Modernization Magnitude: An Interval Measure Applicable to Post- and Pre-Industrial Societies

Trevor D. Denton
657 Water Street, Peterborough, Ontario, CANADA K9H 3N2: dentont@cogeco.ca

An interval measure of modernization is devised, applicable to pre- and post-industrial societies. The modernization of a society denotes the recency of its form of social organization in human (pre-)history. Murdock and Provost' (1973) ordinal markers of pre-industrial modernization are updated to be interval measures observable today. The recency (in years) of marker gradations is not currently observable in prehistory, but marker gradations are observable in databases such as the pre-industrial "Standard Cross-Cultural Sample" and the World Bank's post-industrial "World Development Indicators." The modernization magnitude of a society is defined to be the mean of the standardized, updated, marker variable measures on the society. The new modernization construct and measure may be used for many purposes, including the testing of behavioral theory spanning post- and/or pre-industrial societies.

1. INTRODUCTION

Massive changes in human societies are known to have occurred over the past 12,000 years. Agriculture, writing and currency began. Stratification, division of labor and population density increased. Local communities became more sedentary and grew in size. The volume of goods and services transported swelled. Integration of local communities into hierarchal political jurisdictions became more frequent. Herbert Spencer, Edward Burnett Tylor, Lewis Henry Morgan, Emile Durkheim, Talcott Parsons, George Peter Murdock and Elman Service all wrote about such transitions. More recent transition theorists include Wilbert Moore and development theorists. Levinson and Malone (1980), Lerner (1968) and Inglehart (2001) discuss the ideas of all these writers and give source citations.

Despite the massive changes accompanying the "modernization" outlined above, there is currently no single interval measure which spans post-industrial and pre-industrial societies. Measures of pre-industrial societies (Levinson and Malone 1980:31-37) are ordinal, although Denton (2007a) makes a case that Murdock and Provost's (1973) scale may be treated as interval in linear models. Pre-industrial ordinal measures do not discriminate between post-industrial societies known to differ in degree of modernization. The United Nations (2007) development codes ($d = 1, 2, 3$) and Irwin's (1975) economic efficiency measure do not apply to pre-industrial societies. Irwin's (1975) measure is of narrowed, economic focus.

Marsh (1967) proposed a scale applicable to both post- and pre-industrial societies. Despite good intentions, the pre-industrial portion of his scale is ordinal. The post-industrial portion is interval. Each portion addresses a narrow domain of modernization which is unconnected to that of the other. In addition, the scale was intended to measure cultural complexity, a construct of questionable validity (*infra*).

The goal of this paper is to create a single interval measure which discriminates between societies at all levels of post- and pre-industrial modernization. The history of the construct and its measurement is outlined. Criteria are proposed which a new measure should satisfy. Murdock and Provost's (1973) 10 modernization markers are redefined to be interval constructs measures which discriminate among pre-industrial and post-industrial societies. To these an 11th modernization marker is added – energy consumption. The resulting composite measure may be applied to post-industrial societies, or pre-industrial societies, or both. It has many new practical and theoretical uses.

Three electronic data bases are used.[1] The World Bank's (2007) World Development Indicators (hereafter WDI) is used for United Nations (2007) member countries at 2000 AD – the most recent marker year for which there are sufficient non-missing data. The Standard Cross-Cultural Sample (hereafter SCCS; Murdock and White 1969; Divale 2004) is used for recent pre-industrial societies.[2] The Atlas of Cultural Evolution (hereafter ACE; Peregrine 2003; based on Peregrine and Ember 2001; 2002) is used for prehistoric, archaeological traditions.

2. MODERNIZATION: DEFINITION AND MEASUREMENT TO PRESENT DAY

Anthropologists and sociologists have offered two loose accounts of the vast changes in human social organization known to have occurred over the last 12,000 years – "cultural complexity" and "recency of social organization" (Denton 2007a; 2004a). Citations for writers mentioned below appear in Levinson and Malone (1980), Lerner (1968) and Inglehart (2001).

Cultural Complexity

Herbert Spencer provided the first "explanation" of the enormous changes in human social organization known to have over the past 12,000 years. Spencer posited a general transition in the universe from unspecialized homogeneity to specialized, interdependent parts. Durkheim's division of labor and Adam Smith's occupational specialization seem intuitive examples. From the 1950's through the 1970's Raoul Naroll, Robert Carneiro and Robert Marsh devised scales which successfully discriminated between societies based on what was assumed to be their degree of cultural complexity. Chick (1997) reviews such measures.

Denton (2004a) criticized the construct validity of "cultural complexity." His main criticism was that there is no ontology of causes and consequences for a unit behavioral concept "cultural complexity." No one knows what brings it about. No one knows how heterogeneous parts impact consequences. In addition, while some facets of society appear to become differentiated (e.g., division of labor), others do not (e.g., kinship).

Today, "cultural complexity" remains a construct with many measures, but no ontology. The measures measure something, but absence of ontology suggests that it is not differentiation of interdependent parts. Cultural complexity lingers in the lexicon of the many archaeologists and cultural anthropologists who use it to discriminate between the peoples they study.

Recency of Social Organization

The second "explanation" of the massive, 12,000-year changes in human social organization is "sequential stages of development." Murdock and Provost (1973:379) wrote

> When anthropologists differentiate cultures in terms of their relative [modernization what] they imply … is their status *vis-à-vis* one another with reference to one or more classificatory criteria which have been postulated to correlate with different levels or stages in cultural development.

The stages to which Murdock and Provost (1973:379) alluded are the well known transitions from foraging to agriculture, from bands to states, from absence to presence of writing, from reciprocity to monetary exchange, from small to large populations, from nomadic to sedentary local communities, and the rest of the transitions described by Elman Service and others.

Given the concept which Murdock and Provost (1973) defined, it seems best to label it "modernization." Other writers who used modernization in the same sense of the word include Wilbert Moore, Talcott Parsons and Daniel Lerner. All the scales used to measure pre-industrial "cultural complexity" (Levinson and Malone 1980:31-37) also discriminate (ordinally) between pre-industrial societies based on the recency of their social organization. The subject areas which mark transitions in pre-industrial social organization also mark changes in contemporary societies (Denton 2007a), adding credence to Murdock and Provost (1973).

For modernization, there is the start of a respectable ontology. We make no effort to find the causes and consequences of a unit behavioral concept "differentiation of interdependent parts." Instead, we focus on tracing the diverse causes and consequences of stages of food getting, of political organization, of population, of monetary exchange and the like, stages pinpointed by Murdock and Provost (1973) and others as having evolved in sequence.

Related Notions

Behrman (2001) and Inglehart (2001) outline additional notions of societal transition. Development economics of the former is restricted to contemporary countries. Political attitude change of the latter is too recent to be claimed to be a long range trend (*infra*). The sense in which "modernism" (Ong 2001) has been used bears no relation to the notion "modernization" addressed here.

3. MARKERS OF MODERNIZATION

In what follows, "societies" are the units of analysis and observation. Societies are defined to be politically autonomous geographical territories. SCCS data cases and WDI UN member countries at 2000 AD may be treated as societies (Denton 2007b). The latter databases may also be treated as consisting of the cultures attached to the societies enumerated. ACE units of analysis are archaeological traditions. Measures on ACE modernization markers behave largely as those on societies. [2]

Criteria

Denton (2007a) proposes six criteria which a set of p conceptual variables S_1, S_2, …, S_p should meet if they are to qualify as markers of modernization (i.e., recency of social organization). These six criteria formalize ideas implicit in Murdock and Provost (1973).

1. The marker variables are widely accepted as distinguishing between societies in terms of known, sequential stages of social organization.

2. Each marker variable is (or may be) measured on the same interval scale. For example, if there are p marker variables S_j (upper case), j is a counter running from 1 to p, measures s_j (lower case) of marker variable S_j (upper case) may be transformed so $z_j = (s_{j_j} - \bar{s}_j) / SD_{Sj}$. Since each transformed marker measure z_j has mean zero and unit variance (Spiegel *et al*. 2000), each is scaled the same.

3. The world expectation of the sum of the p marker variables, over all n societies in the world at time t, is a monotonic increasing function of time t following 10,000 BC. [$\xi[Z] = [(1/n)\Sigma_{i=1}^{n}(1/p)\Sigma_{j=1}^{p} z_{ij}$ = f(t) is monotonic increasing 10,000 BC < t < present time]

4. The world expectation of each individual marker variable j is a monotonic, increasing function of the sum of the remaining marker variables. [$\xi[Z_j]$= g([$\xi[Z]$ is monotonic increasing, calculation of $\xi[Z]$ omits z_j]

5. If there are p marker variables of modernization, the expectation of each individual marker variable is a monotonic increasing function of every other marker variable. [$\xi[Z_j]$ = h($\xi[Z_k]$) is monotonic increasing, j = 1, 2, …, p , k = 1, 2, …, p , j ≠ k]

6. The world expectation of each individual marker variable is a monotonic increasing function of time t following 10,000 BC. [$\xi[Z_j] = (1/n)\Sigma_{i=1}^{n} z_{ij}$ = i(t) is monotonic increasing 10,000 BC < t < present time]

The preceding six criteria permit that any of the listed expectations may approach a lower or upper limit (asymptote) as time t approaches 10,000 BC or present day. For example, the distinction between state and non-state political organization discriminates between societies at early stages of modernization. At later stages an asymptote is reached at which all societies become states and the probability P(State) → 1 (approaches 1). As a result, some marker variables may be useful in discriminating societies at early stages, or late stages, but not both. As a set, however, the modernization marker variables will not be useful unless the set discriminates between societies over all stages of modernization from 10,000 BC to present day.

11 Markers of Modernization

In 1973 Murdock and Provost' proposed 10 scales which discriminate between SCCS societies known to differ in their stage of modernization. The 10 scales appear here in Table 1. Each of Murdock and Provost's (1973) 10 scales has five sequential stages ─ 0, 1, 2, 3 or 4. If any society is observed to be at stage 4 (or 3 or 2 or 1) there was an earlier society (not necessarily progenitor) at lower stage 3 (or 2 or 1 or 0). That is the key idea of Table 1.

Peregrine (2003) converted Murdock and Provost's (1973) 5-point scales into archaeologically observable 3-point scales. The latter appear here in Table 2 along with rules (Denton 2004b) for converting the 5-point scales of Table 1 into 3-point scales of Table 2.

The 10 scales in Tables 1 and 2 are not conceptual variables. Each scale consists of stages which first appeared in time in the sequence shown. The stages of subscale 7 (Money) in Table 1 are adjusted in Table 2 to make Table 2 scale 7 applicable to prehistory. Table 1 subscales were designed for recent pre industrial societies. The form of subscale 7 (Money) in Table 1 reflects this intention. Even if the scales of Tables 1 and 2 are not conceptual variables, they may be used to suggest conceptual variables. It is to such matters that we now turn.

The most widely accepted stages of social organization in the contemporary world are the United Nations categories of development. Development codes d = 1, 2, 3 discriminate between countries as being least developed, less developed (excluding least developed) and developed.

Denton (2007a) shows that SCCS measures (of Table 1) and ACE measures (of Table 2) meet all but the second of the six criteria listed above, even though the measures cannot discriminate between contemporary countries. The problem is that Table 1 and 2 measures are ordinal. For UN member countries of the WDI data base at 2000 AD Denton (2007a) provides a set of interval markers in subject areas comparable to those of Tables 1 and 2.

Denton's (2007a) markers appear to meet all six criteria listed above but are pre-industrially unobservable.

Table 1. Component Subscales of Murdock and Provost's (1973) Pre-industrial Modernization Scale

Subscale 1 Writing and Records 4 True writing; records 3 True writing; no records 2 Nonwritten records 1 Mnemonic devices 0 None	Subscale 6 Land Transport 4 Automotive vehicles 3 Animal drawn vehicles 2 Draft animals 1 Pack animals 0 Human only
Subscale 2 Fixity of Residence 4 Sedentary 3 Sedentary; impermanent 2 Semisedentary 1 Seminomadic 0 Nomadic	Subscale 7 Money 4 True money 3 Elementary forms 2 Alien currency 1 Domestically usable articles 0 None
Subscale 3 Agriculture (Intensification) 4 Intensive 3 Primary; not intensive 2 more than 10%; secondary 1 less than 10% food supply 0 None	Subscale 8 Density of Population 4 Greater than 100 persons per sq. mile 3 26 - 100 2 5.1 - 25 1 1 - 5 0 less than 1
Subscale 4 Urbanization (Mean Size of Local Communities) 4 greater than 1000 persons 3 400 - 999 2 200 - 399 1 100 - 199 0 fewer than 100	Subscale 9 Level of Political Integration 4 3 or more administrative levels above local community 3 2 levels 2 1 level 1 Autonomous local communities 0 None
Subscale 5 Technological Specialization 4 At least smiths, weavers and potters 3 Metalwork only 2 Loom weaving only 1 Pottery only 0 None	Subscale 10 Social Stratification 4 3 social classes or castes 3 2 social classes, castes/slavery 2 2 social classes, no castes/slavery 1 Hereditary slavery 0 Egalitarian

Note: Murdock and Provost (1973) define the composite measure m of a society to be the sum $m = \Sigma s_{i=1}^{10}$ of the 10 subscale measures of the society, $0 \leq m \leq 40$. Definitions of subscales are from Divale (2004), V149 - V158 rescaled from 1 – 5 to 0 – 4. More exact definitions appear in Murdock and Provost (1973).

In Table 3 are proposed a single set of 11 modernization markers which, subject to test, may discriminate between prehistoric, recent pre-industrial and post-industrial societies known to differ in stages of social organization. These 11 markers build on Denton (2007a). They are conceptual variables for each of which we may seek empirical indictors to make measures with interval properties. Evidence will be given that these 11 variables satisfy the six criteria listed above for modernization markers. For them we will be able to find recent pre-industrial (SCCS) and WDI coded data. Archaeological and early historic data are discussed below.

Table 2. Murdock and Provost's (1973) Modernization Marker Variables as Adapted by Peregrine (2003) to Make Them Archaeologically Observable

Subscale 1 Writing and Records 2 True writing 1 Mnemonic or unwritten records 0 None	Subscale 6 Land Transport 2 Vehicles 1 Pack or draft animals 0 Human only
Subscale 2 Fixity of Residence 2 Sedentary 1 Seminomadic 0 Nomadic	Subscale 7 Money 2 Currency 1 Domestically usable articles 0 None
Subscale 3 Agriculture (Intensification) 2 Primary 1 ≥10%, secondary 0 None	Subscale 8 Density of Population 2 26+ 1 1 - 25 0 less than 1 person per sq. mi.
Subscale 4 Urbanization (largest settlement) 2 400+ 1 100 - 399 0 Fewer than 100 persons	Subscale 9 Level of Political Integration 2 3 or more 1 1 or 2 above local community 0 Autonomous local communities
Subscale 5 Technological Specialization 2 Metalwork (alloys, forging, casting) 1 Pottery 0 None	Subscale 10 Social Stratification 2 3 or more social classes or castes 1 2 social classes 0 Egalitarian

Notes: (1) In Peregrine (2003) each of the 10 scales occupies the integers 1, 2, 3. Here they occupy the integers 0, 1, 2. Composite measure m is $m = \Sigma_{j=1}^{10} s_j$, $0 \leq m \leq 20$.

(2) Each Murdock and Provost (1973) subscale is a 5-point measure 0, 1, 2, 3, 4. Accordingly, the sum of the 10 subscales ranges from 0 – 40. The (1973) 5-point subscales may be converted to Peregrine's (2003) 3-point scales as follows, where each paired (code(s):code) denotes the equivalence of a Murdock and Provost' (1973) 5-point code(s) to the corresponding Peregrine (2003) 3-point code. Scale 1: (0:0), (1-2:1); (3-4:2). Scale 2: (0:0), (1:1), (2-4:2). Scale 3: (0-1:0), (2:1), (3-4:2). Scale 4: (0:0), (1-2:1), (3-4:2). Scale 5: (0:0), (1-2:1), (3-4:2). Scale 6: (0:0), (1-2:1), (3-4:2). Scale 7: (0:0), (1:1), (2-4:2). Scale 8: (0:0), (1-2:1), (3-4:2), Scale 9: (0-1:0), (2-3:1), (4:2). Scale 10: (0:0), (1-2:1), (3-4:2).

(3) The substantive meaning of each of Peregrine's (2003) Scales is identical to that of the corresponding 3-scale of Murdock and Provost (1973) with the exception of Scale 4 Urbanization which the latter define as, "Mean Size of Local Community.

Table 3 columns 3-4 show that, at any stage of modernization such as forager, horticulture, preindustrial intensive agriculture, and UN development (d = 1, 2, 3), at least six of the 11 column 1 modernization marker variables discriminate between societies known to differ in widely accepted stages of development. For example, markers 2, 5, 6, 8, 9 and 10 discriminate between least modern (forager) societies. Markers 1, 4, 5, 6, 7 and 11 discriminate between the most modern societies where development d = 3.

Table 4 lists rules for converting the updated modernization markers of Table 3 into measures with interval properties on individual societies.

One obvious difference between Tables 1-2 and 3-4 is the addition of marker 11 (S_{11} energy use) in Tables 3-4. Marker S_{11} is added for the following reasons. Energy use provides an additional measure which (subject to test) may discriminate among the most modern societies where development d = 3. There is widespread support for a marker S_{11} of energy use. Irwin (1975) and Marsh (1967) both use energy consumption as a measure of modernization (or related concepts) in contemporary countries. It is shown below that S_{11} satisfies the rest of the six criteria listed above for markers of modernization.

Table 3. Proposed Modernization Markers Applicable 10,000 BC to Present Day

Murdock and Provost's (1973) Table 1 Subscale	Proposed Modernization Marker	Societies Not Discriminated by Marker	Societies Discriminated by Marker
1 Writing and records	S_1 Literacy rate, adult total (% of people ages 15 and older)	All forag. and hort. $S_1 \rightarrow 0$	Pre-indust. ag. to present day
2 Sedentary local communities	S_2 1/(mean no. local commun. locations per 10 years)	All intens. ag. $S_2 \rightarrow 1$	Forag. to pre-indust. ag
3 Intensification of agriculture	S_3 Crop time/(crop time + fallow time)	All forag. $S_3 \rightarrow 0$	Pre-indust. ag. to present day
4 Mean size of local communities	S_4 Urban population (% of total)	All forag. and hort. $S_4 \rightarrow 0$	Pre-indust. ag. to present day
5 Technology	S_5 % employed outside ag.	None.	All
6 Volume of land transport	S_6 Road traffic: (10^9)[(vehicle million metric tons)(km)/(population (1000's))(area)].	None	All
7 Money	S_7 Per capita GDP in constant US$ 2000 AD.	All forag. and hort. $S_7 \rightarrow 0$	Pre-indust. ag. to present day
8 Population density	S_8 Population density (people per sq. km)	All WDI $\xi[S_8] \rightarrow 152.75$	ACE & SCCS data cases
9 Number polit. juris. above local community	S_9 = 1, 0, presence, absence of state polit. org.	All intens. ag. & WDI $S_9 \rightarrow 1$	Forag. to hort.
10 Social classes	S_{10} = 1, 0: presence, absence of classes	All intens. ag $S_{10} \rightarrow 1$	Forag. to hort.
	S_{11} Energy use: Kg of oil equivalent per capita	None	All

Note: $\xi[.]$ denotes an expectation. Forager, horticulture, intensive agriculture, preindustrial, community, political and jurisdiction are abbreviated. See Table 4 for measures of proposed column 2 marker variables. Asymptotes account for the limiting values in column 3.

Modernization markers S_9 (presence/absence of the state) and S_{10} (presence/absence of social classes) are binary random variables. Each is restricted to the values 1 or 0. If a sample of n data cases is measured on S_9 (or S_{10}) the result is a count which is the sum of those data cases where measure s_9 (or s_{10}) = 1. Binary measures may be standardized (*infra*). Ordinal measures cannot be standardized. The justification for binary markers S_9, S_{10} is that, in contrast to ordinal markers, counts may be made of societies where $S = s$, and measures s_9, s_{10} may be standardized.

For UN member countries at 2000 AD the measures of Table 4 column 4 are simply the World Bank's (2007) WDI observations on individual countries, or simple transformations thereof. Some WDI data are missing.[3]

For prehistoric archaeological traditions Peregrine's (2003) ACE data base provides ordinal measures of Table 2 modernization markers. These coded data are not helpful for making measures on some markers in Tables 3-4. While the measures demanded in Table 3 are interval (s_9 and s_{10} are binary), those of Table 2 are ordinal. Moreover, the worldwide expectation $\xi[S_j]$ of each of the 11 marker variables of Table 3 changes over time in ways that cannot currently be reconstructed in (pre-)history. For example, societal S_1 Literacy rate, adult total (% of people ages 15 and older) began at less than 2% (e.g. Mesopotamia) but increased in the millennia following the earliest appearance of writing (Cressy 1981; Gaur 1984). Based on the gradations of Table 2 scale 1, all ACE archaeological traditions where writing is present are coded s_1 = 2.

Rules are shown in Table 4 for converting coded data on recent SCCS pre-industrial societies (Divale 2004) into measures on the 11 modernization markers of Table 3. For each of the 11 markers j = 1, 2, …, 11 these conversion rules assign any given SCCS society i an expectation $\xi[S_{ij}]$ which is justified in the Appendix. The effective, time domain of SCCS societies is narrow. After deletion of those 12 of 186 SCCS data cases which precede 1800 AD, the remaining 174 data cases are in the closed time interval [1800 AD, 1965 AD]. For the measures demanded in Table 4 there are no missing SCCS data.

The modernization markers of Table 4 LH column implement suggestions of Denton (2007a) with the following minor adjustments. S_3 uses Boserup's (1981:19) measure of agricultural intensification. The World Bank (2007) defines road traffic to be "the volume of goods transported by road vehicles, measured in millions of metric tons times kilometers traveled." For reasons stated in the Appendix, S_6 divides the WDI road traffic measure by both population (in 1000's) and country area, and multiplies the result by 10^9. Because it is currently uncertain whether prehistoric life expectancy (Demenay and McNicoll 2003) conforms to the six criteria proposed above for modernization markers, S_8 uses population density. For reasons given in Table 4 Note 6, an upper asymptote is assigned to all WDI countries as density measure s_8. Per capita energy use marker S_{11} is added in Tables 3-4 to the 10 markers derived from Tables 1-2 (*infra*). The World Bank (2007) defines energy use to be "use of primary energy before transformation to other end-use fuels, which is equal to indigenous production plus imports and stock changes, minus exports and fuels supplied to ships and aircraft engaged in international transport" in kg of oil equivalents. SCCS measures s_{11} are outlined in the Appendix.

The validity and reliability of Table 4 measures are considered below.

Table 4. Conversion of ACE, SCCS and WDI Coded Data into Measures of Modernization Marker Variables 10,000 BC to Present Day

Proposed Modernization Marker[1]	Conversion of Data Base Coded Data into Marker Measures		
	Observability in Prehistory	Conversion of Data Base Coded Data into Marker Measures	
		SCCS Societies [1800, 1965 AD]	WDI Countries [2000 AD]
S_1:	Note 2	$\xi[S_1]$ = 20, 0, 0 if Table 2 Scale 1 = 2, 1, 0	S_1 = WDI observation
S_2:	Note 2	$\xi[S_2]$ = 1, 0.05, 0.025 if Table 2 Scale 2 = 2, 1, 0	S_2 = 1 all countries
S_3	Note 2	$\xi[S_3]$ = 0.6, 0.2, 0.01, 0, 0 if Table 1 Subscale 3 = 4, 3, 2, 1, 0	$\xi[S_3 \mid d]$ = 0.7, 0.8, 0.95 Development d = 1, 2, 3[6]
S_4	Note 2	$\xi[S_4]$ = 25, 10, 0, if V63 = 8, 7, 1-6[4]	S_4 = WDI observation
S_5	Note 2	$\xi[S_5]$ = 20, 5, 3 if Table 2 Scale 5 = 2, 1, 0	S_5 = WDI observation
S_6	Note 2	$\xi[S_6]$ = 0.3, .03, 0, 0, 0 if Table 1 Subscale 6 = 4, 3, 2, 1, 0	$S_6 = (10^9)$[WDI observation/[(population(1000's))(area)]
S_7	Note 2	$\xi[S_7]$ = 1035.99, 0, 0 if Table 2 Scale 7 = 2, 1, 0	S_7 = WDI observation (in constant 2000 AD US$)
S_8	Note 2	$\xi[S_8]$ = 750c, 299.95c, 62.45c, 14.95c, 2.95c, 0.5c if V64 = 7, 6, 5. 4, 3, 1-2,[5]	$\xi[S_8]$ = 1942.5, upper asymptote assigned all countries[7]
S_9	Note 3	$\xi[S_9]$ = 1, 1, 0 if Table 2 Scale 9 = 2, 1, 0	S_9 = 1 all countries
S_{10}	Note 3	$\xi[S_{10}]$ = 1, 1, 0 if Table 2 Scale 10 = 2, 1, 0	S_{10} = 1 all countries
S_{11}	Note 2	$\xi[S_{11}]$ = 347.31, 145.13, 100 if Table 1 Scales 5 and 6 = (3 or 4 & 4), (3 or 4 & 3 or 2 or 1), (≤ 2 or 0)	S_{11} = WDI observation

1. Modernization markers defined in Table 3
2. Markers S_1 - S_8, S_{11} are not yet archaeologically observable. They are not historically observable, except at recent times.
3. Markers S_9 - S_{10} are archaeologically observable (Peregrine 2003). They are also historically observable.
4. V63 is from Divale (2004). SCCS # 88, missing data in Divale (2004), is coded $\xi[S_4] = 0$ based on Divale's (2005) code for V152.

(Notes for Table 4, continued)
5. V64 is from Divale (2004). Constant c = 2.59 converts V64 from people per square mile into people per square kilometer. SCCS # 39, 157 are missing data in Divale (2004). They are coded $\xi[S_8]$ = 62.45c, 14.95c based on Divale's (2004) codes for V156.
6. Based on Boserup (1981:19).
7. Population density is an upper asymptote. Between development d = 1, 2, 3 and WDI population density (s_8), $r_S ds_8$ = 0.13885 (n = 180) for which P(Rho = 0) = .0630 (2-tailed test). WDI population density measures and SCCS population measures (Appendix A) appear to be based on different definitions, e.g. total national land area versus area used. WDI countries are assigned the largest SCCS value so s_8 = 1942.5. See Demenay and McNiccol (2003) for the history of population density and related constructs.
Note: For prehistory measurement should start with decisions as to unit of observation. Either archaeological tradition (Peregrine 2003) or society (Denton 2007) might be considered. See the Appendix for justification of column 3 measures on SCCS societies.

Inglehart (2001) aside, neither gender roles nor political attitudes in the developing world today appears to satisfy the six modernization marker criteria suggested above. Democracy is too recent. Long term changes in gender roles may be U-shaped rather than monotonic. In forager societies, and developed countries today, women appear to be accorded greater equality than in societies whose recency of social organization falls between (Levinson and Malone 1980).

4 MODERNIZATION: DEFINITION AND MEASURE

Definition of Modernization: The degree of modernization H of a society is the mean recency of its modernization markers, where the latter meet the six criteria proposed above (from Denton 2007a).

Measure of Modernization: The measure h of the modernization of a society i is $h = (1/11)\Sigma_{j=1}^{11} t_{ij}$, the unweighted mean of the earliest (pre-)historic start times t_j (not the society's start times) of those states s_j (of the 11 marker variables S_j of Table 3 column 2) which characterize the society.

That h is the unweighted mean recency of time in years makes it an interval measure. No absolute zero point is assumed. All 11 times t_j used to calculate h are on the same scale – time. [4]

Modernization measure h and conceptual variable H may be criticized as follows. Measure h and CV (Conceptual variable) H turn on the particular marker variables from which h is calculated. Such a criticism is valid. From Murdock and Provost's (1973) 10 sets of Table 1 stages, different marker variables might be defined in Tables 3-4. For example, instead of S_1 Literacy rate, adult total (% of people ages 15 and older) we might define a marker S_1 People in secondary school as % people of secondary school age. Both markers would satisfy the six criteria listed above. Choosing between them is a matter of judgment. Table 4 provides data for markers S_1 - S_{11}. Markers S_1 - S_{11} satisfy the six criteria proposed above. Nevertheless, when additional data become available, new markers may be found with even greater ability to discriminate between societies known to differ in recency of social organization.

Table 5 shows that the ACE (Peregrine 2003) and WDI (World Bank 2007) data bases permit us to estimate the start times of marker variable states at the extremities of our time domain [10,000 BC, 2000 AD]. As for the rest of Table 5, we await archaeological and historical reconstruction to fill in the blanks (Fagan 2004). In short, we cannot currently estimate interval modernization measure h.

No ontology is claimed for modernization h. If we wish, we may seek an ontology for gradations in each of the 11 marker variables from which h is calculated (Denton 2007a).

If we were able to calculate a measure h of modernization we might use h for all the purposes to which Murdock and Provost's (1973) pre-industrial measure m (the sum of the 10 Table 1 subscales) has been put. Such a measure might be used to describe a single society, or as a candidate predictor in linear models and bivariate correlations (Denton 2007a; 2008).

If m in Table 1 (or 2) is thought of as the sum of 10 sets of stages (or ranks) SCCS 0, 1, 2, 3, 4 (or ACE 0, 1, 2) the sum m may be thought of as an interval measure of summed stage (or rank) of recency of social organization. The difficulty with the latter interval measure is that it places undue burden on the definitions of the stages used.

5. MAGNITUDE OF MODERNIZATION MARKERS: DEFINITION AND MEASURE

As a measure of earliest start time t_j, marker measure s_j is ordinal. Because a single, composite measure of marker variable magnitude may be useful as a predictor of many other behaviors, we turn next to devising a suitable conceptual variable (CV) and measure for it.

Definition of Mean Magnitude of Modernization Marker Standard Deviations: The mean magnitude of modernization marker standard deviations of a society is the unweighted mean of the standard deviations of the 11 Table 4 (LH column) marker variable measures of the society. [4]

Measure of Mean Magnitude of Modernization Marker Standard Deviations: Let s_{ij} be the measure of society i on marker variable j of Table 4 at time t. Let z_{ij} be the z-transform of s_{ij},

$z_{ij} = (s_{ij} - \bar{s}_j) / SD_{sj}$ (Spiegel *et al*. 2000). For any single data case i the mean magnitude of modernization marker standard deviations is $\xi[Z_i] = (1/11)\Sigma_{j=1}^{11} z_{ij})$.

Table 5. Crude Estimates of the Largest (Non-Standardized) Measures s_1 - s_{11} Table 4 Marker Variables in World at 500 Year Intervals 2000 AD to 10,000 BC (t = 0 at 2,000 AD)

	Crude Estimate of Largest Marker Variable State (Table 3)										
Years BP	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}
0	100	1	1	100	100	1042	46,278	1943	1	1	21,429
500		1							1	1	
1000		1							1	1	
1500		1							1	1	
2000	5	1							1	1	
2500	2	1							1	1	
3000	2	1							1	1	
3500	2	1	0.4						1	1	
4000	1	1	0.3						1	1	
4500	1	1	0.3						1	1	
5000	1	1	0.3						1	1	
5500	1	1	0.3						1	1	
6000	1	1	0.3						1	1	
6500	1	1	0.3						1	1	
7000	1	1	0.3						1	1	
7500	1	1	0.3						1	1	
8000	1	1	0.3						1	1	
8500	0	1	0.2						1	1	
9000	0	1	0.2						1	1	
9500	0	1	0.2						1	1	
10,000	0	1	0.2						1	1	
10,500	0	1	0.2						1	1	
11,000	0	1	0.2						1	1	
11,500	0	1	0.2						0	0	
12,000	0	1	0.01						0	0	

Note: Archaeological time in years BP is based on time t = 0 at 1950 AD. To be precise, column 1 times should be adjusted accordingly. At 0 BP scale values are WDI measures calculated as in Table 4 column 4. Measures shown at 0 AD are to the nearest integer. For blank cells pre-historic and historic reconstructions are currently unavailable.

For the 11 markers of Tables 3-4 alternative markers S_1 - S_{11} might be defined. That $\xi[Z_i]$ is based on standardized scores makes it less subject to the criticisms of measure h_i and CV H. The justifications for h_i H apply also to measure $\xi[Z_i]$ and its CV $\xi[Z]$.

Since standard measures z_{ij} of each of the 11 markers Z_j are distributed about mean zero, some standard measures will be positive and others negative. In order to permit logarithmic transformations of mean modernization magnitude it will be useful to define a measure g of size which makes the smallest data case measure equal to 1. The result is a measure g_i on data case i, $g_i = \xi[Z_i] + k = (1/11)\Sigma_{j=1}^{11} z_{ij}) + k$, where k is chosen so the lowest measure of any society in a data set becomes 1 by adding k to each $\xi[Z_i]$. The size of k is determined from data. That $\ln(g_i) \geq 0$ may enhance the usefulness of g as a predictor in models fitted to data.

Standardized measures g_i and $\xi[Z_i]$ may be justified as follows. The 11 marker variables measures s_j of Table 4 are of different scale units. For example, unstandardized marker variable S_1 (percent literate), ranges from 0 to 1000.Unstandardized marker variable S_7 (per capita GDP in constant 2000 AD $US) ranges from 0 to many tens of thousands. Unstandardized marker variable S_{10} (social stratification) ranges from 0 to 1. Use of the 11 standard scores z_j puts each measure on an equal scale. Each mean magnitude measure $\xi[Z_i]$ of a society i is the mean distance (in standard deviations) of the society from the 11 means $\bar{s}_j$ of all n societies. The composite measure $g_i = \xi[Z_i] + k = (1/11)\Sigma_{j=1}^{11} z_{ij}) + k$ is the mean distance (in standard deviations) to which is added a constant k chosen so $\ln(g_i) \geq$ 0.

Suppose we calculate the standardized measure of data case i on marker variable j. Since standard measure $z_{ij} = (s_{ij} - \bar{s}_j) / SD_{sj}$, measure z_{ij} of data case i depends on the mean $\bar{s}_j$ and standard deviation SD_{sj} of scale j calculated over the n data cases in the sample. If data are unequally missing, or there are unequal numbers of data cases over different levels of modernization, mean $\bar{s}_j$ and standard deviation SD_{sj} will impact both $\xi[Z_i]$ and g_i .

One solution to the problem posed in the preceding paragraph is to ensure equal numbers of data cases over different levels of modernization. For ACE data cases the relative mean magnitude of modernization markers may be estimated using Peregrine's (2003) composite measure m − the sum of the 10 3-point scales of Table 2. For SCCS data cases the relative mean magnitude of modernization markers may be estimated using Murdock and Provost's (1973) composite measure m − the sum of the 10 5-point scales of Table 1. For WDI data cases the relative magnitude of modernization markers may be estimated using United Nations (2007) development designations d = 1, 2, 3 (least developed, less developed excluding least developed, and developed). Based on these estimates, data cases might randomly be deleted until there are equal numbers of data cases over mean modernization magnitude measures m and d . Once such equal numbers of data cases, the overall mean $\bar{s}_j$ and standard deviation SD_{sj} might be calculated along with standard scores z_{ij} = $(s_{ij} - \bar{s}_j) / SD_{sj}$. Measures z_{ij} might be assigned to a missing data case i based on s_{ij} , $\bar{s}_j$ and SD_{sj} .

There are pros and cons to adjusting numbers of data cases over modernization prior to standardizing measures. Because the number of WDI data cases coded $d = 2$ is almost twice that of $d = 1$ or 3, deletion of cases will shift the mean $\bar{s}_j$ to the left and increase the size of the standard deviation SD_{sj}. Even if we do not make such an adjustment, societies with different marker variable magnitudes will still be discriminated. In what follows, we will not adjust numbers of societies over relative modernization prior to calculating standard scores z_{ij}.

For the 11 marker variables of Table 4 some WDI data are missing.[3] Numbers of WDI missing data cases are listed below. For the 11 marker variables there are no SCCS missing data.

There at least two solutions to the issue of missing WDI data on the 11 marker variable measures of Table 4. The method of conditional mean imputation (Allison 2002) might proceed as follows. Development $d = 1, 2, 3$ is non-missing for all WDI data cases. Estimate the conditional mean $\bar{s}_j \mid d = (1/q)(\Sigma_{i=1}^{q} s_{ij} \mid d)$ over the q data cases for which measures $s_{ij} \mid d$ of scale j are non-missing, where development $d = 1$, or 2, or 3. Symbol "|" is to be read "given that." For example, $s_{ij} \mid d$ is the measure of scale j on data case i, given that the level of development of data case i is d. To each data case i for which $s_{ij} \mid d$ is missing, assign the mean value $s_{ij} = \bar{s}_j \mid d$.

.A second solution to the issue of missing WDI data is to use what the SCCS and WDI data sets offer. That is the solution used here. If q out of 11 standard scores z_{ij} are non-missing on data case i the measure g_i on data cases i is $g_i = \xi[Z_i] + k = (1/q)\Sigma_{j=1}^{q} z_{ij}) + k$, calculated in such a way that mean $\bar{s}_j$ and standard deviation SD_{sj} are calculated from non-missing data cases and the mean modernization magnitude of a society uses a divisor $(1/q)$ (in $g_i = \xi[Z_i] + k = (1/q)\Sigma_{j=1}^{q} z_{ij}) + k$) which depends on the number q of non-missing standardized marker variable measures z_{ij}, $0 < q \le 11$. Where a standardized measure z_{ij} on marker j for society i is missing, it is excluded from calculation of g_i (and $\xi[Z_i]$) on data cases i.

6. VALISITY AND RELIABILITY

We need to assess the validity and reliability of the 11 marker variable measures s_1 - s_{11} (and z_1 - z_{11}) on SCCS and WDI data cases. We also need to assess the validity and reliability of measures h_i and $\xi[Z_i]$ (or g_i) which are calculated from z_1 - z_{11}. For these purposes we will use the tools of measurement theory (Carmines and Zeller 1979).

Modernization H is defined above. It is an interval version of a widely used construct deployed by Murdock and Provost (1973) and others mentioned above. When adequate archaeological and historic reconstructions become available for the start times needed in Table 5, valid, reliable measures will be available from which to calculate h .

Based on classical test theory, the writer estimates the validity and reliability of $\xi[Z_i]$ to be approximately 0.83. Since the derivation is technical it appears in Footnote 5.

Intercorrelations among z_1 - z_{11} , m , d and $\xi[Z]$ provide additional evidence of validity and reliability. Spearman Rho intercorrelations are calculated in Table 6 because the joint distributions of pairs of z_1 - z_{11} , $\xi[Z]$ are not normal; d is ordinal. Pearson product moments are inappropriate. Since the 174 SCCS data cases and 189 WDI data cases all meet the definition (*supra*) of societies, they are combined in Table 6 into a single sample of size 363 (Denton 2007b; 2008a). Missing WDI cases reduce bivariate sample sizes to those shown in Table 6.

Table 6, along with substantiation given in Denton (2007a), is evidence that scales S_1 - S_{11} of Table 6 meet the six criteria listed above for modernization markers. S_1 - S_{11} are supposed to be behaviorally related. Table 6 shows they are. That SCCS measures and WDI measures are in common use is additional evidence of validity and reliability (Denton 2008a).

Concurrent validity refers to the degree to which a measure correlates with other measures of the same CV. $\xi[Z_i]$ (or g_i) should correlate with both Murdock and Provost's (1973) measure m_i and United Nations (2007) development measures d_i = 1, 2, 3. In Table 6 the requisite Spearman Rho correlations are $r_S mg$ = 0.93 (P(Rho = 0) = <0.0001, n = 174) and $r_S dg$ = 0.68 (P(Rho = 0) < 0.0001, n = 189). Here, the composite measure $0 \leq m \leq 20$ of Table 2 is used. These correlations are evidence of the concurrent validity of $\xi[Z]$ (and g). That the Spearman correlation between g and development d is only 0.68 is, in the writer's opinion, interpretable as evidence that g is a better measure of its subject matter than d . Measure g uses interval (or binary) measures on all 11 marker variables of Table 4. Ordinal development measure d merges countries into groups in which the GDP and other marker measures of some countries coded $d = 2$ are considerably higher than those of some countries coded $d = 3$.

Table 7a examines the distribution of each of the 11 standardized marker variable measures z_{ij} , j = 1, 2, …, 11, from Table 4. In Table 7a the mean of each standard variable is zero but mean modernization magnitude $\xi[Z]$ is 0.076. That the latter mean is not zero is the result of missing WDI data.

For Development d = 1, 2, 3 there are 48, 95 and 46 UN member countries at 2000 AD. They total 189 but are centered on $d = 2$. The result gives d the low index of skewness reported in column 7 of Table 7. The skewness index for SCCS modernization measure m is also low. The latter low index is due to the fact that SCCS data cases are approximately equally distributed over $m = 0, 1, 2, \ldots, 20$.

Table 6. Spearman Rho Intercorrelations: Table 4 Scales S_1 - S_{11}, Modernization m (Table 2), Development d and Modernization Magnitude g

	S_1	S_2	S_3	S_4	S_5	S_6	S_7
S_1	1.0 208						
S_2	0.26 <0.0001 208	1.0 363					
S_3	0.71 <0.0001 208	0.56 <0.0001 363	1.0 363				
S_4	0.75 <0.0001 208	0.37 <0.0001 358	0.89 <0.0001 358	1.0 358			
S_5	0.63 <0.0001 190	0.40 <0.0001 262	0.84 <0.0001 262	0.84 <0.0001 262	1.0 262		
S_6	0.60 <0.0001 176	0.27 <0.0001 200	0.66 <0.0001 200	0.74 <0.0001 200	0.62 <0.0001 196	1.0 200	
S_7	0.44 <0.0001 205	0.37 <0.0001 347	0.65 <0.0001 347	0.66 <0.001 347	0.70 <0.0001 260	0.59 <0.0001 200	1.0 347
S_8	0.65 <0.0001 208	0.59 <0.0001 363	0.86 <0.0001 363	0.80 <0.0001 358	0.78 <0.0001 262	0.62 <0.0001 200	0.58 <0.0001 347
S_9	0.53 <0.0001 208	0.50 <0.0001 363	0.59 <0.0001 363	0.51 <0.0001 358	0.58 <0.0001 262	0.41 <0.0001 200	0.44 <0.0001 347
S_{10}	0.48 <0.0001 208	0.51 <0.0001 363	0.52 <0.0001 363	0.45 <0.0001 358	0.53 <0.0001 262	0.34 <0.0001 200	0.42 <0.0001 347
S_{11}	0.86 <0.0001 195	0.33 <0.0001 303	0.87 <0.0001 303	0.90 <0.0001 303	0.89 <0.0001 255	0.81 <0.0001 199	0.70 <0.0001 298
m	0.65 <0.0001 174	0.64 <0.0001 174	0.76 <0.0001 174	0.18 0.0201 174	0.67 <0.0001 174	0.50 <0.0001 174	0.58 <0.0001 174
d	0.61 0.0001 34	. . 189	1.0 <0.0001 189	0.62 <0.0001 184	0.39 0.0001 88	0.50 0.0091 26	0.70 <0.0001 173
g	0.78 <0.0001 208	0.56 <0.0001 363	0.93 <0.0001 363	0.92 <0.0001 358	0.88 <0.0001 262	0.68 <0.0001 200	0.74 <0.0001 347

Table 6 (continued)

	S_8	S_9	S_{10}	S_{11}	*m*	*d*	*g*
S_8	1.0 363						
S_9	0.67 <0.0001 363	1.0 363					
S_{10}	0.62 <0.0001 363	0.65 <0.0001 363	1.0 363				
S_{11}	0.80 <0.0001 303	0.56 <0.0001 303	0.51 <0.0001 303	1.0 363			
m	0.79 <0.0001 174	0.69 <0.0001 174	0.64 <0.0001 174	0.56 <0.0001 174	1.0 174		
d	. 189	. 189	. 189	0.65 <0.0001 129	. 0	1.0 189	
g	0.88 <0.0001 363	0.69 <0.0001 363	0.64 <0.0001 363	0.91 <0.0001 303	0.93 <0.0001 174	0.68 <0.0001 189	1.0 363

Note: Scales S_1 - S_{11} are the11 modernization marker variables defined in Tables 3-4. Modernization measure m is Murdock and Provost's (1973) measure as implemented by Peregrine (2003) with the 3-point scales of Table 2 rather than 5-point subscales of Table 1. Measures m are restricted to the 174 SCCS data cases. Modernization measure d is UN Development designation d = 1, 2, 3 (least developed, less developed excluding least developed and developed). Measures d are restricted to the 189 WDI data cases. Modernization measure g is that constructed here. Measures g are on all 363 data cases (174 SCCS + 189 WDI). Rows 1 -3 of each cell are the Spearman correlation r_S between a row variable and column variable, the 2-tailed probability P(Rho = 0) and sample size n. Halve the probability shown to get a 1–tailed P(Rho ≤ 0). For row d cells where r_S and P(Rho = 0) are listed as . , all data cases have the same Scale value.

Standardized variable Z_6 (Table 4) is annual volume of goods (in millions of metric tons) transported by road per 1000 people, divided by both population (in 1000's) and land area, all multiplied by 10^9. The range, in standard deviations, is $-2.19 \le z_6 \le 8.171$. An example of the lowest measure z_6 = -2.19 is SCCS ID 2 (Kung). Belgium is the country for which z_6 = 8.171. For z_6 all 174 SCCS data cases are non-missing. Only 26 measures z_{i6} of WDI data cases are non-missing. Nevertheless, Z_6 meets the six criteria listed above for modernization markers. One defect of Z_6 is that it omits volume of services transported.

Standardized measure z_{11} is energy use defined as in Table 4. The range, in standard deviations, is $-0.455 \le z_{11} \le 9.074$. An example of the lowest measure z_{11} = -0.455 is SCCS ID 2 (Kung). Quatar is the country for which z_{11} = 9.074. Iceland is next at z_{11} = 4.658. For USA z_{11} = 3.148. For z_{11} all 174 SCCS data cases are non-missing. For WDI data cases 60 measures z_{11} are missing. That the highest z_{11} goes to an oil-producing nation which the UN codes as development $d = 2$ gives reason to reflect. Qatar, and several other oil producing nations, has large per capita GDP measures z_{i7}. For Qatar, unstandardized per capita GDP (s_{i7}) is missing data in the WDI data base but is estimated elsewhere to be in excess of $17,000 at 2000 AD. The corresponding USA figure is $34,599. Nevertheless, energy use Z_{11} meets the six criteria listed above for markers of modernization. If z_{11} were to be used as a single measure of modernization magnitude, validity would suffer. Composite score $\xi[Z_i]$ is based on 11 measures. The Spearman correlation between $\xi[Z_i]$ based on z_1 - z_{10} and $\xi[Z_i]$ based on z_1 - z_{11} is 0.99689.

Table 7a. Distributions of Marker Measures: Summary Statistics

	$\bar{z}$	SD	Min.	Max.	Median	Skewness	Kurtosis	n
z_1	0	1	-0.661	2.87	-0.561	1.842	2.068	208
z_2	0	1	-2.622	0.385	0.385	-2.217	2.934	363
z_3	0	1	-1.599	1.222	0.480	-0.516	-1.300	363
z_4	0	1	-0.870	2.341	-0.443	0.688	-0.967	358
z_5	0	1	-0.908	1.870	-0.420	0.836	-0.980	262
z_6	0	1	-0.219	8.171	-0.219	6.192	41.390	200
z_7	0	1	-0.457	6.410	-0.304	3.414	12.405	347
z_8	0	1	-1.267	0.866	0.866	-0.335	-1.832	363
z_9	0	1	-1.893	0.527	0.527	-1.374	-0.112	363
z_{10}	0	1	-2.158	0.462	0.462	-1.706	0.915	363
z_{11}	0	1	-0.455	9.074	-0.435	4.220	26.133	303
m	10.195	5.318	0	20	10.500	0.048	-0.778	174
d	1.989	0.707	1	3	2.000	0.015	-0.983	189
$\xi[Z]$	0.075	0.726	-1.183	2.077	0.250	0.070	-5.598	363

Note: For z_1 - z_{11} all 174 SCCS data cases are present; numbers of non-missing WDI data cases may be calculated by subtracting 174 from n in column 9. Unstandardized means and standard deviations are: s_1 (16.274; 29.029), s_2 (0.87514; 0.32424), s_3 (0.53832; 0.33672), s_4 (27.087; 31.140), s_5 (34.644; 34.845), s_6 (26.901; 122.75), s_7 (3081.7; 6739.1), s_8 (1154.3; 909.84), s_9 (0.78237; 0.4132), s_{10} (0.82369; 0.38161), s_{11} (1119.0; 2238.1).

Today there is a cry for reducing greenhouse gas emissions. It may, or may not, turn out that new sources of energy will make energy use continue to satisfy the monotonic increasing properties demanded by the six criteria proposed above for markers of modernization. In the meantime, s_{11} meets all six criteria.

Table 7b. DISTRIBUTION OF 363 DATA CASES OVER MEAN z-SCORE $\xi[Z]$: HISTOGRAM

29	24	28	17	42	25	11	47	53	35	22	10	8	5	2	3	2	$\sum = 363$
								x									
							x	x									
				x			x	x									
				x			x	x	x								
				x			x	x	x								
x		x		x	x		x	x	x								
x	x	x		x	x		x	x	x	x							
x	x	x	x	x	x		x	x	x	x							
x	x	x	x	x	x	x	x	x	x	x	x						
x	x	x	x	x	x	x	x	x	x	x	x	x	x				
x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
-1.1	-0.9	-0.7	-0.5	-0.3	-0.1	0.1	0.3	0.5	0.7	1.9	1.1	1.3	1.5	1.7	1.9	2.1	

Here are examples of how modernization magnitude measure $\xi[Z_i]$ (or g_i) discriminates between societies. SCCS ID 2 (Kung) is one of the societies at the low end of $\xi[Z]$ where the expectation of standardized markers $\xi[Z]$ = -1.183. For SCCS ID 2 (Kung) Table 2 measure $m = 0$. For SCCS ID 43 (a town in Egypt at 1950 AD, perhaps representative of Egypt itself at that time) $\xi[Z]$ = 0.028 and Table 2 measure $m = 19$. At the high end of modernization magnitude, where $\xi[Z]$ = 2.077, is Belgium. Qatar is second at 2.038. Luxembourg (1.865), Denmark (1.839) and Holland (1.831) follow. USA is 1.640. For all three preceding countries development $d = 3$. For contemporary Egypt at 2000 AD $\xi[Z]$ is 0.458 ($d = 2$). For Lesotho $\xi[Z]$ is 0.291 ($d = 1$).

Based on modernization magnitude measure $\xi[Z]$, Table 7b separates the 363 SCCS and WDI societies into groups shown at the top of each histogram column. Histogram columns centered on $-1.1 \leq \xi[Z] \leq 0.1$ sort 176 societies of which 174 are SCCS. Histogram columns centered on $0.1 \leq \xi[Z] \leq 2.1$ sort 198 societies of which 189 are WDI. SCCS and WDI societies are grouped together atop the histogram column centered on $\xi[Z] = 0.1$.

7. MODERIZATION AND MAGNITUDE AS PREDICTOR

Modernization magnitude measure g_i (or $\xi[Z_i]$) may be considered as a candidate predictor in bivariate correlations and linear models. In Table 8 appear three kinship-related binary random variables X_1, X_2, X_3. A society is coded $X_1 = 1$ if the independent family predominates, otherwise $X_1 = 0$. A society is coded $X_2 = 1$ if neolocal postmarital residence predominates, otherwise $X_2 = 0$. A society is coded $X_3 = 1$ if kinship is bilateral, otherwise $X_3 = 0$. Definitions of random variables X_1, X_2, X_3 use definitions for SCCS variables V68, V69, V70 (Divale 2004) as shown in Table 8. Denton (2008a) provides measures x_1, x_2, x_3 on the combined sample of 363 societies — 174 SCCS societies to which are added 189 WDI societies.

Table 8. Logistic Regression Kinship Models Fitted to Combined Sample of 363 SCCS & WDI Data Cases: Predictor Variable g (Modernization Magnitude)

Criterion Variable	$X_1 = 1$ if independent family predominates, otherwise $X_1 = 0$	$X_2 = 1$ if neolocal residence predominates, otherwise $X_2 = 0$	$X_3 = 1$ if bilateral kinship predominates, otherwise $X_3 = 0$
Derivation of Coded Data	$X_1 = 1$ if V68 = (1 OR 2 OR 3 OR 4) ELSE $X_1 = 0$	$X_2 = 1$ if V69 = 5 ELSE $X_2 = 0$	$X_3 = 1$ if V70 = 5 ELSE $X_3 = 0$
Predictor Variable	g grouped (see Remark)	g grouped (see Remark)	g grouped (see Remark)
Fitted Model	$P(X_1 = 1) = [1 + \exp(-\mathbf{z}^T \mathbf{b})]^{-1}$ $\mathbf{z}^T = <1, g, g^2, e^g>$	$P(X_2 = 1) = [1 + \exp(-\mathbf{z}^T \mathbf{b})]^{-1}$ $\mathbf{z}^T = <1, g, g^2, e^g>$	$P(X_3 = 1) = [1 + \exp(-\mathbf{z}^T \mathbf{b})]^{-1}$ $\mathbf{z}^T = <1, g, g^2, e^g>$
Coefficients	$\mathbf{b}^T$ = <8.1029, -9.6287, 3.3440, -0.3801> PR > CHISQU: .0002, <.0001, <.0001, .0006	$\mathbf{b}^T$ = <9.0236, -15.4359, 6.0019, -0.7484, > PR > CHISQU: ..0302, .0003, <.0001, <.0001	$\mathbf{b}^T$ = <14.3213, -15.7334, 4.7417, -0.4599> PR > CHISQU: <.0001
Goodness of fit: PR > CHISQU:	Pearson resid. 0.2043 Dev. resid. 0.1742 H-L statistic 0.2223	Pearson resid. 0.4164 Dev. resid. 0.3621 H-L statistic 0.5400	Pearson resid. 0.2826 Dev. resid. 0.2989 H-L statistic 0.6880
ROC RSquare n	0.713 0.1578 363 (209 coded X_1 =1)	0.918 0.4495 362 (113 coded X_2 =1)	0.761 0.2057 361 (157 coded X_3 =1)

Remark: Model fitting starts with candidate predictors g, g^2, g^3, ln(g), exp(g), $g^{4/3}$, $g^{5/3}$. Score variable selection provided statistical equivalents. The simplest (lowest power) solution was chosen. SCCS definitions and coded data for V68, V69 and V70 appear in Divale (2004). Modernization magnitude measure g is grouped as follows. To $\xi[Z]$ the constant 2.18274 is added so the lowest score g = $\xi[Z] + 2.18274$ becomes 1. In Table 8 g is grouped gr :

If $1.00 \leq g < 1.33$ then group gr = 1.16; If $2.63 \leq g < 2.96$ then group gr = 2.79;
If $1.33 \leq g < 1.65$ then group gr = 1.50; If $2.96 \leq g < 3.28$ then group gr = 3.12;
If $1.65 \leq g < 1.98$ then group gr = 1.82; If $3.28 \leq g < 3.61$ then group gr = 3.45;
If $1.98 \leq g < 2.30$ then group gr = 2.14; If $3.61 \leq g < 3.93$ then group gr = 3.77;
If $2.30 \leq g < 2.63$ then group gr = 2.47; If $3.93 \leq g \leq 4.26$ then group gr = 4.10;

Three logistic regression models (Kutner *et al.* 2004) are fitted in Table 8 with g as a predictor – one model for each of X_1, X_2, X_3. In order to fit a logistic regression model g is grouped in Table 8. All three fitted models show g to be a serviceable predictor.

That g (Table 8) and m (Denton 2007c) both serve as predictors in linear models for the same criterion variables X_1, X_2, X_3 is evidence of convergent validity of g. [6]

In Table 8 the candidate predictor is g which, for any data case i, is $g_i = \xi[Z_i] + k = (1/11)\Sigma_{j=1}^{11} z_{ij}) + k$. We may also consider using individual standardized measures $z_1, z_2, \ldots, z_{11}$ along with, or instead of, g_i. Indeed, we may consider using unstandardized measures $s_1, s_2, \ldots, s_{11}$. Behavioral theory should guide selection of candidate predictors. [7]

8. DISCUSSION

The modernization of a society is defined to be the recency of its form of social organization in the last 12,000 years of human history and prehistory. Tables 1 and 2 list known developmental stages in writing, agriculture, sedentism, monetary exchange, technological specialization, transport of goods, population density and location, stratification and state political organization. Murdock and Provost (1973) pinpointed these stages. The sequencing of stages in Tables 1 and 2 is the sequencing over which subject matter states first appeared in time. In that sense, societies are known to have traveled along a single path. That path continues today.

Six criteria were defined in Tables 3-4 in order to select markers of modernization. Eleven modernization markers resulted. They are either interval or, in the case of markers S_9 and S_{10}, binary. Such evidence as is currently available suggests that all apply from 10,000 BCE to present day (Denton 2007a). At least six of the 11 markers discriminate between societies at any given distance traveled along the modernization path. As a result, composite measures h_i (modernization) and $\xi[Z_i]$ or g_i (mean magnitude of standardized modernization markers) discriminate between societies at all stages of modernization.

Measures h_i and $\xi[Z_i]$ (or g_i) improve on earlier composite measures of modernization. They distinguish between the chronological recency and magnitude of modernization markers. Earlier measures outlined in Levinson and Malone (1980:31-37) neglected the distinction. The measures developed here are interval whereas earlier ones are almost all ordinal. Earlier measures applied to archaeological traditions, or recent pre-industrial societies, or contemporary countries, but none provided an interval measure applicable to all three. Measures h_i and $\xi[Z_i]$ (or g_i) are interval measures which apply to any or all of the era from 10,000 BC to present day, although most measures at times preceding 1800 AD await reconstruction. Validity and reliability are estimated to be $V(\xi[Z]) = R(\xi[Z]) \approx .83$; [5] that of earlier measures is unknown.

Even though composite modernization measures h_i, $\xi[Z_i]$ and g_i appear to improve on earlier measures such as m or d, they are a beginning, not an end. There is a need for archaeological and historic reconstructions of modernization marker measures. A variety of improvements may be considered for the marker variables of Table 4. The 10 stages which Murdock and Provost (1973) pinpointed in Table 1 may be used to suggest markers alternative to those of Tables 3 and 4. Instead of a marker "energy use," a marker such as "energy use in secondary industry" might be considered. No such measure is currently available. Readers have the option of calculating a measure $\xi[Z]$ or g based on markers 1−10, alone. Different solutions might be considered for the calculation of $\xi[Z]$ and g in the presence of missing data. Improvements to behavioral theory are needed. Why do the 11 marker variables S_1 - S_{11} change over time? Denton (1996) attempts a partial answer. What behavioral linkages connect changes in marker variables S_1 - S_{11} to the many other behaviors correlated with them?

$\xi[Z]$ and g are CVs of the overall modernization of societies in the tradition of Murdock, Service and Lerner (*supra*). Irwin's (1975) measure is of a narrowed concept of economic modernization − the efficiency with which economic output (GDP) uses energy input. Here, $\xi[Z_i]$ and g_i give equal weight to measures z_1 - z_{11} from which they are calculated. New behavioral theory of how societies work might suggest new measures of new markers, or new weightings of old measures. In cultural anthropology today there is no satisfactory, comprehensive theory of society. Hence, measures z_1 - z_{11} are equally weighted.

There is a need to improve the precision of measures calculated in Table 4. Rules are given in Table 4 to transform SCCS and WDI coded data into measures s_1 - s_{11} on marker variables S_1 - S_{11}. In Table 4 and the Appendix, rationales are given for the expectations on which the rules of Table 4 are based. The precision of the expectations, however, remains unknown. For such rules the proof of the pudding may lie, in part, in the tasting. Table 6 supports the validity and reliability of the measures used. In Table 8, linear models are successfully fitted to a combined data set of 363 observations of which 174 are SCCS and 189 are WDI. True statistical models result, the premises of which are met by the data to which they are fitted.

The modernization measures proposed here may be applied to contemporary societies alone, or to pre-industrial societies alone, or to both. Applied to both, the measures offer new opportunities to test behavioral theory of much greater range of application than heretofore possible. The measures may put to a variety of concrete uses:

1. To calculate a bivariate correlation $r_S yg$ (or ryg if relevant premises are met) between a measure y and modernization g_i . $\xi[Z_i]$, or h_i, may be considered instead of g_i .

2. To fit a linear model $y = f(.)$ to data with $\xi[Z]$, g or h as a predictor. Subject to availability, the data in 1. and 2. may be recent, historic &/or pre-historic. Predictors z_1 - z_{11}, or s_1 - s_{11}, or subsets thereof, may also be used as candidate

predictors along with, or instead of, $\xi[Z]$ or g. In footnote 7. additional solutions are given for the selection of candidate predictors.

3. To model the present, reconstruct the past and forecast the future using 1. or 2. Denton (2008a) outlines forecast methods using an earlier, less satisfactory modernization measure.

4. To focus efforts in contemporary, historic and pre-historic databases on the collection of improved marker variable measures s_1 - s_{11}.

5. To use $\xi[Z]$, g, z_1 - z_{11} or s_1 - s_{11}, or h in time series models, causal models, stochastic process models, rate-of-change models, etc. when appropriate data become available.

Readers may easily use the measures devised here in their own data analysis. Electronic coded data are available for WDI (World Bank website) and SCCS (*gratis* at World Cultures website; Divale 2004). The means and standard deviations of markers $s_1 - s_{11}$ in Table 7 may be used to calculate standardized measures z_1 - z_{11} from which $\xi[Z_i]$ and g_i may be obtained.

Measures $\xi[Z_i]$, g_i, z_1 - z_{11} or s_1 - s_{11}, and perhaps h_i are enormously useful constructs to deploy in behavioral models. Reasons deserve emphasis. Measures s_1 - s_{11} discriminate between societies known to differ in the distance traveled along stages of modernization. The discriminating marker measures s_1 - s_{11} have massive impacts on many other facets of society. That s_1 - s_{11} increase pair-wise together, lock step, gives composite measure $\xi[Z_i]_i$ (or g_i) an even greater bang for the modeling buck. Modernization is a powerful predictor.

9. NOTES

1. Data analyses were designed and executed by the writer using SAS 8.2 running on a Windows XP-Pro platform.

2. Society is used here in the technical sense of a geographical territory which is relatively politically autonomous. Examples: a Kung band (SCCS # 2), a Yanomamo village (SCCS # 163), a Fur state (SCCS # 29), modern China at 2000 AD. SCCS data cases were originally thought of as cultures (Murdock and White 1969) but may be treated as societies (Denton 2007b; 2008). Denton (2007b) provides a sampling model for data on SCCS and WDI societies. Denton (2007a) shows that Table 2 markers are pair-wise monotonic increasing, Because an archaeological tradition is dropped from the ACE data base following the rise of literacy in it, correlations between ACE markers and time t need special interpretation.

3. In Tables 4-8 missing data are assumed to be missing at random (MAR). Suppose data analysis uses two variables y, x where x is always present and y is sometimes missing. Missing data are defined to be MAR if Pr(y is missing given x) = Pr(y is missing given y, x). MAR cases may be deleted and data analysis proceed as if all cases were present (Allison 2002:4-5).

4. Measures h (modernization) and $\xi[Z]$ (modernization magnitude) give equal weight to each of the 11 marker variables of Tables 4-8. If, instead,

$$h = a_1 t_1 + a_2 t_2 + ... + a1t_{11}, \qquad a_1 \text{ - } a_{11} \text{ are real constants which sum to 1}$$

$$Z^* = a_1 z_1 + a_2 z_2 + ... + a_{11} z_{11}, \qquad a_1 \text{ - } a_{11} \text{ are real constants which sum to 1}$$

then different weights might be given to coefficients $a_1, a_2, ..., a_{11}$. For example, weights might be chosen to maximize the fit to data of a linear model in which h or Z^* is a candidate predictor.

5. Denton (2008b) devises a method for estimating validity and reliability based on classical test theory. The method does not use Pearson product moments. These are inappropriate here because the variables under examination cannot be assumed to be joint normal.

Let zA be the mean of one set of the 11 standardized markers z_1 - z_{11}. Let zB be the mean of a different (non-overlapping) set of standardized markers z_1 - z_{11}. Here, zA and zB are both measures of the same construct "mean standardized modernization marker" the true score of which we will denote t. Assuming (subject to test) the classical test model, a single data case is

$$zA_i = t_i + eA_i, \; \xi[eA_i] = 0, \quad eA_i \text{ is independent of } zA_i, \; zB_i, \; t_i, \; eB_i \qquad (1)$$

$$zB_i = t_i + eB_i, \; \xi[eB_i] = 0, \quad eB_i \text{ is independent of } zB_i, \; zA_i, \; t_i, \; eA_i \qquad (2)$$

Adding and subtracting (1) and (2) over each data case and taking variances over all n data cases

$$VAR[zA + zB] = 4\,VAR[t] + VAR[eA] + VAR[eB] \qquad (3)$$

$$VAR[zA - zB] = VAR[eA] + VAR[eB] \qquad (4)$$

Subtracting (4) from (3) and dividing by 4 we get $VAR[t]$. Based on (1) and (2) the validity V and reliability R of $\xi[Z]$ are

$$V(\xi[Z]) = R(\xi[Z]) = VAR[t] / VAR[\xi[Z]] \qquad (5)$$

Since we have two different estimates zA and zB of $\xi[Z]$ there will be one numerator ($VAR[t]$ from (4)) in (5) but a choice of two denominators — $VAR[zA]$ or $VAR[zB]$. As a result, we will get two different estimates of $V(\xi[Z])$ and $R(\xi[Z])$.

Let zA be $\xi[Z]$ calculated from standardized markers z_2, z_4, z_6, z_7, z_9. Let zB be $\xi[Z]$ calculated from standardized markers $z_1, z_5, z_8, z_{10}, z_{11}$. This partition of markers into two sets zA and zB was chosen as follows. To zA, zB alternate assignment of Table 3 column 4 markers which discriminate foragers until each partition has three markers. Alternate assignment of markers which discriminate societies coded development $d = 3$ until each partition has three markers. Then, z_5 (% employed outside agriculture) was transferred to partition zB based on the belief that z_5 and z_7 (per capita GDP in constant $US) most effectively discriminate between the most development WDI societies coded $d = 3$) and should not both be in zA. Marker z_6, which had been partitioned into both zA and zB, was dropped to achieve non-overlapping partitions.

The null hypothesis that the expectation $\xi[zA - zB]$ of the difference zA - zB is 0 is rejected (t = -3.4326, p > |t| = 0.0007). That population difference $\xi[zA - zB] \neq 0$ may be due to missing data. The null hypothesis that $\xi[zA - zB] = 0$ is accepted if we add to zA the constant 0.0773778. If $VAR[eA] - VAR[eB] = 0$ then d = $(eA - \xi[eA])^2 - (eB - \xi[eB])^2 = 0 = 0$. The null hypothesis that d = 0 is rejected (t = -4.08384, p> |t| <0 .0001). In terms of classical test theory the premises of (1)-(2) are satisfied. Since measures zA and zB differ by constant 0.0773778 they are essentially tau-equivalent (Denton 2008b).

Resulting $VAR[t]/VAR[zA] = 0.95$ and $VAR[t]/VAR[zB]$=0.75. The mean is 0.85 which is an estimate of Validity $V(\xi[Z])$ and Reliability $R(\xi[Z])$ of $\xi[Z]$. Spearman's $r_S zAzB = 0.92$.

We cannot assume that zA, zB are distributed bivariate normal. Neither $\xi[Z]$ (the mean of z_1 - z_{11}), zA (the mean of z_2, z_4, z_6, z_7, z_9) nor zB (the mean of $z_1, z_5, z_8, z_{10}, z_{11}$) meets (SAS PROC UNIVARIATE) tests for normality. The xy-plot of zA, zB is consistent with a straight line but non-constant conditional variance. These remarks also hold if we use a second partition of z_1 - z_{11} into zA, zB which follows next.

Suppose we define zA to be the mean of standardized markers z_1, z_2, z_6, z_7, z_9 and zB to be the mean of standardized markers $z_4, z_5, z_8, z_{10}, z_{11}$. The null hypothesis that the expectation $\xi[zA - zB]$ of the difference zA - zB is 0 is accepted (t = 1.428152, p > |t| = 0.15) so the population means zA and zB identical. Hence, classical test theory premises (1)-(2) are satisfied. If $VAR[eA] - VAR[eB] = 0$ then d = $(eA - \xi[eA])^2 - (eB - \xi[eB])^2 = 0$. The null hypothesis that d = 0 is rejected (t = -4.08384, p> |t| < 0.0001). Hence, gA and gB are tau-equivalent (Denton 2008), rather than parallel, measures of mean standardized modernization marker $\xi[Z]$. Resulting $VAR[t]/VAR[zA] = 0.98$ and $VAR[t]/VAR[zB] = 0.68$ and the mean is 0.83. Spearman's $r_S zAzB = 0.92$.

Each of the preceding estimates $V(zA)$, $V(zB)$ of Validity $V(\xi[Z])$ and Reliability $R(\xi[Z])$ uses five of 11 standardized markers z_1 - z_{11}. Each pooled estimate is 0.83 or 0.85. $\xi[Z]$ uses all 11 markers z_1 - z_{11}. An estimated Validity $V(\xi[Z])$ and Reliability $R(\xi[Z]) \approx 0.83$ seems reasonable. Correction for attenuation (Denton 2008b) may be considered in order to adjust observed correlations between $\xi[Z]$ and other constructs behaviorally related to $\xi[Z]$.

The estimates of validity and reliability calculated in this footnote do not consider confidence intervals. Denton (2008b) gives reasons why the estimates calculated here may be used as guidelines for decision making.

6. For each criterion variable X_1, X_2, X_3 in Table 8 Denton (2008a) fits a logistic regression model using a modernization measure m applicable to all 363 societies in Table 8. The latter measure m is constructed in Denton (2008a). Measure g, constructed here, may be preferred to Denton's (2008a) m. The latter m assigns each WDI country coded d = 1, or 2, or 3 the same score $m \mid d$. Measure $m \mid d$ is based on an "out-of-sample" estimate obtained by fitting a linear model to only 24 SCCS societies.

7. In order to fit a logistic regression model for each of X_1, X_2, X_3 (defined here in Table 8) Denton (2008a) fitted models with candidate predictors which include time t (grouped), grouped interaction terms mt and powers thereof, as well as grouped m. Higher levels of modernization occur at later times. Here, time t and interaction terms were not candidate predictors in the models fitted in Table 8.

If energy use (Table 4 s_{11}) is excluded from the calculation of $\xi[Z]$ and g a simpler, 2-predictor set solution $< g, g^2 >$ arises in all three models of Table 8. Oil producing countries such as Qatar and Bahrain have very high energy use. They also have a recent history of unilinear kinship, extended families and unilocal residence. The writer re-examined his coding of such societies. These countries have had high migrant worker populations which may not be reflected in measures x_1, x_2, x_3 of the countries. In the future, better coded data may (or may not) suggest models simpler than those of Table 8. In the meantime, the models of Table 8 stand.

That 3-predictor solutions are fitted in Table 8 is due to the need for curves which model countries such as Qatar and Bahrain at the upper ranges of $\xi[Z]$ and g. Instead of 3-predictor solutions the following solutions might be considered for Table 8:

a. Lag g behind a kinship variable X so x(t) = f(g (t – 20 years). (No current data).
b. Delete energy use from calculation of g.
c. Create a new Table 4 measure s_{11} of energy use in secondary industry. (No current data).
d. Use a candidate predictor set consisting of both g (calculated from s_1 - s_{11}) and s_{11}.
e. Use additional predictor $\xi[\bar{X}_{adjacent}]$ = "mean $\bar{x}$ of kinship variable X in neighboring countries." This is a solution to Galton's problem suggested by Denton (2007c).
f. Using new data, re-code kinship variable X in oil producing countries such as Qatar and Bahrain. (No current new data).
g. Delete countries such as Qatar and Bahrain as outliers in a solution with fewer predictors.

Solutions a.–g. may increase the usefulness of $\xi[Z]$ and g as predictors.

10. APPENDIX: RULES FOR TRANSFORMING SCCS CODED DATA FROM DIVALE (2004) INTO THE MODERNIZATION MARKER MEASURES S_1 - S_{11} OF TABLE 4

Table 4 column 3 transforms Divale's (2004) coded data into measures of SCCS societies for each modernization marker of Table 4 column 1. The rules given below assume Divale's (2004) measures V149-V158 on SCCS societies are transformed to 5-point format 0, 1, 2, 3 or 4 instead of 1, 2, 3, 4, 5. Where appropriate, the rules of Table 4 transform the transformed 5-point measures into the 3-point measures of Table 2 by applying Note 2 of Table 2.

The strategy of Table 4 column 3 is to assign each SCCS data case the expectation of the group in which the data case falls, given assumptions to be specified below. The goal is to devise SCCS measures of sufficient precision that measures on the 363 combined SCCS and WDI societies will meet tests for the premises of classical test theory outlined in Footnote 5.

S_1 of Table 4 is Literacy rate, adult total (% of people ages 15 and older). In least developed WDI countries at 2000 AD the expectation $\xi[S_1 \mid d = 1]$ in UN member countries coded development $d = 1$ is 50.15%. Gaur (1984) and Cressy (1981) report steadily increasing rates of literacy from early Mesopotamia to 16th century England, albeit rates lower than 50%. For SCCS societies in time interval [1800, 1965] the expected value $\xi[S_1 \mid Table2\_scale1 = 2]$ = 20 seems reasonable.

S_2 (sedentism) of Table 4 is 1/(mean number of local community locations per 10 years). In Table 4 SCCS societies receive the expectations $\xi[S_2]$ = 1, 0.05, 0.025 if Table 2 Scale 2 = 2, 1, 0. The latter expectations assume sedentary communities occupy an average of one location per 10 years, seminomadic communities change locations at a mean frequency of twice per year and nomadic communities change locations at a mean frequency of four times per year.

S_3 of Table 4 is Crop time/(Crop time + Fallow time). This is Boserup's (1981: 19) measure of agricultural intensification. Based on comparing SCCS societies to Boserup's (1981:19) table, SCCS societies receive one of the five expectations $\xi[S_3]$ = 0.6, 0.2, 0.01, 0, 0 if Table 1 Subscale 3 = 4, 3, 2, 1, 0.

S_4 of Table 4 is Urban population (% of total). Here we will use Divale's (2004) codes for SCCS V63 Community Size. Code V63 = 7 denotes a community size in the range 5,000-49,999 and code V63 = 8 denotes a size exceeding 50,000. In least developed WDI countries at 2000 AD $\xi[S_4 \mid d = 1]$ = 29.43. Each WDI observation is based on the definitions of rural/urban of each WDI country enumerated. Of these countries some may use a population threshold of 5000 for urban, others a threshold of 1000. Based on V63, SCCS societies in Table 4 are assigned the expectations $\xi[S_4 \mid V63 = 7] = 25$ and $\xi[S_4 \mid V63 = 8] = 10$. In Table 4 all SCCS societies with community size less than 5000 are coded $\xi[S_4 \mid V63 < 7] = 0$. For nine European countries at 1750 AD Allen (2000 Table 5) estimates the mean % urban population as 18, which is comparable to the expectations of SCCS societies used in Table 4.

S_5 of Table 4 is % of those gainfully employed who are employed outside agriculture. For nine European countries at 1750 AD Allen (2000 Table 5) estimates the mean % employed in agriculture in rural communities as 56. Assuming that an additional 20% of those in cities are employed in agriculturally related occupations, we get estimates of expected percent gainfully employed outside agriculture in SCCS societies $\xi[S_5] = 20, 5, 3$ if Table 2 Scale 5 = 2, 1, 0.

S_6 of Table 4 is Road traffic. WDI (2007) archives measures of "Roads, goods transported (million ton km)" which is defined as "Goods transported by road are the volume of goods transported by road vehicles, measured in millions of metric tons times kilometers traveled." Based on this definition construct S_6 in Table 4 is

$[10^9]$ [(million metric tonnes carried by vehicles)(km traveled)]/[(population in (1000's))(area)]

Division by both population (in 1000s) and country area gives a comparable measure over countries of different population size and land area. Excluding multiplier 10^9, for WDI societies where development d = 1, 2, 3 at 2000 AD, the expectations are $\xi[S_6 \mid d = 1] = (3.33)10^{-10}$ (n = 1), $\xi[S_6 \mid d = 2] = (2.61)10^{-9}$ (n = 4) and $\xi[S_6 \mid d = 3] = (2.51)\ 10^{-8}$ (n = 21). In Table 4 SCCS societies are coded $\xi[S_6] = 0.3, .03, 0, 0, 0$ if Table 1 Subscale 6 = 4, 3, 2, 1, 0.

S_7 of Table 4 is Per capita GDP in constant US$ 2000 AD. Since GDP measures volume of monetary transactions, Table 4 shows per capita GDP = 0 in SCCS where money is absent. There are 21 SCCS societies where true money is present in time interval [1800, 1965]. For these societies the mean time of ethnographic observation is 1942.43. Maddison (2003) estimates world average per capita GDP in 1990 Geary-Khamis dollars as being $1962.in 1940, $1262 in 1900. The World Bank measures 2000 AD per capita GDP in 2000 AD $US. Let $US denote US 2000 AD dollars. Let $G-K denote Geary-Khamis 1990 dollars. We will use a conversion rate based on USA per capita GDP at 2000 AD measured in both $G-K and $US. Conversion from $G-K to $US uses the relation 1 $US = 0.8209 $G-K. A world per capita GDP in $G-K of $1262 at 1900 AD becomes (.8209)(1262) = 1035.99 in constant 2000 AD $US. In Table 4 SCCS societies are translated into per capita GDP in constant 2000 AD $US by the correspondences $\xi[S_7] = 1035.99, 0, 0$ if Table 2 Scale 7 = 2, 1, 0.

S_8 of Table 4 is Population density (people per sq. km.). Denton (2008) uses median values for Divale's (2004) codes (graduated 0, 1, 2, 3, 4) for V64 population density to estimate expected population density in SCCS societies. The expectations so estimated are used in Table 4 where $\xi[S_8]$ = 750c, 299.95c, 62.45c, 14.95c, 2.95c, 0.5c if Divale's (2004) V64 = 7, 6, 5. 4, 3, 1-2, c = 2.59 to convert people per square mile (SCCS) into people per square kilometer (WDI).

In Table 4 measures S_9 and S_{10} are binary so each takes on the values 0 or 1. S_9 of Table 4 is = 1, 0 which denotes presence (1) or absence (0) of state political organization. A state exists if two or more local communities are integrated into a single political organization. S_{10} of Table 4 is = 1, 0 which denotes presence (1) or absence (0) of social classes. Binary markers S_9 and S_{10} are used because true interval measures cannot currently be extracted from World Bank (2007) WDI data. In Table 4 binary measures of SCCS societies are obtained by the translation rules $\xi[S_9]$ = 1, 1, 0 if Table 2 Scale 9 = 2, 1, 0, $\xi[S_{10}]$ = 1, 1, 0 if Table 2 Scale 10 = 2, 1, 0.

S_{11} of Table 4 is per capita energy use defined to be per capita "use of primary energy before transformation to other end-use fuels, which is equal to indigenous production plus imports and stock changes, minus exports and fuels supplied to ships and aircraft engaged in international transport (World Bank 2007)" in kg of oil equivalents. For WDI data cases coded development = 1 mean per capita energy use is 347.31 and the lowest per capita energy use for a country is 145.13. We will use the latter information to estimate mean SCCS values. In Table 1 additional energy will be used in SCCS societies coded subscale 5 (values 4 and 5, smiths and metalworking) and subscale 6 (values 1-4, pack animals to automotive vehicles). If Subscale 5 = (3 or 4) and Subscale 6 = 4 then s_{11} = 347.31. If Subscale 5 = (3 or 4) and Subscale 6 = (3 or 2 or 1) then s_{11} = 145.13; If Subscale 5 $\leq$ 2 or Subscale 6 = 0 then s_{11} = 100.

11. REFERENCES CITED

Allen, Robert C.
2000 Economic structure and agricultural productivity in Europe 1300−1800. *European Review of Economic History* 3:1−25

Allison, Paul. D.
2002 *Missing Data*. Thousand Oaks, CA: Sage.

Behrman, H. R.
2000 Development, Economics. In, *International Encyclopedia of the Social and Behavioral Sciences* (N. J. Smelser and P. B. Baltes, eds.). New York: Pergamon. Pp 3566−3574.

Boserup, Esther
1981 *Population and Technology*. London: Blackwell.

Carmines, Edward G. and Zeller, Richard A.
1979 *Reliability and Validity Assessment*. Thousand Oaks, CA: Sage.
Chick, Gary
1997 Cultural complexity: The concept and its measurement. *Cross-Cultural Research* 31:275−307.
Cressy, David
1981 Levels of illiteracy in England 1530−1730. In, *Literacy and Social Development in the West*. (Harvey J. Graff, ed.). New York: Cambridge University Press. Pp. 105-124.
Demenay, Paul and George McNicoll, Eds.
2003 *Encyclopedia of Population*. New York: Macmillan Reference USA.
Denton, Trevor
2004a Cultural complexity revisited. *Cross-Cultural Research* 38:3−26.
2004b Reconstruction of prehistory by quantitative ethnographic analogy: Data based tests. *World Cultures* 15:2−27.
2007a Behavioral relations for components of recent preindustrial modernization: Quantitative assessment. *Ethnology* 45:229−254.
2007b Unit of observation in cross-cultural research: Implications for sampling and aggregated data analysis. *Cross-Cultural Research* 41:3−31.
2007c Yet another solution to Galton's problem. *Cross-Cultural Research* 41:32−45.
2008a Anthropological forecasts extrapolating correlates of modernization. *World Cultures eJournal* 16(1) Article 4. (http://repositories.cdlib.org/wc/worldcultures/)
2008b Indexes of validity and reliability for cross-societal measures. *Cross-Cultural Research* (in press).
Divale, William T.
2004 Coded variables for the Standard Cross-Cultural Sample. *World Cultures* CD data disk. *World Cultures* 14:194.
Fagan, Brian. M.
2004 *People of the Earth*. 11th edition. Upper saddle River, NJ. Pearson Education.
Gaur, Albertine
1984 *A History of Writing*. London: British Library.
Inglehart, Ronald
2000 Modernization, sociological theories of. In, *International Encyclopedia of the Social and Behavioral Sciences* (N. J. Smelser and P. B. Baltes, eds.). New York: Pergamon. Pp 9965−9971.
Irwin, P. H.
1975 An operational definition of societal modernization. *Economic Development and Culture Change* 23:595−613
Kutner, Michael. H., Christopher. J. Nachtscheim and John. Neter
2004 *Applied Linear Regression Models*. 4th Edition. Boston: McGraw-Hill Irwin.
Lerner, Daniel
1968 Modernization. In, *International Encyclopedia of the Social Sciences* (David. L. Sills, ed.). New York. Macmillan. Pp 386−395.

Levinson, David and Martin. J. Malone
1980 *Toward Explaining Human Culture*. New Haven: HRAF Press.

Maddison, Angus
2003 *The World Economy: Historical Statistics*. Paris: OECD.

Marsh, Robert
1967 *Comparative Sociology*. New York: Harcourt Brace.

Murdock, George Peter and Catarina Provost
1973 Measurement of cultural complexity. *Ethnology* 12:379−392.

Murdock, George Peter and Douglas R. White
1969 The Standard Cross-Cultural Sample. *Ethnology* 8:329− 369.

Ong, Aihwa
2001 Modernity: Anthropological aspects. In, *International Encyclopedia of the Social and Behavioral Sciences* (N. J Smelser and B. Baltes, eds.). New York: Pergamon. Pp 9944−9949.

Peregrine, Peter N.
2003 Atlas of Cultural Evolution. *World Cultures* 14:2-88.

Peregrine, Peter N. and Melvin Ember
2001-2 *Encyclopedia of Prehistory*. Volumes 1–9. New York: Plenum.

Service, Elwin
1964 *Primitive Social Organization: An Evolutionary Perspective*. New York: Random House.

Spiegel, Murray R., John Schiller, and R. Alu Srinivasan
2000 *Theory and Problems of Probability and Statistics*. 2nd Edition. Schaum's Outline Series, New York: McGraw-Hill.

United Nations Population Division
2007 *World Population Prospects: The 2006 Revision*. New York.

World Bank
2007 *World Development Indicators Online* (WDI Online). Washington, DC. (http://web.worldbank.org/data)