

On the behavioral specification and multivariate neural network estimation of cognitive scale economies

Nina Kajiji^{1,*} and Gordon H. Dash, Jr.²

¹ *The NKD-Group, Inc. and University of Rhode Island, Kingston, USA*

² *College of Business Administration, University of Rhode Island, Kingston, USA*

Abstract. It is universally accepted that sovereign economic development is inextricably linked to the efficiency of local education production. In the United States, the task of designing policies to enhance education attainment by students is an effort that is largely reliant upon state-wide and local policy initiatives. Because policy-making skills vary across these disaggregated educational bodies, it has become increasingly important for government leaders at all levels to assist regional and state leaders in the process of identifying effective policy metrics. Using comparative solutions from univariate regression and artificial neural network models, this paper presents augmented evidence that a multivariate artificial intelligence-based model produces more plausible estimates of traditional production elasticity metrics when compared to its parametric alternative. The scale elasticity metrics obtained by solving a multivariate OLS (MOLS) model are compared directly to those generated by solving a multivariate non-parametric multiple objective radial basis function artificial neural network model (MRANN).

Keywords: multiple equation models; neural networks; quantitative policy modeling; education and development

Received September 2012. Accepted January 2013

Introduction

Around the globe, the motivation for public expenditures on education is founded in the belief that the education of children and young adults is fundamental to the future economic growth and a lasting democracy. Contemporary research offers many examples of including education factors in growth models in order to better understand the implications of education policy on macroeconomic performance as well as local and regional socio-economic economics. For example, Barro and Sala-i-Martin (1995) have found that years of secondary and higher education exposure contribute positively toward economic growth. More importantly, the societal benefits that accrue from economic growth tend to spread and manifest themselves as positive attitudes and discipline in workforce productivity along with the ability of the workforce to incorporate new technologies (Lucas, (1988)). Romer (1990) provides evidence that the creation new ideas is a direct function of investment in human capital. Rebelo (1991) added to this line of reasoning by introducing physical capital as an additional input to the production of economic well-being. A number of other studies involving small growing economies have also examined the relationship between the production embedded in human capital and sources of workplace productivity (Haouas and Yagoubi (2005), Schiller, et. al. (2002)). In a slightly different approach, Park (2006) investigated empirically

* *Correspondence:* Nina Kajiji, Department of Computer Science and Statistics, University of Rhode Island, Kingston 02881 USA.
E-mail: nina@uri.edu

the growth implications of a dispersed population in terms of educational attainment levels. The commonality among these studies and others is an agreement that there is a statistically significant link between human productivity and economic growth.

The purpose of this paper is to extend the univariate modeling efforts of Kajiji and Dash (2012) to a multivariate specification. This is a specification that is capable of a joint and simultaneous estimation of the production elasticity metrics that explain variability across multiple educational achievement indexes at the high school level in the State of Rhode Island, USA. Specifically, as in the case of the prior univariate study, the multivariate modeling efforts will target the same two statewide dependent variables – the three-year Rhode Island mathematics achievement index and the three-year English-Language Arts index. Owing to the fact that educational achievement models tend to lack temporal stability and are also difficult to estimate adds to the justification to introduce a comparative approach within a multivariate characterization of educational attainment. The paper proceeds by comparing alternate parametric and nonparametric multivariate models. In addition to the established multivariate ordinary least squares approach this paper introduces a multivariate nonparametric alternative – the K7 radial basis function artificial neural network, or K7-MRANN (MRANN) available in WinORS_{e-AI} (2011). The MRANN algorithm is shown to be an efficient multivariate modeling extension to the Kajiji and Dash univariate K4 RANN.

Modeling production scale economies

Global interest in understanding the implications of educational attainment has been the focus of a number of researchers. Clark (2008) notes that, in the global economy, staged supply chain production that reaches across countries requires each host country to sustain educational attainment to remain part of the chain. In a long-term study of socioeconomically heterogeneous populations in South Africa, Fedderke (2002), found that educational factors of production do matter and that the relationship between inputs and outputs (multiple target variables) was statistically strong even when the variation of such inputs is low. This view is corroborated by Zimba and Tiraboschi (2010) in a study of school-to-work perspective in Sub-Saharan Africa. The collective reasoning is that any change to the factors of production is better than no change at all. One of the key focal points of these studies has been the role of the teacher. Lin (2010) found that both education-level elasticity estimates for teacher quality and teacher years were both positive and significant. Wang and Chao (2008) found positive education spillovers in certain industries in the Taiwanese economy. The authors note that a more educated workforce is capable of achieving a three- (two-) fold increase in productivity for worker's who go on to attain a master's (bachelor's) degree. Taken within the context of this review, these results argue for any effort to improve positive-valued variation in relevant factors of production as a means of improving the short- and long-term benefits to any socio-economy entity. In the following sections we develop a production theoretic model and then estimate scale effects by two different econometric models. That is a comparison between the parametric Multivariate Ordinary Least Squares (MOLS) method and the nonparametric K7-MRANN.

The double-log production function

The double log production functional form is probably best known by the popularization of the Cobb-Douglas (1928) functional form and its representation of the functional relationship of an output to factor inputs can be expressed as:

$$f(x) = A \prod_{i=1}^n x_i^{\beta_i}, \beta_i > 0, i = 1, 2, \dots, n \quad (1)$$

This functional form has the following properties: a) strict monotonic – if $x' > x$ then $f(x') > f(x)$; b) quasi-concavity – $V(y) = [x : f(x) \geq y]$ is a convex set; c) strict essentiality – $f(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n) = 0$ for all $x_i > 0$; d) the set $V(y)$ is finite, nonnegative, real valued and single valued for all nonnegative and finite x . It is also continuous and everywhere twice-continuously differentiable; e) $f(x)$ is homogenous of degree $k = \text{sum of } \beta_i$.

In order to employ the double-log specification (or the more general translog specification) shown in equation 2, it becomes necessary to make some transformation to the zero-value arguments.

$$\ln(y(x)) = \ln A + \sum_{i=1}^n \beta_i \ln(x_i), \beta_i > 0 = 1, 2, \dots, n \tag{2}$$

The transformation of a modified zero-element argument is generally accomplished in one of two ways: a) by replacing the observation with 1.0 so that $\ln(x_i) = 0$ when $x_i = 0$; or, b) replacing the zero with very small values. For a more detailed discussion of zero-valued transformation, see MaCurdy and Pencavel (1986); Jacoby (1992); and Soloaga and Moss (2000). We recognize that these procedures are arbitrary and force the production function to include input quantities that are not actually observed.

Multivariate multiple linear regression modeling (MOLS)

There are at least two reasons why it is important to extend the modeling of educational scale economies to a multivariate method. First, the extant literature has shown that the response variables are correlated (for example, see Fishback and Baskin (1991) and Vanderhart (2006)). Second, multivariate tests provide a way to understand the structure that binds the relations among the factors of educational production across separate response measures. To this end it is possible to: a) uncover the relative importance of two or more educational response measurements and b) determine how the predictor variables contribute to the response structure.

In this research we consider the multivariate extension to the multiple-linear regression model. Specifically, we consider the model relationship between q responses y_1, \dots, y_q and a single set of p predictor variables x_1, \dots, x_p . Each of the q responses is assumed to follow its own regression model, i.e.

$$\begin{aligned} y_1 &= \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \dots + \beta_{p1}x_p + \varepsilon_1 \\ y_2 &= \beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \dots + \beta_{p2}x_p + \varepsilon_2 \\ &\vdots \\ &\vdots \\ y_q &= \beta_{0q} + \beta_{1q}x_1 + \beta_{2q}x_2 + \dots + \beta_{pq}x_p + \varepsilon_q \end{aligned} \tag{3}$$

Restated, the multivariate linear regression model is: $Y = \beta X + \varepsilon$, with $\beta_0 = 0$, $E(\varepsilon) = 0$ and $Var(\varepsilon) = \Sigma$. We use the multivariate OLS model to estimate the linear regression of the double-log production theoretic model to extract multivariate education scale economies.

The multivariate multi-objective MRANN (K7-MRANN)

Unlike the parametric approach of MOLS, the ANN modeling approach does not require *a-priori* knowledge, constants, or assumptions. The ANN approach to modeling inherits many of the attractive elements associated with cognitive reasoning. However, despite the many positive attributes associated with this approach to behavioral econometric modeling, some well-known model disabling pitfalls remain. Notably, both uni- and multivariate-ANNs characteristically suffer from two modeling ills: the “curse” of dimensionality and inflated residual sum of squares. The foundation for how to attack these twin ills in univariate modeling have been attacked successfully by Kajiji and Dash (2012). Additionally, this research also provided the methodology for the use of univariate RANN methods to estimate complex production-theoretic quasi-scale economies based from a double-log generalized form of the educational achievement production function. This paper explicitly proposes extensions to this prior research effort in order to reconcile and accommodate the development of a multivariate modeling specification of the K4-RANN to the K7-MRANN (or, MRANN). To that end the augmented multivariate supervised learning function is stated as:

$$Y = f(X) \tag{4}$$

where Y , the target matrix with q number of outputs, is a function of the input matrix X with p number of inputs (Figure 1).

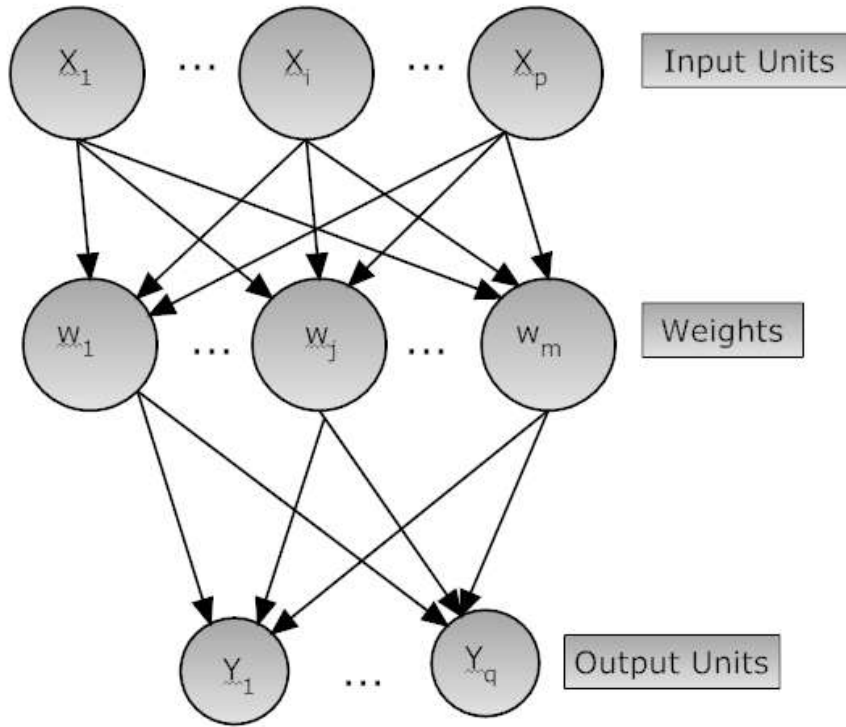


Fig. 1. K7-MRANN Network

The function represented in equation 4 can be restated as:

$$y_l = f(x_i) = \sum_{j=1}^m w_{jl} h_j(x) \tag{5}$$

where, m is the number of basis functions, h is the number of hidden units, w is the weight matrix, and $i = 1..p$ where p is the number of input units; and $l=1..q$ where q is the number of output units.

The flexibility of $f(x)$ and its ability to model many different functions across multiple targets is inherited from the freedom to choose different values for the weight matrix, w . Within the MRANN architecture, the multivariate weight matrix is found through optimization of an objective function. This is equivalent to minimizing the multivariate sum of squared errors (SSE) as measured by:

$$SSE_l = \sum_{i=1}^p (\hat{y}_l - f(x_i))^2 \tag{6}$$

In this paper, we introduce the augmented MRANN hereafter referred to as the K7-MRANN. This modified MRANN extends the univariate minimization equation to its multivariate multi-objective counterpart as stated in equation 6.

$$\underset{k_l}{\operatorname{argmin}} \left[\left(\zeta \sum_{i=1}^p (y_l - f(x_i | \bar{k}_l))^2 + \sum_{j=1}^m k_{jl} w_{jl}^2 \right) \right] \tag{7}$$

For each equation the computationally efficient Bayesian enhanced K7-MRANN algorithm assures that q individual functions are mapped for smoothness and accuracy. In summary, the multivariate K7-MRANN incorporates the algorithmic enhancements evidenced in the univariate predecessor algorithm (K4-RANN). Both the K4-RANN and K7-MRANN have the ability to reconcile the twin evils that deter efficient ANN modeling: data dimensionality and inflated residual sum of. Lastly, we expect to uniformly extend the double-log specification of the educational production function into a multivariate specification; that is, a specification with multiple log transformed educational response terms.

The data

Here, we briefly describe the complete list of the variables that were first considered in the univariate analysis provided by Kajiji and Dash (2012). The data was obtained from the population of high-schools in the state of Rhode Island for the academic year (2004-2005). Predictor variables representing system-wide demographic characteristics to capture assessment results were obtained from the Rhode Island Department of Education (RIDE). At the time of the first research study, data selection for these predictor variables were consistent with the mandates stated in the *No Child Left Behind Act (NCLB) (2001)*. Additional independent variables were obtained from the National Center on Public Education and Social Policy (NCPE) HiPlaces survey; a survey which is administered to students, parents, teachers, and administrators of all public schools in the state of Rhode Island. The predictor variables in both the earlier study and this one represent a subset of those published in the *High Performance Learning Community Assessment: School Improvement Self-Study Surveys* (Felner, (1998, 1987)). The multivariate modeling specification in this paper started by including the 20 predictor variables that were found to be statistically significant in the prior univariate analysis (Kajiji and Dash (2012)). By way of summary, in that analysis (which consisted of two estimated univariate RANN equations) the two dependent (target) variables were: a) *Three year Math Index (MAI)* and b) the *Three year ELA Index (ELAI)* computed as an average of the individual index for academic year (AY) 2003, 2004, and 2005. As previously stated, see Kajiji and Dash (2012) for a complete list of variables and all univariate results.

Model and results

The multivariate education production model replaces the two equation approach required by a univariate specification. As in the prior research paper a simultaneous comparative study is provided by solving a parametric specification using a MOLS specification. For comparative science, an alternative estimation model is solved by applying the MRANN cognitive modeling approach.

$$\text{MOLS Model: } \ln(\text{MAI}, \text{ELAI}) = \beta \ln(\mathbf{X}) + \varepsilon; \text{ where } \varepsilon \sim N(0, \Sigma) \quad (8)$$

$$\text{K7-MRANN Model: } \ln(\text{MAI}, \text{ELAI}) = \mathbf{w} \ln(\mathbf{X}) + \varepsilon; \text{ where } \varepsilon \sim N(0, \Sigma) \quad (9)$$

To achieve efficient solutions, the models in equation 8 and 9 were solved iteratively. Statistically insignificant variables were removed one-at-a-time for the MOLS analysis based on the Wilks' Lambda metric and a review of overall model performance. The optimal parameters of the \mathbf{X} matrix produced by the iterative MOLS procedure are presented in Table 1. The table reports seven statistically significant predictor variables. These include: *Attendance*; *MAI (02-04)*; *ELAI (02-04)*; *Per Pupil Expenditure (PPE) for Classroom Teachers*; *PPE Classroom Materials*; *PPE Special Education*; and *Health Factor 2*. We note that *MAI (02-04)* and *ELAI (02-04)* are three year average index values for AY 2002 through 2004 inclusive. *Health Factor 2* is a composite index that is defined by an individual's health environment (e.g., social TV watching, sleeping habits, dietary quality, etc.)

Table 1. Multivariate Model Solution: MOLS Parameters and K7-MRANN Weights

Variable	Multivariate Analysis (MOLS)				K7-MRANN Analysis	
	Math (MAI)	ELA (ELAI)	Wilks' Lambda	Mult. Sig.	Math (MAI)	ELA (ELAI)
Attendance	0.248*	0.149*	0.745	0.007	0.320	0.232
MAI (02-04)	0.912*	-0.001	0.105	0.000	-0.448	-0.294
ELAI (02-04)	-0.060	0.826*	0.159	0.000	0.092	0.121
PPE Class Teachers	-0.100*	0.00002	0.752	0.008	0.024	0.224
PPE Class Materials	0.018*	0.012*	0.823	0.036	0.514	0.378
PPE Special Ed.	0.039*	0.007	0.677	0.001	0.058	-0.023
Health Factor 2	-0.012*	-0.003	0.793	0.019	0.499	0.450

* Univariate significance at least at the 95% level

A review of the Wilks' Lambda provides evidence that the two most important predictors for an education production model are: PPE Class Materials ($\lambda = 0.823$) and Health Factor 2 ($\lambda = 0.793$). By comparison, relying on the magnitude of the optimal MRANN parameter weights, the nonparametric cognitive MRANN model appears to corroborate the findings of its parametric counterpart.

Upon closer inspection of the comparative MSE matrix results presented in Table 2, the MRANN ELAI model is estimated with a smaller error (0.0085) than the corresponding MAI model (0.0295). These findings are supported by the reported metrics from solving the MOLS model. We note that a positive MRANN covariance (0.0156) indicates that over- (under-) prediction of MAI tends to be accompanied by over- (under-) prediction of ELAI. The same is true for the MOLS model but with an extremely small value of 0.00009; indicating that the MOLS covariance effect is negligible.

Table 2. MSE Matrices

K7-MRANN		MOLS	
MAI	ELAI	MAI	ELAI
0.0295	0.0156	0.00047	0.00009
0.0156	0.0085	0.00009	0.00014

Sections 4.1 through 4.6 of the paper provide an interpretation of policy implications induced by the findings obtained from solving the alternate model specifications. The section begins with an investigation of the *Attendance* effect and continues down to an examination of *Health Factor 2*. Section 4.7 closes our discussion on models and results by presenting a factor-analytic interpretation of latent informational content in model residuals. This latter interrogation is useful to uncover the ability of each alternate model to fully explore related variation among target and predictor variables.

Attendance

The act of attending school is important to individual achievement. Under the MOLS specification the elasticity for attendance reports that for every 10 percent increase in attendance rates, MAI will increase by 2.48 percent. Similarly, for the same percentage change in attendance under the ELAI model, the influence on ELAI performance increases by 1.49 percent.

Standing in contrast are the quasi-elasticity estimates produced under the MRANN modeling approach. Not only are the findings different but the elasticity metric values are more balanced. Again, for a 10 percent change in attendance the respective achievement indexes (MAI and ELAI) increase by 3.20 and 2.32 percent, respectively.

MAI (02-04) and ELAI (02-04)

Smith and Schumacher (2006) provide important evidence that links prior performance in mathematics with the potential for future success in applied disciplines that rely upon fundamental mathematics training. Their findings suggest that it is important for this empirical investigation of Rhode Island high schools to generate positive signed production elasticity estimates. For the results of this study we expect both *MAI (02-04)* and *ELAI (02-04)* will be positively signed and significant in its relationship to both current year MAI and ELAI.

The study finds that under MOLS modeling of Rhode Island data the elasticity estimates from *MAI (02-04)* are high for the MAI target variable but the estimated metric is low and negative for the ELAI target variable. From a policy perspective one can conclude that for a 10 percent change in *MAI (02-04)* expected MAI will increase by 9.12 while ELAI is expected to decline by a modest 0.01 percent. The obvious implication here is that a direct emphasis on mathematics improvement does improve mathematics performance but there is no evidence of a crossover effect to ELAI. Again, the MRANN quasi-elasticity estimates stand in contrast to those generated by MOLS. Under MRANN the respective findings from the *MAI (02-04)* index are (again) more balanced but both elasticity metrics are negative for the two target variables. Thus, for an expected 10-percent change in *MAI (02-04)* the current year MAI achievement index is expected to decline by 4.48 percent and the current year ELAI achievement index is expected to decline by 2.94 percent.

Based on the work of Smith and Schumacher, these findings offer a stark prediction for the future performance of Rhode Island high school students. Ideally, any time series prediction should produce a bullish view of expected student performance. Based on a comparative view of the alternate empirical models, it is clear that school administrators need to review and revise policy to achieve desired future growth. To do nothing assures the likelihood of a continued decline in the MAI dimension at the high school level. The conflicting findings between models for MAI provide enough evidence to encourage further interrogation of mathematics maturity and penetration at the high school level. The MRANN negative elasticity finding for the *MAI (02-04)* achievement index may well be indicative of the syndrome that is often referred to as “teaching to the test.” This syndrome has resulted due to the state-wide curriculum testing mandated by NCLB. As Riffert (2005) explains, the act of teaching to the test limits the teacher’s ability to use multiple creative teaching methods based around their students’ needs. Posner’s (2004) research previously argued that teaching to the test creates an atmosphere where students learn facts and materials without necessarily gaining a real comprehension of the subject matter beyond the ability to answer directed test-questions correctly. Given the strong statistical support for the overall MRANN solution, administrators should assign a reasonable level of credibility to the negatively signed quasi-elasticity metrics of -4.48 and -2.94 percent for MAI and ELAI, respectively. That is, a need for new policy directives is indicated by this finding.

The *ELAI (02-04)* empirical results deserve a similar interpretation under the MOLS model. That is, the MOLS elasticity metrics produce a small negative effect on the current year MAI and a strong positive effect for the current year ELAI. Again, in this dimension MRANN offers a thought-provoking contrast. The MRANN results show that the elasticity estimate for *ELAI (02-04)* is positively signed for both the MAI and ELAI target variables. Contributions from Slavin (2006) offer reconciliation support for this apparent conflict between *MAI (02-04)* and *ELAI (02-04)* metrics. In foresight, Slavin argued for school administrators to uncover the correlated relationship between teaching philosophy and student psychological impact as a means by which to mediate a cornerstone approach to the design and modification of the contemporary high school curriculum.

PPE classroom teachers (PPE/CT)

A recent study by Lavy (2012) reported that spending more money along with increased student interaction time at school on key tasks leads to broad-based increase in academic achievement with no behavioral costs (disciplinary and other distraction based issues). The alternate model findings presented in this paper are mixed as we relate the more narrowly defined role of PPE/CT to the Lavy findings. For a 10 percent change in PPE to support classroom teachers, MOLS reported a declining effect of -1.00 percent in current period MAI and trace increase for ELAI (<0.001 percent). By contrast, the MRANN quasi-elasticity metrics are consistent with the expectation that as teachers interact more and longer with students in a setting with improved funding support there will be a moderate positive effect for each achievement dimension. While this research does not delve into the formal measurement

relationship between PPE and classroom teachers, some additional inference can be drawn by combining the results as presented in section 4.2 (prior year efforts on current year achievement). A plausible argument can be made to reverse year-to-year achievement declines in Rhode Island high school achievement by providing more competitive compensation packages for teachers to encourage greater student-teacher interaction over a longer school week.

PPE classroom materials (PPE/CM)

Under the MOLS specification, the elasticity estimates for PPE/CM are positive for both MAI and ELAI. We find that for every 10 percent increase in PPE/CM, MAI will increase by a modest 0.18 percent and ELAI by a similarly modest 0.12 percent. We also note that the high Wilks' Lambda score indicates that the PPE/CM is one of the most significant predictor variables for explaining educational achievement in both MAI and ELAI. This finding is confirmed and strengthened by the MRANN model solution. The MRANN quasi-elasticity estimates are much higher. For the same expected change of 10 percent in PPE/CM, the reported quasi-elasticity parameters indicate a change in MAI and ELAI achievement of 5.14 and 3.78, percent respectively.

PPE special education (PPE/SE)

Special education expenditures are a significant educational and social issue in public school systems around the country. The current state of study within this dimension is aptly demonstrated in a recent study by Temple and Reynolds (2007). The authors present strong empirical evidence to suggest that in order for a school district to generate cost-benefit superiority in the production of positive economic returns to educational achievement from special education it is more beneficial to offer high-quality preschool programs compared to other types of educational intervention programs. In short, an early start is preferable to late intervention. The MOLS model provides an interesting but supportive view. The MOLS elasticity for PPE/SE reports that for every 10 percent increase in PPE/SE, both MAI and ELAI increase by disappointingly low amounts at 0.39 percent and 0.07 percent, respectively. These findings raise the question about whether pre-school identification of some special education needs based attention would lead to improved student performance as that student moves through the system to the high school level.

Standing in stronger support for the Temple and Reynolds assertion are the MRANN quasi-elasticity findings. As with the MOLS findings, the MRANN quasi-elasticity scale metric for the MAI dimension is positive but low (0.58 percent). However, the low expectations argued by Temple and Reynolds are starkly revealed by the quasi-elasticity for the ELAI dimension. Here, the MRANN model suggests that policymaker's should expect a decline in ELAI achievement of -0.23 percent for the same 10 percent increase in PPE/SE expenditures. In other words, by high school the intervention level is too late to impart a meaningful change in overall achievement.

Health factor 2

Extant research has clearly shown that health related factors have a direct impact on academic achievement. For example, disseminated findings by Murray (2007) clearly show that after controlling for gender, race/ethnicity, and grade level, there is a negative association between at-risk health behaviors and academic achievement among high school students. In the MOLS model solution, both elasticity estimates for the *Health Factor 2* dimension are negative. In policy terms, for every 10 percent increase in the health risk metric, MAI and ELAI will decrease by 0.12- and 0.03-percent, respectively. Strikingly different are the MRANN quasi-elasticity estimates. Both are positive. For every 10 percent increase in the *Health Factor 2* dimension, MAI and ELAI will increase by 4.99- and 4.50-percent, respectively. These conflicting empirical findings immediately call into question the construction of the Health Factor 2 dimension. Before school administrators attempt to fully invoke operation policy based on the reported findings, it would be advisable for education policymakers to review two important statistical issues. First, the loadings on this latent health factor must be confirmed under contemporary social structure to assess their correlation structure of a negative (positive) association with individual student health realities.

In the univariate study involving this same data it was argued that “...it is evident that lack of a proper breakfast, increased TV watching, and lack of sleep are the reasons this factor has a negative impact...” However, under the stronger multivariate modeling results it is important to interrogate whether it could be that a proper breakfast fails to take into consideration home versus school provided early arrival morning breakfast programs. Also, in today’s consumer electronic society one would have to ask if TV watching has been replaced by the ever-present, and possibly more beneficial, positive impacts of technology viewing (tablets, smartphones, etc)? The exercise of validating the construct of this particular health dimension (and possibly other orthogonal contributors) is left to future research.

An examination of estimated residuals

During the discussion that explored alternate model performance and the resultant policy implications gleaned from solving the comparative multivariate specifications, we encountered both corroborative and contradictory findings. This fact alone raises a question about the ability of the two comparative models to fully extract all explanatory variation among variables. To address this issue we turn our attention to the residuals under both parametric and nonparametric specification. In either case, if the model has achieved its appointed task the residual from the model will be devoid of latent and unexplained socio-economic content.

To investigate what, if any, latent content remains in the residuals we apply a principal components analysis (PCA) on the covariance matrix of the residuals. Subsequently, the PCA solution is subjected to a Varimax rotation. The Varimax sub-procedure is an orthogonal rotation of the PCA factor axes to maximize the variance of the squared loadings of a factor (column) on all the variables (rows). This has the effect of differentiating the original variables by extracted factor. That is, each factor will tend to have either large or small loadings of any particular variable. To the policymaker, a Varimax solution yields results which make it straightforward to label the latent content of each extracted factor.

To proceed, the parametric MOLS modeling framework generates a $n \times q$ matrix of residuals as does the non-parametric MRANN modeling application. Owing to the double-log transformation of all variables, the variates are measured in the same units with reported variance terms that are closely related (see table 2). Hence, the appropriate input matrix for the PCA is the covariance matrix of the $n \times (2q)$ residuals. This PCA output options did not invoke a display “fuzz” factor to eliminate low factor loadings. Additionally, all eigenvectors were retained to interrogate the informational content of the alternate model residuals.

Table 3 displays the results of the factor-analytic solution as obtained from SPSS (2010). If either model (MOLS or MRANN) fails to explain all economic variation among the study variables, then latent economic correlation will remain in the model’s residuals. A model that achieved its appointed task will result in a single factor domain where as a failure to achieve full explanatory effects by model simulation would produce multiple latent and easily labeled factor domains. Upon the inspection of table 3, Factor 1 is a single factor that clearly defines the residuals produced by the solution of the K7-MRANN model. That is, the residuals of both target variables (MAI and ELAI) from the MRANN multivariate model specification load on factor 1 at 0.9962 and higher. By contrast, the MOLS residuals spread across factors 2 and 3. This factor-analytic finding is clear evidence that the MOLS procedure could not jointly and simultaneously account for all latent economic content in both MAI and ELAI achievement. At best, the MOLS model results are incomplete in their explanation of model variations. At worst, the MOLS model may exhibit omitted variable bias.

Table 3. Rotated Factor Pattern

	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>
K7 Math	0.9967	0.0309	0.0202	-0.0720
K7 ELA	0.9962	0.0455	0.0192	0.0722
MOLS Math	0.0226	0.1708	0.9850	0.0000
MOLS ELA	0.0465	0.9841	0.1714	0.0004
Weighted Variance Explained	0.0377	0.0002	0.0005	0.0002

The component plot, figure 2, confirms the desired uni-dimensional structure of the MRANN residuals. The figure also shows the two-dimensional orthogonal structure for the MOLS residuals. In conclusion, when combined with the prior analysis of the alternate models policy implications the factor-analytic results provide strong additional statistical evidence to support the K7-MRANN nonparametric cognitive modeling approach when the objective is to estimate a production-theoretic state-wide educational achievement system at the high school level.

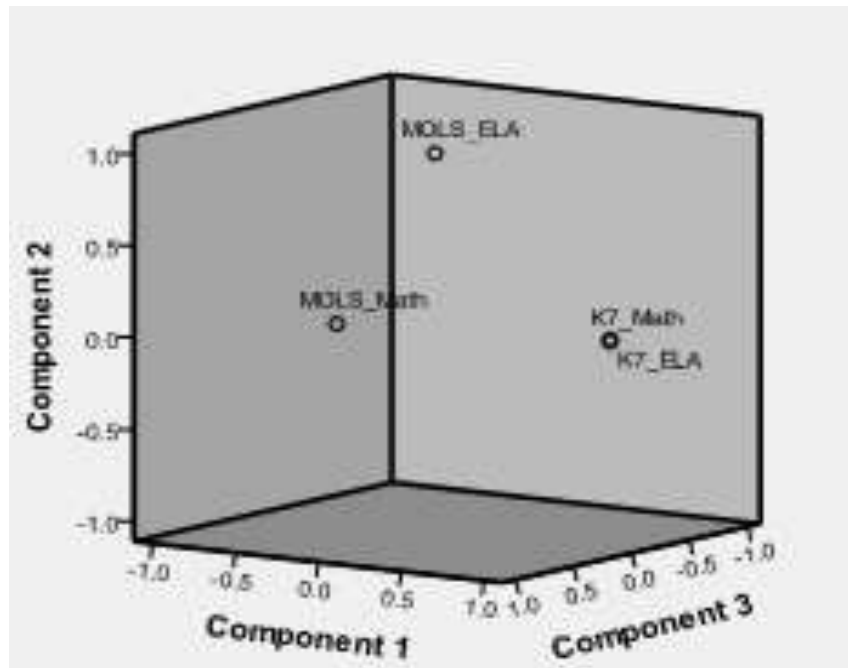


Fig. 2. Component Plot in Rotated Space

Conclusions

The research presented in this paper is guided by the global interest in gaining an authoritative understanding of what factors link educational achievement and economic development. To accomplish the investigation, the primary objective of this paper was directed at extending recent and important modeling efforts. Specifically, to capture joint and simultaneous impacts on multiple state-wide achievement indexes the paper extended univariate modeling specifications to a more robust multivariate structure. To that end, a new multivariate RBF ANN, the K7-MRANN was introduced, solved and compared to the parametric MOLS equivalent model. The comparative modeling results clearly overcame limitations inherent to univariate modeling while also offering better collaborative consistency with extant findings in the literature. For example, under multivariate modeling some policy dimensions, such as *Attendance*, generated consistent results in both parameter sign and magnitude under MOLS and MRANN model specification. However, new evidence was gleaned for other policy dimensions. Notably for dimensions *MAI (02-04)* and *ELA (02-04)*, the extracted elasticity estimates were not consistent. Lastly, some policy dimensions, such as *Health Factor 2*, provided enough contrast between the two alternate econometric models to call into question the structure of the independent predictor variable. Overall, this research provided new and convincing evidence that the multivariate MRANN cognitive approach to elasticity modeling is a more consistent and effective modeling tool than its MOLS counterpart in the estimation of production-theoretic functions of educational achievement at the high school level.

References

- _____. (2010). Ibm Spss Statistics, Version 19, SPSS, Inc., Chicago, IL
- _____. (2011). Winorse-Ai, The NKD Group, Inc., Providence, RI
- Barro RJ, and Sala-i-Martin X (1995). In: *Economic Growth* McGraw Hill pp. 424-432
- Clark DP (2008). Host Country Educational Attainment and Vertical Specialization. *Global Economy Journal* 8(2)
- Cobb CW, and Douglas PH (1928). A Theory of Production. *American Economic Review* 18(Supplement): 139-165
- Fedderke JW, and Luiz JM (2002). Production of Educational Output: Time-Series Evidence from Socioeconomically Heterogenous Populations -- the Case of South Africa, 1927-1993. Technical Report
- Felner RD (1998, 1987). *The High Performance Learning Community Assessment: School Improvement Self-Study Surveys*, National Center on Public Education and Social Policy
- Fishback PV, and Baskin JH (1991). Narrowing the Black-White Gap in Child Literacy in 1910: The Roles of School Inputs and Family Inputs. *Review of Economics and Statistics* 73(4): 725
- Haouas I, and Yagoubi M (2005). Openness and Human Capital as Sources of Productivity Growth: An Empirical Investigation from the Mena Countries. IZA Discussion Paper No: 1461. Institute for the Study of Labor
- Jacoby HG (1992). Productivity of Men and Women and the Sexual Division of Labor in Peasant Agriculture of the Peruvian Sierra. *Journal of Development Economics* 37(265-287)
- Kajiji N, and Dash Jr. GH (2012). Efficient Multiple Objective Neural Network Mapping of State-Wide High School Achievement. *Journal of Applied Operational Research* 4(3): 139-151.
- Lavy V (2012). Expanding School Resources and Increasing Time on Task: Effects of a Policy Experiment in Israel on Student Academic Achievement and Behavior. Working Paper No: w18369. NBER
- Lin T-C (2010). Teacher Quality and Student Performance: The Case of Pennsylvania. *Applied Economic Letters* 17(2): 191-195.
- Lucas RE (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics* 22(3-42)
- MaCurdy TE, and Pencavel JH (1986). Testing between Competing Models of Wage and Employment Determination in Unionized Markets. *Journal of Political Economy* 94(3): 33-39
- Moss CB (2000). Estimation of the Cobb-Douglas with Zero Input Levels: Bootstrapping and Substitution. *Applied Economic Letters* 7(10): 677-679
- Murray N, Low B, Hollis C, Cross A, and Davis S (2007). Coordinated School Health Programs and Academic Achievement: A Systematic Review of the Literature. *Journal of School Health* 77(9): 589-600
- _____. No Child Left Behind Act, (2001)20 U.S.C. 6301 et seq
- Park J (2006). Dispersion of Human Capital and Economic Growth. *Journal of Macroeconomics* 28(3): 520-539
- Posner D (2004). What's Wrong with Teaching to the Test? *Phi Delta Kappan* 85(10): 749
- Rebello S (1991). Long-Run Policy Analysis and Long-Run Growth. *Journal of Political Economy* 99(500-521)
- Riffert F (2005). The Use and Misuse of Standardized Testing: A Whiteheadian Point of View. *Interchange: A Quarterly Review of Education* 36(1-2): 231-252
- Romer PM (1990). Endogenous Technological Change. *Journal of Political Economy* 98(S71-S102)
- Schiller KS, Khmelkov VT, and Wang X-Q (2002). Economic Development and the Effects of Family Characteristics on Mathematics Achievement. *Journal of Marriage and Family* 64(3): 730-742
- Slavin RE (2006). *Education Psychology: Theory and Practice* (8th ed.). Pearson Publishing, Boston, MA
- Smith RM, and Schumacher P (2006). Academic Attributes of College Freshmen That Lead to Success in Actuarial Studies in a Business College. *Journal of Education for Business* 81(5): 256-260
- Soloaga I. The Treatment of Non-Essential Inputs in a Cobb-Douglas Technology: An Application to Mexican Rural Household Level Data. World Bank, Development Research Trade Group
- Temple JA, and Reynolds AJ (2007). The Economics of Early Childhood Education. *Economics of Education Review* 26(1): 126-144
- Vanderhart P (2006). Why Do Some Schools Group by Ability? Some Evidence from the Naep. *American Journal of Economics and Sociology* 65(2): 435
- Wang Y, and Chao C-W (2008). Spillovers and Productivity: The Case of the Taiwanese High-Tech Firms. *Contemporary Economic Policy* 26(2): 248-258
- Zimba M, and Tiraboschi M (2010). Productivity, Investment in Human Capital and the Challenges of Youth Employment in Sub-Saharan Africa: Comparative Developments and Global Responses in the Perspective of School-to-Work Transition. SSRN.