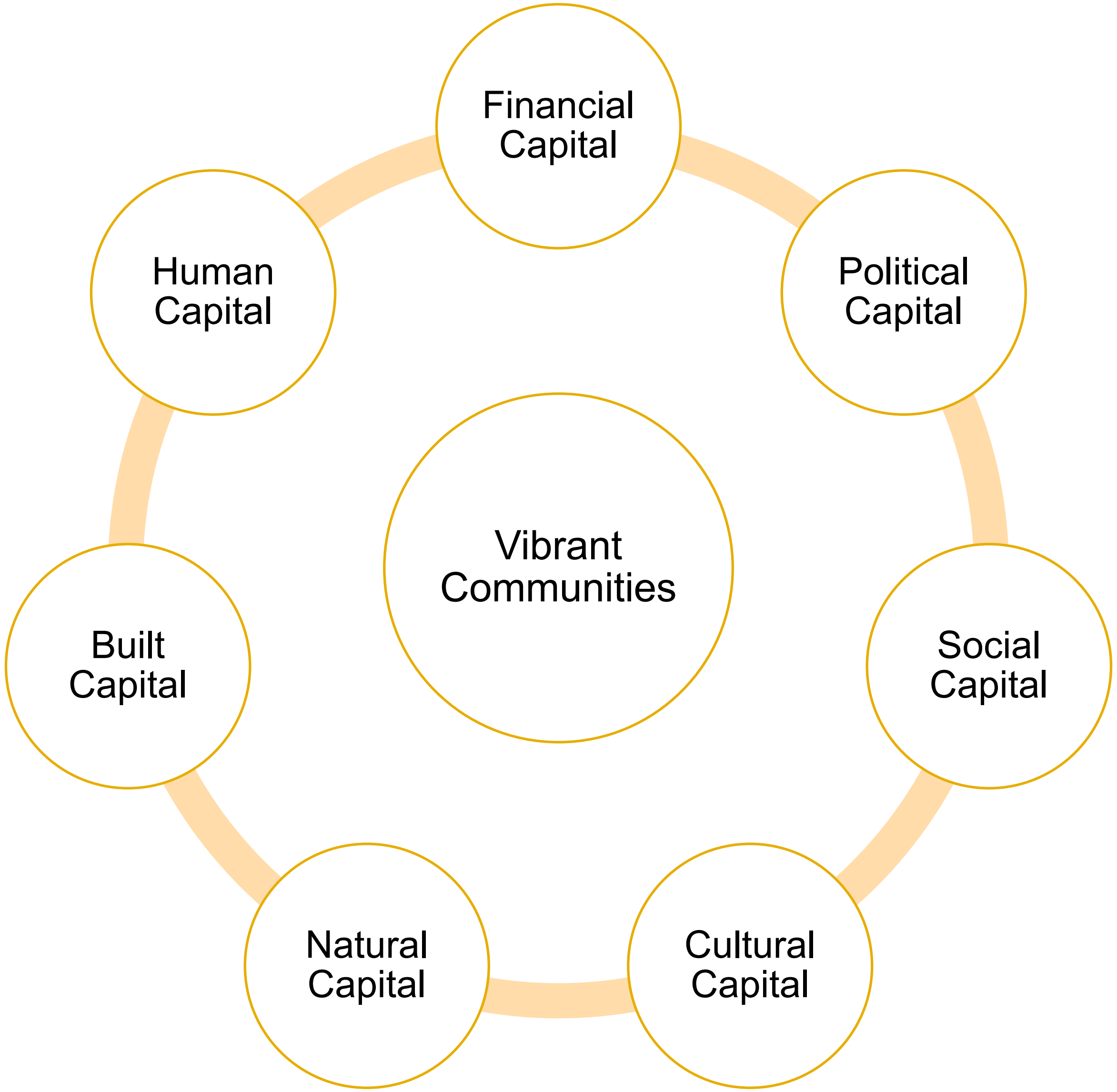
Community Systems Thinking and Modeling Uncertainty

2021 NARSC Presidential Address

Steven Deller
Department of Agricultural and Applied Economics
Division of Extension
University of Wisconsin-Madison

- Notions of the Engage Scholar
- In the spirit of Walter Isard, Community Economic Development is a truly interdisciplinary field of study and work.
- A Systems Thinking Approach
- “Everything Matters” and “Everything is Endogenous”
- Modeling Uncertainty

A Systems Thinking Approach to Community Economic Development

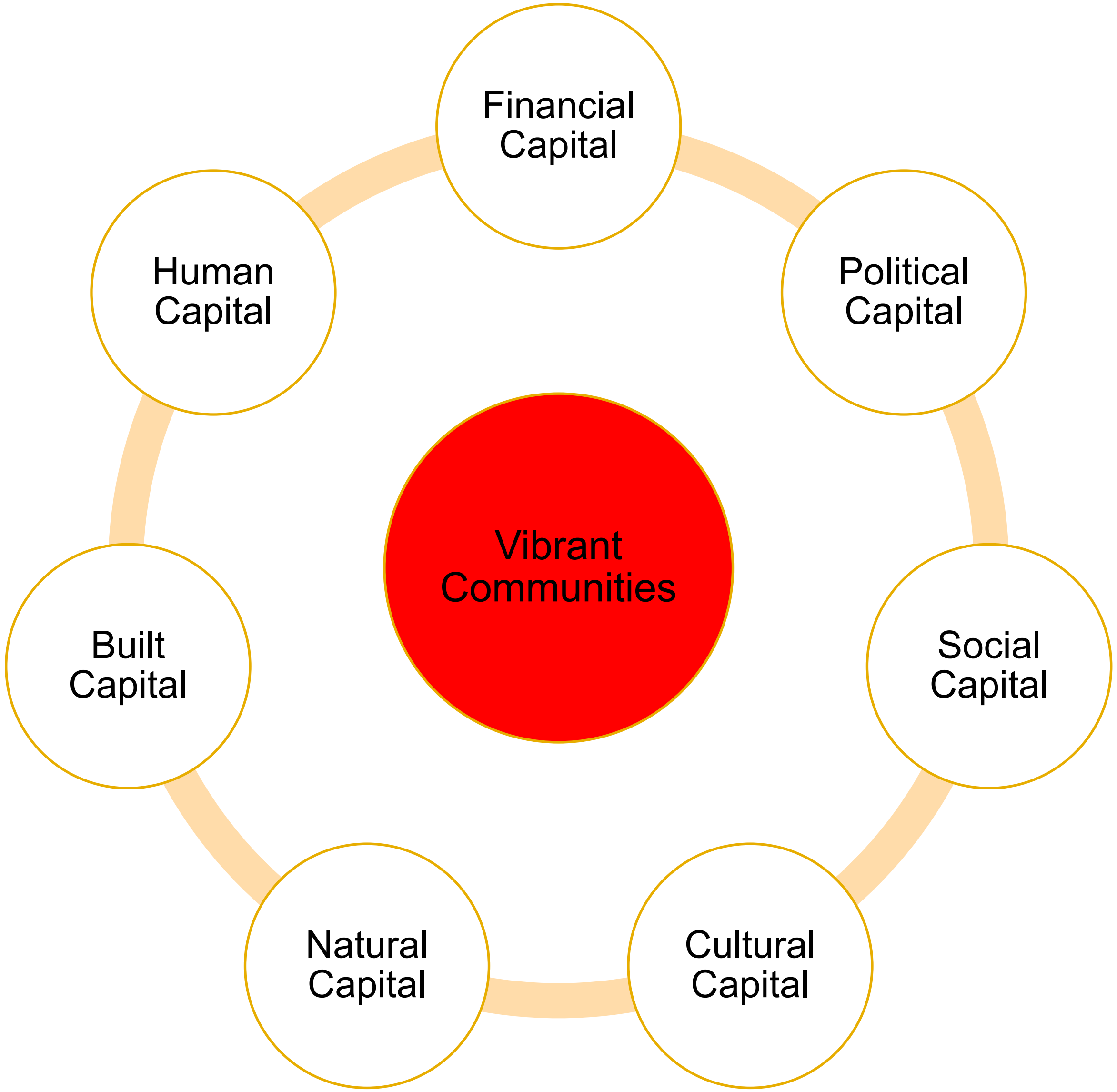


Cornelia Flora
Iowa State University

What are the assets available to the community?

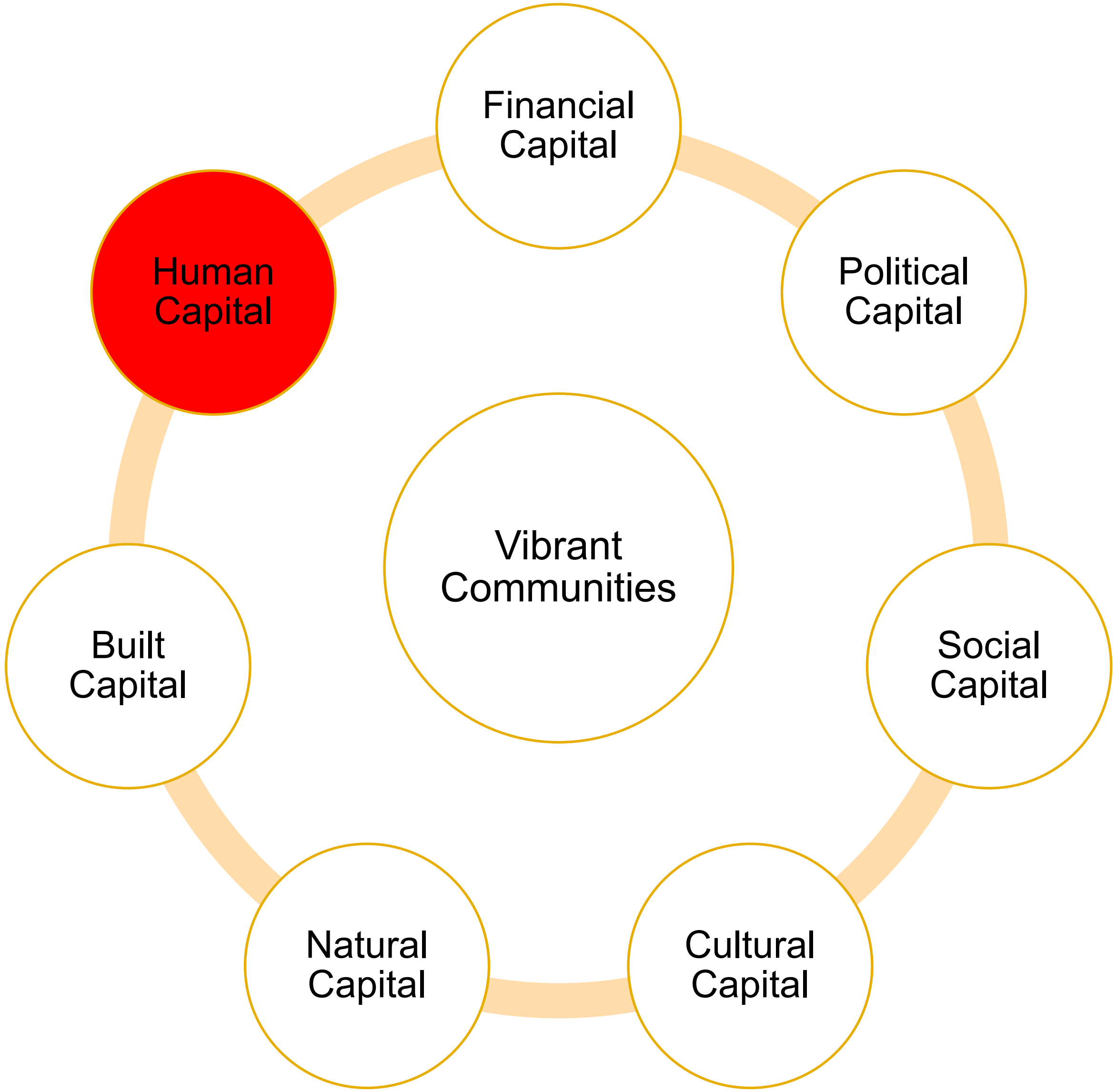
What assets are strength that the community can build upon?

What assets are deficient that need investments?



Vibrant Communities:

- Resilient
- Entrepreneurial
- Innovative
- Forward not backward looking

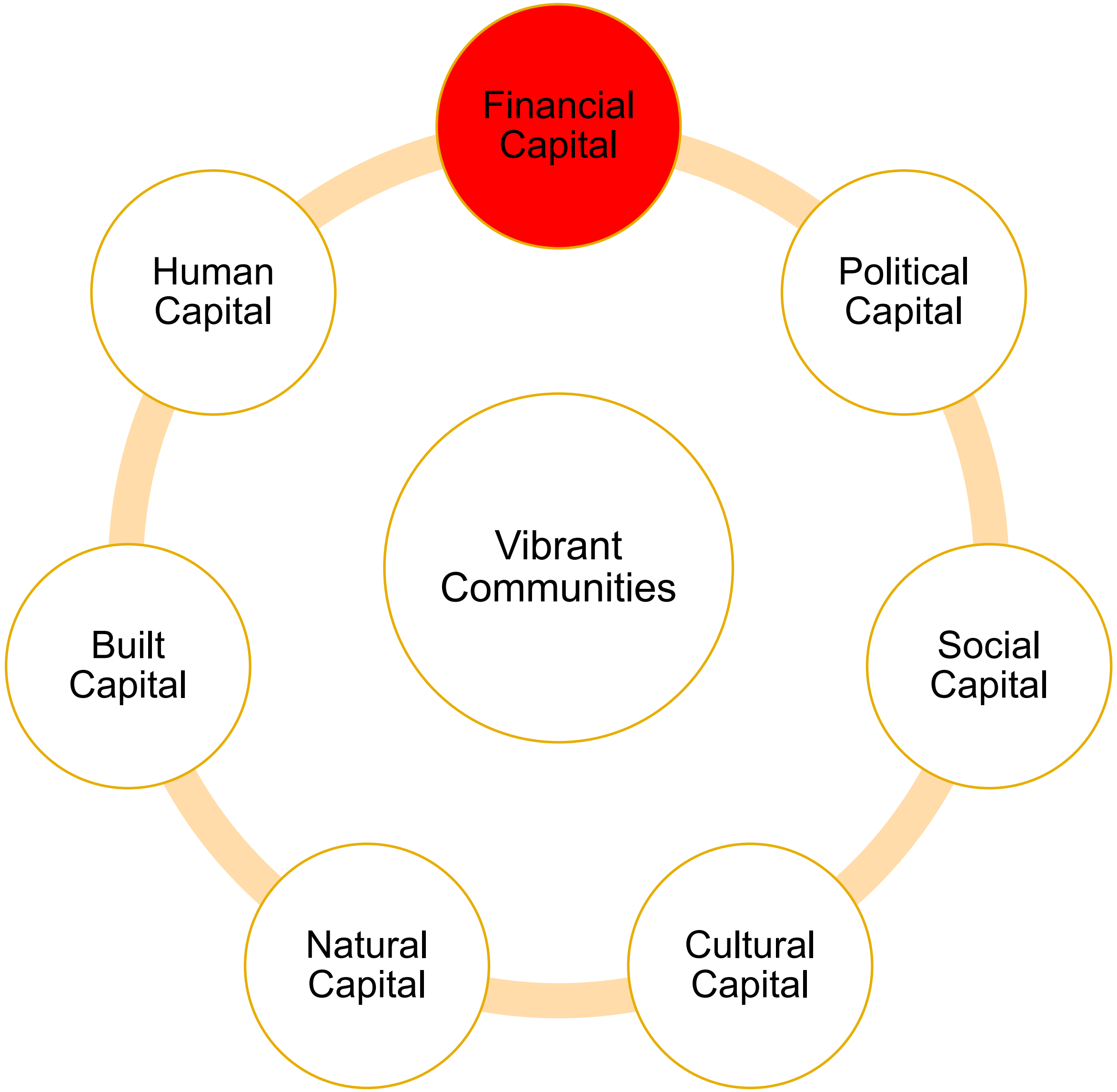


Human capital: The skills and abilities of people, education, problem solving abilities, critical thinking.

How do we measure it?

Educational attainment?

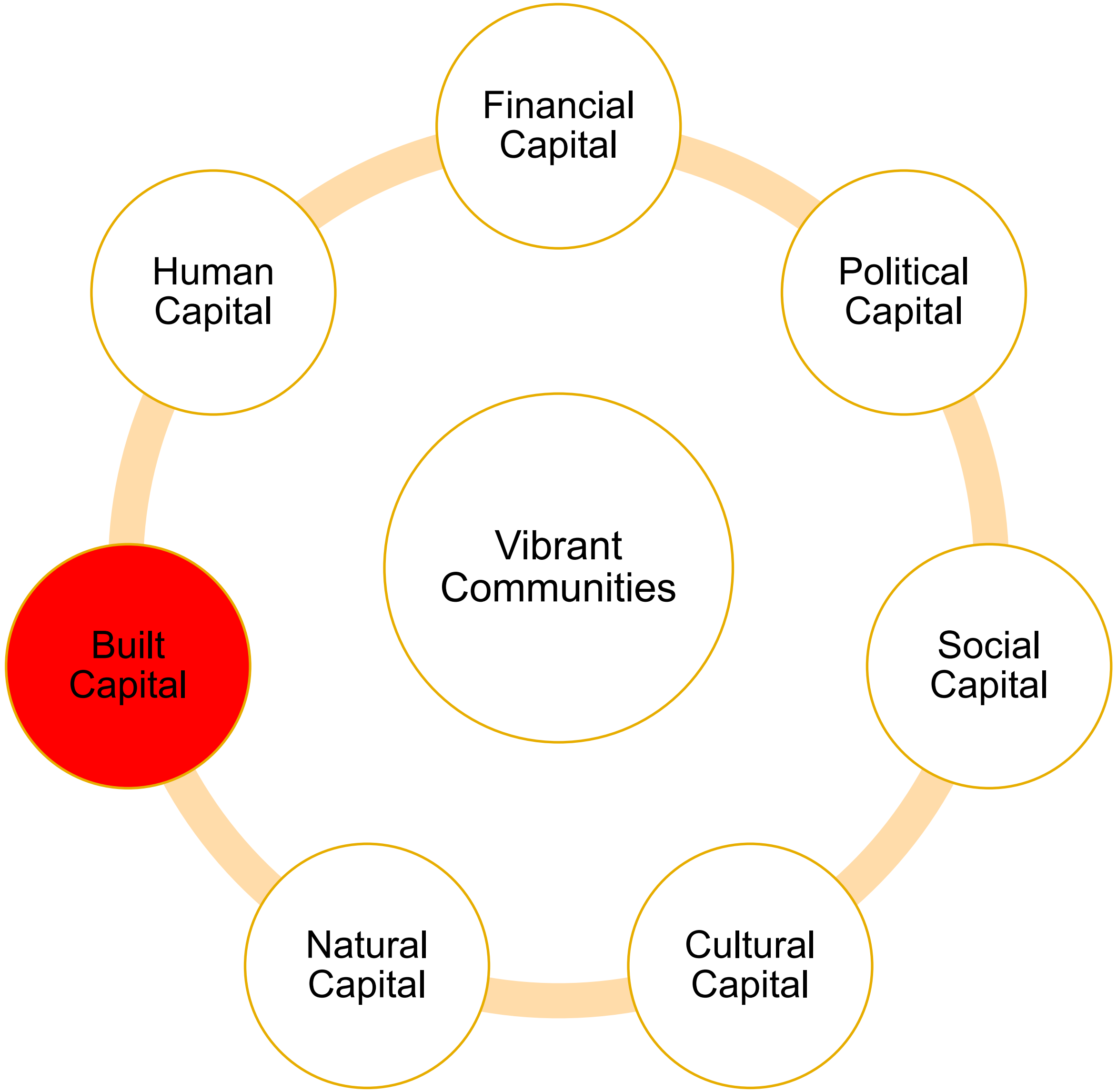
Which measure?



Financial capital: The financial resources available to invest in community capacity building, to underwrite businesses development, to support civic and social entrepreneurship, and to accumulate wealth for future community development.

How do we measure it?

Number of banks?

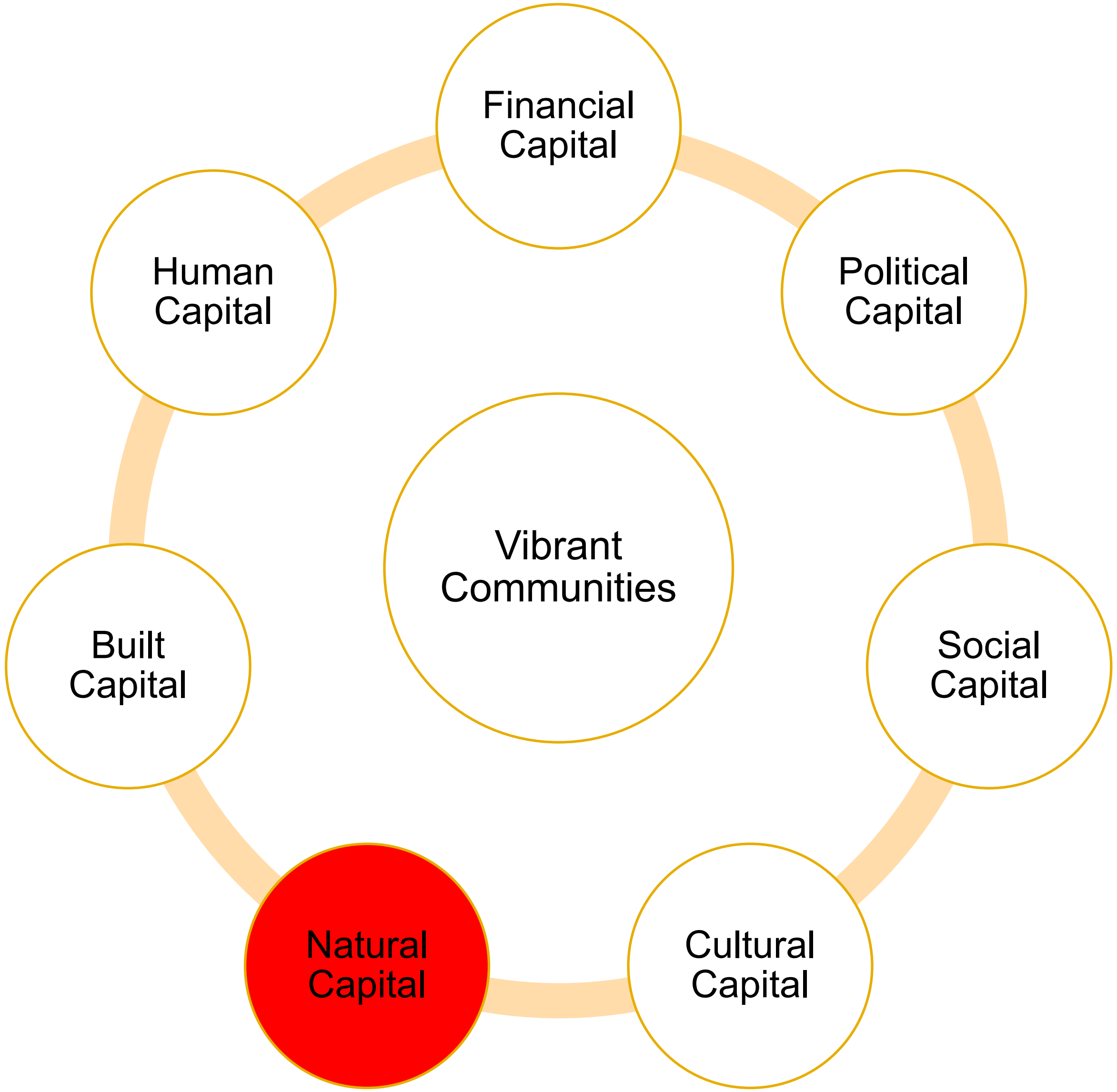


Built capital: The infrastructure that supports the community, such as telecommunications (e.g., broadband), industrial parks, mainstreets, water and sewer systems, roads, etc.

Built capital is often a focus of community development efforts. Why? Tangible, easy to “see” the investments.

How do we measure it?

Broadband, roads, sewer, schools quality-capacity?



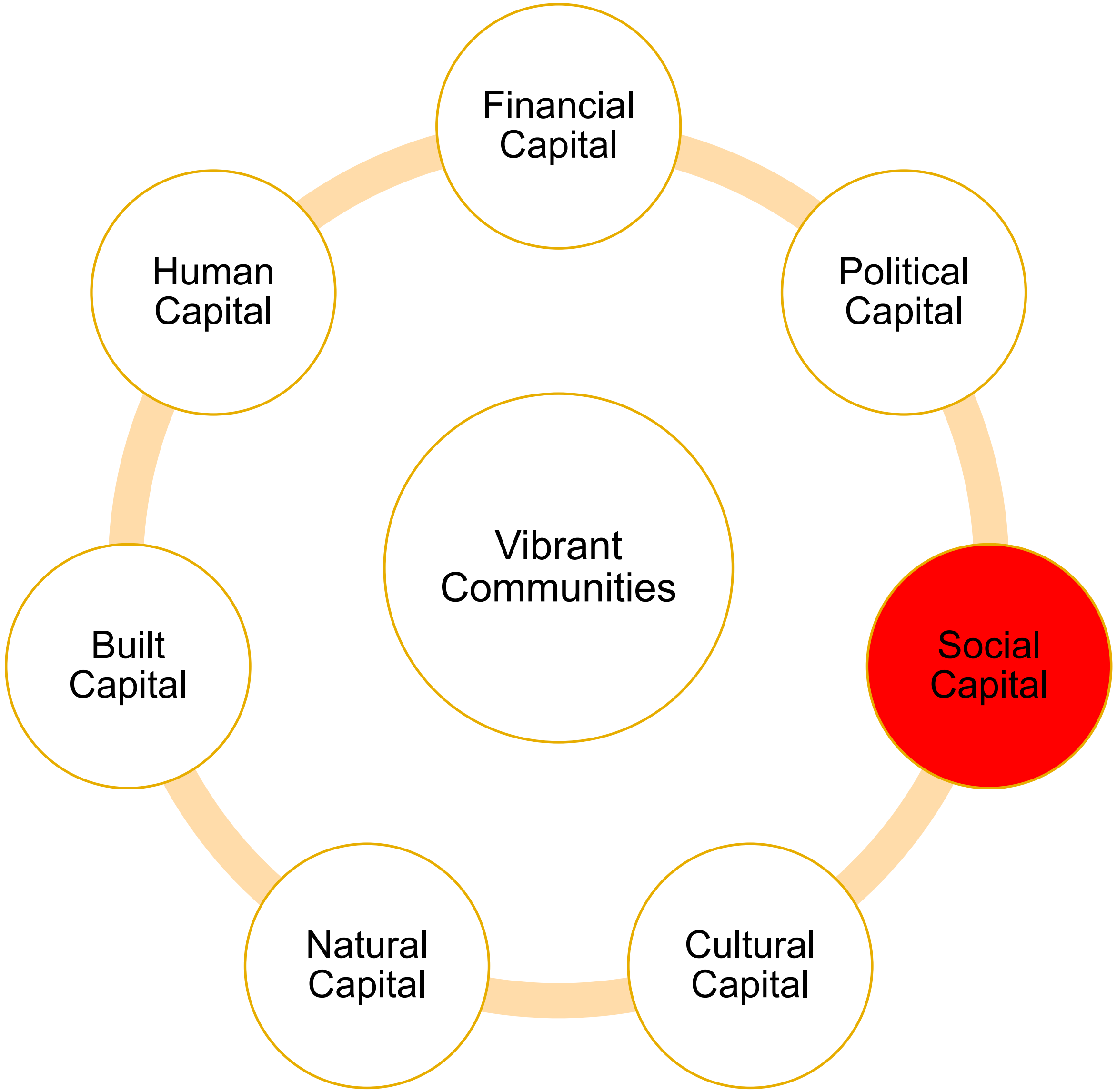
Natural capital:

Those assets that abide in a location, including resources (land), amenities and natural beauty.

Extractive vs Non-Extractive uses of natural capital.

How do we measure it?

Climate, natural resources>?

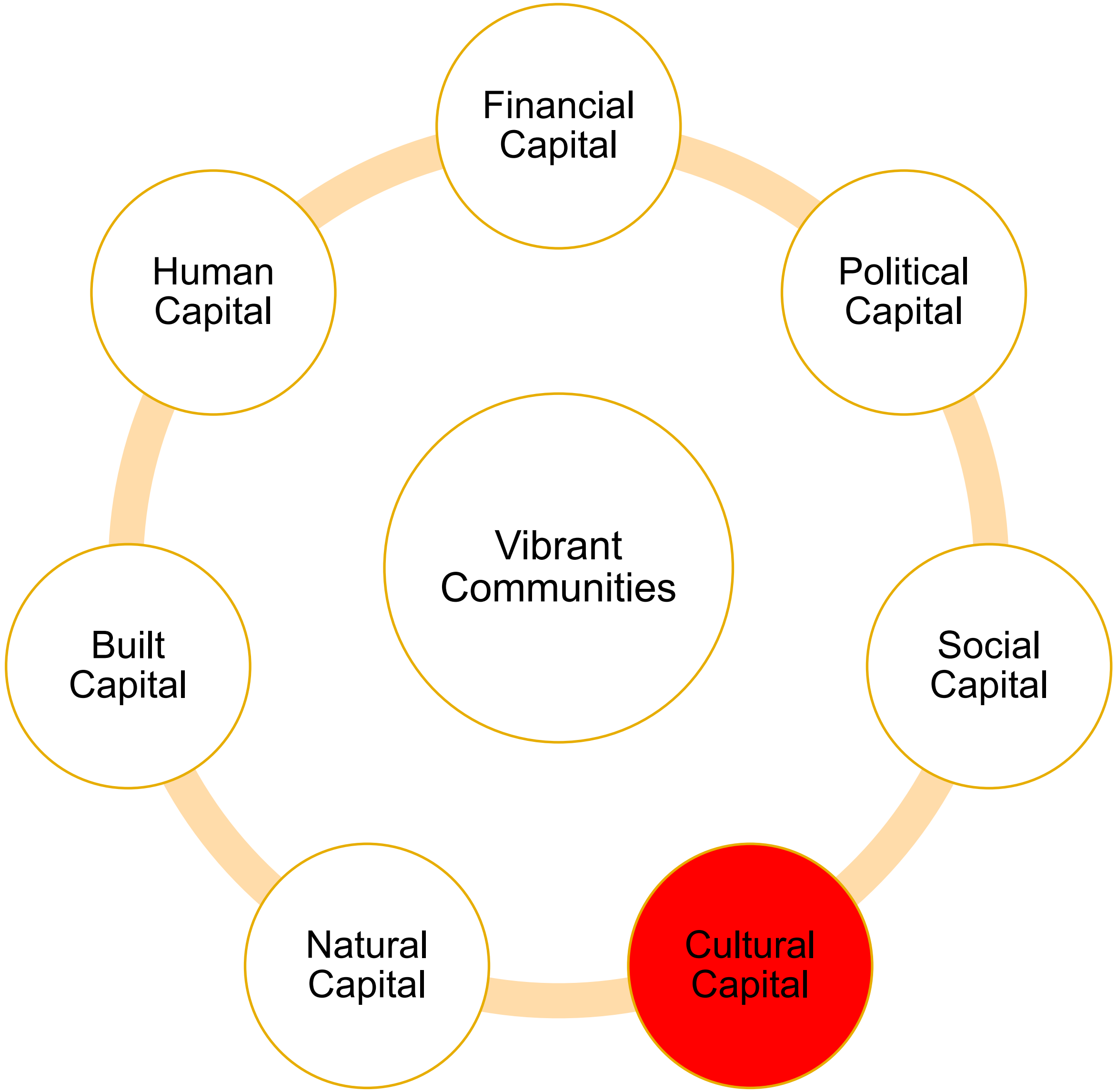


Social capital: Reflects the connections among people and organizations, networks that enable the flow of information.

Bonding and Bridging Social Capital

How do we measure it?

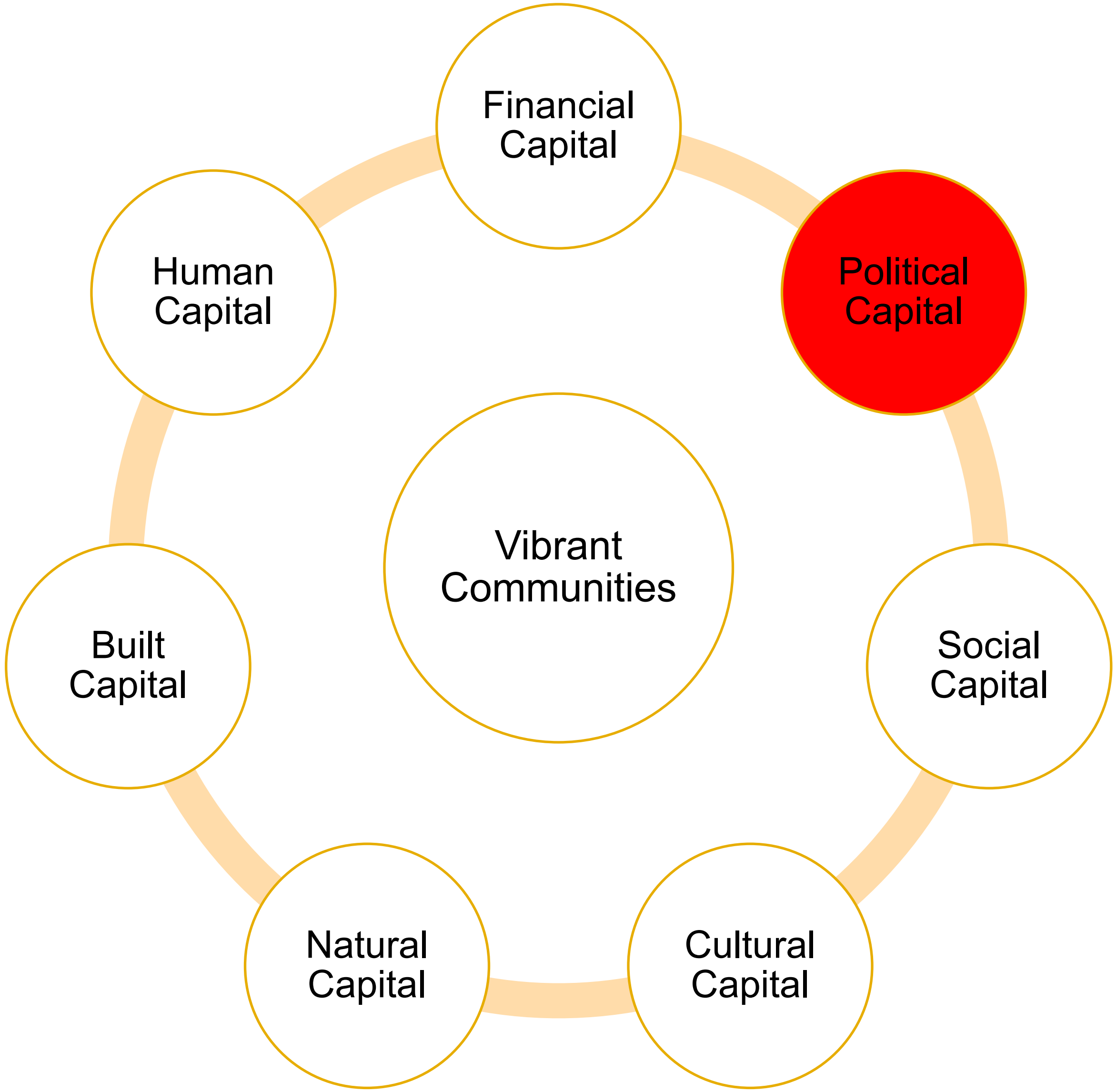
Rupasingha, Goetz & Freshwater Social Capital Index?



Cultural Capital: Reflects the way people "know the world" and how to act within it. The dynamics of who we know and feel comfortable with, what heritages are valued. It influences what voices are heard and listened to and speaks to norms of acceptable behavior.

How do we measure it?

Theaters, museums, arts venues, churches?



Political capital: The ability to influence standards, rules, regulations and their enforcement. Linkages (bridging social capital?) to other units of government and institutions.

How do we measure it?

Political heterogeneity or homogeneity, voting turn out, political organizations?

“Everything Matters” and “Everything is Endogenous”

- Everything matters, everything is interconnected, hence everything is endogenous.
- The current “fad” of looking for the “right” instrumental variable is akin to jumping down the rabbit hole.
- If everything matters and we have multiple ways of measuring each factor, how do we proceed?
- Within the context of “modeling uncertainty”.

One approach is to use principal components, or some variation, to combine several individual variables into a scalar measure of the relevant "capitals".

$$CO = f(\Sigma CC_j,\ PV)$$

$CO \equiv$ Some Community Outcome (e.g., growth, stability, resiliency, etc.)

$CC_j \equiv$ Measures of *j* community capitals

$PV \equiv$ Policy variables of interest.

Land Economics • November 1998 • 74 (4): 541–56

Measuring the Effects of Economic Diversity on Growth and Stability

John E. Wagner and Steven C. Deller

ABSTRACT. *The role of economic diversity in regional stability and growth is examined. Contrary to "conventional wisdom" the empirical lit-*

an internal theoretical inconsistency of jointly pursuing economic growth and stability through the one policy approach of diver-

TABLE 4
EMPIRICAL RESULTS FOR THE GROWTH AND STABILITY MODELS

Variable	Models[a] Growth	Stability
Market	0.000465 (0.27)[b]	0.010115 (0.37)
Labor	−0.006027 (3.20)	0.062298 (2.22)
Tax	0.001034 (0.67)	0.006637 (0.27)
Amenity	0.003097 (1.18)	0.058564 (1.43)
Infrastructure	0.005126 (2.44)	0.093217 (3.03)
Diversification Index	0.018444 (2.15)	−0.157680 (1.70)
Constant	−1.25360 (154.72)	−2.12730 (19.87)
Box-Cox Lambda	0.6300	0.4100
Adj R^2	0.1694	0.2560
F-stat	2.5980	3.696

[a] Measure of growth is the average annual growth rate in state per capita income and the measure for stability is the variance of the average annual unemployment rate for the period 1969–91.

[b] Number in parentheses is the absolute value of the *t*-statistic.

TABLE 3
PRINCIPAL COMPONENT EIGENVECTORS

Variable Block	Eigenvector
Markets	
Income distribution (1980)	0.373896
Percent of the population that is nonwhite (1980)	−.210187
Population (1980)	0.037399
Growth in population (1969–91)	0.233518
Per capita income (1980)	0.052413
Cost of living (1981)	0.009444
Percent of individuals below the poverty level (1979)	0.001064
Percent of children below the poverty level (1979)	−0.041890
Percent of persons over 65 years of age (1979)	−0.406701
Percent of persons living in the region their entire life (1976)	0.768177
Cumulative Variance Explained	*90.85%*
Labor	
Right to work law (1982)	−0.131884
Percent of labor force unionized (1980)	0.454602
Percent of persons with a high school diploma (1980)	−0.411921
Percent of persons with a college diploma (1980)	−0.250696
Average teacher (K–12) salary (1980)	0.456971
Number of doctors per 1,000 persons (1980)	0.216706
Number of prisoners per 1,000 persons (1980)	0.381578
Infant death rate (1980)	0.376870
Cumulative Variance Explained	*77.85%*
Taxes	
Corporate tax rate (1981)	−0.600410
State sales tax rate (1981)	0.659553
Composite effective tax rate (1981)	−0.126624
Gasoline excise tax (1980)	0.434125
Cumulative Variance Explained	*67.45%*
Amenities	
Percent of the regional population classified as rural (1980)	0.500109
Percent of the region's surface area covered by lakes and rivers (1978)	−0.275104
Percent of the population with a fishing license (1978)	0.564072
Percent of the population with a hunting license (1978)	0.596684
Cumulative Variance Explained	*64.03%*
Infrastructure	
Number of public airports per one million persons (1980)	0.582335
Number of private airports per one million persons (1980)	0.517265
Highway density measured as miles of four-lane highway per square mile (1980)	0.627155
Cumulative Variance Explained	*72.86%*

Amer. J. Agr. Econ. 83(2) (May 2001): 352–365

The Role of Amenities and Quality of Life In Rural Economic Growth

Steven

A s
Fiv
usi
int

	Δ Population	Δ Employment	Δ Per Capita Income
Intercept	52.174	74.102	152.007
	(8.511)	(5.730)	(18.949)
Population in 1985	0.00001	0.000.5	−0.0001
	(0.245)	(4.134)	(−1.468)
Employment in 1985	−.0003	−0.001	0.0006
	(−2.714)	(−3.967)	(3.784)
Per capita income in 1985	−0.0004	−0.001	−0.005
	(−2.425)	(−3.833)	(−24.466)
Percent of nonwhite population	−0.049	−0.080	0.192
	(−2.043)	(−1.593)	(6.118)
Percent of population under seventeen	−0.508	−0.965	−0.170
	(−4.740)	(−4.266)	(−1.214)
Percent of population above sixty-five	−0.845	−1.607	−0.449
	(−8.963)	(−8.081)	(−3.637)
Entropy income distribution index	−0.008	0.0004	0.003
	(−4.886)	(0.113)	(1.228)
Household with income under poverty	0.103	−0.154	−0.957
	(0.915)	(−0.648)	(−6.4990)
Unemployment rate	0.182	−0.490	−0.550
	(2.021)	(−2.581)	(−4.671)
Percent high school graduate	0.047	0.185	−0.218
	(1.024)	(1.908)	(−3.629)
Crime rate	0.0005	0.001	−0.0002
	(3.018)	(3.062)	(−0.886)
Number of physicians	−0.004	0.012	0.042
	(−0.970)	(1.393)	(7.572)
Property tax	−0.040	−0.049	−0.061
	(−2.048)	(−1.196)	(−2.391)
Government expenditure	0.00008	−0.00004	−0.0001
	(3.822)	(−0.910)	(−1.995)
Climate	1.763	0.517	0.478
	(6.824)	(0.948)	(1.415)
Developed recreational infrastructure	0.541	1.308	1.018
	(2.772)	(3.174)	(3.984)
Land	0.854	1.491	−0.136
	(3.407)	(2.820)	(−0.414)
Water	0.432	0.046	1.154
	(1.951)	(0.099)	(3.984)
Winter	1.148	1.560	1.039
	(4.003)	(2.578)	(2.768)
N=	2243	2243	2243
F statistic =	48.491	22.817	67.781
Adjusted R^2 =	0.287	0.156	0.3614

Table 4. Principal Component Eigenvectors: Water

Water Variables	Eigenvector
# Marinas	0.4219
# Canoe outfitters, rental firms and raft trip firms	0.3269
# Diving instruction or tours and snorkel outfitters	0.1908
# Guides services	0.4776
# Fish camps, private or public fish lakes, piers and ponds	0.5482
# American Whitewater Association total white water river miles	0.1184
Designated Wild & Scenic River miles: Total 1993	0.1367
National Resources Inventory (NRI) acres in water bodies 2–40 acres, < 2 acres, and >= 40 acres (lake or reservoir)	0.1597
NRI acres in streams < 66′ wide, 66–660′ wide, and >= 1/8 miles wide	−.0364
NRI water body >= 40 acres (bay, gulf, or estuary)	0.2665
NRI wetland acres	0.0654
NRI total river miles, outstanding value	0.1235
Cumulative variance explained	16.84%

Table 5. Principal Component Eigenvectors: Winter

Winter Variables	Eigenvector
Cross-country Ski Areas Association # Xcski firms, and public XCski centers	0.3496
International Ski Service Skiable acreage	0.3206
Federal land acres in counties with > 24″ annual snowfall	0.5233
Agricultural acres in counties with > 24″ annual snowfall	0.1381
Acres of mountains in counties > 24″ annual snowfall	0.5864
Acres of forestland in counties > 24″ annual snowfall	0.3717
Cumulative variance explained	35.93%

Amer J Agr Econ. 2021;1–27.

Rural broadband speeds and business startup

Steven I

[1]Department
Economics, C
Economic De
Wisconsin-M

[2]Department
Oklahoma Sta
Oklahoma, U

TABLE 4 Base model, business startup rate 2014–2015 (spatial error estimator)

Bayesian posterior estimates	All businesses SEM	Construction SEM tobit	Prof, sci and tech SEM tobit	Health and soc serv SEM tobit	Accommodations and food services SEM tobit
Intercept	13.2017***	1.6183***	0.7791***	1.1797***	1.0367***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Lagged Economic Growth Index	1.4587***	0.0180	0.0574*	−0.0183	0.0295
	(0.0001)	(0.3406)	(0.0599)	(0.2973)	(0.2301)
Economic Structure Index	0.0155	−0.3499***	−0.3896***	−0.4739***	−0.4491***
	(0.4650)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Demographic Index	0.1137	−0.0801**	−0.0673**	−0.1020**	−0.0425
	(0.2529)	(0.0481)	(0.0458)	(0.0068)	(0.1715)
Social Capital Index	3.1501***	0.6202***	0.3306***	0.3273***	0.3840***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Asset Index	1.7044***	−0.1283**	−0.2213***	−0.1847***	−0.1114**
	(0.0001)	(0.0003)	(0.0001)	(0.0001)	(0.0009)
Spatial competition of neighbors Inclusive of urban counties	0.2194**	−0.0082	−0.0028	−0.0111*	0.0108*
	(0.0002)	(0.1724)	(0.3538)	(0.0580)	(0.0937)
Spatial lambda (λ)	0.1879**	0.0619	0.3163***	0.1337**	0.0563
	(0.0048)	(0.1476)	(0.0001)	(0.0157)	(0.1694)

Note: Marginal significance (*p*-values) in parentheses.
***: Significant at 99.9%.
**: Significant at 95.0%.
*: Significant at 90.0%.

TABLE 3 Socioeconomic and demographic control factors

	Eigenvector weight
Lagged Growth Index	
Percent change in population 2000–2014	0.5621
Percent change in per capita income 2000–2014	0.3824
Percent change in employment 2000–2014	0.7334
Variance explained	0.5518
Economic Structure Index	
Share of employment in proprietorships	0.7032
Population to employment ratio	0.1214
Economic Diversity Index	0.7005
Variance explained	0.4582
Demographic Index	
Age Index	0.7243
Education Index	−0.6541
Racial Diversity Index	0.1435
Population density	0.1639
Variance explained	0.3383
Asset Index	
Median house value ($000)	0.6943
Student debt interest payment per return with debt	0.1464
Banking concentration (per 10K population)	0.0748
Number of small business bank loans (<$100,000) per capita (1K population)	0.7006
Variance explained	0.4066

While the use of principal components is one approach, are there other approaches that tackle the issue of “everything matters” more directly?

$$CO = f(\Sigma CC_j, PV)$$

We “know” from theory and prior empirical work that CC_j can be vast and complicated, but we are not necessarily interested in how CC_j affects CO we are interested in our policy variables (PV). We only need to control for CC_j.

So, what is the “best model” for controlling for CC_j?

Journal of Economic Literature 2020, 58(3), 644–719
https://doi.org/10.1257/jel.20191385

Model Averaging and Its Use in Economics†

Mark F. J. Steel*

The method of model averaging has become an important tool to deal with model uncertainty, for example in situations where a large amount of different theories exist, as are common in economics. Model averaging is a natural and formal response to model uncertainty in a Bayesian framework, and most of the paper deals with Bayesian model averaging. The important role of the prior assumptions in these Bayesian procedures is highlighted. In addition, frequentist model averaging methods are also discussed. Numerical techniques to implement these methods are explained, and I point the reader to some freely available computational resources. The main focus is on

Steel (2020, page 644)

"The discussion focuses mostly on uncertainty about covariate inclusion in regression models (normal linear regression and its extensions), which is arguably the most pervasive situation in economics."

Imposition of some information criteria in order to select a single "best" model regarded as the true model from which variable parameters are estimated.

Previous research uses determination criteria, such as changes in the equation F statistic, $\bar{R}^2$ or Mallows' C_p statistic, which are tracked across alternative linear regressions for the purpose of identifying a "best" model.

Other potential criteria include the Amemiya criteria (PC), Akaike Information Criteria (AIC), Sawa Bayesian Information Criterion and/or the Schwarz Bayesian Information Criterion (BIC) as well as the Jeffreys-Bayes posterior odds ratio.

Steel (2020) identifies three groupings or classifications around modeling uncertainty within economics:

- Prediction,
- Identifying the factors or determinants driving economic processes, and
- Policy evaluation, where the focus is on assessing the consequences of certain policies.

$$CO = f(\Sigma CC_j,\ PV)$$

Brock, Durlauf, and West (2003) identify <u>three main types of uncertainty</u> that typically need to be considered:

Theory uncertainty. This reflects the situation where economists disagree over fundamental aspects of the economy....

Specification uncertainty. This type of uncertainty is about how the various theories that are considered will be implemented, in terms of how they are translated into specific models.

Heterogeneity uncertainty. This relates to model assumptions regarding different observations. Is the same model appropriate for all, or should the models include differences that are designed to accommodate observational heterogeneity? (GWR anybody?)

Steel (2020, p650) “In line with probability theory, the formal Bayesian response to dealing with uncertainty is to average. When dealing with parameter uncertainty, this involves averaging over parameter values with the posterior distribution of that parameter in order to get the predictive distribution.”

$$\hat{\beta}_z = \sum_{j=1}^{M} \omega_{zj}\beta_{zj}$$

$$\omega_{zi} = L_{zj} \Big/ \sum_{i=1}^{M} L_{zi}$$

L_{zj} is the log likelihood of specific model.

The American Economic Review
Vol. 87, No. 2, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association (May, 1997), pp. 178-183 (6 pages)

I Just Ran Two Million Regressions

By Xavier X. Sala-i-Martin*

Following the seminal work of Robert Barro (1991), the recent empirical literature on economic growth has identified a substantial number of variables that are partially correlated with the rate of economic growth. The basic

An initial answer to this question was given by Ross Levine and David Renelt (1992).[1] They applied Edward Leamer's (1985) *extreme-bounds test* to identify ''robust'' empirical relations in the economic growth

	β_1	β_2	β_3
M_1	1	1	1
M_2	0	1	1
M_3	0	0	1
M_4	1	1	0
M_5	1	0	0
M_6	0	1	0
M_7	1	0	1
M_8	0	0	0

$$\hat{\beta}_z = \sum_{j=1}^{M} \omega_{zj}\beta_{zj}$$

$$\omega_{zi} = {L_{zj}}\Big/{\sum_{i=1}^{M} L_{zi}}$$

The full model space M (possible combinations) is 2^K, if, for example if K=10, then the full model space has a dimension of 1,024.

$$\hat{\beta}_z = \sum_{j=1}^{M} \omega_{zj}\beta_{zj}$$

$$\omega_{zi} = {L_{zj}}\Big/{\sum_{i=1}^{M} L_{zi}}$$

This approach within the model averaging literature could be linked to the “frequentist model averaging (FMA)” literature. Here many alternative weighting schemes are offered: Mallows’ C_p statistic, Amemiya criteria (PC), Akaike Information Criteria (AIC), Sawa Bayesian Information Criterion and/or the Schwarz Bayesian Information Criterion (BIC) as well as the Jeffreys-Bayes posterior odds ratio.

While the “frequentist model averaging (FMA)” is slowing gaining some traction in economics because no priors on the distribution is required and the corresponding estimators are totally determined by data.

The weakness is that there is no theoretical justification for the particular weighting scheme.

While the FMA approach is gaining some traction, there has been an enormous literature on the use of the Bayesian Model Averaging (BMA) approach where the uncertainty on model is considered by setting a prior probability to each candidate model.

Suppose that there is a set of models all of which may be "reasonable" based on the theory for estimating θ from a given data set *y*. Suppose further that a particular parameter θ has a common interpretation across all possible models $M_1, \ldots, M_k$.

Instead of using one single model for making inferences about β, Bayesian Model Averaging constructs $\pi(\beta|y)$, the posterior density of β given the data and is not conditional on any specific model (M_i).

Given the Bayes formula, BMA starts by specifying

- prior probabilities $P(M_j)$ for all models $M_1, \ldots, M_k$ under consideration,
- prior densities $\pi(\beta_j|M_j)$ for all parameters β_j of the model M_j.

Given the prior information on the parameters for a given model, the integrated likelihood $(L_{n,j})$ of model M_j is given by

$$\gamma_{n,j}(y) = \int L_{n,j}(y, \beta_j)\pi(\beta_j\,|M_j)d\beta_j.$$

Here $\gamma_{n,j}(y)$ is also the marginal density of the observed data. Using the Bayes theorem, the posterior density of the model is obtained as

$$P(M_j|y) = \frac{P(M_j)\gamma_{n,j}(y)}{\sum_{j'=1}^{k} P(M_{j'})\gamma_{n,j'}(y)}$$

Notice the overlap here between the FMA and BMA....

$$\omega_{zi} = {L_{zj}}\Big/{\sum_{i=1}^{M} L_{zi}}$$

The posterior inference is based on the models visited by the Markov chain instead of on the complete model space which is untraceable given a large K.

For example, Heather Stephens and I are looking at the drivers of labor force participation rates across four age generations. We look at 43 different variables. Based on our use of BMA the full model space $\mathcal{M}$ is 2^K, specifically K=43 the full model space has a dimension of 8,796,093,022,208. Given that we explore 4 different generational age cohorts of people, and each has a model space of almost 8.8 trillion possible combinations.....

Sala-I-Martin's "I just ran one million regressions" I laugh at your trivial modeling space!

The posterior inference is based on the models visited by the Markov chain instead of on the complete model space which is untraceable given a large *K*.

We can more formally define a neighborhood $nbd(M)$ for each $M \in \mathcal{M}$ (the set of all possible models). From there we can define a transition matrix *q* by setting $q(M \to M') = 0 \; \forall \, M' \notin nbd(M)$ and $q(M \to M') \; \neq 0 \; \forall \, M' \in nbd(M)$. If the chain is currently in state *M*, we can proceed by drawing *M'* from $q(M \to M')$.

M' is accepted with probability

$$\min\{1, \frac{P(M'|y)}{P(M|y)}\}$$

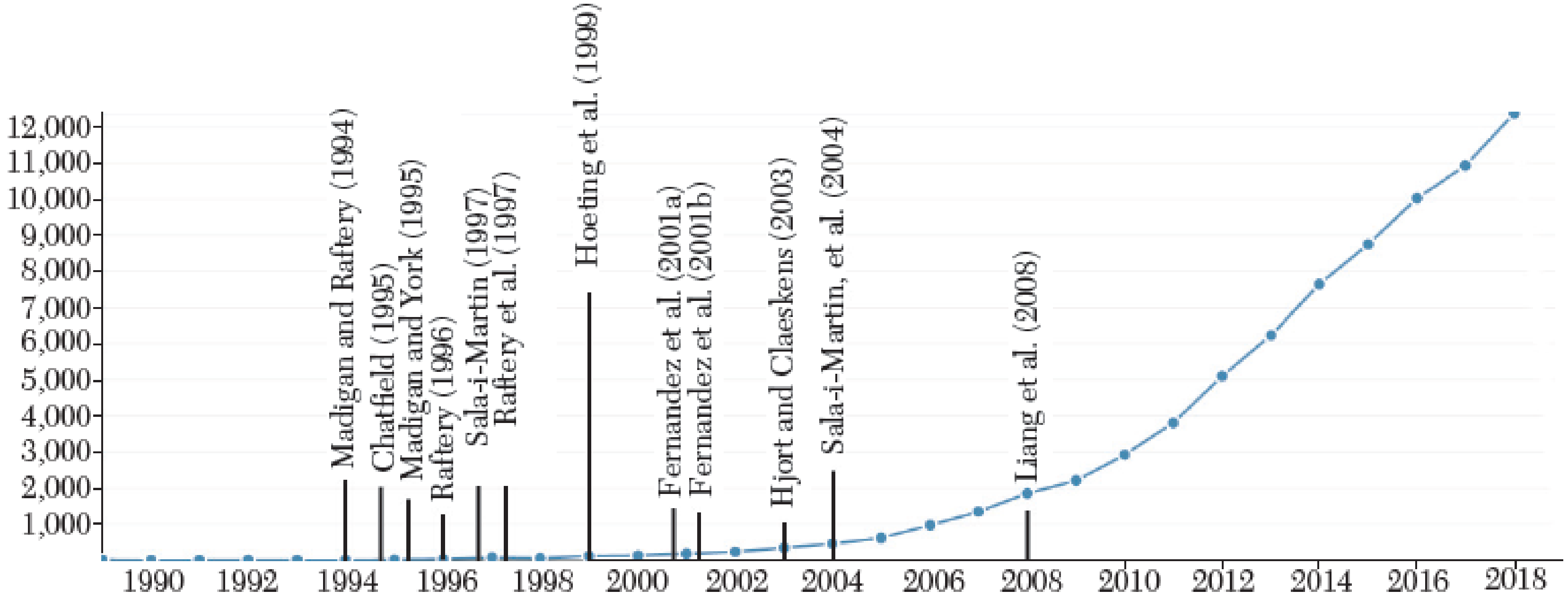


Figure 1. Total Number of Citations to Papers with Topic "Model Averaging" over Years 1989–2018

Note: Papers in economics or statistics journals with at least 250 citations are indicated by vertical lines proportional to the number of citations received.

Source: Web of Science, January 29, 2019.

A Google Scholar search of ["Bayesian model averaging" economics] yielded 14,000 total cites and about 6,600 since 2017.

Geographical Analysis ISSN 0016-7363

Bayesian Model Averaging for Spatial Econometric Models

James P. LeSage,[1] Olivier Parent[2]

[1]McCoy Endowed Chair of Urban and Regional Economics, McCoy College of Business Administration, Department of Finance and Economics, Texas State University—San Marcos, San Marcos, TX, [2]Department of Economics, University of Cincinnati, Cincinnati, OH

We extend the literature on Bayesian model comparison for ordinary least-squares regression models to include spatial autoregressive and spatial error models. Our focus is on comparing models that consist of different matrices of explanatory variables. A Markov Chain Monte Carlo model composition methodology labeled MC[3] by Madigan and York is developed for two types of spatial econometric models that are fre-

$$y = \alpha\iota_n + \rho W y + X_k\beta_k + \varepsilon$$

$$y = \alpha\iota_n + X_k\beta_k + \varepsilon, \quad \varepsilon = \rho W\varepsilon + u, \quad u \sim N(0, \sigma^2 I)$$

	ρ	β_1	β_2	β_3
M_1	1	1	1	1
M_2	1	0	1	1
M_3	1	0	0	1
M_4	1	1	1	0
M_5	1	1	0	0
M_6	1	0	1	0
M_7	1	1	0	1
M_8	1	0	0	0

$$y = \alpha \iota_n + \rho W y + X_k \beta_k + \varepsilon$$

Lesage and Parent (2007) offer three selection criteria to determine the important factors:

- The presence of the variable in the "top model" or model that maximizes $P(M_j)$
- The frequency of the variable in the "top ten models".
- The posterior probability of the individual variable $\pi(\theta_j|M_j)$

Labor Force Participation Rate: SBMA SAR Core Model

Age 55+ Baby Boomers	Vprob	Top	Top 10
Change in Employment 2000 to 2016	0.0522	0	0
Change in Populaiton 2000 to 2016	0.9667	1	10
Population to Employment Ratio 2016	0.9652	1	10
Percent of Employment in Farming 2016	0.9841	1	10
Percent of Employment in Manufacturing 2016	0.1364	0	1
Percent of Employment in Health Care and Social Services 2016	0.6148	0	7
Percent of Employment in Accommodations and Food Services 2016	0.9576	1	10
Percent of Employment in Government 2016	0.9067	1	10
Unemployment Rate 2011	0.7683	0	9
Unemployment Rate 2016	0.5027	1	4
Percent Black or African American	0.2460	0	2
Percent Other Minority (non-white, non-black)	0.1280	0	0
Percent those Age 25 and Over with a High School Diploma (including GED)	0.8023	1	8
Percent those Age 25 and Over with Some College, No Degree	0.0549	0	0
Percent those Age 25 and Over with Bachelor's Degree	0.9678	1	10
Percent of the Population Living in a Rural Place	0.0569	0	0
Population Density (sqr mile)	0.7671	1	9
Percent of Male Population Under Age 18	0.9654	1	10
Percent of the Male Population 65 Years and Older	0.2024	0	3
Percent of the Female Population Age 0-25	0.3014	0	1
Percent of the Female Population 65 Years and Older	0.8165	1	7
Percent of the Population With an Ambulatory Difficulty	0.9383	1	10

Note, we are moving away from the "pure" averaging approach here as was done by Sala-I-Martin toward using the results to variable identification.

Small Bus Econ (2021) 56:189–207
https://doi.org/10.1007/s11187-019-00224-y

Regional income inequality: a link to women-owned businesses

Tessa Conroy · Steven Deller · Philip Watson

Accepted: 18 June 2019 /Published online: 9 August 2019

Abstract We assess how women-owned and operated businesses relate to income inequality at the community level. Using U.S. county-level data within the framework of modeling uncertainty, we employ a spatial Bayesian model averaging approach to identify which

also found meaningful differences in the underlying control variable across our three measures of income inequality. Only a handful of control variables, such as the unemployment rate, rates of college education, and housing costs, are consistent predictors of income

Three-part question:

How does the concentration of women business owners impact community income inequality?

What are the relevant control variables that we need to account for?

Are the results sensitive to different measures of income inequality?

This is an example of Durlauf's policy evaluation, where the focus is on assessing the consequences of certain policies. The "policy" is women entrepreneurs.

Table 1: Spatial Bayesian Modeling Averaging for Income Distribution Measures

	Gini			Theil			Mean to Median Ratio		
	Posterior Probability	Top Model	Top Ten Models	Posterior Probability	Top Model	Top Ten Models	Posterior Probability	Top Model	Top Ten Models
Percent Of Housing Renter-Occupied	0.9816	1	10	0.9633	1	10	0.9797	1	10
Renter-Occupied Housing - Median ($000)	0.9629	1	10	0.9817	1	10	0.9581	1	10
Percent of Population Speak English Less Than "Very Well" (5 years of age and over)	0.2901	0	0	0.5505	0	5	0.0634	0	0
Percent of Population 25 years and Over - 9th to 12th grade, No Diploma	0.1800	0	0	0.7884	1	9	0.1238	0	0
Percent of Population 25 years and Over - High School Graduate (includes equivalency)	0.9612	1	10	0.9605	1	10	0.9567	1	10
Percent of Population 25 years and Over - Some College, No Degree	0.9628	1	10	0.9388	1	10	0.9576	1	10
Percent of Population 25 years and Over - Associate's Degree	0.9633	1	10	0.3864	0	2	0.9582	1	10
Percent of Population 25 years and Over - Bachelor's Degree	0.1714	0	0	0.1466	0	0	0.1823	0	0
Percent of the Population African-American	0.9625	1	10	0.9625	1	10	0.9592	1	10
Percent of the Population Latino	0.1109	0	0	0.5813	1	5	0.1084	0	0
Ethnic Diversity Index	0.4226	0	1	0.9614	1	10	0.0716	0	0
Percent of the Population Under Age 18	0.9630	1	10	0.9634	1	10	0.9419	1	10
Percent of the Population Over Age 65	0.0875	0	0	0.0878	0	0	0.2647	0	0
Population Density	0.1465	0	0	0.0899	0	0	0.6612	1	10
Percent of Employment: Farming	0.0964	0	0	0.1753	0	0	0.3591	0	2
Percent of Employment: Manufacturing	0.8780	1	9	0.5568	1	5	0.7776	1	5
Percent of Employment: Health Care and Social Assitance	0.9595	1	10	0.9621	1	10	0.5573	1	8
Percent of Employment : Tourism Related	0.0913	0	0	0.0719	0	0	0.1236	0	0

This tells us which, out of a wide range of potential control variables, are most consistent with the "underlying data generating process" and are there differences across three different measures of income inequality. What it does not tell us is the direction of these relationships.

Appendix Table A1: Full Specification Results for Gini Coefficient

Gini	Total	Total
Percent Of Housing Renter-Occupied	0.1665 ***	0.1597 ***
	(0.0001)	(0.0001)
Renter-Occupied Housing - Median ($000)	-0.0426 ***	-0.0427 ***
	(0.0001)	(0.0001)
Percent of Population 25 years and Over - High School Graduate (includes equivalency)	-0.1883 ***	-0.1860 ***
	(0.0001)	(0.0001)
Percent of Population 25 years and Over - Some College, No Degree	-0.2078 ***	-0.2245 ***
	(0.0001)	(0.0001)
Percent of Population 25 years and Over - Associate's Degree	-0.2290 ***	-0.2242 ***
	(0.0001)	(0.0001)
Percent of the Population African-American	0.0544 ***	0.0515 ***
	(0.0001)	(0.0001)
Percent of the Population Under Age 18	-0.1473 ***	-0.1488 ***
	(0.0001)	(0.0001)
Percent of Employment: Health Care and Social Assistance	0.0609 **	0.0695 **
	(0.0066)	(0.0016)
Economic Diversity Index	0.4946 ***	0.4770 ***
	(0.0001)	(0.0001)
Share of Total Personal Income: Wages and Salary	0.0564 ***	0.0525 ***
	(0.0001)	(0.0001)
Share of Total Personal Income: Proprietorships	0.0977 ***	0.1023 ***
	(0.0001)	(0.0001)
Share of Total Personal Income: Dividends, Interest and Rental	0.1823 ***	0.1794 ***
	(0.0001)	(0.0001)
Share of Total Personal Income: Transfer Payments	0.1757 ***	0.1574 ***
	(0.0001)	(0.0001)
Religious Adherent Rates	0.0218 ***	0.0230 ***
	(0.0001)	(0.0001)
Share of Employment Women Owned Businesses	0.0556 **	—
	(0.0009)	
Share of Establishments Women Owned Businesses		0.1733 ***
	—	(0.0001)

Marginal Significance or p-values in parentheses.
***: Significant at or above 99.9% level.
**: Significant at 95.0% level.
*: Significant at 90.0% level.

Using the SBMA approach we have high confidence that the model is "correctly" specified.

We have high confidence that the results on the control variables are reliable.

More women business owners, higher inequality.

➔ "push" and "pull" of women starting businesses. Area of more refined research and significant policy implications.

The Quarterly Review of Economics and Finance 64 (2017) 1–11

Contents lists available at ScienceDirect

The Quarterly Review of Economics and Finance

journal homepage: www.elsevier.com/locate/qref

ELSEVIER

CrossMark

Economic diversity, unemployment and the Great Recession

Philip Watson[a], Steven Deller[b,*]

[a] Department of Agricultural Economics and Rural Sociology, 28D Agricultural Science Building, University of Idaho, 875 Perimeter Drive MS 2334, Moscow, ID 83844-2334, United States

[b] Department of Agricultural and Applied Economics, 515 Taylor Hall—427 Lorch Street, University of Wisconsin-Madison, Madison, WI 53706, United States

ARTICLE INFO

Article history:
Received 4 May 2015
Received in revised form
21 November 2016

ABSTRACT

We revisit the relationship between economic diversity and unemployment in light of the economic shock of the Great Recession. Using U.S. county data we test the overall effect of increased industrial diversity on county level unemployment. Allowing for structured spatial spillovers across counties we find evidence supporting the notion that economic diversity within a given county is associated with lower levels of

This is an example of Durlauf's policy evaluation, where the focus is on assessing the consequences of certain policies. The "policy" is women entrepreneurs.

Three-part question:

How does the economic diversity affect unemployment prior to, during and after the Great Recession ?

What are the relevant control variables that we need to account for, but more importantly do they vary over time?

Does the relationship between economic diversity and unemployment rats change over time?

Table 2: Spatial Bayesian Model Averaging Posterior Results

	Unemployment Rate		
	2007	2010	2013
Percent of the Population under Age 18	0.9318	0.4459	0.7588
Percent of the Population over Age 65	0.9543	0.9715	0.9551
Population -- Employment Ratio	0.9572	0.9416	0.9553
Per Capita Income Relative to US Average	0.9553	0.9419	0.9550
Percent of Employment in Goods Production (minus Farming)	0.9554	0.9444	0.9558
Percent of Employment in Service Production	0.9001	0.3789	0.9554
Percent of Employment in Governments	0.9539	0.4271	0.9556
Percent of Households with Income below $20K	0.9523	0.2537	0.9460
Percent of Households with Income above $150K	0.9563	0.9422	0.9548
GINI Coefficient of Income Equality	0.9577	0.9428	0.9555
Per Capita Income from Transfer Payments	0.9551	0.9427	0.9540
Per Capita Income from Dividends, Interest and Rent	0.9569	0.9407	0.9543
Per Capita Proprietor Income	0.9556	0.4653	0.9433
Percent of the Population Latino	0.9441	0.9424	0.9766
Percent of the Population African American	0.9565	0.9411	0.9554
Population Density	0.7882	0.4294	0.6952
Expected Unemployment Rate	0.9763	0.9490	0.9549

Some variables are consistently associated with unemployment rates over all three time periods.

A handful, such as percent of employment in the public sector and services sector, are inconsistent across the three time periods.

Some, such as population density, simply does not matter.

Table 3: Diversity and Great Recession Unemployment (total effect)

	2007	2010	2013
Herfindalh Index (higher values more specialized)	-0.13861	39.22800 **	14.01497
	(0.988)	(0.007)	(0.253)
Percent of the Population under Age 18	4.82395 *	3.42799	-3.85196
	(0.098)	(0.456)	(0.325)
Percent of the Population over Age 65	-3.55676	-13.89941 **	-14.70851 ***
	(0.277)	(0.008)	(0.001)
Population -- Employment Ratio	0.29088	1.81461 ***	1.51402 ***
	(0.204)	(0.001)	(0.001)
Per Capita Income Relative to US Average	-2.68578 *	-11.39457 ***	-7.17084 ***
	(0.067)	(0.001)	(0.001)
Percent of Employment in Goods Production (minus Farming)	2.13880	15.33341 ***	9.68980 ***
	(0.148)	(0.001)	(0.001)
Percent of Employment in Service Production	4.02951 **	8.76709 ***	9.99082 ***
	(0.008)	(0.001)	(0.001)
Percent of Employment in Governments	-0.39968	1.35951	4.17615 *
	(0.820)	(0.625)	(0.092)
Percent of Households with Income below $20K	-1.31063	-9.07596 *	1.66861
	(0.691)	(0.077)	(0.710)
Percent of Households with Income above $150K	7.19975	26.90816 *	28.76415 **
	(0.390)	(0.052)	(0.018)
GINI Coefficient of Income Equality	-5.04035	-7.21488	-11.15642 *
	(0.249)	(0.287)	(0.065)
Per Capita Income from Transfer Payments	0.58164 ***	0.81835 ***	0.75356 ***
	(0.001)	(0.001)	(0.001)
Per Capita Income from Dividends, Interest and Rent	-0.03183	0.59718 ***	0.29669 **
	(0.667)	(0.001)	(0.004)
Per Capita Proprietor Income	0.04841	0.05647	0.10841
	(0.331)	(0.462)	(0.110)
Percent of the Population Latino	-0.54741	2.71071 **	1.70283 **
	(0.350)	(0.004)	(0.045)
Percent of the Population African American	0.30035	1.52044 *	2.04341 **
	(0.552)	(0.058)	(0.003)
Population Density	-0.00002	-0.00005	-0.00001
	(0.751)	(0.649)	(0.975)
Expected Unemployment Rate	0.43112 ***	0.73995 ***	0.55258 ***
	(0.001)	(0.001)	(0.001)

Marginal significance in parentheses.
*** Significant at 99.9% level.
** Significant at 95.0% level.
* Significant at 90.0% level.

We kept all the variables in this estimation of the full model to explore how well the SAR estimation lined up with the SBMA results:

They largely agree.

For example, population density.

Table 6 - Regression Results Herfindalh Index

Dependent Variable	Direct	Indirect	Total
Unemployment Rate 2007	3.56805 *	-3.70666	-0.13861
	(0.051)	(0.655)	(0.988)
Unemployment Rate 2008	5.72350 **	11.39095	17.11444 *
	(0.002)	(0.194)	(0.082)
Unemployment Rate 2009	8.11555 **	24.26320 *	32.37875 **
	(0.003)	(0.076)	(0.033)
Unemployment Rate 2010	10.49337 **	28.73464 **	39.22800 **
	(0.002)	(0.022)	(0.007)
Unemployment Rate 2011	10.34304 **	24.83259 **	35.17562 **
	(0.002)	(0.034)	(0.010)
Unemployment Rate 2012	8.72385 **	14.80207	23.52592 *
	(0.002)	(0.180)	(0.063)
Unemployment Rate 2013	7.97248 **	6.04249	14.01497
	(0.004)	(0.572)	(0.253)
Unemployment Rate 2014	7.64338 ***	8.37360	16.01698 *
	(0.001)	(0.306)	(0.081)

Control variable results supressed.
Marginal significance or p-values in parentheses.
***: Significant at 99.9% level.
** : Significant at 95.0% level.
* : Significant at 90.0% level.

Economic diversity within a county (direct effects) is consistently linked to unemployment rates prior to, during and after the Great Recession.

The effect appears to be the strongest in the "recovery" period.

The spillover effects (indirect) is only relevant in the immediate recovery years.

Some current examples of on-going work using SBMA:

Labor force participation rates with Heather Stephens
(43. Measurements & Methods
Thursday | 4:00 pm-6:00 pm | Mt Wilson)

What are the community characteristics associated with COVID-19 death and infection rates (Stephan Goetz)

Community Characteristics of COVID-19 Death/Infection Rates

Two-part question:

(1) from a wide range of potential community characteristics which ones are most consistent with the underlying data generating process,

(2) does social capital matter and if so what elements of social capital?

Community Characteristics of COVID-19 Death/Infection Rates

Steel's (2020) three broad areas:

- Prediction,
- Identifying the factors or determinants driving economic processes, (or what should be in ΣCC_j) and
- Policy evaluation, where the focus is on assessing the consequences of certain policies (or PV is social capital).

$$COVID\text{-}19 = CO = f(\Sigma CC_j,\ PV)$$

Nine “Blocks” of Characteristics

Income Inequality

Heath Access of Community

Health Characteristics of Community

Ethnic Characteristics of the Community

Education Characteristics of the Community

Age Characteristics of the Community

Poverty Characteristics of the Community

Economic Characteristics of the Community

Social Capital Characteristics of the Community

Consider income inequality:

Two parts: does income inequality help understand COVID-19 and out of the dozens of inequality measures, which is the “right” one to use?

Income Inequality

Gini Index

Median to Mean HH Income

Median to Mean Family Income

Ratio Number of HH Income $15k to $150K

Thiel Index

Income Inequality

Gini Index
Median to Mean HH Income
Median to Mean Family Income
Ratio Number of HH Income $15k to $150K
Thiel Index

Heath Access of Community

Average Daily PM2.5
Percent of Populaiton Food Insecure
Percent of Population Limited Access to Healthy Foods
Percent of Population Uninsured Health Insurance
Number of Hospitals per 10K Population
Number of Pharmacies per 10K Population
Primary Care Physcician per 10K Population
Mental Health Providers per 10K Population
Occupied Nursing Home Beds per 10K Population
Number of Nursing Home Jobs per 10K Population

Health Characteristics of Community

Percent of Adult Reporting Fair or Poor Health
Average Number of Physically Unhealthy Days
Average Number of Mentally Unhealthy Days
Percent of Adults Smokers
Percent of Adults with Obesity
Percent of Adults Physically Inactive
Percent of Adults Reporting Excessive Drinking
Percent of Population Uninsured
Life Expectancy
Percent of Adults with Diabetes

Ethnic Characteristics of the Community

Percent of Population Speak Only English at Home
Ethnic Diversity Index
Percent of the Population Black
Percent of Population Latino

Education Characteristics of the Community

Education Index
Percent Adults Age 25+ with Less Than a High School Degree
Percent Adults Age 25+ with a College Degree (Ass, Bach, Grad)

Age Characteristics of the Community

Age Index
Percent of Population Age 65+
Percent of Population Age 85+
Median Age

Poverty Characteristics of the Community

Family Poverty Rate
Youth Poverty Rate
Poverty Rate Those Age 65+
Working Poverty Rate

Economic Charateristics of the Community	Social Capital Charateristics of the Community
Age 16+ labor Force Participation Rate	Non-religious non-profit organizations p 1,000
Civilian Unemployment Rate	Religious congregations p 1,000
Percent of Workers Commute via Carpool	Violent Crimes p 100,000
Percent of Worker Commute via Public Transportation	Membership Organizations p 1,000
Percent of Wokers Worked from Home	Charitable contributions as share of AGI, middle-class itemizers
Percent of Workers Self-Employed	presidential election GOP minus DEM differenceper_point_diff
Percent of Employment in Arts, Ent., Recreation, Accom, and Food Services	2020 Census Response Rate
Herflndal Index of Economic Diversity	
Percent Households with Earnings Income	
Percent Households with Social Security Income	
Percent Households with Retirement Income	
Percent Households with Cash Public Assistance Income	
Percent Households with SNAP Benefits in the Past 12 Months	

We have a total of 60 variables and if the full model space M (possible combinations) is 2^K, K=60, the full model space is 1,152,921,504,606,850,000.

We estimated each block of potential variables separately.

County Level COVID death and infection rates

Vprob	Death Rate		Infection Rate	
	SAR	SEM	SAR	SEM
Gini Index	0.5498	0.1679	0.1693	0.1675
Median to Mean HH Income	0.2857	0.0855	0.1686	0.1673
Median to Mean Family Income	0.2512	0.1677	0.0870	0.0835
Ratio Number of HH Income $15k to $150K	0.5925	0.0865	0.0843	0.1661
Thiel Index	0.3767	0.1677	0.1723	0.0822

None of our income inequality measures appear to come into the model: income inequality does not appear to affect COVID death or infection rates.

County Level COVID death and infection rates

Vprob	Death Rate		Infection Rate	
	SAR	SEM	SAR	SEM
Average Daily PM2.5	0.8836	0.8789	0.8908	0.9316
Percent of Populaiton Food Insecure	0.8930	0.9074	0.3990	0.8365
Percent of Population Limited Access to Healthy Foods	0.3294	0.4642	0.2273	0.7514
Percent of Population Uninsured Health Insurance	0.8830	0.9522	0.7861	0.8655
Number of Hospitals per 10K Population	0.9332	0.9041	0.1395	0.1881
Number of Pharmacies per 10K Population	0.3487	0.7861	0.6055	0.4052
Primary Care Physcician per 10K Population	0.4427	0.7084	0.1263	0.2463
Mental Health Providers per 10K Population	0.7701	0.8125	0.1922	0.7041
Occupied Nursing Home Beds per 10K Population	0.8848	0.9094	0.8018	0.5990
Number of Nursing Home Jobs per 10K Population	0.5689	0.8962	0.1436	0.5675

- Note that the SAR and SEM specifications tend to “agree”.
- None of these factors pass the posterior probability of the individual variable $\pi(\theta_j|M_j)$ equal to or greater than 0.95. We dropped the threshold to 0.90.
- Note the differences between death and infection rates.
- What seems to matter, air pollution, health insurance, access to hospitals, nursing homes.

County Level COVID death and infection rates

Vprob	Death Rate		Infection Rate	
	SAR	SEM	SAR	SEM
Percent of Adult Reporting Fair or Poor Health	0.9495	0.8792	0.8669	0.9332
Average Number of Physically Unhealthy Days	0.8994	0.8822	0.4376	0.8499
Average Number of Mentally Unhealthy Days	0.7607	0.2944	0.9336	0.8760
Percent of Adults Smokers	0.4521	0.3569	0.5706	0.2644
Percent of Adults with Obesity	0.7330	0.3238	0.8664	0.8807
Percent of Adults Physically Inactive	0.7840	0.7954	0.2946	0.2456
Percent of Adults Reporting Excessive Drinking	0.9017	0.8131	0.2311	0.2593
Percent of Population Uninsured	0.5424	0.5692	0.6464	0.8772
Life Expectancy	0.9012	0.9389	0.8661	0.8781
Percent of Adults with Diabetes	0.8984	0.8821	0.6428	0.5296

- Somewhat surprising that health characteristics of the community population does not come into play to a larger extent.

- The posterior probability of the individual variable $\pi(\theta_j|M_j)$ threshold, if we drop it to 0.85 or 0.80 the variables that enter the model jumps by a lot. Does a certain degree of arbitrariness enter the analysis? Are we back to changes in the equation F statistic, $\bar{R}^2$ or Mallows' C_p Amemiya criteria (PC), Akaike Information Criteria (AIC), Sawa Bayesian Information Criterion and/or the Schwarz Bayesian Information Criterion (BIC)?

County Level COVID death and infection rates

Vprob	Death Rate		Infection Rate	
	SAR	SEM	SAR	SEM
Non-religious non-profit organizations p 1,000	0.8441	0.8465	0.93292	0.8304
Religious congregations p 1,000	0.8483	0.8474	0.63048	0.4428
Violent Crimes p 100,000	0.8962	0.8912	0.86176	0.9139
Membership Organizations p 1,000	0.8471	0.9230	0.84566	0.4667
Charitable contributions as share of AGI, middle-class itemizers	0.9197	0.8495	0.86208	0.8302
Presidential election GOP minus DEM Difference	0.4777	0.5218	0.86286	0.7512
2020 Census Response Rate	0.5509	0.5369	0.86196	0.8271

Does social capital matter, or original question. Well, “it depends” even at the reduced 0.90 only a handful of measures come in. But if we drop to 0.80 a lot of these measures jump in.

Next step: select the relevant control variables (even at a reduced posterior probability threshold), then estimate using SAR and SEM with the social capital measures stepped in.

Concluding Comments

- The study of community economic development is truly interdisciplinary (Isard’s vision of regional science).

- A systems thinking approach helps contextualize the issues.

- Everything matters, we have multiple ways of measuring those “community capitals”, everything is endogenous.

- The notion of “modeling uncertainty” comes to the forefront.

- Bayesian Model Averaging (BMA) is an avenue worthy of farther exploration.

Concluding Comments

Examples of application of BMA in regional science:

Resource Curse: Peren & Braunfels. (2018) *Energy Economics*

Income Inequality: Hortas-Rico & Rios. (2019) *Regional Studies*

Government and Regional Resiliency: Rios & Gianmoena. (2020) *Journal of Policy Modeling*

Human Capital and Regional Growth: Cuaresma, et al. (2018) *Journal of Regional Science*

Predicting Demand for Solar Power: Doubleday, et al. (2020) *IEEE Transactions on Sustainable Energy*

Concluding Comments

Example extensions of BMA:

Instrumental Variable BMA: Oueslati, Salanié, & Wu. (2019) *Journal of Economic Geography*

Panel Data BMA: Desbordes, Koop, & Vicard. (2018) *Economic Modelling*

Stochastic Frontier BMA*:* Makieła & Mazur. (2020) *Econometrics*

Robust Bayesian Meta-analysis (RoBMA): Maier, Bartoš, & Wagenmakers. (forthcoming) *Psychological Methods*

Steven Deller
scdeller@wisc.edu