human capital quality.

In the development accounting literature, the human capital stock is usually computed using average years of schooling as the only input. This approach implicitly assumes that one year of schooling in Ghana is equal to one year of schooling in the U.S. If, however, one year of schooling in high-income countries is more productive relative to one year of schooling in low-income countries, human capital may be able to account for a larger share of income differences than previously thought.

In this paper, I estimate differences in the quality of human capital defined as the increase in cognitive skills obtained from an additional year of schooling. This measure can be directly incorporated into a development accounting framework. I find that there are large differences in the quality of human capital across countries. One year of schooling in the U.S. corresponds to three or even four years of schooling in many developing countries. Moreover, these quality differences are able to account for a considerable share of the variation in income across countries. I find that including the quality of human capital

[1] See e.g. Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), Bils and Klenow (2002) and Caselli (2005).

increases the log-variance of human capital as a fraction of the log-variance of income from 0.04 to 0.26.

How is the quality of human capital estimated? I use an international test scores data set, which has the important feature that the same test was given to two different grades. This allows me to identify differences in the test scores gradient in years of schooling for a cross-section of countries. I define this gradient as the quality of years of schooling. It measures the effectiveness of one year of schooling in country i relative to one year of schooling in the U.S., which I choose to be the numeraire country. The measure can be seen as a conversion factor which adjusts years of schooling to be measured in U.S.-equivalent years of schooling.

This series is then used to evaluate the role of human capital quality in accounting for income differences. To do this I modify a standard accounting framework to include the estimated series of human capital quality. More specifically, I generalize the human capital production function of Bils and Klenow (2002) and calibrate the parameters such that the model is consistent with micro-evidence on Mincerian returns.

We would like the estimated measure of the quality of human capital to reflect the quality of an average worker in the labour force. However, the quality of human capital is estimated based on test scores of students which are not necessarily representative of the labour force. In particular, the earliest test scores data used in this paper is from 1995. Hence, in principle the estimated quality of human capital only reflects the quality of younger cohorts.

How could this bias the main results? If test scores have decreased over time, the cognitive abilities of young workers are low relative to those of the average worker in the labour force. In this case, the quality of human capital will be underestimated. Hence, if test scores have decreased over time in low-income countries relative to test scores in high income countries, the variance of quality of human capital will be biased upwards.

The evidence presented in this paper suggests that this is not an issue. In particular, I show that: 1. Differences in test scores over time are considerably smaller than differences across countries. 2. Changes in test scores over time are not correlated with income per capita.

A limitation of using test scores data is that it is only available for 65 countries and most of these are high-income countries. Hence, the sample of countries for which I can estimate the quality of human capital is not representative of the countries of the world. To deal with this problem I follow Weil (2007) in extending the data using a number of variables which are highly correlated to the quality data, and which are available for 174 countries. These variables are used to predict the quality of human capital for the countries where tests score data is not available. I find the same main conclusion using the small sample consisting only of countries where the quality of human capital is estimated as I do using the extended sample.

This paper is related to a string of contributions which seek to quantify the impact of human capital quality on growth. Within this literature two distinct approaches to the issue at hand can be identified.

The first approach seeks to identify an aggregate effect of human capital quality. This effect could go through many channels; some prominent examples are technology adoption (see e.g. Nelson and Phelps (1966)) and fertility (see e.g. Galor and Weil (2000)). Examples of papers seeking to estimate the aggregate effect are Hanushek and Kimko (2000) and Hanushek and Woessmann (2009).

The second approach focuses on identifying the effect running through individual productivity. Workers with better cognitive skills accomplish more complicated tasks faster and hence produce more. This paper belongs to that literature as do the papers of Hendricks (2002), Schoellman (2011) and Caselli (2005).

Hendricks (2002) uses the wages of immigrants in the U.S. to estimate the quality of human capital in their country of birth. To do this he compares wages levels of immigrants holding constant the level of education. His findings are quantitatively very similar to the standard findings of the literature as seen in e.g. Hall and Jones (1999) and Klenow and Rodriguez-Clare (1997), that is, human capital accounts for a relatively small fraction of income differences.

Schoellman (2011) also uses U.S. immigrant data to estimate differences in the quality of human capital, but he uses a different methodology than Hendricks (2002), and he reaches a different conclusion. Schoellman (2011) estimates a separate Mincer regression for each country of origin using the U.S. wage data and interprets the slope estimates as the quality of education pertaining to the respective countries. He includes these estimates in a development accounting framework similar to the framework used in the present paper, and finds that this increases the fraction of income explained by human capital considerably. Quantitatively, his results are very close to mine.

An issue with using immigrant data to infer the quality of human capital of source countries is that of selection. It is clear that immigrants are not selected randomly out of the population of the source country, however, it is unclear how the selection occurs and what the consequences are for the estimates of the quality of human capital. Another issue, pointed out by Friedberg (2000), is that education is not perfectly transferable across borders. Knowledge obtained in foreign countries may be valued less because of specificities in e.g. norms and institutions in the schooling system. Hence, the Mincerian return for a Dane in the U.S. might differ from the return of a Dane in Denmark. If such barriers are relatively higher for migrants coming from developing countries, the variance of the quality of human capital could be overestimated.

In the test scores data used in this paper, these issues are less likely to be a problem. The participating students are selected to be representative of the entire student population, and around 5000 students are tested in each country. Furthermore, to mitigate barriers of language and culture all test questions are back-translated and based on an international curriculum representative of all participating countries.

Caselli (2005) uses micro-evidence on the wage return to test scores as well as international test scores data to account for the effect quality of human capital on development. He finds that quality differences are relatively unimportant in explaining cross-country income differences. A drawback with this methodology is that the test scores data used to estimate the returns and the cross-country data on test scores are not comparable. They are based on different tests and different samples. Hence, it is very difficult to directly translate differences in international test scores into differences in productivity across countries.

The paper proceeds as follows. The next section estimates the quality of human capital. In section 3, these estimates are used in a development accounting exercise. Section 4 investigates the role of potential biases arising from using the test scores of younger cohorts to estimate the average quality of human capital. The final section concludes.

2 Estimating the quality of human capital

This section falls in four subsections. The first subsection presents the data, the second the methodology. The third subsection contains the main estimation results, and the final subsection expands the estimated human capital quality data to a larger number of countries.

2.1 The data

Trends in Math and Science Study (TIMSS) is a series of science and math tests conducted in schools in a number of countries in the years 1995 to 2007 by the International Association for the Evaluation of Educational Achievement (IEA). In each of the years 1995, 1999, 2003 and 2007, four different tests - a science test in primary school, a math test in primary school, math test in secondary school and a science test in secondary school - were administered in an unbalanced panel of countries. In the following, these four classifications will be denoted as *test types*.

The tests were assigned to a large number of students in each country (usually over 5000 students per test). Furthermore, great care was taken in constructing the tests so that they matched an international curriculum, and not just the curriculum of one country such as the U.S. In 1995, the same test was given to different grades which, as alluded to above, is invaluable to the identification of the quality of human capital.

In all of the TIMSS programs, each student is given a multiple choice test where the answers are ranked according to correctness. The grading of the tests is done separately for each test type, and is based on item response theory (IRT) which is a method used to convert answers into a test score. This conversion method is designed such that the resulting test scores are placed on a certain predetermined metric. In TIMSS, it is decided that the pooled sample of test scores from students of all countries in 1995 should have a mean of 500 and a standard deviation of 100. A detailed description of the method is given in Chapter 11 of TIMSS (2007).

Even though different tests were constructed from year to year some of the questions were repeated, which allows the IEA to temporally link the scaling of test scores such that all of the scores are placed on the 1995 metric. Thus, test scores are comparable over time.

Table 1 provides an overview of the availability of the data. All of the test scores data used in this paper are country averages. The maximum number of participating countries in one type of test is 46 (secondary math test, 2007), and the total number of countries which participated in at least one test is 65.

Table 1: Categorization of the Neolithic Transition Across Countries

Crop system	Region	Center	Cereals	Number of countries
Maize	Americas	The Balsas River Valley, Mexico	Yes	18
Rice and foxtail millet	Asia	Yangtze and Yellow Rivers, China	Yes	19
Sorghum and pearl millet	Sub-Saharan Africa	The Sahel, Africa	Yes	37
Wheat and barley	West Asia, North Africa and Europe	The Levantine Corridor, Isreal	Yes	72
Tropical tubers and fruits	South East Asia, South America and Caribbean	Various centers	No	15
Total				161

Notes: The table accounts for the origins of the Neolithic Transition across countries. The first column shows the main crops involved in the Neolithic Transition. The second and third column show, respectively, the region(s) and the center of origin of the crop system. In the fourth column it is stated whether cereals were part of the Neolithic Transition or not. The fifth column shows the number of countries pertaining to that cathegory.

The main source of test scores data is TIMSS (2008). This publication includes all the data shown in Table 1 with the exception of the data for 3rd and 7th graders from the 1995 round. TIMSS (1997a,b) includes the data for 3rd and 7th graders, but the test scores from this publication are placed on a slightly different metric. Fortunately, since both publications contain data on 4th and 8th graders from the 1995-round, it is possible to rescale the results for the 3rd and 7th graders from the TIMSS (1997a,b) metric to the TIMSS (2008) metric. I use linear regression to do this for each of the four test types; the details are in the appendix.

The final data set consists of 1170 observations of four different test scores from 65 countries from four different years. The correlation between test types is high. In 1995, the year with the most observations, the average correlation between two test types is 0.82. Figure 1 shows the rescaled test scores over time for mathematics tests of 8th graders.

Figure 1: Log of TIMSS test scores for mathematics in 8th grade

Notes: The TIMSS test scores have been rescaled to the metric of TIMSS (1997a,b), see the main text for details.

At the high end of the spectrum we find East Asian countries such as Korea, Japan and Singapore. The Western industrialized countries lie in the middle, while low-income countries such as Colombia, South Africa and Ghana are amongst the countries with the lowest test scores.

2.2 The model and the empirical specification

To estimate the quality of human capital I assume that the test score $T_{k,i,s,t}$ of test type k, in country i, at grade s and in year t is determined by the following production function:

$$T_{k,i,s,t} = f\left(s \times q_i, \Gamma_k\right), \tag{1}$$

where q_i is the quality of years of schooling, and Γ_k is a vector of parameters. For example, $T_{msec,US,8,2003} = f(8 \times q_{US}, \Gamma_{msec})$ is the 8th grade math test score of the U.S. in 2003. It is produced with 8 years of schooling and the quality of U.S. education as input and the production function parameters are those of the math secondary school test type.

The goal is to use a specific functional form for (1) to estimate q_i and Γ_k jointly using the test scores data. This produces an estimated series for q_i for the 65 different countries participating in the TIMSS.

The set of parameters Γ_k are varying across test types. Thus, in effect, (1) reflects four different test scores production functions, one for each test type. Since the test scores are not comparable across test types, that is, one point on the math primary scale does not correspond to one point of the science secondary scale, we cannot use the same production function for all test types.

In all four production functions, however, I let the same q_i enter as input. This is because we wish to estimate the general level of human capital quality in the country. This seems like a reasonable assumption given that the test scores are closely correlated across test types, as was shown above. Furthermore, q_i is assumed constant over time, since we wish to estimate the overall quality of human capital ignoring idiosyncratic changes in test scores over shorter periods. As will be evident in the robustness section, this assumption appears to be plausible since the variation of test scores over time is substantially smaller than the variation across countries.

(1) assumes that there are no differences in the quality of human capital evaluated at zero years of schooling. Such differences could be captured by adding a country-specific constant term to (1). Differences in e.g. parental education or health could generate cross-country variation in the cognitive ability measured at zero years of schooling. Unfortunately, given that we only have data for two years of schooling within each test type, it is not possible to add a constant term capturing such effects. Furthermore, to estimate

the quality of human capital at zero years of schooling one would have to include cognitive tests of persons who did not go to school, and such data is difficult to find for a large cross-section of countries.

Which functional form should we choose for (1)? Since the physical capacities of the brain places an upper bar on the stock of knowledge which can be accumulated, it seems plausible that the test scores production function exhibits decreasing returns. Hence, I will use the following functional form which potentially satisfies this assumption:

$$T_{k,i,s,t} = \beta_k \left(s \times q_i\right)^{\gamma_k}, \tag{2}$$

where β_k and γ_k are a production function parameters, which, as noted, are allowed to vary across test types. If γ_k is estimated to be below one, (2) exhibits decreasing returns to the input $s \times q_i$.

I estimate (2) using the following empirical specification:[2]

$$\ln T_{k,i,s,t} = \ln \beta_k + \gamma_k \left[\ln s + \sum_{j=1}^{65} D_{ji} \ln q_i\right] + e_{k,i,s,t}, \tag{3}$$

where D_{ij} is a country dummy which is one if $j = i$, and $e_{k,i,s,t}$ is an error term. (3) is estimated using non-linear least squares. Even though the number of participating countries is 65 and each country potentially could have had up to 18 different test scores, the total number of observations is only 544 since some of countries participated in only one or two years.

[2] (2) could also be estimated using OLS as:

$$\ln T_{k,i,s,t} = \ln \beta_k + \gamma_k \ln s + \sum_{j=1}^{65} D_{ji} x_{k,i} + e_{k,i,s,t},$$

where $x_{k,i} = \gamma_k \ln q_i$. The estimates of q_i could then be inferred from the estimates of $x_{k,i}$ and γ_k. However, this method would not give us the standard deviations of the estimated q_i's directly as estimation output.

Note, that it is only possible to estimate (3) since we have test scores data for the same test type at different grades. If we only had data for one year there would be no variation in s. In this case, since β_k, γ_k and q_i are parameters, there would be no variation in the right-hand side of (3).

To identify β_k and the q_i's separately it is necessary to fix one of the q_i's. I choose to set the q_i of U.S. to one ($q_{US} \equiv 1$). Thus, q_{US} acts as a numeraire allowing us to interpret $1/q_i$ as the years of education it takes for the average student in country i to learn as much as the average student in the U.S. learns in one year. $s \times q_i$ is then denoted quality-adjusted years of schooling or U.S.-equivalent years of schooling. The multiplication sign is included to underline that $s \times q_i$ is the product of two separate variables, but will be suppressed from now on as will the index i.

2.3 Results

Estimating (3) results in a data set of q's spanning 65 countries. The estimated q's and their standard errors are shown in the appendix. The mean standard error of the estimated q's is 0.04 and it does not exceed 0.07 for any one country. Table 2 shows the estimated parameters for the test scores production functions.

Table 2: Timing of the Neolithic Transition and climatic barriers to diffusion of crops

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Years since the Neolithic Transition (in 1000s)					
Latitudinal distance (1000 km)	-0.477***		-0.409***	-0.345***	-0.243**	-0.286***
	(0.072)		(0.070)	(0.081)	(0.102)	(0.103)
Longitudinal distance (1000 km)	-0.066		-0.092	-0.117	-0.104	-0.123
	(0.077)		(0.073)	(0.074)	(0.084)	(0.092)
Difference in altitude (km)	-0.497**		-0.386*	-0.771***	-0.860***	-0.563*
	(0.223)		(0.213)	(0.230)	(0.324)	(0.326)
Difference in seasonality		-5.992***	-4.017***	-3.534***	-4.104***	-3.688***
		(1.118)	(1.000)	(1.140)	(1.331)	(1.235)
Absolute latitude				-19.41**	-24.55*	-34.22**
				(9.741)	(12.890)	(14.140)
Longitude				1.916	3.923	6.317
				(5.760)	(7.611)	(7.037)
Altitude (km)				0.530***	0.331	0.285
				(0.177)	(0.221)	(0.217)
Seasonality				-1.054	-0.514	-0.879
				(0.844)	(1.067)	(1.112)
Island dummy						43.720
						(124.400)
Number of domesticable plants						-32.970
						(36.620)
Number of domesticable animals						43.720
						(124.400)
Suitability of climate for agriculture (standardized)						445.9***
						(147.300)
Land area (standardized)						0.00148**
						(0.001)
Observations	139	139	139	139	98	98
R-squared	0.85	0.81	0.87	0.88	0.89	0.91
Crop and continent dummies	Yes	Yes	Yes	Yes	Yes	Yes

Summary: The table shows that longer latitudinal distance and larger difference in seasonality of precipitation between a country and the center of origin of agriculture is associated with later adoption of agriculture. This is evidence in favour of the theory that climatic differences slow down the speed of diffusion of crops and hence delay the timing of the Neolithic Transition.

Notes: Latitudinal distance is shortest distance in 1000s of km to the line of latitude passing through the center of origin of agriculture. Longitudinal distance is the shortest distance to the line of longitude passing through the center of agriculture and measured at the latitude halfway between the points. Difference in altitude is the absolute difference between the altitude of the country and the altitude of the center of agriculture. Difference in the seasonality of precipitation is the absolute difference between the seasonality of the country and that of the center of agriculture, where the seasonality of precipitation is the covariance between montly temperature and precipitation. See the data section for definition of all other variables. The p-value of the F-test of jointly excluding all crop dummies is less than 0.01% in all specifications. The p-value of the F-test of jointly excluding continent dummies is less than 0.01% in all specifications. Standard errors in parentheses. * = significant at the 10% level, ** = significant at the 5% level. *** = significant at the 1% level.

The standard errors of all parameters are low compared to the point estimates. Furthermore, the estimated β_k and γ_k are relatively constant over test types. γ_k is around 0.5 and significantly different from 1 for all test types. This confirms our prior that the returns to test scores are decreasing in inputs.

Figure 2 below shows a scatter plot of log GDP per worker and the quality of human capital for the 41 countries for which both data series were available.

Figure 2: Log of GDP per worker in 2007 and quality of human capital.

Notes: The quality of human capital is estimated using test scores data, see the main text for details on how this is done.

The figure shows that q varies considerably across countries going as low as 0.25-0.30 in Yemen and Ghana. In a country where $q = 0.5$ the average student achieves in two years of schooling what the average student in the U.S. achieves in one year. There is a strong positive relationship between q and log income per worker. Regressing on log income on q yields a slope of 1.74 with an R^2 of 0.25.

An interesting finding is that for many Eastern Asian countries the quality of human capital is relatively high compared to the level of GDP per capita. The q of e.g. Singapore, Japan, South Korea and Japan is around 1.2 implying that one year of schooling generates 20% more knowledge as in the U.S.

Another interesting group of outliers consists of the oil-producing countries Saudi Arabia, Qatar, Kuwait and Oman. They all have relatively a high GDP per worker but a low quality of human capital.

2.4 Extending the data set of quality-adjusted years of schooling

The data set created in the previous section consists of only 65 countries with an over-representation of high-income countries. Therefore, this section extends the sample of q's. The method used to do so is taken from Weil (2007).

First, the quality of education estimated in the previous section is regressed on a set of variables. Second, the predicted values from this regression are then used as q's for 107 countries which do not have estimated data for q.

I use the following three variables which are strongly correlated to the quality of human capital: GDP per worker (in 2007, from Penn World Table 7.0), average years of schooling (in 2005, from Barro and Lee (2010), and the population density (population per 100 m^2 in 2000, extracted online from UNdata[3]). I also include region dummies[4].

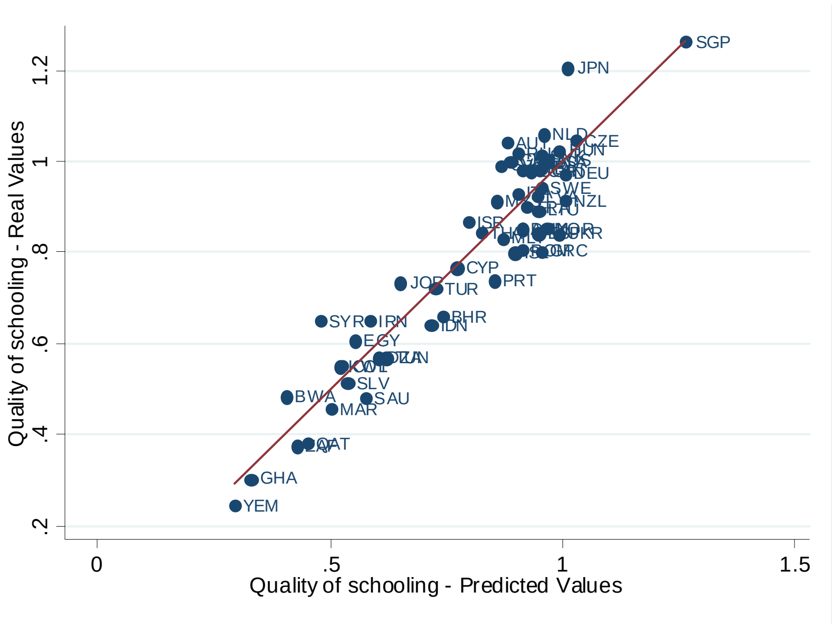
Regressing the estimated q on these variables for the 56 countries for which all data is available yields an R^2 of 0.87. The details of the regression are shown in the appendix.

[3] http://data.un.org/

[4] I use the region dummies provided by Barro and Lee (2010).

Figure 3 plots the actual values against the predicted values. It shows that the regression line provides a reasonable fit, also in the case of countries with lower q.

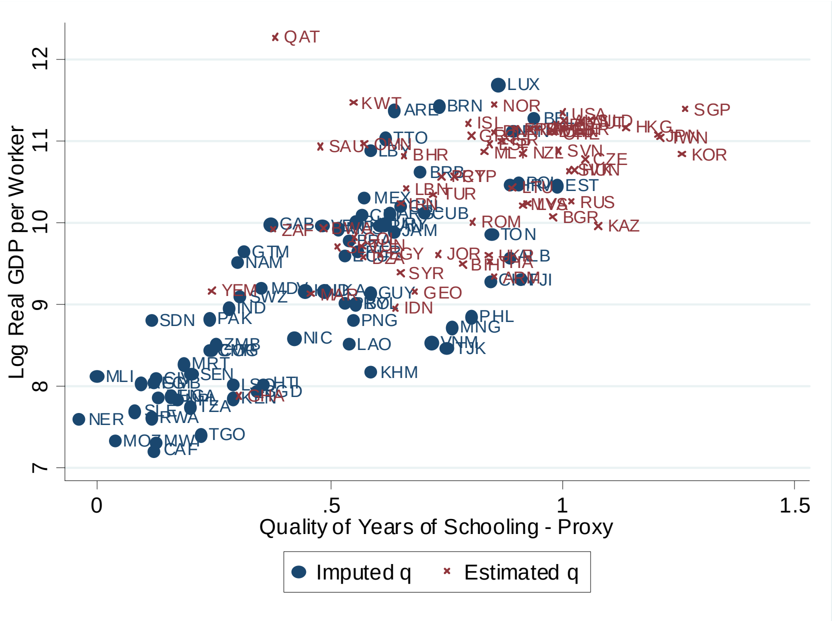
Figure 3: Predicted and actual values of quality of schooling.



Notes: The actual values of the quality of schooling are estimated using test scores. Predicted values are based on regressing actual values on a number of proxy variables. See the main text for more details.

The above regression results are now used to predict q's for 74 countries resulting in a data set containing q for 139 countries shown in Table A1 in the appendix. Figure 4 plots log income per worker against q for the full data set.

Figure 4: Log of GDP per worker, 2007 and quality of years of schooling - full sample.



Notes: The quality of schooling is estimated using international test scores data. For countries where the test scores data is not available the quality of human capital has been imputed. See main text for details

As in Figure 2, there is a strong positive relationship between log GDP per worker and q and a lot of dispersion in the quality of human capital. For countries with a q below 0.25 (the q of Yemen), data is extrapolated and should only be used for further analysis with caution[5]. As is evident from the figure this is the case for quite a lot of developing countries. Hence, all decomposition results below are shown for both the sample of extrapolated q's and the baseline sample of estimated q's. The main results are similar for the two samples.

[5]For Mali and Nigeria the imputed q was slightly below zero. For both countries, I replace q by zero.

3 Quality-Adjusted Years of Schooling and World Income Differences

The estimated differences in the quality of human capital across countries of the world are substantial. However, can they account for the large income differences observed? To answer this question we need to incorporate the quality measure in a development accounting framework. I rely on the standard framework used by e.g. Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), Bils and Klenow (2002) and Caselli (2005).

The first step is to generate human-capital stocks. I generate two series, one which is based on the assumption that there are no differences in the quality of human capital. This series is based entirely on average years of schooling and thus follows the usual methodology from the literature. The second series is computed incorporating the quality data estimated above.

The second step is to use the two series in an accounting framework. I use two tools which are commonplace in the literature of development accounting. The first is decomposition of the log-variance of income; the second is decomposition of differences in income percentiles. Using either of these tools gives the same result which is that incorporating differences in the quality of human capital substantially increases the fraction of income differences explained by the model.

3.1 Construction of human-capital stocks

To generate human capital stocks I follow Schoellman (2011) who generalize the framework of Bils and Klenow (2002) to include quality of human capital. The backbone of the model is the Cobb-Douglas production function[6]:

$$y = Ak^{\alpha}h^{1-\alpha}, \tag{4}$$

where y is output per worker, A is total-factor productivity, k is physical capital per worker, and h is human capital per worker. To use (4) to account for differences in output we need data for y, k and h. A is computed as a residual under the assumption that $\alpha = 1/3$. For y and k I use data from 2008 from PWT 7.0[7]. The computation of the cross-country series for h proceeds as follows.

I assume that the human capital production function is:

$$h = e^{\phi(s,q)}, \tag{5}$$

where s is years of schooling, q is quality of human capital, and $\phi(s, q)$ is given by

$$\phi(s, q) = \frac{\theta}{1-\psi}(sq)^{1-\psi}. \tag{6}$$

For $q = 1$ this collapses into the production function used by Bils and Klenow (2002). Data for s is taken from Barro and Lee (2010) and data for q is taken from above. I calibrate θ and ψ using evidence on the return of log wages to years of schooling. The basic methodology is similar to that of Bils and Klenow (2002) and Schoellman (2011).

The standard reference for data on Mincerian returns is Psacharopoulos (1994) who surveys a list of micro studies. In this data, the returns to years of schooling is consistently higher for developing countries. However, a newer survey from Banerjee and Duflo (2005)

[6] This analysis does not take into account the effect of human capital on income levels through externalities as e.g. technology. Altough such effects potentially could be important, there is no relliable estimates of the magnitude of such effects, so they are left out of the analysis.

[7] k is computed by using the perpetual inventory method described in Caselli (2005).

concludes that, on average, the return to years of schooling is around 0.1 for both developing and developed economies. Hence, I assume that the return to years of schooling is constant at 0.1.

To calibrate θ and ψ first assume that markets are competitive implying that the wage of the individual worker is given by his human capital times the return on human capital in the country. Hence, the Mincerian returns MR are given by

$$MR = \frac{\partial \ln h}{\partial s} = \theta s^{-\psi} q^{1-\psi}, \tag{7}$$

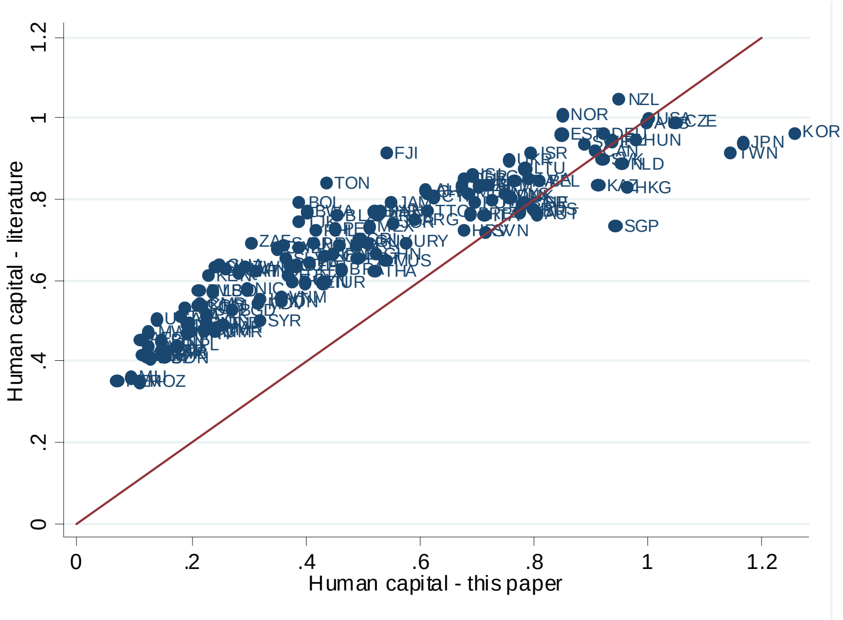
Now insert $MR = 0.1$, take logs and rearrange to obtain:

$$\ln s = \frac{\ln \theta - \ln 0.1}{\psi} + \frac{1-\psi}{\psi} \ln q. \tag{8}$$

Estimates of θ and ψ can be backed out from the estimated constant term and slope one gets from regressing $\ln s$ on $\ln q$. To do this, I take data for q from above and data for s from Barro and Lee (2010). I use the sample of 65 estimated q's, however, using the extended sample of q's does change not calibration results markedly. 58 of these countries have data for s. For this sample I regress $\ln s$ on $\ln q$ yielding an estimated constant and slope of 2.36 and 0.51, respectively. The corresponding standard errors are 0.027 and 0.062, respectively, and R^2 is 0.55. Backing out the implied parameters of the human capital production function yields $\psi = 0.66$ and $\theta = 0.48$.

With these calibrated values for ψ and θ we can use (8) to compute a series for quality-adjusted human capital. As a benchmark, I also computed a human capital series under the assumption that $q = 1$, implying that $h = e^{0.1s}$. Figure 5 below plots the two different measures of human capital against each other.

Figure 5: Human capital as estimated in the literature vs. human capital adjusted for quality differences.



Notes: Both measures are relative to human capital in the U.S. See the main text to get a description of how the human capital variables are constructed.

In the figure, I have normalized the human capital measures such that U.S. human capital is equal to one. The figure also shows a 45^o line. Accounting for the quality differences increases the variance of human capital. Human capital computed using only years of schooling varies from 0.4 to around 1 - a factor 2.5. Quality-adjusted human capital varies from around 0.1 to 1.2 - a factor 12.

3.2 Accounting results

How large a share of income differences can the generated human capital stocks account for? To find out first take logs and variance of (4) yielding

$$var[\ln y] = var[\ln A] + var[\ln x] + 2Cov[\ln A, \ln x], \tag{9}$$

where $x = k^{\alpha}h^{1-\alpha}$ is GDP per capita predicted by the pure input-factors model. This formulae forms the basis of the variance decomposition exercise found in e.g. Caselli (2005)[8]. With data for y, k and h, A can be computed as a residual using (4). Hence, we have all the data needed to compute the terms in (9).

An alternative way to evaluate the role of various factors in explaining income differences is to compare income ratios at different percentiles. Define V_f as the f percentile of V. For instance, y_{90}/y_{10} is the ratio of income at the 90 percentile to income at the 10 percentile. $(h_{90}/h_{10})/(y_{90}/y_{10})$ and $(x_{90}/x_{10})/(y_{90}/y_{10})$ measure, respectively, the fraction of the income ratio explained by human capital alone and the fraction of the income ratio explained by the model.

Table 3 below shows the results of the decomposition of log-variance of income and income ratios.

[8] Another way to decompose income differences, based on Hall and Jones (1999) and Klenow and Rodriguez-Claire (1997), is to write (4) in terms of capital-output ratio instead of capital-labour ratio. This tends to put more weight on human capital differences and less weight on physical capital differences. Hence, the increase in the fraction income differences explained by the model is even larger. See the appendix for results.

Table 3: Decomposition of income differences

Sample:	Only estimated q's (N = 47)		Imputed and estimated q's (N = 116)	
Includes quality of human capital:	No	Yes	No	Yes
	(1)	(2)	(3)	(4)
Var(lny)	0.64	0.64	1.65	1.65
Var(lnh)	0.04	0.19	0.06	0.48
Var(lnx)	0.15	0.30	0.38	0.82
Var(lnh)/Var(lny)	0.06	0.29	0.04	0.29
Var(lnx)/Var(lny)	0.23	0.46	0.23	0.50
(Var(lnx)+Cov(lnA,lnx))/Var(lny)	0.46	0.64	0.46	0.67
y_{90}/y_{10}	5.62	5.62	26.93	26.93
h_{90}/h_{10}	1.61	2.80	1.96	7.24
x_{90}/x_{10}	2.74	3.99	5.10	11.70
$(h_{90}/h_{10})/(y_{90}/y_{10})$	0.29	0.50	0.07	0.27
$(x_{90}/x_{10})/(y_{90}/y_{10})$	0.49	0.71	0.19	0.43

Summary: This table decomposes income differences into different components. It shows that accounting for the quality of human capital increases substantially the fraction of income differences explained by the model.

Notes: y is GDP per capita, h is human capital, x is GDP per capita predicted by the factors-only model. Subscript 90 indicates the 90% percentile and subscript 10 indicates the 10% percentaile. The first two columns show the results for the small sample of countries where data for q is estimated from test scores data. The third and fourth columns show the results for the large sample of countries, where q is imputed. See section on details of how data on q is estimated and imputed. In the first and third columns, human capital is computed using only years of schooling. In the second and fourth columns human capital is computed using quality of human capital and years of schooling.

The two first columns show the results for the small sample consisting only of countries for which estimated q's are available. Columns 3 and 4 show the results for the extended sample consisting of countries for which q is either estimated or imputed. Columns 1 and 3 show the results for the model where there are no differences in q, whereas columns 2 and 4 show the results for the model where differences in q are taken into account.

First of all, the table confirms what Figure 5 illustrated: accounting for quality differences increases the variance of human capital. Moreover, this increase appears to be substantial. For both samples $var(\ln h)/var(\ln y)$ increases substantially when the quality of human capital is taken into account. In the full sample, this figure increases from 0.04 to 0.26. Schoellman (2011) who uses immigrant wage data to compute quality-adjusted human capital finds that $var(\ln h)/var(\ln y) = 0.26$, a number which is identical to mine. Although he uses a different methodology and different data to estimate the quality of human capital, he reaches the same conclusion as this paper.

Turning to the fifth row, $var(\ln x)/var(\ln y)$, which is Caselli's (2005) preferred measure of model success, increases substantially when human capital quality is included. For both samples $var(\ln x)/var(\ln y)$ increases from around 24% to 47%[9].

Another thing to note from this row is that, in the baseline model with no differences in the quality of human capital, $var(\ln x)/var(\ln y) = 0.24$. This is substantially smaller than the 0.39 reported by Caselli (2005). Since the model and method of decomposition are the same, these differences can be attributed to the differences in the data used[10]. To check that the choice of data does not make a difference to the main results I redid the calibration and variance decomposition using Caselli's (2005) data set. Using this data does not change the main conclusion. The results are given in the appendix.

As shown in the sixth row, adding a covariance term increases the fraction of income differences explained by both models substantially. However, the main result that the quality of human capital explains a large share of income differences holds through.

Using income ratios as done in the bottom 5 rows produces similar conclusions as those obtained by looking at variance decomposition. For both samples, the fraction of income differences explained by the model increases substantially when the quality of human capital is included.

For the large sample, $(h_{90}/h_{10})/(y_{90}/y_{10})$ is equal to, respectively, 0.07 and 0.23 in the model without and with differences in the quality of human capital. Hence, my finding for

[9] Unfortunately, it is not possible to compare these numbers to the findings of Schoellman (2011) since he does not report results involving income predicted by the model x.

[10] Caselli (2005) uses earlier versions of Penn World Table and the Barro-Lee years of schooling data. He also uses data from an earlier year (1995) and the sample is different. Although the precise reasons for this reduction in the fraction of the variance explained by the baseline model are interesting, it is beyond the scope of this paper to investigate this matter further.

the model with no differences in human capital matches exactly those of Hall and Jones (1999) who also find that $(h_{90}/h_{10}) / (y_{90}/y_{10}) = 0.07$. As for the model where quality differences are accounted for, my findings are again very close to those of Schoellman (2011) who finds that $(h_{90}/h_{10}) / (y_{90}/y_{10}) = 0.21$.

To sum up the main findings, accounting for quality differences seems to increase the explanatory power of the model substantially. This conclusion holds through for different samples and different methods of decomposition.

4 Trends in test scores and representativeness of students

As noted above, I use student test scores from 1995 and later to infer the quality of human capital of the entire labour force in 2007. However, the cognitive abilities of older cohorts could differ from those of younger cohorts. If, for instance, test scores have increased over time, the quality of human capital will be underestimated.

If, for instance, test scores historically have increased faster in poor countries than in rich countries, the estimated differences in human capital quality across the world may be underestimated. Accordingly, in this case the accounting results should be viewed as lower bound estimates of the contribution from human capital quality. Of course, if quality has risen faster in ex ante rich countries, human capital quality differences may be overestimated.

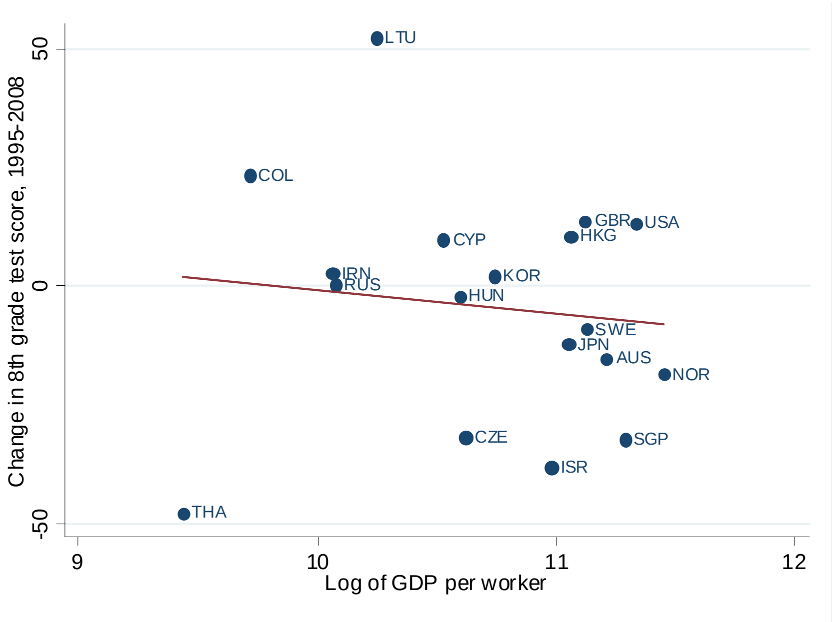
As another example, suppose human capital quality has risen worldwide at a fairly uniform speed. Then human capital quality may be overestimated in richer countries where the population is more mature.

Hence, it is useful to examine whether test scores tend to change over time, and, if so, at a differential speed in rich and poor countries. The following empirical evidence suggests that: 1. The variation in test scores over time is low relative to the variation across countries. 2. Changes in test scores over time are not systematically related to income per capita.

To show the first point I focus on test scores of 8th graders since these are available for most countries and most years. I first compute the average math and science test score for the 8th grade. To find the across-time standard deviation I then compute, for each country, the standard deviation of the test score in 1995 and 2007. The average across-time standard deviation across the 19 countries where data is available is 11. To find the cross-country standard deviation, I first compute the average test score over the years from 1995-2007 ignoring missing observations. The standard deviation of the average test scores across the 62 countries which have at least one observation from 1995-2008 is 61. The standard deviation across the 19 countries which has data for both 1995 and 2007 is 48. It is clear that although the test scores are not completely constant over time the cross-country variation dwarfs the cross-time variation.

To show the second point I look at the change in the average math and science test score for the 8th grade from 1995 to 2007. Figure 1 below plots this change against GDP per capita in 2007.

Figure 6: Log GDP per worker and change in test scores.



Notes: The 8th grade test scores is computed as the average between a math test score and a science test score. The change measured in absolute terms, that is, in points on a scale from 100 to 500.

The slope coefficient is -4.87 and has a t-value of -0.47. Thus, it does not seem that the change in test scores over time is correlated to the level of economic development. If anything, the correlation is negative. In this case, we would expect the quality of human capital to be overestimated in low-income countries, which implies that the analysis above underestimates the importance of human capital quality in accounting for income differences. But the evidence presented in this section suggests that the biases resulting from using students of recent cohorts to infer the average quality of human capital of the labour force should be relatively small and not change the main findings.

5 Conclusion

With a few exceptions, the development accounting literature has so far implicitly assumed that education is equally effective across countries. That is, one year of education in the U.S. correspond to one year of education in Ghana. This paper challenges this assumption.

I use test scores data to estimate the differences in human capital quality, and find that there are large differences in the quality of education throughout the world. In many developing countries the average student needs two years of schooling or more to gain knowledge corresponding to what the average student in the U.S. learns in one year.

Including the quality of human capital in a development accounting exercise increases the fraction of the variance of income explained by the model by around 0.25. The fraction of income differences explained by human capital alone is around 0.26, which is very close to the findings of Schoellman (2011), who uses immigrant data to estimate the quality of human capital.

While the methodology used produces a measure of the overall quality of human capital across countries it has one limitation. It does not illuminate the fundamental causes of quality differences. Future work should concentrate on quantifying the relative importance of parental input, teacher quality, health etc. in explaining cross-country differences in the quality of human capital.

Appendix

A Rescaling of Test Scores

First, four separate linear regressions are run, one for each test type, with the TIMSS (1997a,b) version of the 1995 test scores as left-hand side variables and the TIMSS (2008) version of the 1995 test scores as right-hand side variables:

$$T^{97}_{k,i,95} = B_k + b_k T^{08}_{k,i,95} + \varepsilon_i, \tag{10}$$

where $k = mpri$, $spri$, $msec$, $ssec$ indexes test types and i indexes countries. In all four regressions, the correlation is strong with R^2's in the range of $0.85 - 0.93$. The next step is to use the regression results to convert the TIMSS (2008) version of the test scores from 1999, 2003 and 2007 into the scale of TIMSS (1997a,b). This is done by replacing the TIMSS (2008) values of test scores from 1999, 2003 and 2007 with the predicted values from the regressions, i.e.

$$\hat{T}^{97}_{k,i,t} = \hat{B}_k + \hat{b}_k T^{08}_{k,i,t}. \tag{11}$$

B Table of the Quality of Human Capital for a Cross Section of Countries

Table B: The quality of human capital by country

Country	Iso	q	Std. dev.	q, imputed
Afghanistan	AFG			0.09
Albania	ALB			0.88
Algeria	DZA	0.57	0.03	0.57
Argentina	ARG			0.63
Armenia	ARM	0.85	0.03	0.85
Australia	AUS	1.00	0.03	1.00
Austria	AUT	1.04	0.04	1.04
Bahrain	BHR	0.66	0.04	0.66
Bangladesh	BGD			0.34
Barbados	BRB			0.69
Belgium	BEL			0.94
Belize	BLZ			0.55
Benin	BEN			0.13
Bolivia	BOL			0.55
Bosnia and Herzegovina	BIH	0.79	0.06	0.79
Botswana	BWA	0.48	0.03	0.48
Brazil	BRA			0.54
Brunei Darussalam	BRN			0.73
Bulgaria	BGR	0.98	0.04	0.98
Burundi	BDI			0.09
Cambodia	KHM			0.59
Cameroon	CMR			0.24
Canada	CAN	0.98	0.04	0.98
Central African Republic	CAF			0.12
Chile	CHL			0.65
China	CHN			0.84
Colombia	COL	0.55	0.03	0.55
Congo	COG			0.24
Costa Rica	CRI			0.57
Croatia	HRV			0.88
Cuba	CUB			0.70
Cyprus	CYP	0.77	0.03	0.77
Czech Republic	CZE	1.05	0.03	1.05
Côte d'Ivoire	CIV			0.13
Denmark	DNK	0.85	0.04	0.85
Ecuador	ECU			0.53
Egypt	EGY	0.61	0.04	0.61
El Salvador	SLV	0.51	0.03	0.51
Estonia	EST			0.99
Fiji	FJI			0.91
Finland	FIN			0.89
France	FRA	0.90	0.05	0.90

Country	Iso	q	Std. dev.	q, imputed
Gabon	GAB			0.37
Gambia	GMB			0.12
Georgia	GEO	0.68	0.04	0.68
Germany	DEU	0.97	0.04	0.97
Ghana	GHA	0.30	0.03	0.30
Greece	GRC	0.80	0.03	0.80
Guatemala	GTM			0.31
Guyana	GUY			0.58
Haiti	HTI			0.36
Honduras	HND			0.44
Hong Kong	HKG	1.14	0.04	1.14
Hungary	HUN	1.02	0.03	1.02
Iceland	ISL	0.80	0.03	0.80
India	IND			0.28
Indonesia	IDN	0.64	0.03	0.64
Iran, Islamic Republic of	IRN	0.65	0.02	0.65
Ireland	IRL	0.99	0.04	0.99
Israel	ISR	0.87	0.03	0.87
Italy	ITA	0.93	0.04	0.93
Jamaica	JAM			0.64
Japan	JPN	1.20	0.04	1.20
Jordan	JOR	0.73	0.04	0.73
Kazakhstan	KAZ	1.08	0.07	1.08
Kenya	KEN			0.29
Korea, republic of	KOR	1.26	0.04	1.26
Kuwait	KWT	0.55	0.03	0.55
Lao People's Democratic Republic	LAO			0.54
Latvia	LVA	0.92	0.03	0.92
Lebanon	LBN	0.66	0.04	0.66
Lesotho	LSO			0.29
Libyan Arab Jamahiriya	LBY			0.58
Lithuania	LTU	0.89	0.03	0.89
Luxembourg	LUX			0.86
Malawi	MWI			0.13
Malaysia	MYS	0.91	0.04	0.91
Maldives	MDV			0.35
Malta	MLT	0.83	0.06	0.83
Mauritania	MRT			0.18
Mauritius	MUS			0.51
Mexico	MEX			0.57
Mongolia	MNG			0.76
Morocco	MAR	0.46	0.03	0.46
Mozambique	MOZ			0.04
Namibia	NAM			0.30
Nepal	NPL			0.17
Netherlands	NLD	1.06	0.04	1.06
New Zealand	NZL	0.92	0.03	0.92
Nicaragua	NIC			0.42
Norway	NOR	0.85	0.03	0.85
Oman	OMN	0.57	0.04	0.57
Pakistan	PAK			0.24
Panama	PAN			0.61
Papua New Guinea	PNG			0.55

Table B, continued: The quality of human capital by country

Country	Iso	q	Std. dev.	q, imputed
Paraguay	PRY			0.53
Peru	PER			0.56
Philippines	PHL			0.80
Poland	POL			0.90
Portugal	PRT	0.74	0.03	0.74
Qatar	QAT	0.38	0.03	0.38
Romania	ROM	0.81	0.03	0.81
Russian Federation	RUS	1.02	0.03	1.02
Rwanda	RWA			0.11
Saudi Arabia	SAU	0.48	0.04	0.48
Senegal	SEN			0.20
Serbia	SER	0.84	0.05	0.84
Sierra Leone	SLE			0.08
Singapore	SGP	1.26	0.04	1.26
Slovakia	SVK	1.01	0.04	1.01
Slovenia	SVN	0.99	0.03	0.99
South Africa	ZAF	0.38	0.03	0.38
Spain	ESP	0.84	0.05	0.84
Sri Lanka	LKA			0.49
Sudan	SDN			0.11
Swaziland	SWZ			0.31
Sweden	SWE	0.94	0.04	0.94
Switzerland	CHE	0.98	0.05	0.98
Syria	SYR	0.65	0.05	0.65
Taiwan	TWN	1.21	0.05	1.21
Tajikistan	TJK			0.75
Tanzania, United Republic of	TZA			0.20
Thailand	THA	0.84	0.03	0.84
Togo	TGO			0.22
Tonga	TON			0.85
Trinidad and Tobago	TTO			0.62
Tunisia	TUN	0.57	0.03	0.57
Turkey	TUR	0.72	0.05	0.72
Uganda	UGA			0.16
Ukraine	UKR	0.84	0.04	0.84
United Arab Emirates	ARE			0.64
United Kingdom	GBR	1.00	0.03	1.00
United States	USA	1.00		1.00
Uruguay	URY			0.62
Venezuela, Bolivarian Republic of	VEN			0.48
Viet Nam	VNM			0.72
Yemen	YEM	0.25	0.02	0.25
Zambia	ZMB			0.25
Zimbabwe	ZWE			0.35

C Regression output for sample extension

Table C: Results of proxy regression

	Dependent variable: estimated quality of human capital
GDP per worker	-0.088
	(0.069)
Years of schooling	0.497
	(0.075)
Population density	0.242
	(0.064)
Region dummies	Yes
N	56
R2	0.870

Summary: This table shows the results of the regression used to extend the sample of estimated q's.

Notes: GDP per worker is from Penn World Table v 7.0, average years of schooling of the population above 15 years is from Barro and Lee (2010), Population density is from (www.). Coefficients are standardized. Standard errors are in paranthesis.

D Robustness to choice of data

A surprising result found in Table 3 of the main section is that the baseline model without differences in the quality of human capital is only able to account for a fraction 0.23 of income differences. This number is considerably higher in the literature e.g. in Caselli (2005) who gets 0.4. Given that the Barro-Lee data set as well as the Penn World Table data has been revised several times it is perhaps not so surprising that differences occur. However, it does leave us with the question of whether the main results are robust to using an earlier version of the data.

To show that the main results are not affected by the choice of data I redo the construction of human capital stocks and the development accounting using Caselli's (2005) data for years of schooling, GDP per capita and the capital stock. The accounting results are shown in Table D below. The table shows that the main results persist. For both samples the explanatory power increases considerably when human capital quality is added to the model.

Table D: Decomposition of income differences - using Caselli's (2005) data

Sample:	Only estimated q's (N = 41)		Imputed and estimated q's (N = 89)	
Includes quality of human capital:	No	Yes	No	Yes
	(1)	(2)	(3)	(4)
Var(lny)	0.42	0.42	1.13	1.13
Var(lnh)	0.07	0.17	0.08	0.26
Var(lnx)	0.22	0.34	0.45	0.67
Var(lnh)/Var(lny)	0.16	0.41	0.07	0.23
Var(lnx)/Var(lny)	0.53	0.80	0.40	0.59
(Var(lnx)+Cov(lnA,lnx))/Var(lny)	0.88	1.08	0.73	0.84
y_{90}/y_{10}	3.79	3.79	18.28	18.28
h_{90}/h_{10}	2.04	2.76	2.17	4.15
x_{90}/x_{10}	2.98	3.99	6.64	10.20
$(h_{90}/h_{10})/(y_{90}/y_{10})$	0.54	0.73	0.12	0.23
$(x_{90}/x_{10})/(y_{90}/y_{10})$	0.79	1.05	0.36	0.56

Summary: This table decomposes income differences into different components. It shows that the main results are robust to using the data set of Caselli (2005) for years of schooling, GDP and the physical capital stock.

Notes: y is GDP per capita, h is human capital, x is GDP per capita predicted by the factors-only model. Subscript 90 indicates the 90% percentile and subscript 10 indicates the 10% percentaile. The first two columns show the results for the small sample of countries where data for q is estimated from test scores data. The third and fourth columns show the results for the large sample of countries, where q is imputed. See section on details of how data on q is estimated and imputed. In the first and third columns, human capital is computed using only years of schooling. In the second and fourth columns human capital is computed using quality of human capital and years of schooling.

E Robustness to method of decomposition

In this appendix, I show that qualitatively the main results persist if one uses the alternative method of decomposition based on Hall and Jones (1999) and Klenow and Rodriguez-Clare (1997). In this case, the per capita production function is written in terms of the capital-output ratio instead of the capital-labour ratio.

The main reason for doing this is that in a neoclassical growth model, larger TFP or more human capital per worker increases the capital labour ratio in steady state. Hence, differences in TFP or human capital might be wrongly attributed to differences in capital per worker. To deal with this issue we can write GDP per capita in terms of the capital-output ratio, which is unaffected by changes in TFP in a neoclassical growth model.

In order to do this, first rewrite (4) to get

$$y = A^{\frac{1}{1-\alpha}} \left(\frac{k}{y}\right)^{\frac{\alpha}{1-\alpha}} h. \tag{12}$$

In this case, I define $z \equiv (k/y)^{\alpha/(1-\alpha)}\, h$ to be the contribution from input factors and $B \equiv A^{\frac{1}{1-\alpha}}$ to be the contribution from TFP. To decompose the variance take logs and variance to get:

$$var\left[\ln y\right] = var\left[\ln A\right] + var\left[\ln z\right] + 2Cov\left[\ln B, \ln z\right] \tag{13}$$

This yields an expression identical to (9) with z replacing x.

The results of the variance decomposition based on (13) are shown below. Table D shows that the results are robust. For both samples the explanatory power increases considerably when human capital quality is added to the model.

Table E: Variance decomposition - using capital-output ratio

Sample:	Only estimated q's (N = 47)		Imputed and estimated q's (N = 116)	
Includes quality of human capital:	No	Yes	No	Yes
	(1)	(2)	(3)	(4)
Var(lny)	0.64	0.64	1.65	1.65
Var(lnz)	0.05	0.21	0.13	0.60
Var(lnz)/Var(lny)	0.08	0.33	0.08	0.37
(Var(lnx)+Cov(lnB,lnz))/Var(lnz)	0.19	0.47	0.19	0.51

Summary: This table decomposes income differences into different components. It shows that accounting for the quality of human capital increases substantially the fraction of income differences explained by the model.

Notes: y is GDP per capita, h is human capital, z is GDP per capita predicted by the factors-only model using the capital-output ratio. B is TFP computed as a residual. The first two columns show the results for the small sample of countries where data for q is estimated from test scores data. The third and fourth columns show the results for the large sample of countries, where q is imputed. See section on details of how data on q is estimated and imputed. In the first and third columns, human capital is computed using only years of schooling. In the second and fourth columns human capital is computed using quality of human capital and years of schooling.

References

Banerjee, A. V., and Duflo, E. (2005). "Growth Theory Through the Lens of Development Economics." Handbook of Economic Growth, Volume 1A. Edited by P. Aghion and S. N. Durlauf. Elsevier Science, North-Holland Publishers.

Barro, R., and Lee, J., (2010). "A New Data Set of Educational Attainment in the World, 1950-2010". NBER Working Paper No. 15902.

Bils M. and P.J. Klenow (2000). "Does Schooling Cause Growth?" The American Economic Review 90(5), 1160-1183.

Caselli, F. (2005). "Accounting for Cross-Country Income Differences." Handbook of Economic Growth, Volume 1A. Edited by P. Aghion and S. N. Durlauf. Elsevier Science, North-Holland Publishers.

Card, D. (1999). "The Causal Impact of Education on Earnings." Handbook of Labour Economics, Volume 3. Edited by O. Ashenfelter and D. Card.

Friedberg, R. M. (2000). "You Can't Take It with You? Immigrant Assimilation and the Portability of Human Capital". Journal of Labor Economics 18(2), 221-251.

Galor, O. and Weil, D. N. (2000). "Population, Technology and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." The American Economic Review 90 (4), 806-828.

Gollin, D., Lagakos, D. and Waugh, M. (2011). "The Agricultural Productivity Gap in Developing Countries." Working paper available at

https://sites.google.com/site/davidlagakos/home/research.

Hall, R.E., Jones, C.I. (1999). "Why do some countries produce so much more output per worker than others?" The Quarterly Journal of Economics 114 (1), 83–116.

Hanushek, E.A., Kimko, D.D. (2000). "Labor-Force Quality and the Growth of Nations." The American Economic Review 90(5), 1184-1208.

Hanushek, E.A, Woessman, L. (2009). "Do Better Schools Lead to more Growth? Cognitive Skills, Economic Outcomes and Causation." NBER Working Paper 14633.

Hendricks, L. (2002). "How Important Is Human Capital for Development? Evidence from Immigrant Earnings." The American Economic Review 92 (1), 198-219.

Klenow, P. J. and Rodriguez-Clare, A. (1997). "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?". NBER Macroeconomics Annual 12(1997), 71-103.

Manuelli, R.E. and Seshadri, A. (2006) "Human Capital and the Wealth of Nations." Presented at USC FBE Dept. Macroeconomics & International Finance Workshop.

Nelson, R. R. and Phelps, E. S. (1966). "Investments in Humans, Technology Diffusion and Economic Growth." The American Economic Review 56(1), 69-75.

Psacharopoulos, G. (1994). "Returns to Investment in Education: A Global Update." World Development 22(9), 1325–1343.

Schoellman, T. (2011) "Education Quality and Development Accounting". Forthcoming in Review of Economic Studies.

TIMSS (1997a), "TIMSS Highlights from the Primary Grades." Third International Mathematics and Science Study. Boston College.

TIMSS (1997b), "Highlights of Results from TIMSS." Third International Mathematics and Science Study. Boston College.

TIMSS (2008), "Highlights From TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth and Eighth-Grade Students in an International Context." U.S. Department of Education and National Center for Education Statistics.

UNESCO (2006) "Education for All Global Monitoring Report - Literacy for Life." UNESCO publishing.

Weil, N. D.(2007). "Accounting for the Effect of Health on Economic Growth." The Quarterly Journal of Economics 122(3), 1265-1306.

Chapter 2: Literacy and the Quality of Human Capital in Developing Countries*

Nicolai Kaarsen†

September 2011.

Abstract

This paper estimates the quality of human capital for 34 developing countries. Whereas existing studies have derived quality indices based on either student test scores or immigrant earnings in the U.S. labour market, the present study takes a novel approach by estimating human capital quality from national representative household surveys on years of schooling and literacy. The approach has the significant advantage that quality indices fully reflecting the current labour force can be obtained; in contrast, student test scores only reliably speak to younger individuals in a population, and the human capital quality of people who emigrate to the U.S. may not be representative of the source population. A development accounting analysis of cross-country income differences reveals that the inclusion of human capital quality doubles the contribution from human capital and lowers the contribution from TFP by ten percentage points.

*I would like to thank Carl-Johan Dalgaard for helpful comments and suggestions.

†Contact: University of Copenhagen, Øster Farimagsgade 5B, Building 26, DK-1353 Copenhagen K, Denmark. Nicolai.kaarsen@econ.ku.dk.

1 Introduction

What are the causes of the large differences in income per capita across countries? The literature of development accounting seeks to answer this question. It does so by quantifying the relative importance of a number of proximate factors, such as physical capital and human capital. Hall and Jones (1999) and Caselli (2005) find that these input factors account for some 40% of the cross-country variation in income; the rest remains unexplained.

The present paper investigates what role differences in human capital plays in accounting for income differences. The standard metholology in the development accounting literature uses years of schooling along with micro estimates of the effect of schooling on individual productivity. However, a number of scholars have argued that the time spent in the education system does not fully capture the differences in cognitive abilities across countries (e.g. Hanushek and Kimko (2000), Hendricks (2002), Hanushek and Zhang (2009), Manuelli and Seshadri (2006, 2009), Schoellmann (2011)). The main critique is that years of schooling is only one of many input factors in the formation of productive cognitive skills. Other inputs, such as school and teacher quality or health, are unaccounted for. As a results, the role of human capital in accounting for income differences may be understated.

In the present paper, I seek to estimate the differences in the *quality* of human capital, defined as the cognitive ability at a given level of education. I set up a framework to estimate this variable for a sample of 34 developing countries. I find that there are large differences in the quality of human capital across this sample of countries. One year of schooling in several of the more developed economies corresponds to three or four years of schooling in the group of countries with lowest schooling quality. Furthermore, I incorporate these differences into a development accounting framework and find that the log-variance of human capital stocks doubles compared to the benchmark model with no differences in the quality of education. In my baseline sample, the fraction of income explained by the model increases from 0.34 to 0.44.

To estimate the quality of human capital I compile a cross-country micro data set of literacy and years of schooling consisting of 652945 representative household members. I then estimate the quality of human capital as the literacy gradient in years of schooling; intuitively, human capital quality is lower in places where fewer individuals manage to become literacy for years of schooling given.

The quality of human capital can affect development through several channels. It can affect the individual labour productivity directly by making workers faster and better at performing tasks, or it may affect economic development through aggregate channels such as fertility and technology adoption[1]. I contribute to an emerging literature focusing on accounting for the effect running through individual productivity. Other papers in this literature are: Hendricks (2002), Schoellman (2011), Kaarsen (2011) and Gollin, Lagakos and Waugh (2011).

Compared to Hendricks (2002), Schoellman (2011) and Kaarsen (2011) the value added of the present paper is that the sample used is representative of the labour force of the country in question. In the three papers in the literature, the data used to infer the quality of human capital is not necessarily representative and hence the quality estimates could be biased.[2]

Hendricks (2002) and Schoellman (2011) use immigrant wage data from the U.S. to estimate differences in the quality of education across source countries. A main problem with this is that immigrants are not randomly selected from the source country and thus

[1]Important papers seeking to identify the aggregate impact from human capital quality are Hanushek and Kimko (2000) and Hanushek and Woessmann (2009).

[2]Hanushek and Zhang use representative survey data on schooling and literacy test scores to estimate the quality of schooling for a sample of 13 OECD countries. However, the goal of this paper is not to account for income differences, but to estimate quality-adjusted Mincerian returns and compare them to existing returns.

not representative of the labour force. If the human capital quality of immigrants differs from that of the average worker the results could be biased. Another issue is that human capital is not necessarily perfectly transferable across nation borders[3]. This implies that the returns to schooling will be lower for an immigrant in the U.S. than in the source country, which could bias the results.

Kaarsen (2011) uses student test scores to estimate the quality of education across countries. In this case, the issue is that the test scores data is of a recent date and thus not informative of the cognitive abilities of older cohorts in the labour force who went to school in the period prior to the years where the tests were administered. If the quality of education has changed over time the quality of human capital will differ between younger and older cohorts and the estimates will be biased.

Much like the present study, Gollin, Lagakos and Waugh (2011) use literacy rates to obtain estimates of the quality of human capital in developing countries. In this sense, their approach is very close to mine. However, they use a different methodology and, more importantly, concentrate on explaining the cross-sectoral differences within countries.

How do the results of the present paper compare to those of the existing literature? Hendricks (2002) finds that, compared to existing studies where human capital is constructed using only years of schooling, accounting for the quality of human capital does not increase the fraction of income explained by the model. This contrasts the findings of Schoellman (2011) and Kaarsen (2011). In both papers, the ratio of the log-variance of human capital to the log-variance of income is around 0.25; a large number compared to the around 0.05 which is usually found in models which do not incorporate the quality of human capital.

In the present paper, the ratio of log-variance of human capital to log-variance of income is 0.15 when quality differences are taken into account. Hence, the results are close to

[3] A point made by Friedberg (2000)

those of Schoellman (2011) and Kaarsen (2011), although, quantitatively, the contribution of human capital quality is smaller in the present paper. A possible explanation for this deviation is that both papers in the literature use a different sample than the one used in this paper.

The paper proceeds as follows. The next section estimates the quality of human capital. In the third section, I incorporate these estimates in a development accounting framework. The fourth section estimates the quality of human capital for different cohorts to gauge the change in quality over time. The final section concludes.

2 Estimating the quality of human capital

2.1 Data

I estimate the cross-country differences in the quality of human capital using data on literacy and years of schooling. For this purpose, I compile a micro data set from Multiple Indicator Cluster Survey (MICS) data sets for 34 developing countries. MICS is a series of surveys conducted by the United Nations Children's Fund to investigate health, education and other environmental and social indicators in developing countries. The participating households are selected to represent the population of the country in question.

MICS uses three different questionnaires: one for all household members, one for women and one for children under five years. I use the household data set since this reflects the working-age population in each country. I also choose to use the second round of MICS since this has data for literacy and the level of educational attainment. This round of surveys was carried out in 2000.

The country data sets are available on www.childinfo.org. I merge these to get a final data set of 1394461 household members. I define the working-age population to be individuals between 16 and 65 years old. This narrows the sample down by around 50% to 717596 observations. Around 9% of these have missing data for either literacy or years of schooling and are also removed. The remaining 652945 observations comprise my main sample.

The data consists of a number of variables of which I only use the ones pertaining to years of education, age and literacy. I construct years of schooling using data on the level of schooling attended (primary, secondary, higher etc.), and the number of years completed on that level. The appendix compares the average years of schooling across countries to the same variable constructed by Barro and Lee (2010). I find that the two measures are closely correlated.

Literacy is determined by a question about the household member's ability to read a letter or newspaper. There are three different response options, "does not read at all", "reads with difficulty" and "reads easily". The interpretation of the "reads with difficulty" category could differ from country to country. To avoid this issue I choose to combine "reads with difficulty" and "reads easily" into one category denoted "reads".

Table 1 provides descriptive statistics for each country. The number of observations included in the first column are the number of individuals in my base sample. The second column shows the average number of completed years of schooling. This figure varies substantially between the countries in the sample from 2.2 years in Senegal to over 11 years in Azerbaijan. The third column shows the fraction of persons who can read. Again the variability across countries is high. For some countries the fraction who can read is as low as around 30% and for others it is close to 100%.

Table 1: Descriptive Statistics

Country	Observations	Average years of schooling	% literate	% with < 6 years
Albania	13030	10.0	0.99	0.11
Angola	13770	3.6	0.67	0.72
Azerbaijan	14620	11.3	0.99	0.03
Bolivia	9294	8.2	0.92	0.36
Bosnia and Herzegovina	21770	9.5	0.99	0.15
Burundi	6919	3.2	0.59	0.83
Cameroon	11790	6.7	0.75	0.34
Central African Republic	39190	3.6	0.51	0.68
Comoros	9132	4.1	0.69	0.61
Cote d'Ivoire	23210	3.0	0.44	0.75
DR Congo	25390	5.8	0.71	0.48
Equatorial Guinea	8936	6.0	0.91	0.46
Gambia	11630	3.7	0.38	0.68
Guinea-Bissau	14370	3.3	0.44	0.79
Iraq	49020	5.8	0.73	0.43
Kenya	19000	6.7	0.82	0.31
Lao	18740	3.9	0.70	0.72
Lesotho	16800	6.0	0.88	0.41
Moldova	20980	9.3	0.99	0.03
Myanmar	67240	4.3	0.82	0.68
Niger	10290	2.3	0.29	0.82
Rwanda	9583	3.9	0.70	0.68
SaoTome	6247	5.5	0.93	0.55
Senegal	25690	2.2	0.38	0.74
Sierra Leone	12060	2.7	0.31	0.74
Sudan	69990	4.2	0.58	0.62
Suriname	8582	7.2	0.91	0.28
Swaziland	10690	6.9	0.84	0.34
Tajikistan	13330	9.7	0.98	0.04
Togo	11260	3.2	0.49	0.71
Uzbekistan	11980	10.0	0.99	0.03
Venezuela	9282	7.3	0.97	0.20
Vietnam	19470	6.9	0.90	0.36
Zambia	19660	6.8	0.79	0.30
Mean	19204	5.8	0.73	0.47
Standard deviation	15367	2.5	0.22	0.26

Notes: Observations is number of household members for which average years of schooling and literacy data is available. Average years of schooling is average years of completed schooling for the main sample. % literate is the fraction of household members which are able to read. % with < 6 years is the fraction of household members who completed less than 6 years of schooling.

The final column reports the fraction of individuals who completed less than six years of schooling. Almost all household members who completed sixth grade or above have achieved literacy[4]. For these individuals there is not enough variation in literacy to estimate differences in the quality of human capital. For some countries the fraction of individuals who completed less than sixth grade is very low. In these cases, the estimated quality of human capital is only based on a small fraction of the labour force. However, for many countries a high fraction of individuals have completed less than sixth grade, and we can expect the estimated quality of human capital to be a more precise indicator of the quality of the labour force. Furthermore, Section 3 shows that excluding countries where a low fraction of individuals completed less than sixth grade from the analysis does not change the main conclusions of this paper.

2.2 Estimating the Quality of Human Capital

This subsection estimates the quality of human capital. The literacy measure L_{ij} of individual i in country j takes on the discrete values 0 and 1 corresponding to "does not read" and "reads", respectively. Assume that this is a function of an unobserved index, C_{ij} measuring cognitive abilities. Specifically, when C_{ij} surpasses the threshold λ, $L_{ij} = 1$. Hence, L_{ij} is given by:

$$L_{ij} = \begin{cases} 0 & for \quad C_{ij} < \lambda \\ 1 & for \quad C_{ij} > \lambda \end{cases} \tag{1}$$

[4]The literacy rate for those who completed any grade beyond the fifth is 96%. For the group who completed any grade beyond the fifth the litaracy rate is 99%.

Assume furthermore that cognitive abilities C_{ij} depend on years of schooling s_{ij} and quality of schooling q_i:

$$C_{ij} = \rho + q_i s_{ij} + \varepsilon_{ij}, \tag{2}$$

where ε_{ij} is an individual-specific error term. (1) and (2) lead to the following limited dependent variable specification:

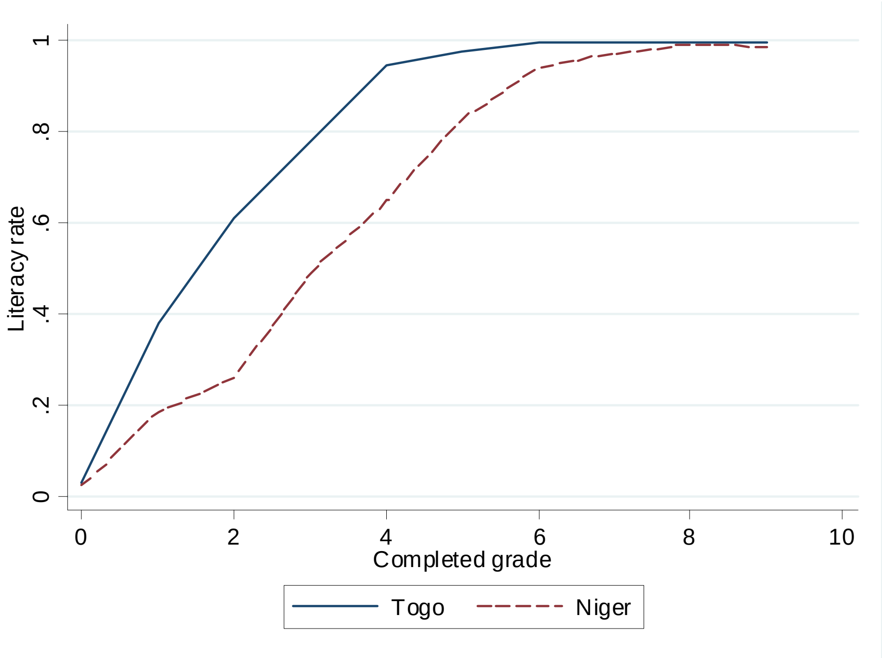
$$P\left(L_{ij} = 1 | s_{ij}\right) = f\left(\rho + q_i s_{ij}\right) \tag{3}$$

This equation can be estimated using logit or probit depending on the assumed functional form of $f(.)$. In the baseline case, I use logit although using probit yields very similar results, as is shown in the appendix[5].

To give an intuitive feel of how the estimation procedure works Figure 1 below shows a plot of literacy by grade for Togo and Niger. In both countries a high fraction of workers completed only sixth grade or less. The average completed years of schooling is 3.2 in Togo versus 2.3 in Niger. The figure shows that for a given grade there are large differences in the literacy rate obtained. At zero years of schooling the literacy rate is virtually zero in both countries. However, the rate of increase in literacy is much faster in Togo. At second grade 60% of the workers in Togo have achieved literacy. Meanwhile, it takes four years of schooling in Niger to obtain the same level of literacy. The estimated quality of human capital measures the speed at which literacy increases with years of schooling.

[5] The pseudo-R^2 is a little higher in the case of the logit estimation, which is why I use this as baseline.

Figure 1: Literacy rate and completed years of schooling for Togo and Niger



Notes: See the section 2 for data source.

The figure also shows that individuals who completed 6th grade and above achieve close to 100% literacy. In this case, we cannot use literacy rates to infer differences in cognitive abilities. Hence, in estimating (3) I follow Gollin, Lagakos and Waugh (2011) in excluding individuals who completed sixth grade or above.

Using logit to estimate (3) results in a series of estimated q's for the 34 countries. Table 2 shows the estimated q's and the standard errors both normalized such that the average q is 1.

Table 2: Quality of Human Capital

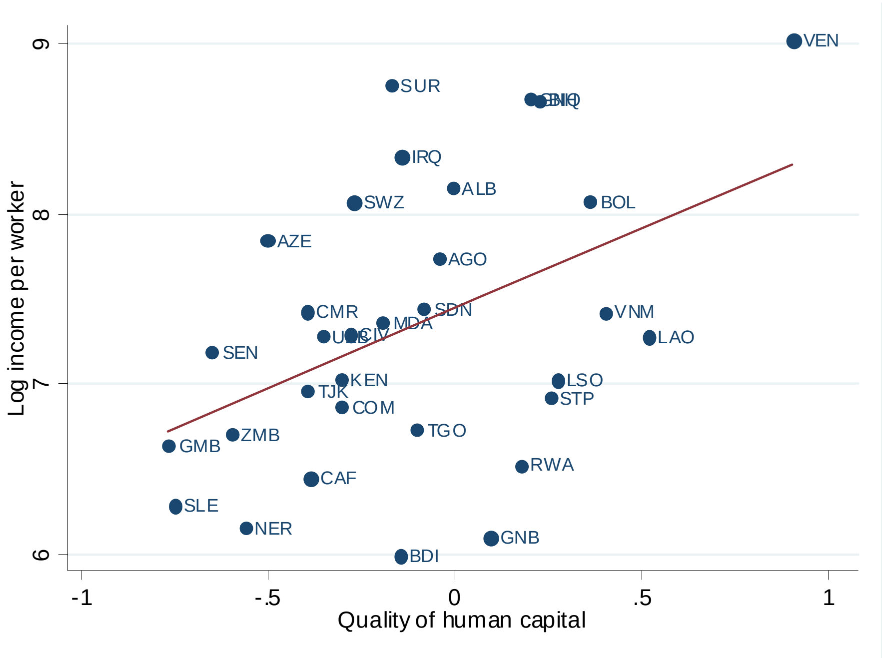
Country	q	Std dev of q
Albania	0.995	0.024
Angola	0.956	0.011
Azerbaijan	0.603	0.010
Bolivia	1.439	0.035
Bosnia and Herzegovina	1.251	0.030
Burundi	0.863	0.014
Cameroon	0.672	0.010
Central African Republic	0.679	0.005
Comoros	0.738	0.015
Cote d'Ivoire	0.753	0.009
DR Congo	0.631	0.005
Equatorial Guinea	1.223	0.025
Gambia	0.465	0.008
Guinea-Bissau	1.099	0.020
Iraq	0.866	0.007
Kenya	0.737	0.008
Lao	1.680	0.029
Lesotho	1.315	0.023
Moldova	0.825	0.017
Myanmar	2.830	0.046
Niger	0.570	0.010
Rwanda	1.193	0.021
SaoTome	1.290	0.031
Senegal	0.521	0.005
Sierra Leone	0.471	0.008
Sudan	0.921	0.007
Suriname	0.844	0.015
Swaziland	0.763	0.012
Tajikistan	0.674	0.013
Togo	0.903	0.016
Uzbekistan	0.702	0.017
Venezuela	2.479	0.136
Vietnam	1.498	0.029
Zambia	0.550	0.005
Mean	1.000	0.020
Standard deviation	0.523	0.023

Notes: The quality of human capital is estimated as the years of schooling gradient in literacy and normalized to a mean of one. See the main text for further information about the data and estimation method.

The quality of human capital varies considerably from country to country. In the low end of the spectrum lies several Sub-Saharan African countries with a q of around 0.5. Compared to these countries the q of Bolivia, Lao and Vietnam are around three times as high. Venezuela has the highest quality of human capital with a q of 2.5. The standard errors of the estimated q's seem fairly small compared to the point estimates and compared to the variation across countries.

Figure 2 plots log of GDP per worker in 2000 taken from Penn World Table v 7.0 versus log of the quality of human capital.

Figure 2: GDP per worker and quality of human capital



The relationship between income and quality is significant and positive[6]. However, the correlation is far from perfect. Some countries, such as Suriname, Equatorial Guinea, Iraq and Swaziland have a quality of human capital which is lower than predicted by log income per worker. Interestingly, these countries all have a high dependence on natural resources. Another group of countries with a low quality of human capital relative to income are former socialist countries such as Bosnia and Hercegovina, Azerbaijan and Albania.

Excluding workers with secondary education or higher underlines a potential problem with the estimation procedure: We only estimate the quality of human capital of individuals who completed less than six years of schooling. This estimate is used to infer the quality of human capital of the average worker. This is problematic if the q of workers who studied more than five years is different than the q of the average worker. In particular, if high-ability individuals study more years the estimate of q will be biased downwards. This is more likely to be a problem in high-income countries where a low fraction of individuals complete less than six years of schooling. In this case, the bias would work against the main result that quality differences are large.

However, to make sure that selection is not an issue to the main results, I redo the accounting exercise below using different samples of countries. In particular, I exclude countries where the fraction of the population who completed fifth grade is below a certain threshold. This should reduce a potential bias since the remaining countries have a more precise estimates of q. I experiment with different values for this threshold and find that the main results persist. In all cases, the variance of human capital increases considerably once the quality of human capital is taken into account.

[6] Venezuela appears to be an outlier. The relationship is still significant at a 10% level when this country is excluded.

3 Income and the Quality of Human Capital

It seems that there are non-negligible differences in the quality of human capital. However, do these quality differences translate into larger differences in human capital? To find out this section constructs quality-adjusted human capital stocks and compares them to benchmark human capital stocks computed using only years of schooling as an input. I find that including the quality of education significantly increases the variance of human capital. Moreover, the fraction of income differences accounted for by the model increases.

3.1 Quality-Adjusted Human Capital Stocks

The approach used in this section is a generalization of the standard framework used in Bils and Klenow (2002). Assume that income per capita y_i is given by a Cobb-Douglas production function:

$$y_i = A_i k_i^{\alpha} h_i^{1-\alpha}, \tag{4}$$

where A_i is technology, k_i is physical capital per worker and h_i is human capital per worker given by:

$$h_i = e^{\phi(s_i q_i)}. \tag{5}$$

I take the human capital production function $\phi(s_i, q_i)$ from Schoellman (2011):

$$\phi(s_i, q_i) = \frac{\theta}{1-\psi} (s_i q_i)^{1-\psi}. \tag{6}$$

This is a generalization of the production function used in Bils and Klenow (2002). To construct human capital stocks we need to calibrate θ and ψ. I calibrate these parameters using evidence on Mincerian returns. The most common reference for Mincerian returns in the literature on development accounting is Psacharopoulos (1993). He surveys the micro literature on returns to years of schooling and finds that the returns are relatively higher in low-income countries. However, in a newer survey, Banerjee and Duflo (2005) find that, on average, Mincerian returns are around 0.1 for developing as well as developed countries. This is the number I use to calibrate (6).

The Mincerian returns are defined as the marginal product of log wages with respect to years of schooling. Assuming a competitive labour market implies that for the individual worker the wage is proportional to his human capital. Hence, Mincerians returns are:

$$MR_i = \frac{\partial \ln h_i}{\partial s_i} = \theta s_i^{-\psi} q_i^{1-\psi}. \tag{7}$$

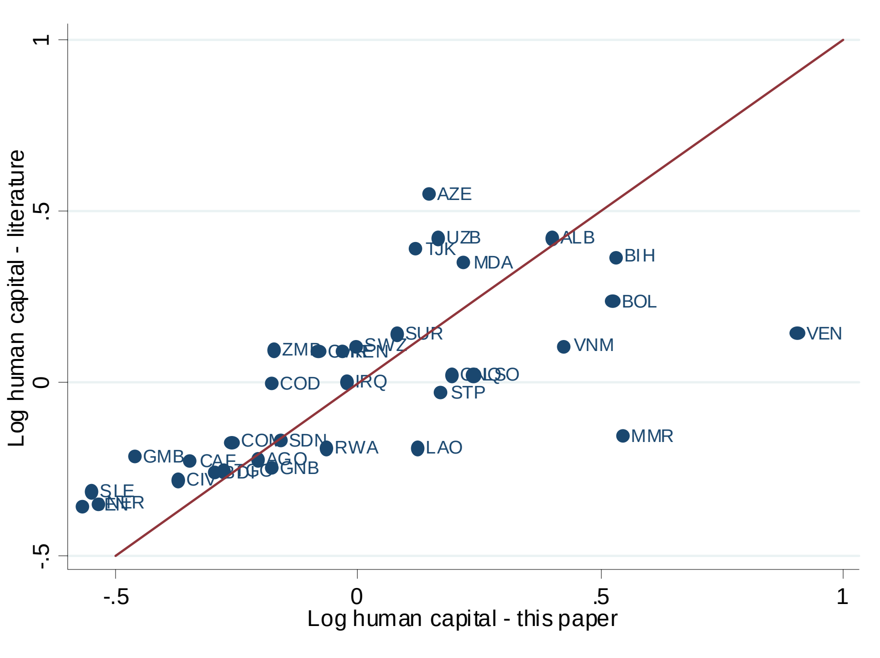
We wish to calibrate θ and ψ such that the Mincerian returns predicted by the model deviates as little as possible from 0.1. A way to do this is to insert $MR_i = 0.1$ into (7) and estimate ψ and θ using nonlinear least squares (NLS). The NLS procedure chooses ψ and θ such that the sum of the squared deviations between 0.1 and $\theta s_i^{-\psi} q_i^{1-\psi}$ is minimized. Doing so yields $\psi = 0.56$ and $\theta = 0.23$[7]. The corresponding standard errors are 0.068 and 0.027, respectively.

With data for s_i and q_i and with ψ and θ calibrated a series for quality-adjusted human capital can be computed using (5) and (6). I also compute an alternative measure of human capital using the standard assumption that there are no differences in the quality of education. This series is computed using $h_i = e^{0.1 s_i}$. Figure 3 below plots the two

[7] Bils and Klenow (2002) uses a very similar approach. Setting $q_i = 1$ in (7) yields the expression for Mincerian returns used by Bils and Klenow (2002). They calibrate the elasticity parameter by regressing $\ln MR_i$ on $\ln s_i$ and backing out ψ from the estimated slope. This amounts to picking the ψ which minimizes the sum of squares of $\ln MR_i - \ln\left(\theta s_i^{-\psi}\right)$, which is log of Mincerian returns obtained in the data minus log of Mincerian returns predicted by the model. Their calibrated value for ψ is not directly comparable to mine since they do not include the quality of human capital. Schoellman (2011) uses a different methodology to calibrate the same production function. His baseline estimate is $\psi = 0.5$ which is very close to my calibration.

alternative series against each other:

Figure 3: Cross-country plot of average human capital constructed using only years of schooling data and human capital adjusted for quality using literacy.



Both series are normalized such that the average of log human capital of the series in question is equal to 0. The figure shows that the variance of human capital increases once quality is taken into account.

For some countries, such as Vietnam, Bolivia and Myanmar, accounting for the quality of human capital substantially increases human capital. Venezuela appears to be an outlier with a large increase in human capital mirroring the large estimate of quality of human capital seen in the previous section. I show below that the main results are robust to using samples where I exclude a number of high-income countries, amongst them Venezuela.

For another group of countries, mostly in Sub-Saharan African, accounting for quality differences implies a decrease in human capital. Furthermore, Azerbaijan, Uzbekistan, Tajikistan and Moldova, all former socialist countries, appear to have a low quality of schooling despite a relatively high quantity of schooling. However, for these countries only a small fraction of workers have completed less than six years of schooling. Hence, because of the above-mentioned selection issue, the quality of human capital for these countries is more likely to be biased downwards. Since human capital not adjusted for quality is high in these countries this would tend to bias the variance of human capital downwards, and hence work against the main finding that human capital differences increase once quality is taken into account.

Table 3 below shows the variance of human capital for the model where only years of schooling are included and the model where differences in human capital quality are accounted for, respectively. For the latter model I also show the results for different parametrizations to check for robustness. My baseline calibration is $\psi = 0.55$ and $\theta = 0.23$ as mentioned above. What should be the upper bound for these parameters? As shown above the standard errors for both parameters are relatively low. However, it turns out that even widely varying calibrations yields similar results. To illustrate this I choose as upper and lower bounds, $\psi = 0.8$ and $\psi = 0.30$, respectively. In each case, I insert the value for ψ in (7) and pick θ by minimizing the sum of squares of 0.1 minus predicted MR to ensure that the calibration is consistent with an average Mincerian return of around 0.1.

Table 3: The variance of human capital

	Literature	This paper		
		Ψ = 0.30	Ψ = 0.55	Ψ = 0.80
Var(ln(h))	0.042	0.106	0.105	0.092

Summary: The table shows that the log-variance of human capital doubles when the quality of human capital is taken into account. It also shows that the variance of human capital changes little with various calibrations of the human capital production function.

Notes: The leftmost column shows the variance of log human capital constructed using only years of schooling as input. The three right-most columns show the variance of log human capital constructed incorporating the quality of human capital.

The table shows that the log-variance of human capital doubles when quality-differences are incorporated into the model. Moreover, the variance of human capital of the model is relatively insensitive to choice of ψ. Hence, all results are from now on based on the calibration where $\psi = 0.55$.

3.2 Development accounting

Taking into account the quality of human capital increases the variance of human capital. But what role do differences in the quality of human capital play in accounting for income differences amongst developing countries? To find out I decompose income differences into contributions from different factors of input. Rewriting (4) yields:

$$\frac{y_i}{\bar{y}} = \frac{A_i}{\bar{A}} \left(\frac{k_i}{\bar{k}}\right)^{\alpha} \left(\frac{h_i}{\bar{h}}\right)^{1-\alpha}, \tag{8}$$

where an upper bar denotes the average across countries. To compute the terms in (8) we need data for h_i, y_i, and A_i. I get data for k_i and y_i from PWT 7.0[8]. As for h_i I

[8]To compute the capital stock I use the perpetual inventory method described in Caselli (2005).

use the two different series constructed above. The first is computed under the standard assumption that there are no differences in the quality of human capital, the second is computed incorporating the quality estimates from above.

Table 4 shows the contribution of input factors to income differences for various countries.

Table 4: Development accounting - income ratios

Includes the quality of human capital			No		Yes	
Country	y	k^{α}	$h^{1-\alpha}$	A	$h^{1-\alpha}$	A
Burundi	0.19	0.39	0.87	0.56	0.83	0.59
Niger	0.23	0.51	0.82	0.54	0.71	0.62
Zambia	0.39	0.71	1.10	0.50	0.90	0.61
Senegal	0.64	0.76	0.82	1.03	0.70	1.20
Angola	1.11	0.98	0.89	1.26	0.88	1.28
Swaziland	1.54	0.96	1.11	1.44	1.00	1.61
Bolivia	1.56	0.86	1.21	1.5	1.37	1.32
Albania	1.67	1.35	1.37	0.90	1.27	0.97
Iraq	2.02	1.12	1.04	1.73	0.99	1.82
Suriname	3.07	1.9	1.14	1.42	1.05	1.53
Venezuela	3.99	1.42	1.14	2.46	1.71	1.64
Average, all 32 countries	1.00	1.00	1.00	1.00	1.00	1.00
Std. dev., all 32 countries	0.97	0.37	0.14	0.49	0.23	0.42

Summary: This table documents that accounting for the quality of human capital increases the role of human capital in accounting for income differences.

Notes: The figures are computed relative to the average of all 27 countries for which data for h, k and y is available. The two columns in the middle show the results for the model where human capital is computed using only years of schooling. The two rightmost columns show the results for the model where years of schooling and the quality of human capital are included.

The table shows the relative contributions of factors of input to differences in GDP per worker. All variables are measured relative to the average of the 27 countries for which data for k, h and y is available. The two first columns show, respectively, GDP per worker and the contribution of capital per worker. The two middle columns show, respectively, the contributions of human capital and TFP for the model where the quality of human capital is not accounted for. The two right-most columns show, respectively, the contributions of human capital and TFP for the model where human capital is accounted for.

The first thing to notice is that although most of the countries are developing countries there are fairly large differences in GDP per worker. In the lower range lie Sub-Saharan countries where income is as low as one fourth or one fifth of the average. In the upper range, a number of countries have income per capita which is several times larger than the average.

A general finding is that, for countries with incomes below the average, human capital relative to the average decreases when quality is taken into account. Moreover, for most countries where income is higher than the average, human capital relative to the average increases. This should not come as a surprise since the previous section showed that the quality of human capital is positively correlated to GDP per capita.

The residual productivity term A can be used to assess the overall ability of the two input factors h and k to account for income differences. The closer A is to one the better are the proximate factors at explaining income differences. In many cases, though not in all cases, including the quality of human capital brings A closer to one. The standard deviation of A falls from 0.49 to 0.42 indicating overall the explanatory power of the model increases.

Interestingly, the TFP of Iraq and Swaziland increases and thus becomes further away from one when quality is accounted for. Apparently these countries are relatively rich but have a relatively low quality of human capital. As explained above, in these countries, natural resources contribute considerably to GDP which is a possible explanation for this finding.

Another way to asses the role of human capital quality in explaining income differences is to decompose the variance of log GDP. To do this first define $x_i = k_i^{\alpha} h_i^{1-\alpha}$. This is income as predicted by input factors k_i and h_i. Use this definition in (4), take logs and the variance to obtain:

$$var\left[\ln x_i\right] = var\left[\ln A_i\right] + var\left[x_i\right] + 2cov\left[\ln A_i, \ln x_i\right]. \tag{9}$$

Table 5 uses this equation to decompose the relative contributions of the various factors. All of the results reported in Table 5 are computed using relatively few countries, and hence should be interpreted with caution.

Table 5: Development accounting - variance decomposition

Sample	Base sample (N=27)		Medium sample (N =24)		Small sample (N=18)	
Includes quality of human capital:	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)
Var(lny)	0.72	0.72	0.52	0.52	0.55	0.55
Var(lnh)	0.04	0.11	0.03	0.07	0.01	0.06
Var(lnh)/var(lny)	0.06	0.15	0.06	0.14	0.03	0.11
Var(lnx)/Var(lny)	0.34	0.44	0.31	0.4	0.31	0.4
(Var(lnx)+Cov(lnA,lnx))/Var(lny)	0.53	0.59	0.48	0.53	0.49	0.53

Summary: This table shows that accounting for the quality of human capital increases the fraction of income differences explained by the model.

Notes: y is GDP per capita, h is human capital, x is GDP per capita predicted by the factors-only model. The first two columns show the results for the base sample of countries where data for y, h and k is avilible. The third and fourth column excludes countries where the fraction of workers who completed 5th grade or less is below 30%. The fifth and sixth column excludes countries where the fraction of workers who completed 5th grade or less is below 40%.

The first two columns of table 5 show the results for the base sample consisting of all countries for which data is available. Column one show the results for the model where human capital is computed using only years of schooling as input, that is, the benchmark model from the literature, and column 2 shows the results for the model where the quality of human capital is included. The table confirms the main finding of Table 3, Table 4 and Figure 2 which is that the variance of human capital increases when quality is incorporated into the model. $var(\ln h)/var(\ln y)$ increases from 0.06 to 0.15. Compared to Kaarsen (2011) and Schoellman (2011) the increase is smaller. They both get a $var(\ln h)/var(\ln y)$ of 0.26 in the model where the quality of human capital is included. However, given that they look at a broader sample of countries it is not surprising that the results deviate.

The two bottomn rows show the fraction of income differences accounted for by the model. The third row uses the measure preferred by Caselli (2005) while the fourth row uses the measure preferred by Klenow and Rodriguez-Clare (1997). In both cases, accounting for the quality of human capital increases the fraction of the log-variance of income explained by the model. Quantitatively, the increase is smaller than the one found in Kaarsen (2011)[9].

The two columns in the middle show the results for the sample where data is available and countries where the fraction of workers who studies less than six years is below 30%. In the two rightmost columns, I further exclude countries where the fraction of workers who completed less than six years of schooling is less than 40%. As mentioned above, I do this because using literacy data in principle only allows me to estimate the quality of schooling for those who completed lower grades. Hence, the results could be biased if the quality of human capital of workers with higher education differs systematically from workers with lower education.

Changing the sample does not seem to change the main results. In both cases the increase in the variance of human capital and the fraction of income differences explained by the model is comparable to the numbers obtained using the baseline sample.

The main finding of this section is that the quality of human capital is a non-negligible factor in the determination of income differences across the sample of low-income countries considered. Quantitatively, the differences in the quality of human capital do not explain as large a fraction of income differences as in two recent papers in the literature. However,

[9]Schoellman (2011) does not report $var(\ln x)/var(\ln y)$.

the results are in the same ballpark. Moreover, the studies in the literature use a broader sample of countries which could explain why the results differ.

4 The Quality of Human Capital over Time

An advantage of using household survey data is that it is possible to estimate the quality of human capital across time. To do this I estimate for each country five different literacy gradients of years of schooling; one for each of the 10-year cohorts from the ages 16-65. In particular, I estimate the following equation:

$$P\left(L_{ijc} = 1 | s_{ijc}\right) = \rho + q_{ic} s_{ijc} + \varepsilon_{ij} \tag{10}$$

where c indexes cohorts. This equation generalizes (3) by allowing the quality of human capital to differ between cohorts. I estimate (10) using logit which yields five estimates of the quality of human capital for each country. To give an overall picture of how the quality of human capital changes over time Table 6 reports a summary of the output.

The upper part of Table 6 shows the quality of human capital across cohorts. The first column shows the results using (10) while the second column adds age and age squared as controls. Age could have several effects on the quality of human capital not related to the changes in quality over time. For example, as workers grow older they may tend to forget knowledge obtained in school generating a negative correlation between quality and age. If quality is positively correlated with longevity, the fraction of low-quality individuals will be smaller in older cohorts. If this effects dominates, age will be positively correlated with quality. Controlling for age filters out such effects. However, if there is a general trend towards higher quality of human capital this could also be captured by controlling for age.

Table 6: The quality of human capital across cohorts

Controls for age	No	Yes
Cohort	Average quality of Human Capital	
16-25	0.95	0.91
26-35	0.96	0.96
36-45	1.01	1.02
46-55	1.03	1.05
56-65	1.05	1.06
Average	1	1
Standard deviation across countries	0.51	0.51
Average across-cohort standard deviation	0.13	0.13

Summary: The table shows that compared to the large differences in the quality of human capital across countries the variability across cohorts is small.

Notes: The table shows the average quality of human capital for 34 countries by different cohorts. In the first column the quality of human capital is estimated as the literacy gradient in years of schooling with no controls. The second column uses the same approach but adds age and age squared as controls. The standard deviation across countries is the standard deviation across countries of the average over cohorts. The average across-cohorts standard deviation is the average over cohorts of the standard deviation across cohorts.

Both columns show that the quality of human capital is increasing in age of cohorts indicating that overall the quality of human capital is decreasing over time[10]. However, the differences between cohorts are relatively small. The increase over time seems to be slightly smaller when age is controlled for.

How large is the variability of the quality of across time versus the differences across countries? To answer this question I compute two additional measures reported in the lower part of Table 6: The standard deviation of the average quality across countries and the average standard deviation of quality across cohorts. To compute the first measure, I first

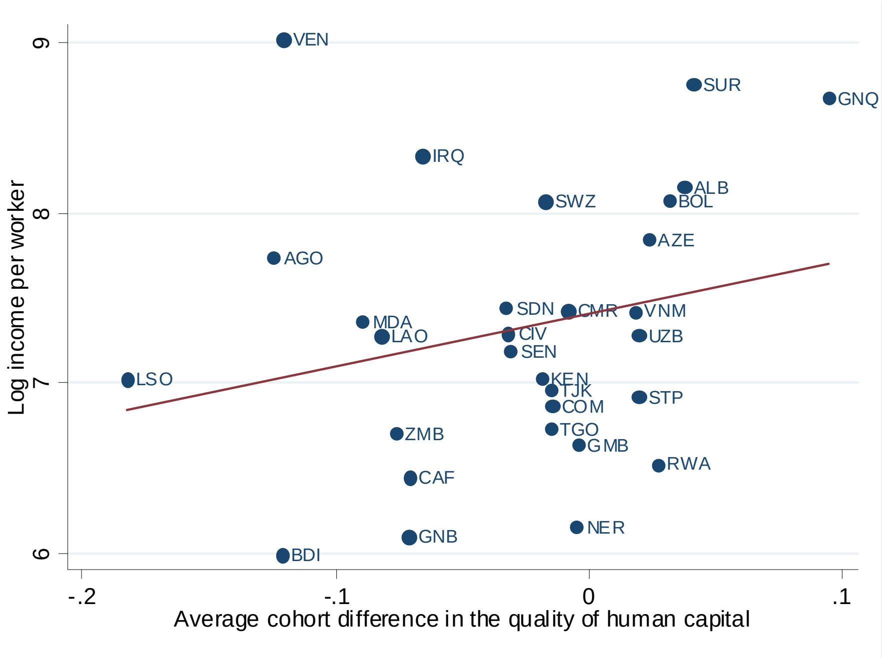
[10]To evaluate whether the quality of human capital for different cohorts are significantly different from each other I estimated a pooled version of (10) where q vary between cohorts but not between countries. This yields five q's, one for each cohort. The joint hypothesis that $q_{16-25} = q_{26-35} = q_{36-45} = q_{46-55} = q_{56-65}$ is rejected at a 0.1% level of significance. For all possible combinaitons of pairs of q's I also test the hypothesis that the two q's are equal. These tests are all rejected with a p-value lower than 0.01% except for the test of $q_{46-55} = q_{56-65}$. The p-value of this test is 30%.

compute the average quality of human capital across cohorts for each country and then take the standard deviation of this series. To compute the second measure, I first compute the standard deviation across cohorts for each country and then compute the average of this series. In both columns the standard deviation across countries is around four times as large as the standard deviation across cohorts. This echoes the finding of Kaarsen (2011) which is that changes in test scores across time are small relative to differences in test scores between countries.

Another interesting question is whether changes in the quality of human capital across time is correlated with income. To answer this question I first compute the difference between the quality of human capital of cohort i and the cohort younger than cohort i. This produces four measures of the change in quality. For each country, I then take the average of these four variables to get an overall measure of the change over time. Figure 4 plots the average difference between old and young cohorts versus log income.

The figure shows that there is a positive correlation between income and the average cohort difference in the quality of human capital indicating that the quality of human capital of richer countries relative to poorer countries is increasing over time. However, the coefficient of the estimated regression line is insignificant and has an R^2 of 0.09. Hence, it does not seem that there is a strong relationship between income and changes in the quality of human capital in this sample. This finding is similar to that of Kaarsen (2011) who finds that changes in test scores over time are uncorrelated to income per capita in a sample consisting mostly of developed economies.

Figure 4: Log income per worker and average difference in the quality of human capital between old and young cohorts



5 Conclusion

The canonical way of measuring human capital within and across countries uses years of schooling to measure cognitive skills. However, cognitive skills obtained from one year of schooling may differ across countries. In this paper, I use a cross-country micro data set to estimate the quality of human capital defined as the literacy gradient in years of schooling.

Comparable papers in the literature estimate the quality of human capital using either immigrant data or test scores data. A main issue with both approaches is that the data used is not necessarily representative of the labour force of the country in question. I avoid this problem by using a sample of representative household members.

My estimates show that there are considerable differences in the quality of human capital across 34 developing countries, and that these differences can account for a non-negligible fraction of observed income differences. The log-variance of human capital doubles when quality differences are accounted for. Furthermore, I include the quality of human capital in a development accounting framework and find that this increases the fraction of income explained by the model by around 0.1 percentage points. Two recent papers in the literature use a similar methodology to estimate the quality of human capital. Compared to the these papers the effect estimated in this paper is smaller but in the same ballpark. A potential explanation for this is that both papers use a broader sample than the one considered in the present paper.

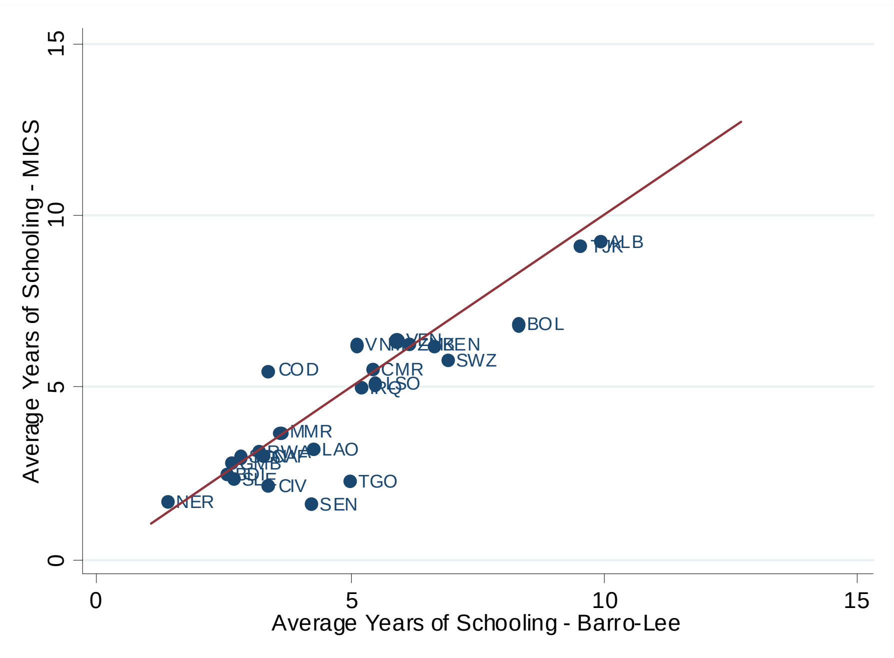
The present paper underlines the need for policy-makers to direct their efforts towards increasing the quantity as well as the quality of education in developing countries. Furthermore, it highlights the necessity for future academic studies to investigate further the relationship between the quality of human capital and economic development.

Appendix

A Comparison of MICS Data to Other Data

In the literature of development accounting the most commonly used data for average years of schooling is taken from Barro and Lee (2010). To ensure that the years of schooling data used in this paper is consistant with the Barro-Lee data Figure A1 below plots the two data series against each other. The figure clearly shows that there is a strong correlation between the two series, this is confirmed by a OLS regression which yields a slope of 0.93 and an R^2 of 0.82.

Figure A: Average years of schooling - MICS and Barro-Lee



B Accounting Results - Probit

In the baseline case I use logit to estimate (3). This approach uses the logistic distribution function. Another option is to use probit which uses the standard normal distribution function. As a baseline, I use logit since the goodness-of-fit as measured by the pseudo-R^2 is a little higher in this case. As a robustness check, I estimate (3) using probit and use the estimated q's from this regression to redo the calibration and variance decomposition. The results of the variance decomposition are given in Table A1 below. The results are very similar to the results of Table 5.

Table A1: Development accounting - probit results

Sample	Base sample (N=27)		Medium sample (N =24)		Small sample (N=18)	
Includes quality of human capital:	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)
Var(lny)	0.72	0.72	0.52	0.52	0.55	0.55
Var(lnh)	0.04	0.09	0.03	0.07	0.01	0.05
Var(lnh)/var(lny)	0.06	0.13	0.06	0.13	0.03	0.09
Var(lnx)/Var(lny)	0.34	0.43	0.31	0.39	0.31	0.39
(Var(lnx)+Cov(lnA,lnx))/Var(lny)	0.53	0.59	0.48	0.53	0.49	0.52

Summary: This table shows that accounting for the quality of human capital increases the fraction of income differences explained by the model.

Notes: y is GDP per capita, h is human capital, x is GDP per capita predicted by the factors-only model. The first two columns show the results for the base sample of countries where data for y, h and k is avilible. The third and fourth column excludes countries where the fraction of workers who completed 5th grade or less is below 30%. The fifth and sixth column excludes countries where the fraction of workers who completed 5th grade or less is below 40%.

References

Banerjee, A. V., and E. Duflo (2005). "Growth Theory Through the Lens of Development Economics." Handbook of Economic Growth, Volume 1A. Edited by P. Aghion and S. N. Durlauf. Elsevier Science, North-Holland Publishers.

Barro, R., and Jong-Wha L, April 2010, "A New Data Set of Educational Attainment in the World, 1950-2010." NBER Working Paper No. 15902.

Bils, M., and Klenow, P.J. (2000). "Does Schooling Cause Growth?". The American Economic Review 90(5), 1160-1183.

Caselli, F. (2005). "Accounting for Cross-Country Income Differences". Handbook of Economic Growth, Volume 1A. Edited by P. Aghion and S. N. Durlauf. Elsevier Science, North-Holland Publishers.

Card, D. (1999). "The Causal Impact of Education on Earnings". Handbook of Labour Economics, Volume 3. Edited by O. Ashenfelter and D. Card.

Friedberg, R. M. (2000). "You Can't Take It with You? Immigrant Assimilation and the Portability of Human Capital". Journal of Labor Economics 18(2), 221-251.

Gollin, D., Lagakos, D. and Waugh, M. (2011). "The Agricultural Productivity Gap in Developing Countries". Working paper available at

https://sites.google.com/site/davidlagakos/home/research.

Hall, R.E., Jones, C.I. (1999). "Why do some countries produce so much more output per worker than others?". The Quarterly Journal of Economics 114 (1), 83–116.

Hanushek, E.A., Kimko, D.D. (2000). "Labor-Force Quality and the Growth of Nations". The American Economic Review 90(5), 1184-1208.

Hanushek, E.A, Woessman, L. (2009) "Do Better Schools Lead to more Growth? Cognitive Skills, Economic Outcomes and Causation". NBER Working Paper 14633.

Hanushek, E.A., Zhang, L. (2009). "Quality-Consistent Estimates of International

Schooling and Skill Gradients". Journal of Human Capital 3(2), 107-143.

Hendricks, L. (2002). "How Important Is Human Capital for Development? Evidence from Immigrant Earnings". The American Economic Review 92 (1), 198-219.

Kaarsen, N. (2011) "Cross-Country Differences in the Quality of Human Capital". Chapter 1 of this thesis.

Klenow, P. J. and Rodriguez-Clare, A. (1997) "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?". NBER Macroeconomics Annual 12(1997), 71-103.

Manuelli, R.E. and Seshadri, A. (2006) "Human Capital and the Wealth of Nations". Presented at USC FBE Dept. Macroeconomics & International Finance Workshop.

Psacharopoulos, G. (1994). "Returns to investment in education: A global update". World Development 22(9), 1325–1343.

Schoellman, T. (2011) "Education Quality and Development Accounting". Forthcoming in Review of Economic Studies.

Chapter 3: Climatic Barriers to the Diffusion of the Neolithic Transition*

Nicolai Kaarsen†

September 2011.

Abstract

This paper investigates the causes of the worldwide differences in the timing of the Neolithic Transition. I find that, in most cases, the transition was caused by diffusion of crops from four different centers of cereal-based agriculture. Hence, the timing of the Neolithic Transition was determined by the speed of the spread of these crops. I then proceed to test two prominent, yet hitherto untested theories that explains the speed of diffusion. The first is that latitudinal distance is more inhibiting to diffusion of crops than is longitudinal distance. The second is that seasonal variation in precipitation is a barrier to diffusion of crops. I find support for both hypotheses.

*I would like to thank Jeanet Bentzen, Carl-Johan Dalgaard, Oded Galor, Louis Putterman, David Weil and seminar participants at Brown University, University of Copenhagen for helpful comments and suggestions.

†Contact: University of Copenhagen, Øster Farimagsgade 5B, Building 26, DK-1353 Copenhagen K, Denmark. Nicolai.kaarsen@econ.ku.dk.

1 Introduction

Around 10000 years ago the first humans started subsisting on agriculture instead of hunted animals and gathered plants. This marked a turning point in the history of mankind dubbed the Neolithic Transition. Population densities increased and humans started living in sedentary communities instead of scattered hunter-gathering bands. Around 4-6000 years later civilization in the form of cities, writing and religion flourished in the same regions where agriculture first arose (Bairoch (1988), Modelski (2003), Morris (2010)).

In his seminal work, Diamond (1997) argues that differences in the timing of the Neolithic Transition played a major role in shaping the contours of world history. The basic idea is that the transition to agriculture triggers a self-perpetuating process of technological advancement and population growth. The earlier the transition, the more time a society has to grow and flourish. Hence, the timing of the Neolithic Transition becomes an important determinant of long-run economic development.

Recently, this theory has received a lot of attention in the economics literature, and there is mounting evidence that the timing of the Neolithic Revolution is an important determinant of historical as well as contemporary economic development.[1] But why did some societies develop agriculture earlier than others? This is the question the present paper seeks to answer.

I find that, in most cases, the Neolithic Transition occured as a result of diffusion of crops from four pristine areas of cereal domestication. Hence, the speed of the spread of crops from these areas becomes an important determinant of the timing of the transition

[1] Galor and Ashraf (2010) find that earlier transition to agriculture is positively correlated with population density and level of technology in year 1500. Olsson and Hibbs (2005), Putterman (2008) and Putterman and Weil (2010) test Diamond's (1997) theory on contemporary data. They find that earlier agriculture is asociated with higher income per capita today. Putterman and Weil (2010) also show that inequality is related to the diversity of the agricultural history of people within a country. Galor and Moav (2007) show that the timing of the Neolithic Revolution is positively correlated with contemporary life expectancy

to agriculture. Using cross-country data I then test and find support for two prominent, but previously untested theories related to crop diffusion[2].

The first is Diamond's (1997) theory that crops spread slower along latitudinal lines than longitudinal lines. The second is Bellwood's (2005) theory that differences in the seasonality of rainfall slow down diffusion of crops. Quantitatively, I find that these theories are about equally important in terms of explaining the variation in the timing of the Neolithic Transition[3]. Diamond (1997) and Bellwood (2005) present anecdotal evidence in favour of these theories, however, as far as I know, this paper is the first to test them systematically using worldwide data and quantify their relative importance.

The idea that the Neolithic Transition spread faster across longitudes than across latitudes is particular important to test since it is a crucial component of Diamond's (1997) hypothesis of why European civilizations flourished compared to societies in other parts of the world. In particular, he argues that the east-west orientation of the axes of the Eurasian continent implied that the diffusion of agriculture occured relatively fast. In contrast, the axes of the American and African continents are predominantly north-south orientated implying that diffusion of agriculture was more slow in these regions. He also argues several post-Neolithic technologies and crops spread faster across longitudes granting Eurasian societies a further advantage.

Why is latitude and seasonality of rainfall particularly important to diffusion of crops? Plants are genetically programmed to grow in a particular climate. Taking a crop to a different climate requires a time-consuming process of adaptation. Hence, diffusion of crops is curbed by large climatic differences.

[2] As to the question of why the initial agriculture arose in the centers, the reader is referred to Weisdorf (2005) for an extensive survey of the literature. This paper describes the diffusion of agriculture once the transition has taken, and I remain agnostic as to the question of why the first farmers started to grow crops.

[3] I also test another of Diamond's (1997) theories, namely the one that differences in altitude curb diffusion of crops. Although the evidence is not as strong as in the case of the two main hypotheses tested, I also find support for this theory.

Absolute latitude affects the seasonal distribution of temperature, sunlight hours, and, to some degree, precipitation. An example of latitude inhibiting diffusion of crops is the slow spread of maize-based agriculture from the subtropical areas in Mexico to the temperate areas in North America (Diamond (1997, ch. 10)).

Within latitudes the seasonal precipitation pattern can vary considerably, and if differences are large, diffusion is curbed. An example of this is the slow diffusion of wheat and barley from the Middle East to South Asia and East Asia (Bellwood (2005, ch. 12)). Wheat and barley are cereals adapted to dry summers and rainy winters; the type of climate seen in Europe, the Middle East and North Africa. East of the Indus valley in Pakistan, the climate is characterized by rainy monsoonal summers and dry winters, which explains the slow diffusion of wheat and barley from the Middle East to South Asia.

Bellwood (2005) presents a number of case studies to investigate the climatic determinants of the rate of spread of agriculture. The average rate of spread suggested by Bellwood (2005) is of similar magnitude as the rate of spread implied by the cross-country data used in this paper.

Olsson and Hibbs (2005) and Putterman (2008) seek to validate another of Diamond's (1997) theories. They show that the date of transition is related to a number of biogeographic variables, such as the number of potentially domesticable plants and animals in the region and the size of the continent. Most of the variables relate to differences in the probability of independent agriculture, whereas I analyse the barriers to diffusion speed. As a result, it is not surprising that the results of the present paper, which relates to diffusion of agriculture conditional on its discovery, are robust to controlling for the factors suggested by Olsson and Hibbs (2005) and Putterman (2008)

The paper proceeds as follows. The next section describes the main hypotheses and sets up the empirical specification. The third section describes the data and documents that the Neolithic Revolution in 146 out of 161 of the countries in the base sample was caused by the spread of crops from four centers of agriculture. In the fourth section, I present the main regression output and the results are compared to the usual findings in the literature. The final section concludes.

2 Background

The purpose of this section is to explain why latitude and seasonality of precipitation should be considered as important determinants of the diffusion of the Neolithic Transition. First, some necessary preliminaries are presented. I then explain the main theories and the empirical specification used to test them.

2.1 Preliminaries: Cultivation and domestication of plants

Domesticated crops are defined by Bellwood (2005, p. 5) as plants

> "...that show recognisable indications of morphological change from the wild phenotype, attributable to human interference in the genotype through cultivation".

The process of domestication began when humans started cultivating, that is planting, protecting, harvesting and sowing, a wild plant. They thereby inadvertedly selected for certain traits changing the genetic structure of the plant over time. An important such trait is loss of seed dispersal. Plants which hold on to their seeds for a longer period are more likely to be harvested and resown. Thus, over time, cultivation changes the genes of the plant until the point where they no longer disperse the seeds. Other domestication traits or markers are larger seeds and more frequent and synchronized germination.

Archeaobotanists are able to distinguish whether these markers are present by analyzing archaeological plant remains. This allows researchers to distinguish between domesticated crops and gathered wild plants. Through radiocarbon dating of plant material it is possible to determine, approximately, the year of domestication of crops in different areas around the world. Bellwood (2005), Smith (1995) and Diamond (1997) review the archaeobotanical literature on plant domestication in a worldwide perspective.

As of today, there is evidence of around 7 independent centers of domestication of crops around the world. From these centers agriculture spread to other areas either through expansion of agriculturalists (demic diffusion) or through adoption by existing hunter-gatherers (cultural diffusion).

Four of these centers gave rise to agriculture based on cereals. Bocquet-Appel and Bar-Yosef (2008) note that systems of agriculture based on cereals spread faster and to larger areas than agriculture based on tubers such as potatoes, manioc or sweet potatoes. A potential explanation is that, in general, cereals have high yields, are easy to collect and can be stored for long periods of time (Heiser (1973, p. 73)). As will be evident in section 3, this is supported by cross-country empirical evidence. In the data section, I document that allmost all countries grew cereals when they made the transition from hunter-gathering to agriculture.[4]

[4] I do not claim that non-cereal crops in general are unimportant for economic development (see e.g. Nunn and Qian (2011) who document that introduction of the potato in the Old World had a positive effect on population density and urbanization). Only that non-cereal crops played a minor role in the transition from hunter-gathering to farming.

2.2 Theory: The Spread of Domestication and Barriers to Diffusion

In this paper I test two prominent hypotheses of crop diffusion. The first is Diamond's (1997) theory that crops spread slower across latitudes than across longitudies. The second is Bellwood's (2005) theory that differences in the seasonality of precipitation slow down crop diffusion. Additionally, I test Diamond's (1997) idea that crop diffusion is hindered by differences in altitude. In the following, I will explain the evolutionary mechanisms behind these hypotheses.

Why is latitudinal distance important to crop diffusion? The forces of natural selection ensures that optimal growth occurs in places where the climate resembles that of the plants' natural habitat. For example, germination is triggered by temperature and sunlight, and subsequent growth depends on a combination of precipitation, sunlight and temperature. If the yearly sequence of sunlight, temperature and precipitation differs from the optimal sequence, plant growth will be inhibited. For instance, some plants have minimum temperature requirements for germination which may not be fulfilled in colder climates.

When plants are grown as crops, farmers inadvertedly harvest individuals which are fast growing, eschewing those which do not germinate or grow slowly. Mutations generating faster plant growth and germination will then be favoured, and over time selection will make sure that the genes of the crops are taylored optimally to the climate it grows in. This means that while it may not be possible to move a crop to a different climate from one day to another, it is possible to gradually move a crop, allowing the genetic structure the time it takes to adapt. Latitude determines the yearly sequence of sunlight and tempature and, to some degree, precipitation. Consequently, it summarizes a number of climatic factors which affect diffusion speed.

Another potentially important factor emphasized by Diamond (1997) is differences in altitude. Differences in altitude can reflect climatic differences and mountains representing physical barriers, both of which could affect the diffusion speed.

Within latitude the yearly pattern of precipitation can vary considerably, and Bellwood (2005) stresses that this is another important barrier to crop diffusion. An example is the case of the slow spread of wheat and barley from the Indus valley and southward. It took these crops around 1500 years to spread from the fertile crescent in the Middle East to the Indus Valley in modern day Pakistan. From there on, however, it took another 3000 years to move the short distance southward and eastward to India. The spread was very slow because India, as most of South Asia and South-East Asia, has a monsoonal rainfall regime with rainy summers and dry winters, while wheat and barley were adapted to winter rain.

There is also reason to believe that oceans could have inhibited crop diffusion. In particular, travelling over larger bodies of water was impossible until the invention of ocean-going vessels. On the other hand Bellwood (2005) notes that once a society had the technologies for long-distance sea travelling, agriculture could spread very fast across the sea indeed.

I use the following specification to test the main theories:

$$T_i = \alpha_0 + \alpha_1 \Delta lat_i + \alpha_2 \Delta lon_i + \alpha_3 \Delta alt_i + \alpha_4 \Delta season_i + \alpha_5' D_i + \alpha_6' X_i + \varepsilon_i, \tag{1}$$

T_i is thousands of years since the transition to agriculture, Δlat_i and Δlon_i are, respectively, the latitudinal and longitudinal distance to the center of origin of agriculture, Δalt_i is the difference in altitude to the center of origin, $\Delta season_i$ is the difference in rainfall seasonality relative to the center of origin, X_i is a set of controls, D_i is a set of dummies and ε_i is a country-specific error term.

In order to understand the logic of this specification, observe that the timing of arrival of the Neolithic Transition at some particular location or country, depends on two things: (1) When was the crops that formed the basis of the transition at location i domesticated for the first time? (2) How fast did the crop spread from their source of origin?

The first component arguably depends on factors such as the number of species (plants, animals) that could be domesticated in particular regions around the world. A simple and parsimonious way of capturing such regional differences in plant and animal availability is to add crop dummies, which capture the individual instances of the independent Neolithic Transition. Hence, in terms of the specification above, crop dummies are included in D_i. In addition, a full set of regional fixed effects are included.

The second source of variation, the speed of diffusion from center of discovery, is the main subject of the present study. In particular, the first main hypothesis is that longer latitudinal distance to the center of agriculture has a negative effect on year of transition and that the effect is stronger than the negative effect of longitudinal distance; that is, that $\alpha_1 < 0$ and $\alpha_1 < \alpha_2$. An additional hypothesis is that differences in altitude slow down diffusion; i.e., that $\alpha_3 < 0$. The second main hypothesis is that a larger difference in the seasonality of precipitation delays diffusion; i.e., that $\alpha_4 < 0$.

The next section deals with the question of how to measure the distance variables and difference in seasonality of precipitation, and presents additional control variables.

3 Data

This section describes the data and falls into three subsections. The first subsection describes the data on the earliest adoption of agriculture and documents that, in most cases, the Neolithic Transition originated from four centers of cereal agriculture; the second subsection describes in detail how the distance to the center of agriculture and difference in rainfall seasonality variables are created; the final subsection deals with all other data.

3.1 Main data: Crop dummies

Where did the Neolithic Revolution originate and how did it spread? The cross-country data from Putterman (2008) on the date of Neolithic Revolution forms the basis of the empirical analysis in this paper. He collects data from surveying the archaeobotanical literature. The date recorded in years BP is the earliest occurence of agriculture within the modern-day borders of the country. However, the data of Putterman (2008) does not indicate where the Neolithic Revolution in a particular country originated from.

To answer this question I construct a number of crop dummies based on Putterman's (2008) online appendix and a number of alternative sources. These dummies are based on the earliest crops found in the country, and can thus be used to infer where agriculture diffused from. If e.g. rice or foxtail millet were the earliest crops found, I conclude that the Neolithic Transition originated from the East Asian center of agriculture in China between the Yellow river and the Yangtze river, where rice and foxtail millet were domesticated.

I try as far as possible to cross-check my findings from Putterman's (2008) data appendix with various papers and books from the archaeobotanical literature. In some cases, Putterman (2008) does not state which crop(s) forms the basis of the Neolithic Transition. For these countries, I rely entirely on sources from the archaeobotanical literature. The details are given in the appendix, and the results of this categorization are shown in the table below.

Table 1: Categorization of the Neolithic Transition Across Countries

Crop system	Region	Center	Cereals	Number of countries
Maize	Americas	The Balsas River Valley, Mexico	Yes	18
Rice and foxtail millet	Asia	Yangtze and Yellow Rivers, China	Yes	19
Sorghum and pearl millet	Sub-Saharan Africa	The Sahel, Africa	Yes	37
Wheat and barley	West Asia, North Africa and Europe	The Levantine Corridor, Isreal	Yes	72
Tropical tubers and fruits	South East Asia, South America and Caribbean	Various centers	No	15
Total				161

Notes: The table accounts for the origins of the Neolithic Transition across countries. The first column shows the main crops involved in the Neolithic Transition. The second and third column show, respectively, the region(s) and the center of origin of the crop system. In the fourth column it is stated whether cereals were part of the Neolithic Transition or not. The fifth column shows the number of countries pertaining to that cathegory.

The table shows that in 146 out of the 161 countries included in Putterman's (2008) data set the Neolithic Revolution was based on cereals originally domesticated in four different centers of agriculture. Agriculture without cereals only appeared in tropical areas in the Amazons, in South East Asia and in Papua New Guinea.[5] In these cases, the main stables were a variety of tubers and fruits. Since these tubers and fruits were only responsible for the Neolithic Revolution in a small fraction of countries in the sample, the subsequent analysis is limited to the 146 cereal-based domestications.

3.1.1 Main data: Distance to crop centers

For each of the four cereal crop systems, I now find the coordinates of the place where the crops were first domesticated according to the archeaobotanical literature. In all four cases, there appears to be consensus about the region of earliest domestication. However,

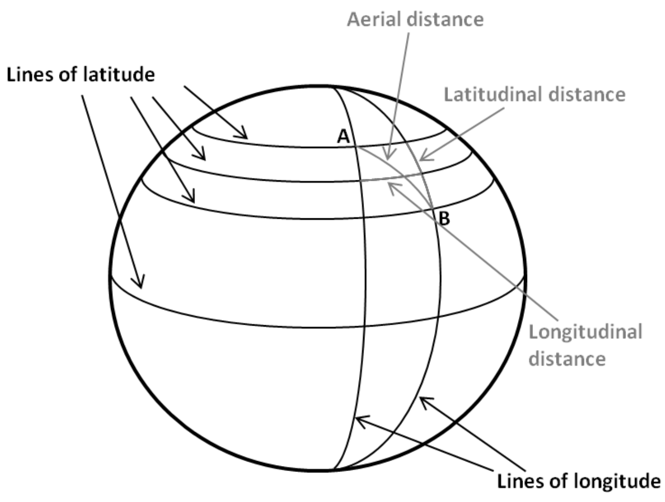
[5] In the case of South East Asia, agriculture was imported from mainland East Asia, however, it was not based on rice or millet as in the rest of Asia, but based on tubers and other tropical crops. It seems that, as the Neolithic Transition diffused south into tropical areas, rice and millet were abandoned in favour of other crops acustomed to the different climate.

the precise site is in many cases subject of some debate. If there are several candidate sites, I use the coordinates of the site situated halfway in between the two sites furthest away. A detalied account of the sites is given in the appendix. I also run robustness checks using different sites as centers of domestication. In practice, for all four centers the alternative sites are relatively close to the respective baseline sites. Consequently, it is no surprise that the results to follow are not sensitive to changing the location of the individual centers.

For each country, I then calculate three different measures of the distance in km to the site where the crops emerged that formed the basis of its own Neolithic Transition: the geodesic or aerial distance, the latitudinal distance and the longitudinal distance.[6]

An illustration of the three distance measures is given in the figure below.

Figure 1: Latitudinal distance, longitudinal distance and aerial distance.



The aerial distance between A and B is the shortest distance. Latitudinal distance

[6]I compute all distance measures in km, which is necessary since the length of one degree longitude is not the same across the world. Computing the distance in degrees would then make it difficult to interpret and compare regression coefficients. All distance measures are calculated in Stata using the geodist module, which is based on Vincenty's (1975) formulae.

is defined as the shortest distance from the line of latitude of the country to the line of latitude of the site of origin. Note, that the shortest distance between two lines of latitude does not change as you move vertically around the earth. However, as for the distance between to lines of longitude, this does not hold. In this case, the distance will depend on the latitude; as you move away from the equator the distance becomes smaller until it reaches zero at the poles where the two lines of longitude converge. Hence, I define the longitudinal distance as the distance between the two lines of longitude evaluated at the latitude coordinate half-way in between the points.[7]

As an example of the calculations of the distance measures consider the case of Denmark. The first agriculture in Denmark was based on wheat and barley originally domesticated in the Fertile Crescent area in the Middle East. The earliest evidence of wheat and barley in the Fertile Crescent is found in a number of sites in the Levant corridor in Israel. The Gesher site lies in the middle of this area, and I consider this to be the founder site of the Fertile Crescent crops. The aerial distance to this site from the geodesic center of Denmark is 3263 km. The latitudinal distance measured as the distance from the line of longitude passing through Gesher to the line of longitude passing through the geodesic center of Denmark is 2599 km. The longitudinal distance is measured as the distance from the latitudinal line passing through Gesher to the latitudinal line passing through Denmark, computed at the latitude halfway in between the two points. I compute it to 2049 km.

Another potential explanatory factor emphasized by Diamond (1997) is difference in altitude. This could curb diffusion since altitudinal differences are correlated with climatic differences. Moreover, altitude differences could reflect physical barriers, such as mountains and hills, which could slow down diffusion. To control for this I computed the altitudinal

[7] I also tried to compute the longitudinal distance using the latitude coordinates of the site of origin of agriculture. This produces a measure very closely correlated to my preferred longitudinal distance measure.

distance in meters from the country to the country of origin of agriculture. The altitude of the country is measured as the altitude averaged over grid cells.[8]

As for the measure of rainfall seasonality, we need a measure that simultaneously captures differences in the timing of precipitation (i.e., does rain fall in the summer or in the winter?) and differences in the variation (i.e., how large is the variation throughout the year?). I have not been able to find such measure in the literature. Hence, I use a novel measure of seasonality computed as the covariance between monthly rainfall and temperature.[9] Countries where precipitation falls predominantly during summer months have a positive correlation between rain and temperature whereas the opposite holds for countries where precipitation falls in the winter.[10] When there is little or no variation in seasons, the covariance will be close to zero. Accordingly, to obtain a measure of how seasonally synchronized country i is to the place of origin of the individual crops, I compute the absolute difference between the covariance between montly rainfall and temperature in country i and the covariance between montly rainfall and temperature in the country of origin.[11]

Since the yearly distributions of temperature and precipitation are partly determined by latitude, one could think that latitudinal distance and differences in the seasonality of rainfall are correlated. If this is the case, including the difference in seasonality will not add

[8] The data for altitude is taken from Andersen et. al. (2011).

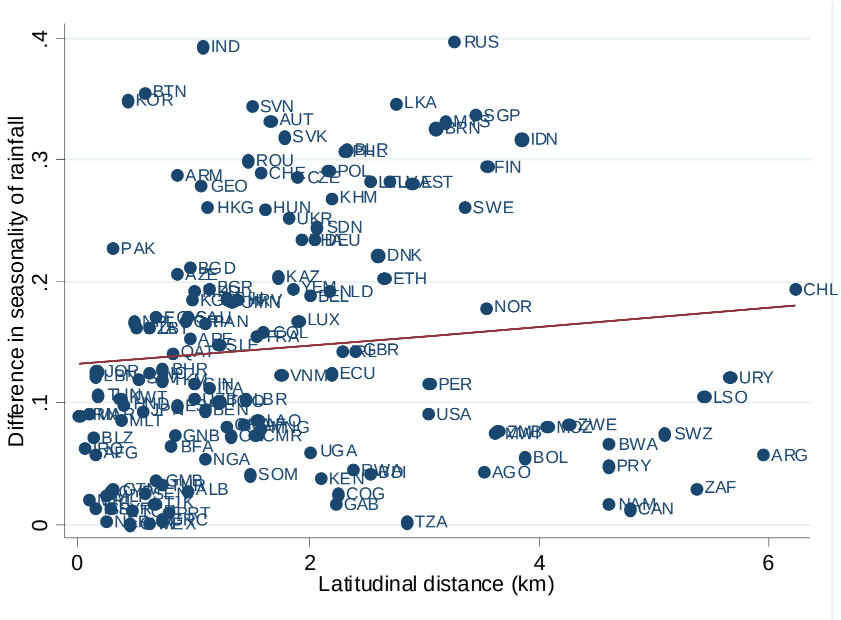
[9] The monthly rainfall and temperature data used to calculate the measure of seasonality is taken from Mitchell et. al. (2003). They provide cross-country data for temperature and precipitation for each month in the years 1961-90. For each monthly series, I average over years. This produces 24 cross-country series of temperature and precipitation, one for each month. These series are used to compute the seasonality measure. Unfortunately it is not possible to get the prehistoric data needed to compute the seasonality measure. Using a recent data is a potential source of measurement error in seasonality if the climate has changed since prehistoric times. However, as long as climatic changes are relatively uniform across countries this should make little difference to the *difference* in seasonality.

[10] It is not possible to use, say, the amount of precipitation in June, July and August as a measure of the amount of summer rain since these months are winter months in the Southern Hemisphere.

[11] In the case of the African crops and the Fertile Crescent crops, the general area of the center of domestication spans three respectively two countries. In these cases, I define the seasonality of the center as the average seasonality of the respective countries spanning the area.

much to our understanding of the differences in the timing of the Neolithic Transition. To check this Figure 1 below plots the difference in seasonality of rainfall versus the latitudinal distance for the countries in the base sample used in the regression in Section 4.

Figure 2: Cross-country plot of latitudinal distance in km to origin of agriculture and the difference in seasonality of rainfall compared to the origin of agriculture.

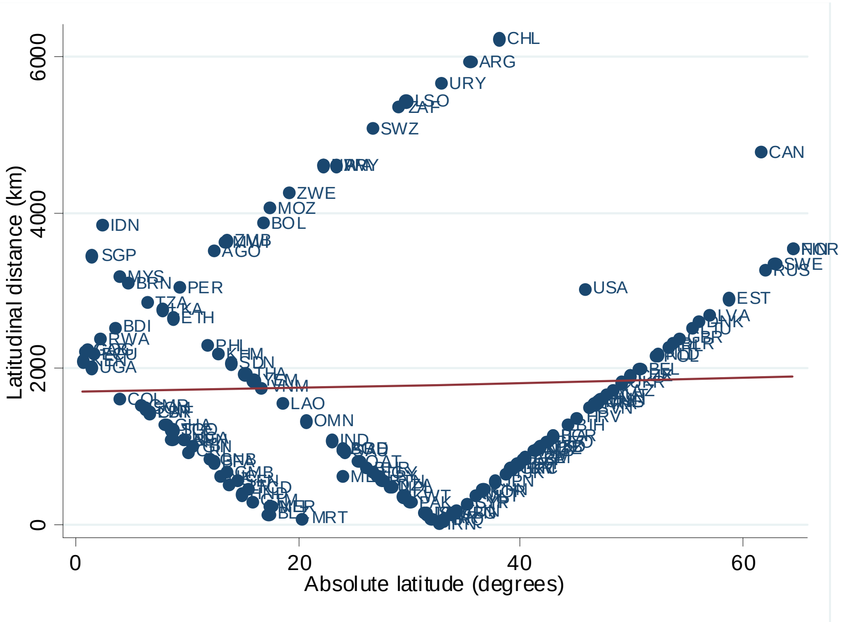


Notes: See the main text for a precise definition of latitudinal distance and difference in seasonality. The regression line has an insignificant slope of 0.0078 and an R^2 of 0.013.

As seen in the figure, the regression line has a slightly positive slope, however, the coefficient is insignificant. Moreover, there seems to be plenty of variation in difference of seasonality of rainfall not captured by latitudinal distance.

Another point of concern is that the latitudinal distance is correlated to absolute latitude. Since absolute latitude is known to be correlated with contemporary measures of development (see e.g. Bloom and Sachs (1997)), it could be that the latitudinal distance just captures underlying climatic factors related to absolute latitude. To check this I plot the latitudinal distance versus absolute latitude.

Figure 3: Cross-country plot of latitudinal distance in km to origin of agriculture and absolute latitude in degrees



Notes: See the main text for a precise definition of latitudinal distance. Absolute latitude is the shortest distance in degrees to the equator. The regression line has an insignificant slope of 2.98 and an R^2 of 0.001.

The figure shows that within some groups of countries there is a linear relationship between latitudinal distance and absolute latitude. This is not surprising. For instance, for all of the countries lying to the North of the center of wheat and barley domestication in Israel, the latitudinal distance will be increasing in absolute latitude. However, the opposite holds for Sub-Saharan African countries south of the equator. Across all of the countries in the sample there is no significant relationship between the two variables.

3.2 Geographic controls

I include a number of control variables to ensure that the parameter estimates of main concern are not tainted by confounding factors.

I construct an island dummy which is equal to one if the country is not connected with any of the continents. The effect of this variable on the timing of the Neolithic Transition could run in both directions. It may inhibit dispersion of agriculture since a society has to develop the maritime technologies needed to travel over oceans. On the other hand, for a society with capacity to travel fast across oceans, diffusion can be speedy. Bellwood (2005, ch. 12) points out that in some cases, such as the case of the Austronesian expansion in Polynesia, crops spread very fast over water, while in other cases, such as the case of England, water is a barrier to crop diffusion.

Olsson and Hibbs (2005) show that a number of biogeographic variables are related to the timing of the Neolithic Transition, and include these variables as controls.

"Number of domesticable plants" and "number of domesticable animals" measure the domestication potential of the flora and fauna in the vicinity of the country. Diamond (1997) states this as one of the main determinants of the timing of the transition. Australia and South Africa have good climatic and geographic conditions for crop growth, however, early societies in these areas failed to develop independent agriculture simply because there were no easily domesticable species.

Furthermore, Olsson and Hibbs (2005) include the variables "climate" and "size of country". These are variables not pertaining directly to Diamond's (1997) theories, however, there are other reasons to expect that they may have an effect on the timing of the Neolithic Transition.[12] "Climate" is the climatic suitability for agriculture based on the Köppen climate zones. It takes on discrete values from 1 to 4, where 1 is least pro-

[12] In the robustness tests in the appendix I include the variables "orientation of continent axis" and "size of continent" also taken from Olsson and Hibbs (2005).

ductive and 4 is most. Better conditions for agriculture could make an early transition more likely. The size of the country is included since a larger landmass of the country could make an early transition more probable. Moreover, the size of a country may matter for the likelihood that early remains from the Neolithic period are excavated and thus recorded. Hence, another motivation for adding land area in the regression is to control for a potential excavation bias.

4 Cross-country empirical evidence

This section tests the theories of crop diffusion described above. The first subsection contains the main regression results which suggest that latitudinal distance and differences in the seasonality of rainfall are important barriers to diffusion of agriculture. In the second subsection, I estimate the speed of crop diffusion and compare it to benchmark findings from the archaeobotanical literature.

4.1 Main results

This subsection tests the main hypotheses of the paper that diffusion of crops are inhibited by latitudinal differences and differences in the seasonality of precipitation. The result is that longer latitudinal distance and larger differences in seasonality significantly decreases years since first adoption of cereals. These results are robust to controlling for a number of geographic and climatic variables. In the appendix, I show that the results are robust to including additional controls and to using alternative sites of origin of agriculture.

For convenience, I repeat the empirical specification given in (1):

$$T_i = \alpha_0 + \alpha_1 \Delta lat_i + \alpha_2 \Delta lon_i + \alpha_3 \Delta alt_i + \alpha_4 \Delta season_i + \alpha_5' D_i + \alpha_6' X_i + \varepsilon_i, \qquad (2)$$

T_i is thousands of years since transition to agriculture, Δlat_i and Δlon_i are, respectively, the latitudinal and longitudinal distance in thousands of km to the center of origin, Δalt_i is the difference in altitude to the center of origin measured in km and $\Delta season_i$ is the difference in rainfall seasonality to the center of origin. D_i is a set of crop and continent dummies, X_i is a vector of controls and ε_i is an error term.

My base sample includes all countries where the Neolithic Transition originated from the four centers of cereal domestication and where data is available. The regression results are shown in Table 2 below. The appendix shows added variable plots for the two main variables.

Table 2: Timing of the Neolithic Transition and climatic barriers to diffusion of crops

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Years since the Neolithic Transition (in 1000s)					
Latitudinal distance (1000 km)	-0.477***		-0.409***	-0.345***	-0.243**	-0.286***
	(0.072)		(0.070)	(0.081)	(0.102)	(0.103)
Longitudinal distance (1000 km)	-0.066		-0.092	-0.117	-0.104	-0.123
	(0.077)		(0.073)	(0.074)	(0.084)	(0.092)
Difference in altitude (km)	-0.497**		-0.386*	-0.771***	-0.860***	-0.563*
	(0.223)		(0.213)	(0.230)	(0.324)	(0.326)
Difference in seasonality		-5.992***	-4.017***	-3.534***	-4.104***	-3.688***
		(1.118)	(1.000)	(1.140)	(1.331)	(1.235)
Absolute latitude				-19.41**	-24.55*	-34.22**
				(9.741)	(12.890)	(14.140)
Longitude				1.916	3.923	6.317
				(5.760)	(7.611)	(7.037)
Altitude (km)				0.530***	0.331	0.285
				(0.177)	(0.221)	(0.217)
Seasonality				-1.054	-0.514	-0.879
				(0.844)	(1.067)	(1.112)
Island dummy						43.720
						(124.400)
Number of domesticable plants						-32.970
						(36.620)
Number of domesticable animals						43.720
						(124.400)
Suitability of climate for agriculture (standardized)						445.9***
						(147.300)
Land area (standardized)						0.00148**
						(0.001)
Observations	139	139	139	139	98	98
R-squared	0.85	0.81	0.87	0.88	0.89	0.91
Crop and continent dummies	Yes	Yes	Yes	Yes	Yes	Yes

Summary: The table shows that longer latitudinal distance and larger difference in seasonality of precipitation between a country and the center of origin of agriculture is associated with later adoption of agriculture. This is evidence in favour of the theory that climatic differences slow down the speed of diffusion of crops and hence delay the timing of the Neolithic Transition.

Notes: Latitudinal distance is shortest distance in 1000s of km to the line of latitude passing through the center of origin of agriculture. Longitudinal distance is the shortest distance to the line of longitude passing through the center of agriculture and measured at the latitude halfway between the points. Difference in altitude is the absolute difference between the altitude of the country and the altitude of the center of agriculture. Difference in the seasonality of precipitation is the absolute difference between the seasonality of the country and that of the center of agriculture, where the seasonality of precipitation is the covariance between montly temperature and precipitation. See the data section for definition of all other variables. The p-value of the F-test of jointly excluding all crop dummies is less than 0.01% in all specifications. The p-value of the F-test of jointly excluding continent dummies is less than 0.01% in all specifications. Standard errors in parentheses. * = significant at the 10% level, ** = significant at the 5% level. *** = significant at the 1% level.

The first column shows that there is a significant and negative effect of latitudinal distance to center of origin on local transition year. Moreover, the coefficent on longitudinal distance is smaller in absolute value than the coefficient on latitudinal distance and the coefficient on altitudinal distance is significant and negative. This is in accordance with Diamond's (1997) hypotheses that crops spread slower across latitudes than across longitudes and that differences in altitude pose a barrier to diffusion.

The second column shows that difference in seasonality of precipitation has a negative and significant effect on year of transition. This affirms Bellwood's (2005) hypothesis that cereals spread slower when differences in the seasonality of rainfall increase.

In column 3, I include all four measures in one regression. Again the coefficients on latitudinal distance and difference in seasonality are significant. Furthermore, the estimated effect of latitude is larger in absolute value than the estimated effect of longitude and the coefficient on difference in altitude is negative and significant on a 10%-level.

A potential concern is that latitudinal distance is correlated with absolute latitude which could affect the timing of the Neolithic Transition through other channels than the one we seek to identify. As shown in Section 3 there seems to be little correlation between latitudinal distance and absolute latitude so this should not be the case. However, to make sure that this is not an issue I control for absolute latitude in column 4. For the same reason I also include longitude, altitude and seasonality of precipitation. The coefficients on latitudinal distance and difference in seasonality remain significant and the magnitudes are relatively unchanged compared to column 3. Moreover, the estimated effect of difference in altitude becomes stronger.

Interestingly, the coefficient on absolute latitude is significant throughout all specifications. This is interesting since measures of contemporary development are usually positively correlated to this variable (see e.g. Bloom and Sachs (1997)); but in prehistory development was apparently greater close to the equator. See also Ashraf and Galor (2011) who document higher population density closer to the equator as of 1500.

Column 6 includes an island dummy and a number of variables from Olsson and Hibbs (2005).[13] The island dummy is potentially an important determinant of the timing of the Neolithic Revolution. As mentioned above, Bellwood (2005) provides examples of fast as well as slow diffusion across bodies of water. In column 6 the island dummy is positive, but insignificant.

Having more domesticable plants and animals should increase the probability of an early independent domestication. However, in column 6 both variables have insignificant coefficients. At first this may seem odd since Olsson and Hibbs (2005) and Ashraf and Galor (2011) find that they are strong predictors of the timing of the Neolithic Transition. However, as explained in Section 2, the crop dummies effectively root out all variation caused by early independent domestication. In this light, it is not so surprising that neither of the variables have significant coefficients.

I also control for two other variables constructed by Olsson and Hibbs (2005). Climate reflects the climate-related potential for agriculture. In places where it is easier to grow crops, one would expect an earlier transition. This prior is confirmed by the regression results which show that climate has a positive and significant effect on year of adoption. Land area is the size of the country measured in 1000 of Ha. A larger country implies a larger probability of early agriculture. As expected, this variable enters with a positive and significant coefficient.

Which of the climatic barriers dominate in terms of economic significance? For the full specification shown in column 6, the standardized regression coefficients of latitudinal distance and difference in seasonality are, respectively, -0.20 and -0.17. Hence, we can conclude that the two measures of climatic barriers are of roughly equal importance. The

[13] The appendix shows the results of including two other variables used by Olsson and Hibbs (2005) which are "size of continent" and "orientation of continent". The results are robust to including these variables.

difference in altitude has a standardized regression coefficient of -0.12 indicating that it is slightly less important than the two other barriers.

A point of concern is that there might be measurement error in the coordinates of the centers of agriculture. Such errors would affect the construction of the main distance variables for a large number of countries. Hence, I checked for robustness of the main results using alternative centers of agriculture. In all cases, the results are very similar to those reported in Table 2. For further details about the robustness check and regression results, I refer the reader to the appendix.

Furthermore, the results are robust to including other geographic controls such as ruggedness and distance to coast. The results of these robustness checks are also in the appendix.

To summarize the main findings: The cross-country evidence presented is in favour of two prominent hypotheses of diffusion of the Neolithic Transition. The first is Diamond's (1997) theory that crops spread slower across latitudes than across longitudes. The second is Bellwood's (2005) theory that differences in seasonality of precipitation curb diffusion of crops. Additionally, I find weak support for Diamond's (1997) theory that differences in altitude delay diffusion of agriculture. These results are robust to controlling for a number of climatic and geographic variables.

4.2 The speed of diffusion: Comparison to the literature

The previous subsection established that the main hypotheses tested in this paper finds support in the data. Another interesting question is whether the speed of diffusion implied by the cross-country data set used is comparable to usual findings in the archaeobotanical literature. Bellwood (2005 p.273) surveys the literature and reports the speed of diffusion in km per year for a number of studies. He finds that pure longitudinal spreads can be as fast as 10 km per year. The lower bound for spread rates is 0.33 km per year in areas with major climatic barriers. In the intermediate cases, the spread rates are around $3-5$ km per year.

These numbers are not directly comparable to the estimated coefficients from Table 2. To get a comparable estimate of the speed of diffusion I estimate the following specification:[14]

$$T_i = \beta_0 + \beta_1 \Delta aerialdist_i + \beta_2 \Delta aerialdist_i \times direction_i + \beta_3' C_i + \epsilon_i, \tag{3}$$

where $\Delta aerialdist_i$ is the aerial distance from country i to the center of origin of the cereal first domesticated in the country and $direction_i$ is the direction of the spread of agriculture, where 0 corresponds to a pure latitudinal spread and 1 corresponds to a pure longitudinal spread.[15] C_i is a set of crop dummies and ϵ_i is an error-term.

$\Delta aerialdist_i$ and $\Delta aerialdist_i \times direction_i$ are closely correlated to Δlat_i and Δlon_i.[16] The only reason to use these variables is that they provide estimates of the speed of diffusion which are directly comparable to the evidence provided by Bellwood (2005).

If $direction_i = 0$ the spread is purely latitudinal, that is, country i is situated on the same line of latitude as the center of origin. Inserting $direction_i = 0$ into (3) shows that, in this case, the effect of $\Delta aerialdist_i$ on T_i is β_1. Accordingly, $-1/\beta_1$ measures the speed of a pure latitudinal spread in km per year.

If $direction_i = 1$ the spread is purely longitudinal, that is, country i is situated on the same line of longitude as the center of origin. In this case, the effect of $\Delta aerialdist_i$ on

[14] The evidence I wish to compare to is anecdotal of nature and hence does not control for additional variables. Hence, I also choose not include controls in this specification. I do, however, include region dummies since they capture differences in the year of transition at the center of agriculture.

[15] This variable is computed as the angle between the line of longitude at the center and the line of the shortest distance from the center to country i divided by 90^o. Thus, if this angle is 0^o, direction is 0 and the spread is pure latitudinal, and if the angle is 90^o direction if 1 and the spread is pure longitudinal.

[16] Regressing $\Delta aerialdist_i$ on Δlat_i and Δlon_i yields an R^2 of 0.98, while regressing $\Delta aerialdist_i \times direction_i$ on Δlat_i and Δlon_i yields an R^2 of 0.99.

T_i is $\beta_1 + \beta_2$. Hence, the speed of a pure longitudinal spread in km per year is given by $-1/(\beta_1 + \beta_2)$.

I estimate (3) using the data set of 139 countries included in the baseline regression in Table 2. The estimated coefficients are $\hat{\beta}_1 = -0.63$ and $\hat{\beta}_2 = 0.49$. The corresponding standard errors are 0.10 and 0.16, respectively. How does these estimates compare to the empirical evidence from the archaeobotanical literature? The pure latitudinal speed and pure longitudinal speed implied by these estimates are 1.59 km per year and 7.14 km per year, respectively. In the intermediate case, the rate of spread is $-1/\left(\hat{\beta}_1 + 0.5\hat{\beta}_2\right) = 2.60$ km per year. These estimates are a bit lower but around the same order of magnitude as those provided by Bellwood (2005).

5 Conclusion

There is mounting empirical evidence that the differences in income per capita observed today have deep historical roots. One important historical determinant of economic development is the timing of the Neolithic Revolution. An earlier transition to agriculture gives a society a head start in a process of cumulative increases in population density and technology. This is supported by contemporary as well as pre-colonial empirical evidence.

This paper investigates empirically the causes of differences in the timing of the transition to agriculture. In most of the cases, the Neolithic Transition was initiated by cereal crops originating from four different hearths of cereal domestication. Hence, the speed of the diffusion from these centers is a crucial determinant of the timing of the Neolithic Revolution.

There are two prominent but previously untested theories in the literature on the spread of agriculture. The first is Diamond's (1997) theory that latitudinal spreads occurred faster than longitudinal spreads; the second is Bellwood's (2005) theory that larger differences in the seasonality of rainfall reduce the speed of diffusion. I use a cross-country data set to test and find support for both hypotheses. The data also supports a third theory which is that differences in altitude curb the spread of agriculture.

Furthermore, I use the data to estimate the speed of diffusion of, respectively, pure latitudinal and pure longitudinal spreads. I find that these estimates are broadly in line with evidence on the speed of diffusion from the archaeobotanical literature.

A Appendix: Data documentation

This appendix documents the coding of the data which forms the basis of the empirical analysis. The first subsection accounts for the coordinates of the centers of cereal agriculture used as a basis for calculating the distance measures and the difference in seasonality. I also describe the alternative sites used to check for robustness. The second subsection documents, for each country, the crops involved in the Neolithic Transition. In this case, my main source is the online data appendix of Putterman (2008) (OAP), which also documents the data for timing of the Neolithic Transition. Since OAP does not provide detailed information about the centers of agriculture I use a number of alternative sources in the first subsection.

A.1 Centers of Domestication

To document the initial centers of agriculture I rely mainly on Smith (1995), Diamond (1997) and Bellwood (2005) who all present thorough surveys of the literature on the early history of agriculture. However, I also supplement these surveys by more recent findings. For each of the four cereal crop systems, I describe the coordinates and seasonality regime of the centre of domestication.

When it comes to cereal agriculture in Africa there seems to be agreement that the cereals sorghum and pearl millet were domesticated somewhere in the Sahel, which is a stretch of savannah south of the Sahara (see (Bellwood 2005, p. 97), Smith (1995, pp. 107), Diamond (1997, p. 388)). What are the coordinates of the earliest domestication of these crops? Both Bellwood (2005, p. 104) and Smith (1995, p. 107) agree that pearl millet was first domesticated around 3500 years ago at Dhar Thichitt in Mauritania ($18.50^o N$, $-9.50^o E$). However, Smith (1995) reports that sorghum was domesticated even earlier at 4000 BP at the Adrar Bous site in Niger ($19.52^o N$, $1.25^o E$) , whereas Bellwood (2005) reports that it was not domesticated until 2000 years ago. In Putterman's (2008) data, the earliest transition in Subsaharan Africa occurs in Niger at 4000 years ago lending support to the first hypothesis.

There does not seem to be agreement about which of these sites is the founder site of cereal agriculture in Sub-Saharan Africa. Meanwhile, there is a general agreement that sorghum and pearl millet was domesticated somewhere in the Sahel area close to these sites. Hence, I choose to use the coordinates of the location halfway in between Dhar Thichitt and Adrar Bous. The coordinates of this site are $19.52^o N$, $1.25^o E$. This places the centre of Subsaharan cereal agriculture in Mali. As a measure of seasonality, I compute the average of the seasonality of Mali, Mauritania and Niger which is 50. As alternative centers I use the coordinates of Dhar Thitchitt and Adrar Bous and the seasonality of Mauritania and Niger, respectively.

In the Americas, maize was the main staple behind the spread of farming. The traditional view has been that maize was first domesticated in the Balsas River Valley in southern Mexico (Smith (1995, pp. 157)). The discussion in Bellwood (2005, pp. 146) confirms that maize indeed is from this region but shows that there is some disagreement as to the exact geographical origin[17]. The latest research (see e.g. Ranere et. al. (2009) or Piperno et. al. (2009)) confirms that maize is from the Balsas River valley. They document that the earliest findings of domesticated maize is from the Xihuatoxtla shelter in the Balsas River Valley, and I use the coordinates from this site ($18.37^o N, -99.41^o E$).

[17]Interestingly, as Bellwood (2005, pp. 146) notes, there also seems to be a general discussion about the chronology of maize. The traditional view as held by Smith (1993) is in favour of late maize domestication around 4000 years BP. This is also the view taken by Putterman (2005) who places the date for domestication of maize in Mexico at 4100 BP. However, the more recent evidence in Ranere et. al. (2009) and Piperno et. al. (2009) suggests that it was domesticated much earlier, 9000 years BP. This date is backed by Matsouka et. al. (2002) who use genetic methods to infer the chronology of maize domestication.

While changes in the chronology could have interesting impacts on the results of e.g. Putterman (2008) and Galor and Ashraf (2010), such endeavours are beyond the scope of the present paper.

As seasonality of the center I use the seasonality of Mexico which is 141. I was not able to find other sites rivalling the early age of the Xihautoxtla site; hence I do not use any alternative sites for maize.

In Asia and Oceania, rice and millets were the primary cereals behind the spread of agriculture originating from the Yangtze River and the Yellow River in China, respectively. Bellwood (2005, p. 119) reports that a single center of domestication seems most likely. Lu et. al. (2010) state that the earliest account of millet is at the Cishan site $(36.34^o N, 114.06^o E)$. Zhao (2010) places the earliest rice domestication at the caves of Xianrendong and Diaotonghuan $(28.72^o N, 116.98^o E)$ further south. These sites are also mentioned in Bellwood (2005, pp. 114) as some of the earliest sites with clear evidence of domesticated cereals. He also reports on several early sites situated somewhere in between the two abovementioned sites where a mix of domesticated millets and rice evolved. This suggests that rice and millets originated somewhere in the area south of the Yellow River and North of the Yangtze River. Hence, I choose to use the coordinates of the point halfway in between the Cishan site and the Xianrendong site $(32.53^o N, 115.52^o E)$. As for seasonality I use the number for China which is 323. The alternative sites used in the robustness checks are the abovementioned Cishan site and Xianrendong site.

As for the Fertile Crescent crops wheat and barley, there seems to be consensus that it arose in the Levantine corridor in present day Israel. Smith (1995) reports that from within a period of 300 years from 10000 BC to 9700 BC domesticated wheat and barley at the sites of Gesher, Aswad, Netiv Hagdud, Gilgal and Jericho. These sites all lie within a narrow north-south oriented band called the Levantine corridor. Furthermore, they all lie within some 200 km of each other. I choose the site of Gesher $(32.60^o N, 35.56^o E)$ in Israel since this lies midways in between the site furthest to the north (Aswad in Syria) and the site furthest to the south (Jericho in Israel). As alternative sites I use the northernmost site Aswad $(33.40^o N, 36.55^o E)$ and the southernmost site Jericho $(31.86^o N, 35.46^o E)$. For seasonality the I use the average of Israel and Syria (-181) as a benchmark. For the two sites alternative I use, respectively, the seasonality of Israel (-194) and Syria (-168).

A.2 Crops involved in the spread of the Neolithic Revolution

This subsection documents the nature of the agricultural transition for a cross-section of countries. For each country, I record the crops involved in the Neolithic Transition. The crops define which center agriculture diffused from.

I base most of the coding on the online appendix of Putterman (2008) (OAP). For some countries where I could not find adequate documentation for the nature of the spread in OAP I consulted secondary sources. In this case, the documentation for the coding is given below. In some countries it is not possible to find secondary sources simply because there are no excavations. In these cases, I follow OAP in using data for neighboring countries.

A.2.1 Pearl Millet and Sorghum: Africa

There is considerable uncertainty about the details of the early agricultural history of Africa mostly due to the paucity of excavations. This is also reflected in the data appendix of OAP, where several of the dates are based on evidence from neighboring countries. For countries where I could not find any evidence of the crop system, I rely on evidence from neighboring countries.

As for the African countries north of the Sahara, that is, Egypt, Libya, Sudan, Algeria, Morocco and Tunisia, as well as one country south of the Sahara, Ethiopia, the Neolithic Transition was based off the Western Asian crop package. This is suggested in OAP and confirmed by Bellwood (2005, p. 97).

In almost all of Africa south of Sahara, except for Ethiopia, it turns out that the earliest agriculture involved either sorghum, millet or a combination of the two. The following account of the earliest crops in the countries of the Sahel, West Africa and Central Africa is based entirely on OAP:

The Sahel: pearl millet was the earliest crop in Mauritania. With no concrete evidence for neighboring Senegal, Gambia and Guinea-Bissau I code these countries as pearl millet. In Mali pearl millet and sorghum was the earliest crop. In Burkina Faso it was pearl millet[18]. In Niger, the earliest crop found was sorghum.

West Africa: In Ghana the earliest agriculturalists appear to have grown teff, sorghum and pearl millet. I also set these crops to Cote D'Ivoire. Evidence suggest that agriculture moved eastwards towards Guinea, Sierra Leone and Liberia. There are no excavations from these sites so I set them to pearl millet and sorghum as these were the crops grown in the neighboring countries.

In Uganda and Kenya agriculture started based on sorghum and millet. There are no excavations in Somalia so I set agriculture to sorghum and millet.

The Bantu farmers was a group migrating from Cameroon and were responsible for the expansion of agriculture to south Africa, that is, the modern day countries of Democratic Republic of Congo, Rwanda, Burundi, Tanzania, Malawi, Mozambique, Botswana, south Africa, Zambia, Zimbabwe, Swaziland, Lesotho, Cameroon, Equatorial Guinea, Republic of Congo, Gabon, Angola, Namibia, Central African Republic, Botswana and south Africa. According to Bellwood (2005, p. 109) the economy of the Bantu farmers was based on sorghum, pearl and finger millet. According to Putterman, millet was the earliest crop in Malawi, Mozambique and Botswana. As for the rest of the south African countries OAP has no direct account of the first crops introduced. Hence, based on the Bellwood (2005)

[18] Putterman reports that the earliest site in Burkina Faso if Ti-n-Akof, but not the crop. However, Breunig and Neumann (2002) attest that the crop at this site was pearl millet.

and the findings of the Bantu expansion, I code the crops of these countries to sorghum and millet.

As for Madagascar, OAP suggests that the earliest agriculture occurred as a result of the so-called Austronesian expansion from South East Asia. Since this expansion was based on a selection of tropical tubers and roots I place Madagascar in this category. This maybe somewhat surprising finding is confirmed by Diamond (1997, pp. 387).

A.2.2 Maize: Americas

In most areas of the Americas maize formed the base of the Neolithic revolution. There seems to be agreement that there were two other independent centers of agriculture one in the Amazon rainforest and one in Eastern North America where a range of tropical non-cereal crops, most notably manioc, was grown. However, maize preempted the spread of these crops in most of the countries in this sample.

OAP reports that the Neolithic Revolution of the Central American countries of Belize, Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica all spread from Mexico; hence, I code them as maize-based.

In South America, Panama, Peru, Ecuador and Colombia Maize was involved in the first agriculture according to OAP. As for Venezuela, Guyana, Suriname, French Guiana and Brazil: OAP declare that agriculture was first based on manioc. Bellwood (2005, pp. 165) is indecisive about whether manioc preceded maize in Amazon Rainforest. I code them as manioc, since this works against the finding of the present paper that cereals were responsible for the Neolithic Revolution in most cases. As for the Caribbean islands, (Puerto Rico, Dominican Republic, Haiti, Trinidad and Tobago, Granada, Barbados, Jamaica, Bahamas, Cuba), OAP reports that agriculture probably originated from mainland tropical areas and was based on manioc.

For Bolivia, Chile, Argentina, Paraguay and Uruguay, OAP does not report the type of agriculture. For Uruguay, OAP uses the Los Ajos site to date the Neolithic Revolution. Iriarte et. al. (2004) confirm that agriculture at this site was maize-based. According to Pohl et al. (2007) a site in southern Peru very close to Bolivia had maize. Hence, I code the Neolithic Transition of Bolivia as Maize based. OAP uses Fiedel (1992) to date the Neolithic revolution of Chile and Argentine. Fiedel (1992, p. 186) states that the early agriculture introduced into these countries was maize. Based on this, I also code Uruguay as maize-based.

As for North America, OAP attests that the earliest agriculture in the USA and in Canada was maize-based.

A.2.3 Rice and Foxtail Millet: Asia and Oceania

The main cereal crops in Asia and Oceania were rice and millet both domesticated in China in the region around the Yangtze and Yellow rivers. OAP is not very explicit in accounting for the crops involved in the spread of the Neolithic Transition, however, he does state that agriculture spread from the homeland in China southward to Taiwan, Hong Kong, Myanmar (Burma), Bhutan, Bangladesh, Vietnam, Laos, Thailand, Cambodia, Malaysia, Brunei, Philippines, Singapore, Brunei and Indonesia . Diamond (1997, p. 344) writes that as the Neolithic transition spread South towards tropical climate the crop package began to rely increasingly on tropical root and tree crops. This is in tune with Bellwood (2005, p. 139) who writes that there is "..."We have no evidence that any of them grew rice beyond the islands of the Philippines and Borneo". However, up until that point, rice was based on crops originated in East Asia. Based on this I record the Neolithic Transition in all of the abovementioned countries as rice and millet-based.

As noted by Bellwood (2005) and Diamond (1997) the agriculture spread further to

New Zealand, the Mariana Islands, and Papua New guinea. Early agriculture in these countries was based on tropical roots and tubers.

OAP states that the Neolithic transition in Korea was based on millet. He also notes that it spread from China to Japan and Mongolia, so I code these as millet- and rice-based. As for Sri Lanka and Nepal, OAP does not report the crops involved. I set it to rice since this was the main crop of early agriculture in the neighboring areas.

A.2.4 Wheat and Barley: Europe, the Middle East and South Asia

The record of the dispersal of the Fertile Crescent crop system based on wheat and barley is well documented. The accounts of Smith (1995, pp. 92) and Bellwood (2005, pp. 67) show that, for all of the countries spanning the area of from the Middle East to the Indus Valley in India to Russia and all of Europe, the Fertile Crescent crops was involved in the Neolithic Transition.

B Robustness

In this appendix section, I first show that the main results are robust to including additional controls, and then show that the main results are robust to using alternative centers of origin as a basis for calculating the distance variables.

As additional controls I include "ruggedness" and "average distance to coast" taken from Nunn and Puga (2010). These measures could influence the speed of the spread by physically making traveling easier or more difficult. Furthermore, I control for two additional variables taken from Olsson and Hibbs (2005):

"Orientation of continent axis" is defined as the continent's length in longitudinal degrees divided by the continent's length in latitudinal degrees. This variable is included to test the theory also tested in this paper that crops diffuse faster across latitudes than across longitudes. If this theory holds, agriculture should spread faster in continents that are more east-west oriented.

"Size of continent" measures the size of the landmass of the continent. If the country is not connected to a continent it measures the size of the country. A larger continent should imply an earlier transition simply because the law of large numbers increases the probability of invention of agriculture.

Table B1 shows that the main results are robust to including these variables.

Table B1: Robustness - additional controls

	(1)	(2)	(3)
	Dependent variable: Years since the Neolithic Transition (in 1000s)		
Latitudinal distance (1000 km)	-0.286***	-0.274**	-0.284**
	(0.103)	(0.104)	(0.117)
Longitudinal distance (1000 km)	-0.123	-0.123	-0.122
	(0.092)	(0.092)	(0.093)
Difference in altitude (km)	-0.563*	-0.586*	-0.605*
	(.326)	(.328)	(.355)
Difference in seasonality	-3.688***	-3.528***	-3.521***
	(1.235)	(1.241)	(1.261)
Absolute latitude	-0.0342**	-0.0370**	-0.0361**
	(0.014)	(0.014)	(0.016)
Longitude	0.006	0.007	0.008
	(0.007)	(0.007)	(0.008)
Altitude (km)	0.285	0.279	0.271
	(0.217)	(0.218)	(0.376)
Seasonality	-0.879	-0.843	-0.897
	(1.112)	(1.120)	(1.164)
Island dummy	-0.471	-7.253	-7.763
	(0.390)	(7.263)	(7.477)
Number of domesticable plants	-0.033	-0.014	-0.013
	(0.037)	(0.039)	(0.040)
Number of domesticable animals	0.044	0.017	0.017
	(0.124)	(0.136)	(0.137)
Suitability of climate for agriculture (standardized)	0.467***	0.473***	0.445***
	(0.154)	(0.154)	(0.164)
Land area (standardized)	0.286**	0.312**	0.349**
	(0.130)	(0.132)	(0.150)
Orientation of axes (standardized)		-0.231	-0.231
		(0.314)	(0.319)
Size of continent (standardized)		-2.010	-2.164
		(2.370)	(2.437)
Average distance to coast			-0.148
			(0.393)
Ruggedness			0.022
			(0.124)
Observations	0	0	0
R-squared	-32.97	-0.12	-0.1
Crop and continent dummies	Yes	Yes	Yes

Summary: The table shows that the main results are robust to including additional control variables.

Notes: Latitudinal distance is shortest distance in 1000s of km to the line of latitude passing through the center of origin of agriculture. Longitudinal distance is the shortest distance to the line of longitude passing through the center of agriculture and measured at the latitude halfway between the points. Difference in altitude is the absolute difference between the altitude of the country and the altitude of the center of agriculture. Difference in the seasonality of precipitation is the absolute difference between the seasonality of the country and that of the center of agriculture, where the seasonality of precipitation is the covariance between montly temperature and precipitation. See the data section and the appendix text for definition of all other variables. The p-value of the F-test of jointly excluding all crop dummies is less than 0.01% in all specifications. The p-value of the F-test of jointly excluding continent dummies is less than 0.01% in all specifications. Standard errors in parentheses. * = significant at the 10% level, ** = significant at the 5% level. *** = significant at the 1% level.

A point of concern is that there is uncertainty about the precise location of the center of origin of the four cereal crops. Since the location of the centers of origin is used to compute the main distance variables, the difference in seasonality and the difference in altitude this could affect the main results. I show that the main results are robust to using alternative centers as a basis for domestication. In particular, the uncertainty applies to the cases of the East Asian, African and Fertile Crescent centers. In the case of the Central American center of cereal domestication, I was not able to find sites as old as the site used in the benchmark case. The description and coordinates of the alternative sites are given in the data documentation data section of this appendix. I recalculate the distance measures and seasonality difference using the methodology described in the Section 3 of the main text. I run one regression per alternative site. The results are shown in the Table B2 below.

The table shows that the main results are robust. Changing the sites of origin seems to have minimal influence on the main parameters of interest. This is not surprising since, in all cases, the baseline sites, geographically, are relative close the the alternative sites.

Table B2: Robustness - alternative centers of domestication

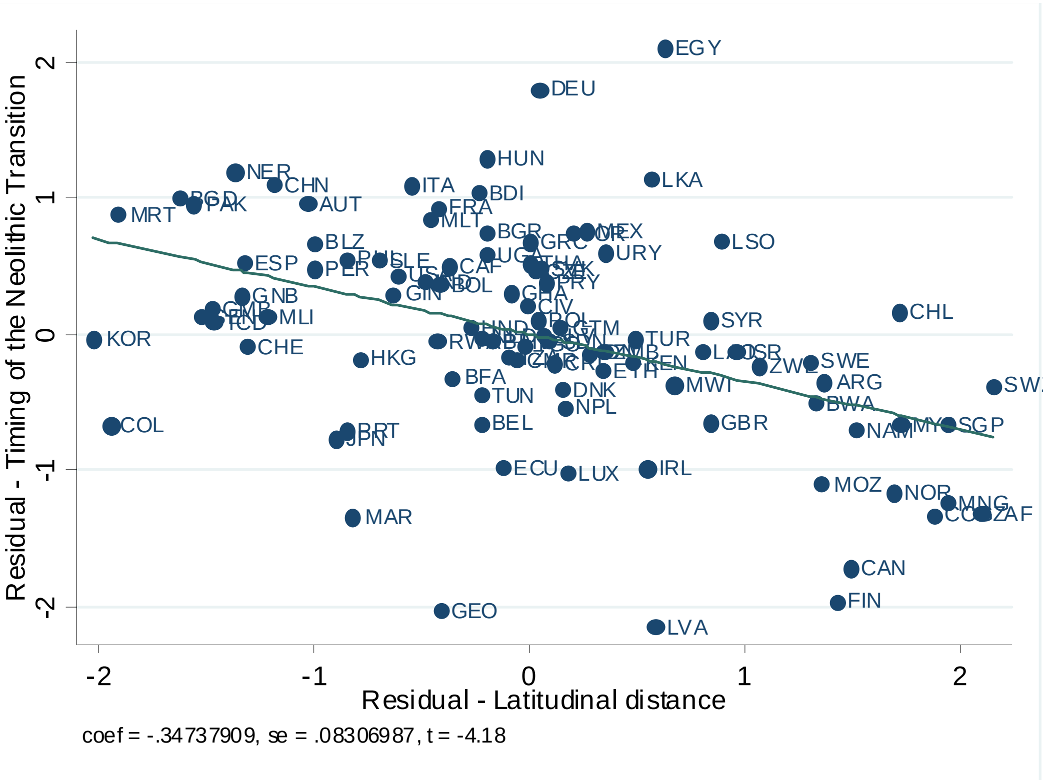
Alternative center for:	Africa		Asia		Europe	
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Years since the Neolithic Transition (in 1000s)					
Latitudinal distance	-0.244**	-0.272**	-0.253***	-0.239**	-0.286***	-0.275***
(1000 km)	(0.108)	(0.104)	(0.095)	(0.094)	(0.102)	(0.102)
Longitudinal distance	-0.138	-0.120	-0.130	-0.125	-0.136	-0.132
(1000 km)	(0.087)	(0.092)	(0.091)	(0.092)	(0.082)	(0.083)
Difference in altitude (km)	-0.581*	-0.572*	-0.573*	-0.584*	-0.540*	-0.562*
	(0.322)	(0.325)	(0.328)	(0.330)	(0.324)	(0.324)
Difference in seasonality	-4.045***	-3.834***	-3.767***	-3.805***	-3.759***	-3.723***
	(1.272)	(1.261)	(1.239)	(1.248)	(1.245)	(1.217)
Absolute latitude	-0.0377**	-0.0346**	-0.0379***	-0.0399***	-0.0340**	-0.0354**
	(0.014)	(0.014)	(0.013)	(0.013)	(0.014)	(0.014)
Longitude	0.009	0.007	0.005	0.004	0.006	0.006
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Altitude (km)	0.321	0.279	0.273	0.265	0.297	0.291
	(0.220)	(0.216)	(0.218)	(0.219)	(0.215)	(0.215)
Seasonality	-0.649	-0.763	-0.684	-0.589	-0.920	-0.794
	(1.122)	(1.121)	(1.098)	(1.091)	(1.102)	(1.106)
Island dummy	-0.412	-0.463	-0.456	-0.479	-0.431	-0.418
	(0.389)	(0.388)	(0.396)	(0.395)	(0.389)	(0.391)
Number of domesticable plants	-0.030	-0.030	-0.031	-0.030	-0.037	-0.036
	(0.036)	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)
Number of domesticable animals	0.068	0.044	0.044	0.046	0.060	0.061
	(0.126)	(0.124)	(0.124)	(0.127)	(0.123)	(0.124)
Suitability of climate for agriculture (standardized)	0.439***	0.447***	0.468***	0.467***	0.470***	0.465***
	(0.153)	(0.155)	(0.155)	(0.156)	(0.153)	(0.154)
Land area (standardized)	0.293**	0.287**	0.279**	0.283**	0.283**	0.288**
	(0.129)	(0.130)	(0.131)	(0.131)	(0.129)	(0.129)
Observations	98	98	98	98	98	98
R-squared	0.91	0.91	0.91	0.91	0.91	0.91
Crop and continent dummies	Yes	Yes	Yes	Yes	Yes	Yes

Summary: The table shows that longer latitudinal distance and larger difference in seasonality of precipitation between a country and the center of origin of agriculture is associated with later adoption of agriculture. This is evidence in favour of the theory that climatic differences slow down the speed of diffusion of crops and hence delay the timing of the Neolithic Transition.

Notes: Latitudinal distance is shortest distance in 1000s of km to the line of latitude passing through the center of origin of agriculture. Longitudinal distance is the shortest distance to the line of longitude passing through the center of agriculture and measured at the latitude halfway between the points. Difference in altitude is the absolute difference between the altitude of the country and the altitude of the center of agriculture. Difference in the seasonality of precipitation is the absolute difference between the seasonality of the country and that of the center of agriculture, where the seasonality of precipitation is the covariance between montly temperature and precipitation. See the data section for definition of all other variables. The p-value of the F-test of jointly excluding all crop dummies is less than 0.01% in all specifications. The p-value of the F-test of jointly excluding continent dummies is less than 0.01% in all specifications. Standard errors in parentheses. * = significant at the 10% level, ** = significant at the 5% level. *** = significant at the 1% level.

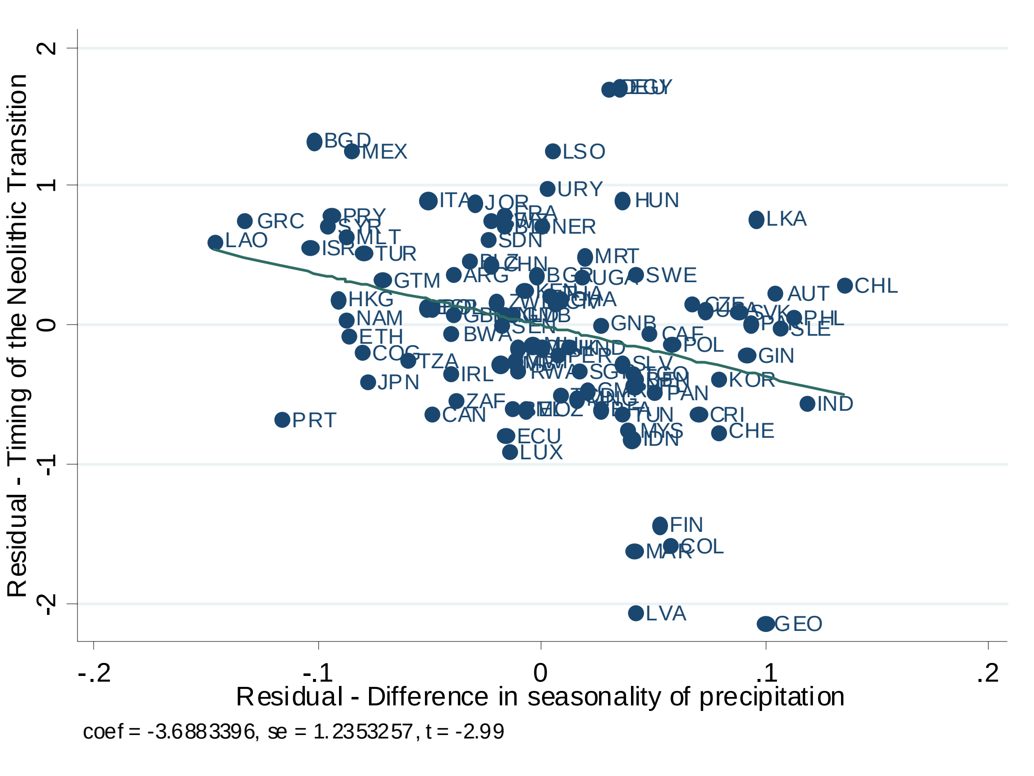
C Appendix: Added variable plots

Figure: A1: The partial effect of latitudinal distance on the timing of the Neolithic Transition



Notes: The figure is an added variable plot, showing the partial effect of the latitudinal distance to the center of agriculture on the timing of the Neolithic Transition controlling for number of dummies and other variables; see column 5 of Table 2 for a full specification. The sample consists of 98 countries.

Figure: A2: The partial effect of seasonality of precipitation on the timing of the Neolithic Transition



Notes: The figure is an added variable plot, showing the partial effect of the difference in seasonality to the center of agriculture on the timing of the Neolithic Transition controlling for number of dummies and other variables; see column 5 of Table 2 for a full specification. The sample consists of 98 countries.

References

Ashraf, Q. and Galor, O. (2011) "Dynamics and Stagnation in the Malthusian Epoch." Forthcoming in the American Economic Review.

Andersen, T. B., Dalggard, C, L. and Selaya, P. (2011). "Eye Disease and Development". University of Copenhagen, Department of Economics Discussion paper 11-22.

Bairoch, P. (1988). "Cities and Economic Development. From the Dawn of History to the Present." Translated by Christopher Braider. University of Chicago Press. Chicago.

Bellwood, P. (2005). "First Farmers. The Origins of Agricultural Societies." Blackwell Publishing.

Bloom, D. E. and Sachs, J. D. (1998). "Geography, Demography, and Economic Growth in Africa". Brookings Papers on Economic Activity, 1998 (2), 207-295.

Bocquet-Appel, J. and Bar-Yosef, O. (2008). "The Neolithic Demographic Transition and Its Consequences." Springer.

Diamond, J. (1997). "Guns, Germs, and Steel. The Fates of Human Societies." W. W. Norton & Co.

Fiedel, S. (1992). "Prehistory of the Americas." Cambridge University Press. Cambridge, UK.

Heiser, C. (1973). "Seed to Civilization". WH Freeman and Co.

Iriarte, J., Holst, I., Marozzi, O., Listopad, C., Alonso, E., Rinderknecht, A. and Montana, J. (2004). "Evidence for Cultivar Adoption and Emerging Complexity During the Mid-Holocene in the La Plata Basin". Nature 432 (2), 614-617.

Lu, H., Zhang, J. ... Li, Q. (2010). "Earliest Domestication of Common Millet (Panicum Miliaceum) in East Asia Extended to 10,000 Years Ago". Proceedings of the National Academy of Sciences 106(16), p. 7367–7372.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M., New, M., (2003). "A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100)". Journal of Climate: Submitted.

Modelski, G. (2003). "World Cities, -3000 to 2000." FAROS 2000. Washington DC.

Morris, I. (2010). "Why the West Rules - For Now." Farrar, Straus, and Giroux.

Nunn, N. and Qian, N. (2011) "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment." The Quarterly Journal of Economics 126, 596-650.

Nunn, N. and Puga, D. (2010). "Ruggedness: The Blessing of Bad Geography in Africa." Forthcoming in the Review of Economics and Statistics.

Olsson, O. and Hibbs, D. A. (2005). "Biogeography and Long-Run Economic Development." European Economic Review 49 (4), p. 909-938.

Putterman, L. (2008). "Agriculture, Diffusion and Development: Ripple Effects of the Neolithic Revolution." Economica 75 (300), p. 729-748

Piperno, D. R., Ranere, A. J., Holst, I. Iriarte, J. Dickau, R. "Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico." Proceedings of the National Academy of Sciences 106(13), p. 5019-5024.

Pohl, M., Piperno, D. R., Pope, K. O. and Jones, J. G. (2007). "Microfossil Evidence for Pre-Columbian Maize Dispersals in the Neotropics from San Andres, Tabasco, Mexico". Proceedings of the National Academy of Sciences 104(16), p. 6870–6875.

Ranere, A. J., Piperno, D. R., Holst, I., Dickau, R. and Iriarte, J. (2009). "The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico." Proceedings of the National Academy of Sciences 106(13), p. 5014-5018.

Smith, B. (1995). "The Emergence of Agriculture." Scientific American Library, New York.

Vincenty, T. (1975). "Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations." Survey Review. 22(176) p. 88-93.

Zhao, Z. (2010). "New data and new issues for the study of origin of rice agriculture in China." Archaeological and Anthropological Sciences 2(2), p. 99-105.