The Knowledge Structure of Individuals and Cohorts in American Knowledge Economy*

by

Lingxin Hao
Department of Sociology
Johns Hopkins University
3400 N. Charles Street
Baltimore, MD 21218
Tel: (410) 516-4022
Fax: (410) 516-7590
Email: hao@jhu.edu

and

Siri Warkentien
Department of Sociology
Johns Hopkins University
3400 N. Charles Street
Baltimore, MD 21218
Email: siriwarkentien@gmail.com

September 2012

* Prepared to be presented at the 2013 Annual Meetings of Population Association of America, New Orleans, Louisiana, April.

The Knowledge Structure of Individuals and Cohorts in American Knowledge Economy

Abstract

This paper seeks to describe and explain the dynamics of knowledge structure along the individual timeline and the cohort timeline in American knowledge economy. Three advances in our investigation include: (1) a theory-based 3-dimension knowledge categorization that extracts the essence of numerous degree fields and occupation fields; (2) a conceptual model that integrates the interdependent institutional transformations of higher education and industry and the impact of immigration; and (3) an application of the latent transition analysis (LTA). LTA first abstracts the 3-dimension categorization for fields at each level of degrees and occupation into two latent classes. It then analyzes each transition among the latent classes, from which we can identify the knowledge structure throughout all transitions. The analysis uses the uniquely rich data on degree and occupation fields for a large, nationally representative sample of about 100,000 college graduates from the National Survey of College Graduates (NSCG 2003).

The Knowledge Structure of Individuals and Cohorts in American Knowledge Economy

Extended Abstract

Introduction. A *knowledge economy* can be defined as an economy based on knowledge-intensive activities in production and services that contribute to an accelerated pace of technical and scientific advance, as well as rapid obsolescence (Powell and Snellman 2004). This notion is conceptually clearer than the notion of “postindustrial economy” that only points out timing but not the nature of the economy. Since the post-WWII period, especially the 1980s, the U.S. economy has transformed to a knowledge economy, in which the scientific workforce plays a central role. The literature has examined the rise of the knowledge economy from several perspectives, such as how new industries have demanded workers with different knowledge (Powell and Snellman 2004), how research and development (R&D) activities have increased (Boroush 2008), and how study fields have been expanded and determine earnings (Robst 2007). Less known, however, is how the knowledge structure embodied in individuals and cohorts has evolved. For example, individuals may pursue a traditional field in college and add a new frontier field for the Master’s degree, a “discipline-plus” model increasingly adopted by universities. The influx of skilled immigrants adds complexity to the individual knowledge structure. At the cohort aggregate level, we may see gradual changes or sharp shifts along the cohort timeline. These movements of the population, or more precisely, the scientific workforce, will help better understand the rise of American knowledge economy.

Objectives. The central objective of this paper is to describe and explain the dynamics of knowledge structure along the individual timeline and the cohort timeline in American knowledge economy. To this end, we make several advances in our investigation. First, we establish a theory-based dimensional knowledge categorization that extracts the essence of numerous degree fields and occupation fields. Second, we develop a conceptual model by integrating the interdependent institutional transformations of higher education and industry and the impact of immigration. From this framework we derive testable hypotheses regarding knowledge complexity of individuals and the uneven distribution of this complexity by cohorts and nativity. We apply the latent transition analysis (LTA) to test our hypotheses. LTA first abstracts the 3-dimension categorization for fields at each level of degrees and occupation into a small number of latent classes. It then analyzes each transition among the latent classes, from which we can identify the knowledge structure throughout all transitions. The analysis uses the uniquely rich data on degree and occupation fields for a large, nationally representative sample of college graduates from the National Survey of College Graduates (NSCG 2003).

Human Capital and Theory-Based 3-Dimension Fields of Degree and Occupation. Classical human capital theory distinguishes between general human capital measured by educational attainment and specialized human capital obtained through work experience in specific occupation and industry (Becker 1975; Tam 1997). Educational attainment captures only the "vertical" dimension and ignores the "horizontal" dimension of study fields (Charles and Bradley 2002). We expand specialized human capital to include human capital acquired from fields of study and occupation, which are numerous and changing over time. Past research examined selected popular fields or collapses fields into natural science, engineering, social science and humanity, which fail to capture within-category change. To address this limitation, we add a things-people dimension to the existing hard-soft and basic-applied classification in sociology of science to capture the production vs. service economic activities.

The three dimensions have theoretical rationales. *Mathematization* is used to describe the extent to which mathematics are used in organizing the body of knowledge with considerable precision and rigor, fostering consensus, and ultimately making a science "hard" (Storer 1967, 1972). *Practicality* is used to identify an applied field which has derives from the corresponding basic field but is concerned with the practical world (Biglan 1973). *Interactivity* helps distinguish between people-related service and things-related production (Spenner 1983; Marini et al 1996). These rationales guide us to classify 141 fields of education and 128 occupation fields: the level of mathematization classifies hard vs. soft, the level of practicality classifies basic vs. applied, and the object of interactivity classifies things vs. people.

The Conceptual Model. We consider two lines of theoretical thought – institutional and demographic forces – in examining the knowledge structure of individuals and cohorts, whose timeline unfolds under specific institutional environments and demographic movements.

Interdependent Transformation of Academic and Industrial Institutions. Using an institutional perspective, we highlight institutional change, particularly interdependent institutional transformation. To start, governmental policy may significantly shift the direction of relevant institutions. By allowing universities to own intellectual property of innovations, the Bayh-Dole Act of 1980 created long-lasting opportunity to bring university and industry institutions closer. As a result, a "double transformation" of university and industry has been happening (Moore, et al. 2011). The increasing streams of funding from industries have favored hard over soft disciplines, shifted research priorities from fundamental questions to practical issues, and preferred research activities dealing with things to people. The availability and volume of funding lead universities to adjust their strategies in supporting the expansion of selected fields. At the same time, university research has increasingly adopted industry-oriented language and culture. The result is the movement of university research towards industrial research. In the other direction, the orientation of policy makers has been scientized to base public policy (e.g., FDA's regulations on new drugs) on scientific evidence. This pushes industries to move toward universities. The institutional consequence is the uneven convergence of university research and industry research, favoring hard, applied, and things fields. The social consequence is differentiated labor market rewards between workers with different field knowledge. The university and industry institutional environment and the anticipated labor market reward have reshaped the knowledge structure of individuals and cohorts, tilted toward hard, applied, and things fields overtime.

Impact of Skilled Immigration. Among demographic forces gender is the most researched in the science literature (e.g., Grogger and Eide 1995; Xie and Shauman 2003). The role of skilled immigration in the changing knowledge structure has less been studied, however. We argue that because skilled immigration effectively brings to the United States the global economy consisting of both industrial and knowledge economies, its impact may be countervailing to the changing pattern of gender and race/ethnicity within the country. The flows of labor are path dependent on the historical and geopolitical relations between the United States and sending countries (Massey et al. 1993). In this context, skilled immigration is dependent on the affinity of the education systems between the origin and the host. The higher education systems in former British or American colonial societies (e.g., India, the Philippines, Nigeria, and Jamaica) share common education organizational structure and instructional language (Altbach 1998). Other developing countries are also influenced by the universalism in education (Meyer et al. 1997). When study fields are concerned, the developing world adopts the developmental strategy and highly values technology for manufacturing production. Immigrants are more likely to specialize in hard, applied, and things fields than their native-born counterparts. As a result, the inflows of high-skilled immigrants increase the pool of workers with stable hard, practical, and things-oriented fields. The immigration policy shift in 1990 has facilitated large volumes of incoming skilled immigrants, slowing the change in knowledge structure of American scientific workforce.

Hypotheses. Integrating the institutional and demographic arguments, we derive two sets of testable hypotheses. Individual decisions on what field to pursue for the Bachelor's degree and advanced degrees and what occupation field to enter are influenced by both institutional and demographic forces. The first set of hypotheses concerns individuals within birth cohorts under the same broader institutional and demographic context. Because the intensive knowledge-based activities dealing with both production and services, we expect to see a tendency of transition toward degree fields that are harder and more applied in advanced degrees (H1a) as well as occupation fields that are harder and more applied (H1b). Within birth cohorts, we also expect to see the transition tendency in H1a and H1b to be weaker for immigrants than natives (H1c and H1d, correspondingly). The second set of hypotheses concerns inter-cohort patterns. Between cohorts, we expect to see a sharper shift after 1990 given the university-industry convergence at a more mature stage (H2a and H2b correspondingly). This inter-cohort shift is expected to be weaker among immigrants (H2c and H2d correspondingly).

Data and Methods. The project draws on data from the National Survey of College Graduates (NSCG 2003) that provide unique education, employment and demographic information for a large probability sample of 100,402 college graduates in 2003. Three steps of analysis address the knowledge structure of degree(s) and current occupation conditional on the level of degrees – Bachelor's, Master's, and doctoral. Table 1 shows the three analytic samples' size and degree level distribution by birth cohorts and nativity. The cohort pattern of degree levels reflects more of the age pattern but the nativity pattern shows the higher attainment among the foreign born. Because individuals are not only self-selected but also influenced by the institutional environment in deciding on the level of degree, we will address this potential sample selection bias by estimating an ordered logit model of the three levels of degree as a function of birth cohort, nativity, gender, race/ethnicity, and parental education (the latter is an instrumental variable), from which we create inverse Mill's ratio to be included in the substantive analysis of knowledge structure.

The dependent variables are three indicators of hard-soft, basic-applied, and things-people, characterizing degree and occupation fields. For example, computer engineering is hard, applied, and things, whereas sociology is soft, basic, and people. The key explanatory variables are 4 birth cohorts (born before 1949, 1949-58, 1959-68, and after 1968) and immigrant status. The analysis controls for gender and race/ethnicity (white, black, Hispanic, and Asian).

We use latent transition analysis (LTA) to test our hypotheses. The key difference between LTA and the econometric transition analysis (stayer-mover models) is that LTA allows for a finite number of distributions of transitions as opposed to assuming one overall distribution (Collins and Lanza 2010). LTA identifies latent class membership at multiple time points as well as models transitions from one time point to the next, constituting a longitudinal mixture model (Nylund 2007). In our study, LTA first identifies latent classes for the 3 dimensions of fields rather than the typology $2^3 = 8$ manifest classes. LTA then estimates the transition patterns from one latent class of fields to a next latent class of fields. These data reduction methods can effectively highlight the distinctly different transition paths as shaped by institutional and demographic forces.

The diagram in Figure 1 shows the model applied to people with the highest degree at the Master's level. Fitting an LTA model provides sets of latent status prevalences (the proportion falling into a latent class at each time point), item-response probabilities (for manifest items conditional on membership in a given latent class), and transition probabilities (for change in latent class membership over time). As commonly practiced in LTA, we constrain the item-response probabilities to be equal across time in our analysis.

Preliminary Results. Present below are results from preliminary LTA analysis, which is based on the Master's degree sample respondents without correcting for potential sample selection bias. Thus these results should be read as empirical associations and not for hypotheses testing.

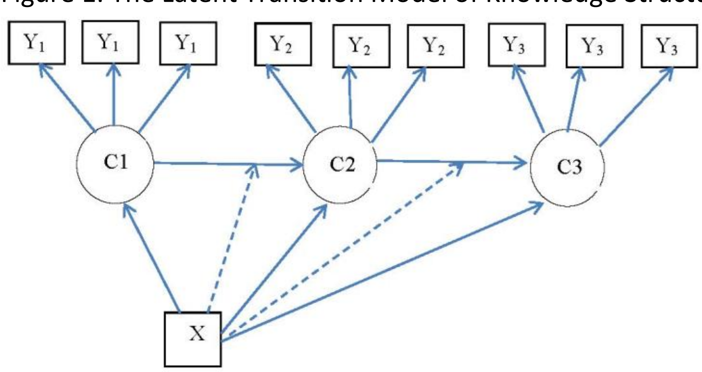
Two latent classes at each educational stage and occupation specified S – harder (0.77), more applied (0.64), and dealing with things (1.00), – and NS – softer (0.91), more applied (0.91), and dealing with people (1.00). The proportion of respondents falling in S or NS is freely estimated. Table 2 shows that, among respondents with Master's, the proportion in S decreased from 0.55 at Bachelor's to 0.40 at Master's, and remained stable at occupation. The prevalence distribution is uneven, however. Men exhibited an increase in S from Master's to occupation whereas immigrants exhibited an increase in S from Bachelor's to Master's. The total prevalence trend appears to be dominated by the immigrant pattern. These bivariate descriptive patterns provide us hints to further examine the partial patterns from the LTA estimates.

Table 3 shows the estimates for the covariates on the membership in the S latent class, again among respondents with Master's. Men, Asians, and later-born are more likely to be engaged in fields that are harder, more applied, and dealing with things; women, blacks and Hispanics, and early-born are more likely to be engaged in fields that are softer, more applied, and dealing with people, all else constant.

The knowledge structure of individuals and cohorts can be overviewed from the model-based estimates of transition patterns. The transitions that lead to changing prevalences of the S vs. NS latent classes provide a good description of the knowledge structure dynamics. The preliminary analysis results in Table 4 show that transitioning from S to NS (S-NS-NS (.13) and S-S-NS (.09) totaling .22 of respondents) is more common than from NS to S (NS-S-S (.02) and NS-NS-S (.04) totaling .06 of respondents). A higher proportion of respondents remain in NS (.36) than in S (.27) throughout their education and into their occupation. When transitions do occur, we see more shifts from S to NS in the education process (from bachelor's to master's), while more shifts occur from NS to S between master's and occupation.

Analysis to Be Completed. The final analysis will (1) correct for sample selection bias; (2) add the specification of the influence of covariates on transitions; (3) estimate a nested set of LTA models with various interaction terms to directly test each hypothesis; and (4) perform sensitivity analysis for cohort definition by birth year vs. Bachelor's admission or graduation year, as well as first occupation vs. current occupation.

Figure 1. The Latent Transition Model of Knowledge Structure



Note: This model is an example for individuals with the highest degree at the Master's level. X is a vector of observed covariates; Y_1, Y_2, Y_3 are the observed indicators for hard-soft, basic-applied, and things-people dimensions of fields; C_1, C_2, C_3 are the latent classes of fields of the Bachelor's degree, Master's degree, and occupation. The dotted arrows specify that the covariates affect the transitions from C_1 to C_2 and from C_2 to C_3.

Table 1. Level of Degree by Birth Cohort and Nativity

Total Sample		Proportion						
Level of Degree	n	Total	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Native	Immigrant
Bachelor's	42,516	0.574	0.469	0.557	0.597	0.656	0.610	0.434
Master's	24,494	0.331	0.385	0.345	0.306	0.298	0.321	0.367
Doctoral	7,083	0.096	0.146	0.099	0.096	0.046	0.069	0.200

Table 2. Descriptive distribution of covariates by latent class for respondents with Master's as highest degree

	Bachelor's		Master's		Occupation	
Covariate	S	NS	S	NS	S	NS
Male	0.66	0.38	0.67	0.44	0.70	0.42
White	0.70	0.76	0.68	0.76	0.68	0.76
Black	0.05	0.10	0.04	0.09	0.04	0.09
Hispanic	0.05	0.07	0.05	0.07	0.05	0.07
Asian	0.20	0.06	0.24	0.07	0.23	0.07
Born before 1949	0.18	0.20	0.17	0.20	0.15	0.21
Born 1949-1958	0.31	0.38	0.31	0.37	0.30	0.37
Born 1959-1968	0.33	0.26	0.33	0.28	0.35	0.27
Born after 1969	0.18	0.16	0.19	0.16	0.20	0.16
Immigrant	0.32	0.12	0.36	0.14	0.35	0.15
Prop. in latent class	0.55	0.45	0.40	0.60	0.39	0.61

Note: The conditional probabilities for latent classes are as follows: S=hard (.77), applied (.64), things (1.00); NS=soft (.91), applied (.91), people (1.00). For example, given membership in the NS-class for BA, the probability of a "soft" educational field is .91.

Table 3. Predictors of membership for BA, MA, and occupation-level latent classes for respondents earning Master's as highest degree

	Bachelor's		Master's		Occupation	
Covariate	S		S		S	
Male	1.062	**	0.448	**	0.898	**
Black	-0.645	**	-0.435	**	-0.376	**
Hispanic	-0.478	**	-0.475	**	-0.262	**
Asian	0.477	**	0.373	**	0.505	**
Born 1949-1958	-0.036		0.016		0.190	**
Born 1959-1968	0.307	**	0.084	#	0.516	**
Born after 1969	0.306	**	0.153	**	0.533	**
Immigrant	0.894	**	0.696	**	0.254	**
Bachelor's latent class S	—		2.540	**	—	
Master's latent class S	—		—		2.290	**

NOTE: S is characterized as primarily hard-applied-things and NS is primarily soft-applied-people.
** p<.01; * p<.05; # p<.10

Table 4. Unconditional and conditional transition probabilities among respondents with master's as highest degree

Unconditional transition probabilities		Conditional transition probabilities		
NS - NS - NS	0.36	Probability transition to …	… MA latent status	
NS - NS - S	0.04	Conditional on BA latent status	S	NS
NS - S - NS	0.02	S	0.65	0.35
NS - S - S	0.02	NS	0.10	0.90
S - NS - NS	0.13		… Occupation latent status	
S - NS - S	0.06	Conditional on MA latent status	S	NS
S - S - NS	0.09	S	0.72	0.28
S - S -S	0.27	NS	0.17	0.83