Gaussian Process Regression with Bayesian Model Averaging (GPR-BMA): Possibilities for Social Capital

Jacob Dearmon

Oklahoma City University

jdearmon@okcu.edu

October 17, 2014

Outline

1. Previous Social Capital Research
2. Stylized Facts
3. Possible Research Extensions
4. GPR-BMA
5. Bayesian Information Criterion (BIC)
6. Applications: Simulation
7. Applications: Boston Housing Dataset
8. Applications: Fernández Ley Steel Dataset
9. Concluding Remarks: Addressing the Stylized Facts of Social Capital

Dearmon and Grier (2009)

Per-capita income panel regressions.

	Dependent variable: ln(RGDPPC)					
	(OLS)	*(2SLS)*	*(OLS)*	*(2SLS)*	*(OLS)*	*(OLS)*
ln(*RGDPPC*) (lagged)	0.903 ***	0.910 ***	0.901 ***	0.910 ***	0.904 ***	0.895 ***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)
ln(*Edu*)	0.005	0.024	−0.027	−0.006	−0.014	0.106 *
	(0.027)	(0.028)	(0.029)	(0.030)	(0.028)	(0.054)
$\ln(n+g+d)$	−0.060	−0.066	−0.040	−0.046	−0.034	−0.027
	(0.048)	(0.050)	(0.048)	(0.048)	(0.046)	(0.046)
ln(*Inv/GDP*)	0.184 ***	0.135 ***	0.182 ***	0.128 ***	0.307 ***	0.179 ***
	(0.028)	(0.034)	(0.027)	(0.033)	(0.051)	(0.027)
ln(*Trust*)			0.048 ***	0.049 ***	0.197 ***	0.123 ***
			(0.017)	(0.017)	(0.055)	(0.031)
ln(*Trust*) × ln(*Inv/GDP*)					0.086 ***	
					(0.030)	
ln(*Trust*) × ln(*Edu*)						0.067 ***
						(0.024)
td2	0.113 ***	0.114 ***	0.116 ***	0.117 ***	0.118 ***	0.112 ***
	(0.028)	(0.030)	(0.027)	(0.027)	(0.026)	(0.026)
td3	0.036	0.036	0.054 *	0.055 *	0.056 **	0.053 *
	(0.028)	(0.029)	(0.028)	(0.029)	(0.027)	(0.027)
td4	0.066 **	0.064 **	0.077 ***	0.075 ***	0.075 ***	0.079 ***
	(0.027)	(0.027)	(0.026)	(0.027)	(0.025)	(0.025)
Nobs	119	119	119	119	119	119
R^2	0.989	0.988	0.989	0.99	0.990	0.990

Notes: Standard errors are in parentheses.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

Social Capital Research- Dearmon and Grier (2009)

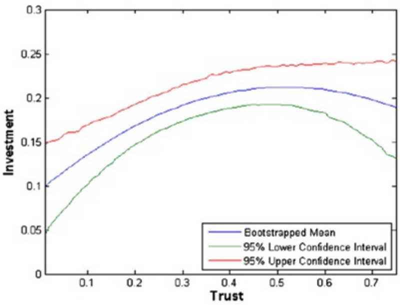
Trust and Development

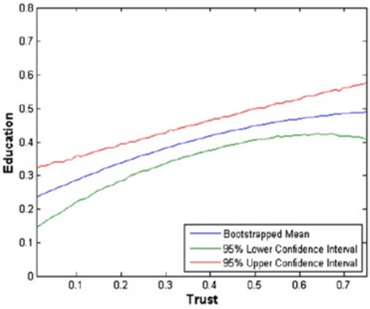
- Column 3- *Trust*
  - 1 std increase in *Trust* increases *RGDPPC* by 2.4%
- Column 5- *Trust*, Interaction with *Inv/GDP*
  - 1 s.d. increase in *Trust* increases *RGDPPC* by 2.8%
  - 1 s.d. increase in *Inv/GDP* increases *RGDPPC* by 7.4%
    - Increasing *Trust* by 1 s.d. will increase *Inv/GDP*'s impact to 8.6%
- Column 6- *Trust*, Interaction with *Edu*
  - 1 s.d. increase in *Trust* increases *RGDPPC* by 3%
  - 1 s.d. increase in *Edu* increases *RGDPPC* by 1.1%
    - Increasing *Trust* by 1 s.d. will increase *Edu*'s impact to 2%

Social Capital Research- Dearmon and Grier (2011)

Trust and the accumulation of physical and human capital

- Human and physical capital are endogenous
- Modeled jointly using 3SLS
- Trust has a nonlinear effect on outcomes





Social Capital Research

Stylized Facts

1. **Nonlinear relationship**
   - Nonlinear relationship between trust and economic outcomes
2. **Levels Matter**
   - Trust's effect depends on the level of other variables
3. **Marginal Effects**
   - Implies that trust's marginal effect may differ across variables, countries, or groups

Social Capital Research

Research Extensions

1. **Nonlinear Relationship**
   - **Current**: Nonlinear relationship specified by researcher
   - **Proposed**: Technique should identify the unknown nonlinear relationship
2. **Explanatory Variables**
   - **Current**: One set of explanatory variables is chosen
   - **Restriction**: True set of explanatory variables is unknown
   - **Proposed**: Use larger set of candidate explanatory variables, address model uncertainty
3. **Marginal Effects**
   - **Current**: Restricted by chosen nonlinear specification
   - **Proposed 1**: Marginal effects based on estimate of unknown nonlinear function that accounts for model uncertainty
   - **Proposed 2**: Marginal effects are localized and differ by observation

GPR-BMA: Based on work with Tony E. Smith (UPenn)

Gaussian Process Regression

- A nonparametric technique that identifies an unknown nonlinear function
- Produces localized marginal effects that differ by observation
- Can easily capture non-separable behavior

Bayesian Model Averaging

- Allows for a large number of candidate explanatory variables
- Provides a natural measure of statistical relevance
- Addresses model uncertainty

Gaussian Process Regression

GP Equations

1. Stochastic Process: $(y(x) : x \epsilon X)$ where $x = (x_1, .., x_k)$
2. G.P. Prior: $y(x) \sim \mathcal{G}.\mathcal{P}.[0, c(x, x')]$
3. Covariance Function: $c_\omega(x, x') = v \cdot \exp(\frac{-1}{2\tau^2} \cdot \sum_{j=1}^{k} (x_j - x_j')^2)$
   - Hyperparameters: $\omega = (v, \tau)$
4. Measurement Error: $\tilde{y}(x) = y(x) + \varepsilon, \varepsilon \sim N(0, \sigma^2)$
   - Hyperparameters: $\theta = (\omega, \sigma^2)$
5. Distribution: $\tilde{y} \sim MVN(0_n, K_\theta(\tilde{X}))$
   - where: $K_\theta(\tilde{X}) = c_\omega(\tilde{X}, \tilde{X}') + \sigma^2 I$ and ˜ denotes training sample
6. Marginal Likelihood:
   $(\tilde{y}|\tilde{X}, \theta) = (2\pi)^{\frac{-n}{2}} \cdot det[K_\theta(\tilde{X})]^{\frac{-1}{2}} \exp(\frac{-1}{2} \tilde{y}'(K_\theta(\tilde{X}))^{-1} \tilde{y})$

Bayesian Model Averaging

Model Averaging Equations

1. Model Vector: $\delta = (\delta_1, ..., \delta_k)$
   - where $\delta_k = 1$ if x_k is included, 0 otherwise
1. Prior for δ : $p(\delta) = p(\delta|q)p(q)$
   - where q is the model size
1. Prior for Model Size: $p(q) = \frac{\lambda(1-\lambda)^{q-1}}{(1-(1-\lambda)^k)}$ where $q = 1, .., k$

Gaussian Process Regression with Bayesian Model Averaging

GPR-BMA

1. Joint Posterior: $p(\delta, \theta | \tilde{y}, \tilde{X}) = p(\tilde{y} | \tilde{X}, \theta, \delta) p(\theta) p(\delta)$
2. Gibbs Sampling on Conditional Posterior
   - $p(\theta | \delta, \tilde{y}, \tilde{X})$ use Hamiltonian Monte Carlo
   - $p(\delta | \theta, \tilde{y}, \tilde{X})$ use Metropolis Hastings
3. For Metropolis Hastings step
   - Change a single element of δ using a birth-death step
   - Use the following acceptance ratio: $r = \frac{p(\tilde{y}|\tilde{X},\theta,\delta^*)}{p(\tilde{y}|\tilde{X},\theta,\delta)} \frac{p(q^*)}{p(q)} R$

Theoretical Results

Key GPR Results

1. Predictive distribution is multivariate normal
2. Conditional Mean: $E(y_l|x_l, \tilde{y}, \tilde{X}) = c_\omega(x_l, \tilde{X})(K_\theta(\tilde{X}))^{-1}\tilde{y}$
3. Conditional Variance:
   $var(y_l|x_l, \tilde{y}, \tilde{X}) = c_\omega(x_l, x_l) - c_\omega(x_l, \tilde{X}) * (K_\theta(\tilde{X}))^{-1} * c_\omega(\tilde{X}, x_l)$

Key GPR-BMA Results

1. Model Probability (N denotes num. of draws): $\hat{p}(\delta|\tilde{y}, \tilde{X}) = \frac{N(\delta)}{N}$
2. Variable Inclusion Probability for x_k: $\hat{p}(\delta_k = 1|\tilde{y}, \tilde{X}) = \frac{N_k}{N}$
3. Prediction at observation l:
   $E(y_l|x_l, \tilde{y}, \tilde{X}) = \frac{1}{N}\sum_{i=1}^{N} E(y_l|x_l, \tilde{y}, \tilde{X}, \theta_i, \delta_i)$
4. Marginal effect at observation l for variable j:
   $\frac{\partial(y_l|x_l,\tilde{y},\tilde{X})}{\partial x_{lj}} = \frac{1}{N}\sum_{i:\delta_{ij}=1} \frac{\partial c_{\omega i}}{\partial x_{lj}}(K_{\theta i}(\tilde{X}))^{-1}\tilde{y}$

Gaussian Process Regression with Bayesian Model Averaging- Extension

Bayesian Information Criterion

- Significantly reduces computational time
- Eliminates the need for HMC
- Evaluate only once per new model drawn rather than for each draw of theta
- Can be used where the number of parameters change with model size eliminating the need for a computationally expensive reversible jump process
  - Anisotropic covariance function can be be used: $c_\omega(x, x') = v \exp(\frac{-1}{2} \sum_{j=1}^{k} \frac{(x_j - x_j')^2}{\tau_j^2})$

Techniques

Methods

- **Method 1**: GPR-BMA isotropic
- **Method 2**: GPR-BIC-BMA isotropic
- **Method 3**: GPR-BIC-BMA anisotropic

Metrics

- **Metric 1**: Variable Inclusion Probability
  - App. to Social Capital: Given a set of candidate explanatory variables, what is trust's variable inclusion probability?
- **Metric 2**: Estimate of the Unknown Nonlinear Function and M.E.'s
  - App. to Social Capital: Given a certain model, what is the unknown function that maps the chosen explanatory variables to economic growth?
- **Metric 3**: Localized Marginal Effects
  - App. to Social Capital: How does trust's marginal effect vary across countries?

Simulation

Simulation-Setup

- $y = 5x_1^3 + 10x_1^2 - 5x_2^2 + x_1x_2 + u$
- $u = \rho W u + e, e \sim N(0, I)$
- $E[\frac{\partial y}{\partial x_1}] = 15x_1^2 + 20x_1 + x_2$
- $E[\frac{\partial y}{\partial x_2}] = -10x_2 + x_1$

Variable Inclusion Probabilities

Variable	*Average SEM-BMA*	*Average SAR-BMA*	*Average GPR-BMA*	*Average GPR-BIC-BMA (Isotropic)*	*Average GPR-BIC-BMA (Anisotropic)*
x_1	0.427	0.213	1	1	1
x_2	0.275	0.192	1	1	1
x_3	0.241	0.203	0	0	0
x_4	0.215	0.152	0	0	0
x_5	0.234	0.192	0	0	0
Time (secs.)	1	7	549	33	16
Number of Draws	500	500	500	500	500
Num. of Obs.	150	150	150	150	150

Prediction- RMSE

Variable	*Average GPR-BMA*	*Average GPR-BIC-BMA (Isotropic)*	*Average GPR-BIC-BMA (Anisotropic)*
y	2.094	2.566	1.375
ME x_1	3.102	4.091	2.666
ME x_2	0.873	1.214	0.675
Time (secs.)	549	33	16
Number of Draws	500	500	500
Num. of Obs.	150	150	150

Boston Housing

Background

- Harrison and Rubinfeld 1978; 506 observations- census tract level

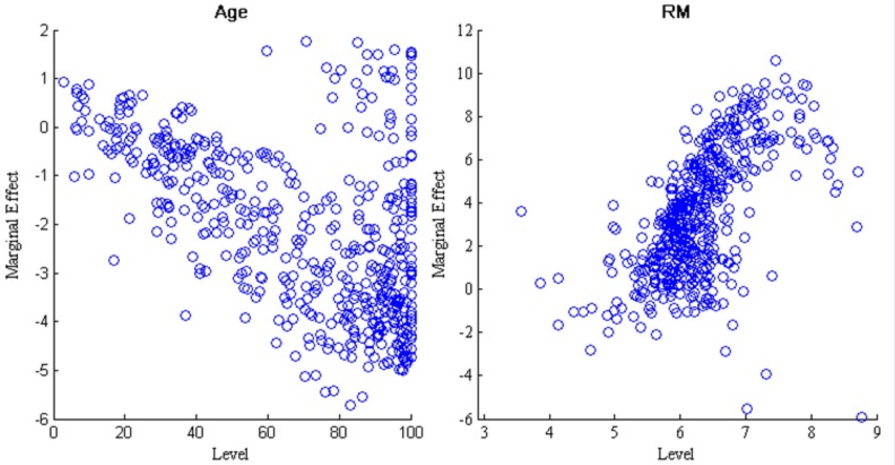
Variable Inclusion Probabilities

Variable	*GPR-BIC-BMA (Isotropic)*	*GPR-BIC-BMA (Anisotropic)*
CRIM	0.997	0.997
ZN	0.000	0.014
INDUS	0.881	0.482
CHAS	0.000	0.028
NOX	0.999	0.997
RM	1.000	1.000
AGE	1.000	1.000
DIS	1.000	1.000
RAD	1.000	0.131
TAX	1.000	1.000
PTRATIO	0.040	0.882
B	1.000	1.000
LSTAT	1.000	1.000
LAT	0.866	0.032
LONG	0.079	1.000

Average Marginal Effects

Variable	*Average GPR-BIC-BMA (Isotropic)*	*Average GPR-BIC-BMA (Anisotropic)*
CRIM	-0.69	-0.51
ZN	0.00	0.00
INDUS	0.25	-0.23
CHAS	0.00	0.00
NOX	-0.68	-0.25
RM	3.50	3.41
AGE	-2.19	-2.10
DIS	-1.35	-2.57
RAD	2.00	0.11
TAX	-2.64	-2.73
PTRATIO	-0.02	-0.56
B	-0.66	-0.09
LSTAT	-2.59	-2.37
LAT	0.61	0.02
LONG	-0.10	-1.46

Localized Marginal Effects



Fernández Ley Steel

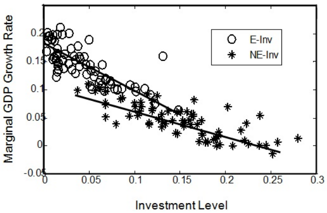
Background

- JAE 2001: 72 countries, 42 can. exp. var.; iso. GPR-BMA results

Localized Marginal Effects for Investment

Country	E_Inv	NE_Inv	Country	E_Inv	NE_Inv
Malawi	0.211	0.092	*Algeria*	0.145	0.016
Cameroon	0.202	0.089	*Brazil*	0.144	0.053
Kenya	0.201	0.060	*Chile*	0.143	0.056
Tanzania	0.199	0.082	*Panama*	0.141	0.040
Nigeria	0.199	0.076	*Mexico*	0.138	0.045
Madagascar	0.198	0.104	*Costa Rica*	0.136	0.041
Ethiopia	0.196	0.110	*Argentina*	0.136	0.066
Uganda	0.195	0.099	*Taiwan*	0.135	0.042
Zimbabwe	0.191	0.069	*Portugal*	0.133	0.038
Congo	0.190	0.054	*Uruguay*	0.132	0.055
Zaire	0.189	0.098	*Venezuela*	0.129	0.039
Ghana	0.188	0.086	*Spain*	0.128	0.033
Senegal	0.187	0.084	*Cyprus*	0.124	0.006
Zambia	0.178	0.013	*India*	0.123	0.035
Philippines	0.174	0.081	*Greece*	0.120	0.004
Pakistan	0.174	0.069	*United Kingdom*	0.119	0.038
Haiti	0.169	0.097	*Hong Kong*	0.117	0.043
Morocco	0.164	0.073	*South Korea*	0.117	0.038
Thailand	0.164	0.061	*Ireland*	0.116	0.011
Bolivia	0.161	0.063	*Italy*	0.113	0.013
Tunisia	0.160	0.061	*Denmark*	0.110	0.019
Botswana	0.160	0.045	*Australia*	0.108	0.010
Turkey	0.159	0.056	*Belgium*	0.107	0.008
Honduras	0.159	0.054	*Sweden*	0.107	0.002
Sri Lanka	0.158	0.062	*Austria*	0.105	0.039
El Salvador	0.155	0.077	*Germany*	0.102	0.000
Malaysia	0.155	0.025	*Israel*	0.102	0.019
Paraguay	0.154	0.078	*Canada*	0.102	0.022
Peru	0.154	0.070	*Switzerland*	0.101	0.001
Jordan	0.153	0.045	*Netherlands*	0.101	0.007
Guatemala	0.153	0.052	*France*	0.100	0.007
Dominican Republic	0.153	0.061	*United States*	0.098	0.023
Nicaragua	0.152	0.047	*Norway*	0.097	0.000
Colombia	0.148	0.056	*Finland*	0.091	-0.015
Ecuador	0.146	0.028	*Japan*	0.075	0.001
Jamaica	0.146	0.046	*Singapore*	0.064	0.023

Localized Marginal Effects for Investment



Concluding Remarks

Future Research

1. Identify a large set of candidate explanatory variables for development and social capital
2. Allow for an unknown nonlinear relationship and model uncertainty
3. Identify how the marginal effect of trust varies across countries
4. Draw targeted policy conclusions based on marginal effect differences