Assessing Factors that Influence Women's Participation in the Invention Ecosystem

Michelle Saksena with Nicholas Rada, Katherine Black and Lisa Cook

ASSA 2022

1/8/2022

UNITED STATES PATENT AND TRADEMARK OFFICE

uspto

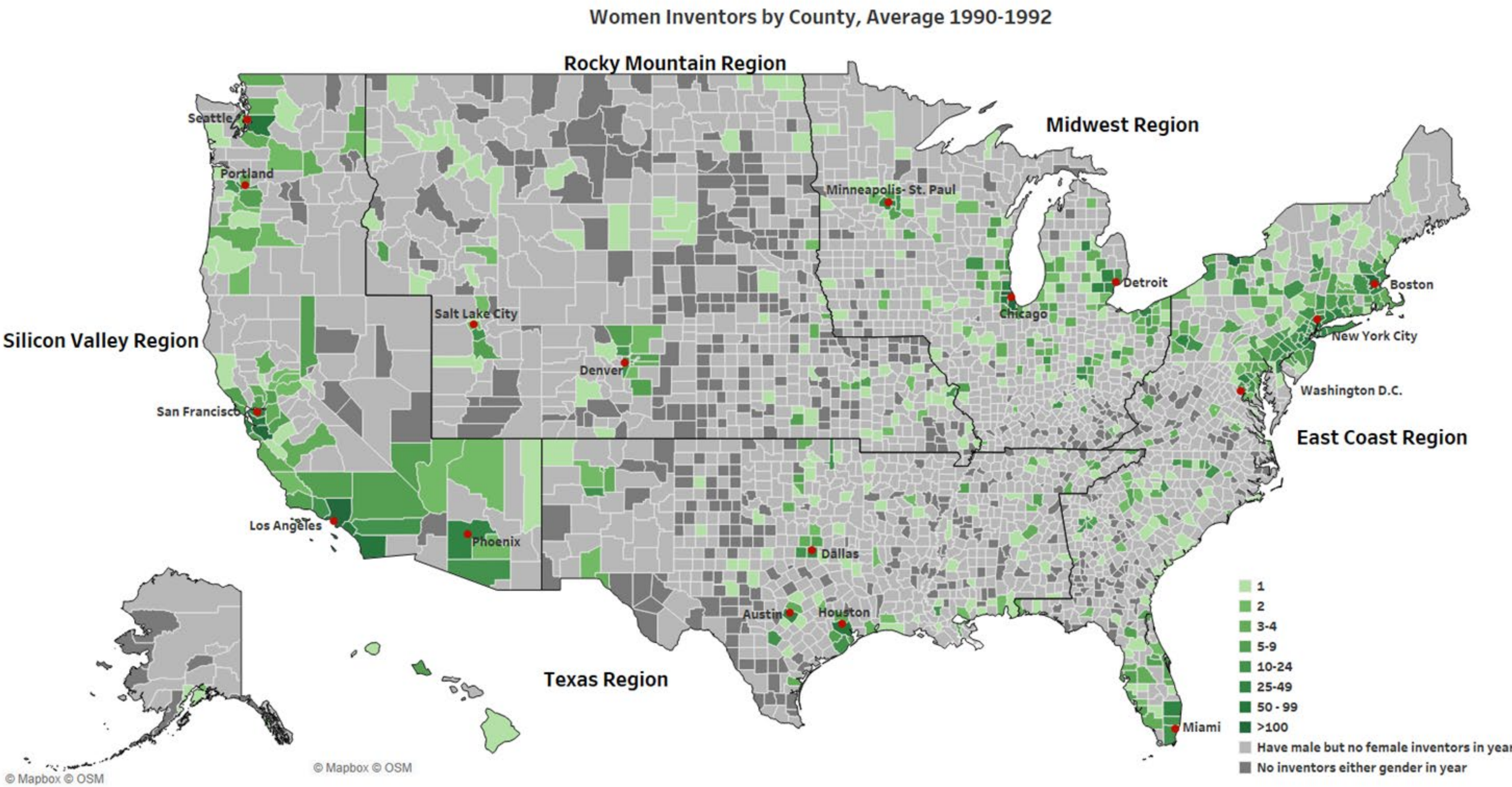
Background

- In 2019 women made up only 13% of all inventor-patentees in the United States (Toole et al. 2019).
- If women were to patent at the same rate as men, commercialized patents could rise by 24% and per-capita GDP could increase by 2.7%. (Bell et. al 2019 and Hunt 2016)
- Gender diversity boosts the inventive process in essential ways:
  1. unique female experiences and viewpoints help inform, and thus improve, the quantity and quality of innovation (Milli and Williams-Baron et al., 2016; OECD 2018; Xie et al., 2020)
  2. gender diversity expands research into under examined topics thereby filling overlooked technology gaps (Koning et al., 2021)
  3. women contribute social shrewdness that increases team cooperation and productivity (Xie et al., 2020)

Questions

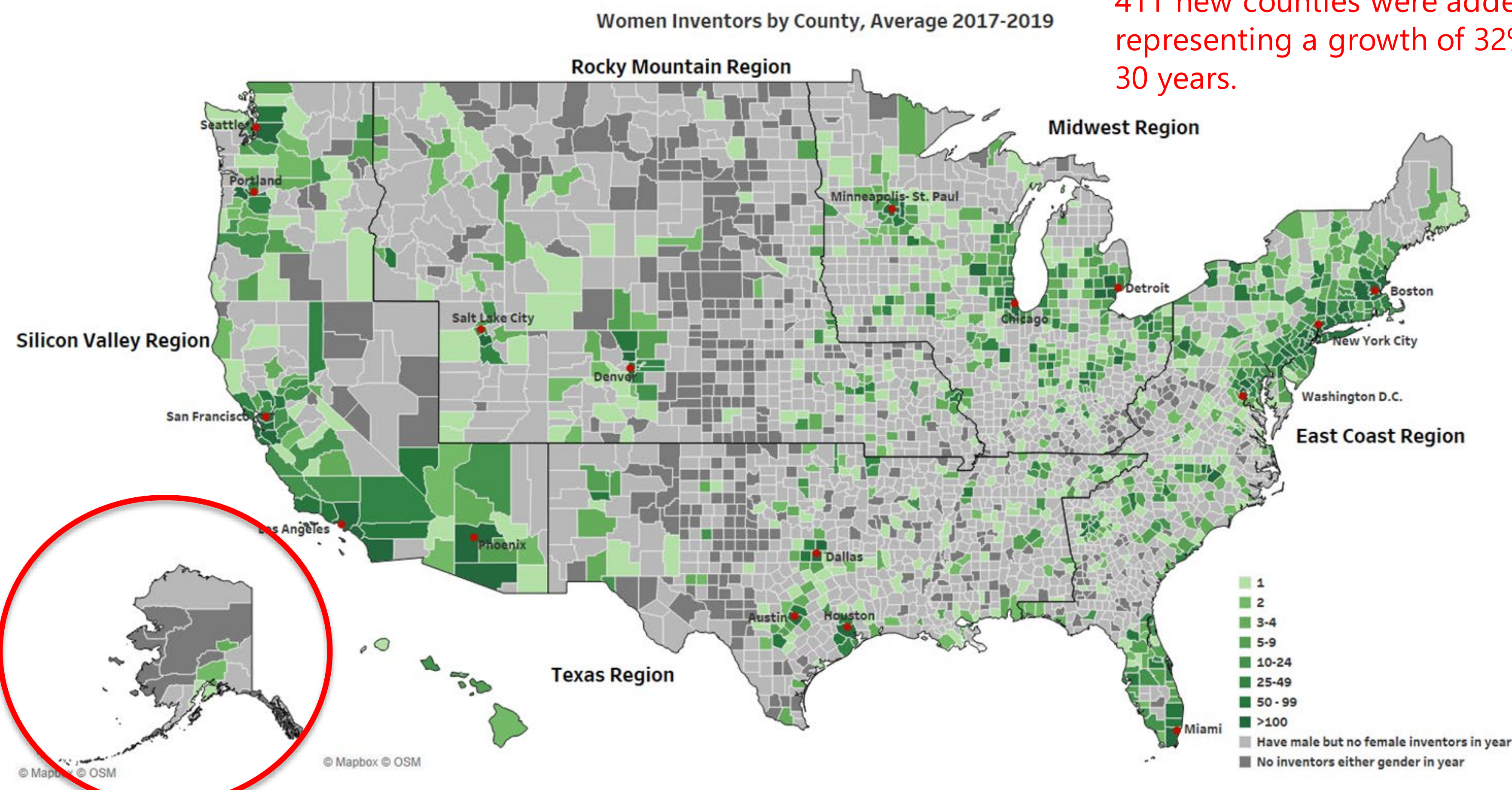
1. Where are the women inventors?
2. Does an environment with highly educated women help a county's chance of having its very first woman inventor-patentee?
3. How does team size affect the propensity for women to patent?
   - Team size proxies capital investment in a technology development (Breitzman and Thomas, 2015)

Past: Number of women inventors



Present: Number of women inventors

411 new counties were added representing a growth of 32% in 30 years.



Differences in counties with and without women inventors that patent

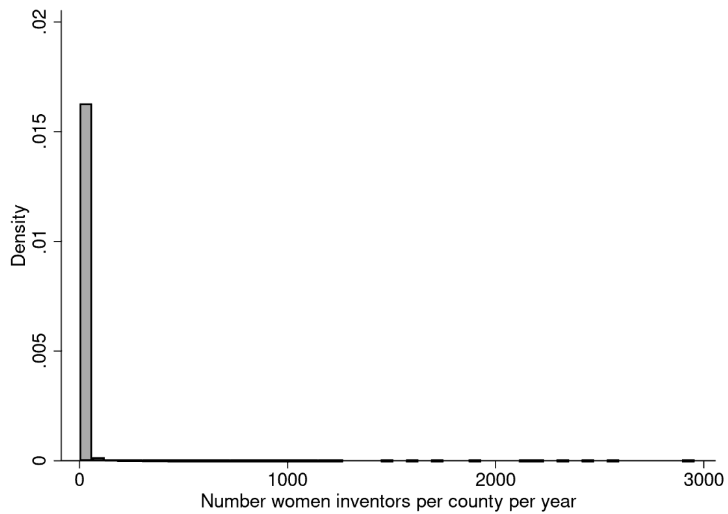
	Counties with women inventors = 0					Counties with women inventors > 0				
	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
Number of women inventors	-	-	-	-	-	29,801	15	67	1	2,956
Labor force	34,145	16,481	16,508	261	416,540	29,801	120,298	250,215	460	5,121,584
Per capita income ($)	34,145	27,615	10,474	7,096	175,998	29,801	33,672	13,570	9,798	230,141
Number of women with...										
Bachelor's degrees	34,145	1,310	1,572	17	33,362	29,801	14,556	33,026	34	761,572
Master's degrees	34,145	521	664	0	22,797	29,801	6,252	14,227	3	287,419
PhDs	34,145	42	70	0	1,483	29,801	700	1,854	0	40,577

Controlling for population density, there are 38% more highly educated women in counties with women inventors that patent than in counties without them.

Methods

Zero-inflated negative binomial model (ZINB)?

- Y is a count variable
- Used when our Y variable (women inventor counts) is overdispersed; $\mu = 5$ $\sigma = 39$
- Used when there are a lot of zeros in the data



ZINB

- The intuition behind the Zero Inflated Negative Binomial model is that there is a precursor process that determines whether a county is (structural-) zero or non-zero.
    - Zero-counties= {observational zeros, structural zeros}
- Zero/non-zero process is determined by a logit model:
    - In our model this process is characterized by the number of highly educated women (bachelors, masters and PhDs). Specifically, how it influences a county's probability of having its first female inventor.
- Once a county is determined to be non-zero, the Negative binomial process takes over to determine the number of women inventors.
    - In our model the negative binomial process is a function county-level labor-force, per-capita income, team size, # of all male teams and technology field concentrations and indicator variables.

Empirical specification

Logit:

$$\lambda_{it} = \exp(\boldsymbol{Z_{it}}\Delta) = \exp\left(z'_{itj}\delta\right)$$
$$= \exp(\delta_0 + \delta_{fbachelors} fbachelors_{it} + \delta_{fmasters} fmasters_{it}$$
$$+ \delta_{fphd} fphd_{it} + \sum_{c=2}^{8} \beta_c \%cpc_{cit} + \sum_{c=1}^{8} \beta_c dcpc_{cit}$$
$$+ \delta_t YearFE_t + \delta_l StateFE_l)$$

Negative Binomial

$$\mu_{it} = \exp(\mathbf{X}_{it}\mathbf{B}) = \exp(x'_{itj}\beta) = \beta_0 + \beta_{LF}LF_{it} + \beta_{PCI}PCI_{it} + (\beta_{TS}TS_{it} + \beta_{TS^2}TS_{it}^2) + \beta_{AM}AM_{it} + \sum_{c=2}^{9}\beta_c\%cpc_{cit} + \sum_{c=1}^{9}\beta_c dcpc_{cit} + \beta_t Year_t + \beta_{state}State_l)$$

- where subscript *i* refers to counties, *t* refers to time (1991-2019), and *j* = *Labor force (LF), per capita income (PCI), Team size (TS), number of all male teams (AM), technology field concentration (%cpc), technology field indicators (dcpc)*.

Pooled, national summary statistics

Source: *County-level data from* ***BEA****,* ***BLS*** *and* ***PatentsView*** *(1990-2019)*

	N	Mean	Std. Dev.	Min	Max
Number of women inventors	29,801	15	67	1	2,956
Labor force	29,801	120,298	250,215	460	5,121,584
Per capita income ($)	29,801	33,672	13,570	9,798	230,141
Team size	29,801	3	1	1	23
Number of all male teams	29,801	149	578	0	21,415
% CPC technology concentration					
Human necessities	29,801	22%	21%	0%	100%
Performing operations; transporting	29,801	20%	18%	0%	100%
Chemistry; metallurgy	29,801	12%	15%	0%	100%
Textiles; paper	29,801	1%	6%	0%	100%
Fixed constructions	29,801	5%	11%	0%	100%
Mechanical engineering; lighting; heating; weapons; blasting engines or pumps	29,801	10%	14%	0%	100%
Physics	29,801	17%	16%	0%	100%
Electricity	29,801	13%	15%	0%	100%
CPC technology indicators					
Human necessities	29,801	85%	35%	0%	100%
Performing operations; transporting	29,801	84%	36%	0%	100%
Chemistry; metallurgy	29,801	68%	47%	0%	100%
Textiles; paper	29,801	28%	45%	0%	100%
Fixed constructions	29,801	58%	49%	0%	100%
Mechanical engineering; lighting; heating; weapons; blasting engines or pumps	29,801	71%	45%	0%	100%
Physics	29,801	76%	43%	0%	100%
Electricity	29,801	69%	46%	0%	100%
Number of women with…					
Bachelor's degrees	29,801	14,556	33,026	34	761,572
Master's degrees	29,801	6,252	14,227	3	287,419
PhDs	29,801	700	1,854	0	40,577

Economic & Team Size vars

% of patents by technology field in each county

% percent of counties with at least one patent by technology field

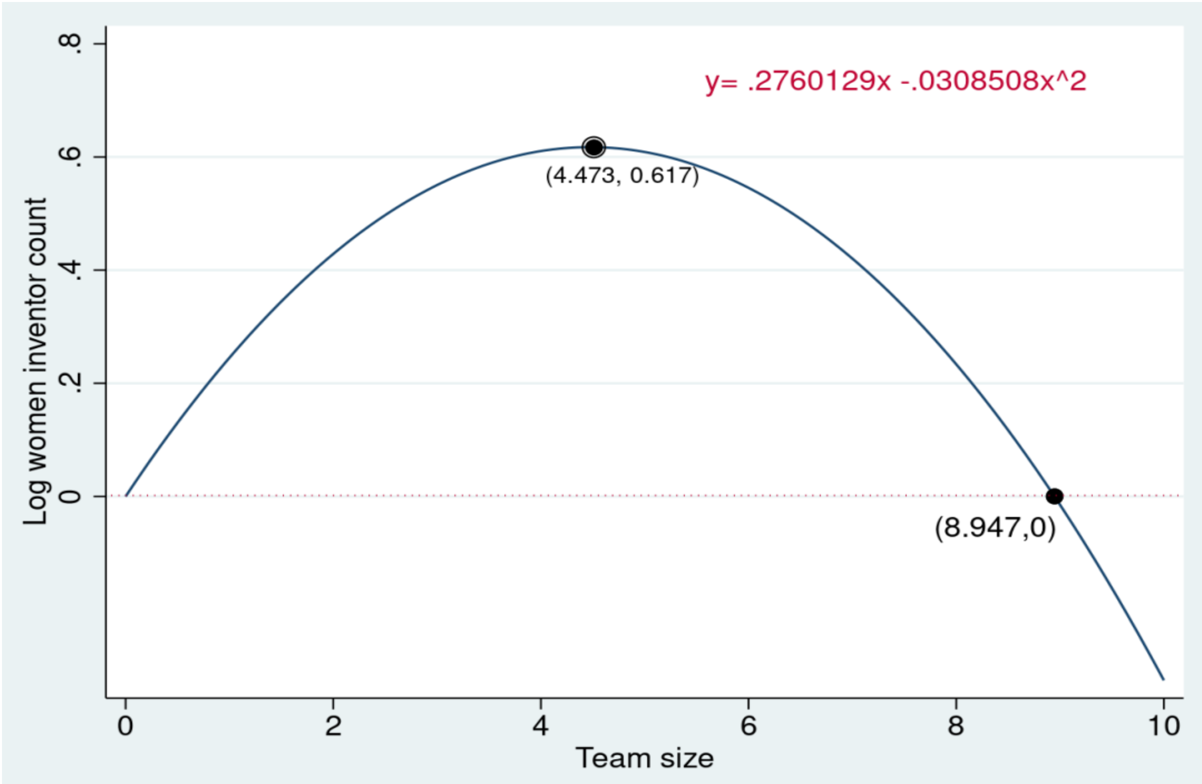
Women's graduate levels

Results:

National Model			
Variable	β	exp(β)	p-value
Logit			
Bachelors	-0.00046	0.999537	0.07
Masters	-0.00077	0.999231	0.20
PhD	-0.00482	0.995192	0.04
Negative Binomal			
Economic variables			
Labor force	1.64E-06	1.000002	0.00
Per capita income (USD)	0.000019	1.000019	0.00
Inventor team variables			
Team size	0.276013	1.317865	0.00
Team size squared	-0.03085	0.96962	0.00
Number of all male teams	0.000681	1.000681	0.02
% of counties with Cooperative Patent Classification (CPC) concentrations			
Performing operations; transporting	-0.9668	0.3803	0.00
Chemistry; metallurgy	0.669494	1.953248	0.00
Textiles; paper	-1.47264	0.229321	0.00
Fixed constructions	-2.06514	0.1268	0.00
Mechanical engineering; lighting; heating; weapons; blasting engines or pumps	-1.68458	0.185522	0.00
Physics	0.321866	1.3797	0.09
Electricity	-0.01247	0.987607	0.95

Results summary

- Team size has non-linear relationship



uspto

Results: Regional Model

Regional model		
	exp(β)	p-value
Negative Binomial		
Labor force		
East coast	1.000002	0.00
Midwest	1.000002	0.42
South	1.000002	0.72
Mountain	1.000004	0.00
West Coast	1.000002	0.38
Per Capita Income (USD)		
East Coast	1.000015	0.00
Midwest	1.000015	0.49
South	1.000015	0.39
Mountain	1.0000023	0.00
West Coast	1.000015	0.63
Team size		
East Coast	1.212372	0.00
Midwest	1.212372	0.45
South	1.097086213	0.00
Mountain	1.134009487	0.06
West Coast	1.364214319	0.07
Number of all male teams		
East Coast	1.001505	0.00
Midwest	1.001505	0.35
South	1.001505	0.18
Mountain	1.001505	0.74
West Coast	1.000282663	0.00

Results summary

- Doubling the number of female college graduates in a county that has never had a women inventor would increase the county's chance of hosting its first women inventor by 60%.
- A county adding one female PhD has 10X the effect of increasing the county's chance of having its first female inventor than that of adding one female with a Bachelor's degree.
- Larger teams sizes up to 4.5 would increase the likelihood of women inventors. The average team size is currently 2.7.
- Increasing team size in the Silicone Valley region has the highest potential for increasing women's inventors.

Additional slides

Math (See appendix for detail)

- We substitute 3 into 1 and 2 and add them all together and take logs to get a log-likelihood function
- ZINB essentially nests the negative binomial distribution of >0 counties, Γ(.), inside the logistic distribution, F(.). Several math steps later, we get the log-likelihood function…

Pr of observing a 0 county ~ F (logit)

Pr of k women inventors in a non-zero county ~ Γ(μ,α) (Negative binomial)

$$ln\mathrm{L} = \sum_{i \in S} w_i \ln\{F_i + (1 - F_i)\, p_i^m\} + \sum_{i \notin S} w_i \{\ln(1 - F_i) + \boldsymbol{\Gamma}(m + y_i) - \boldsymbol{\Gamma}(y_i + 1) - \boldsymbol{\Gamma}(m) + m \ln p_i + y_i (1 - p_i)\}$$

where $m = \frac{1}{\alpha}$; $p_i = \frac{1}{1+\alpha\mu_i}$

Negative binomial: Regional model

$$\mu_{it} = \exp(x'_{itj}\beta)$$
$$= \exp(\beta_0 + (\beta_1 LF_{it} + \beta_2 LF_{it} * R_{r-1})$$
$$+ (PCI_{it}\beta_3 + \beta_4 PCI_{it} * R_{r-1}) + (\beta_5 TS_{it} + \beta_6 TS_{it} * R_{r-1})$$
$$+ \sum_{c=2}^{8} \beta_{(c+5)} \% cpc_{(c+5)it} + \sum_{c=1}^{8} \beta_{(c+14)} dcpc_{(c+14)it}$$
$$+ \beta_t YearFE_t + \beta_l StateFE_l$$

where USPTO region subscript *r* = *Eastern, Midwest, South, Mountain, Western.*;

Table 3. Regional model preliminary results

Regional model		
	exp(β)	p-value
Negative Binomial		
Labor force		
East coast	1.000002	0.00
Midwest	1.000002	0.42
South	1.000002	0.72
Mountain	1.000004	0.00
West Coast	1.000002	0.38
Per Capita Income (USD)		
East Coast	1.000015	0.00
Midwest	1.000015	0.49
South	1.000015	0.39
Mountain	1.0000023	0.00
West Coast	1.000015	0.63
Team size		
East Coast	1.212372	0.00
Midwest	1.212372	0.45
South	1.097086213	0.00
Mountain	1.212372	0.06
West Coast	1.212372	0.07
Number of all male teams		
East Coast	1.001505	0.00
Midwest	1.001505	0.35
South	1.001505	0.18
Mountain	1.001505	0.74
West Coast	1.000282663	0.00

% of counties with Cooperative Patent Classification (CPC) concentrations		
Performing operations; transporting	0.3763327	0.00
Chemistry; metallurgy	2.218898	0.00
Textiles; paper	0.3363812	0.00
Fixed constructions	0.1533107	0.00
Mechanical engineering; lighting; heating; weapons; blasting engines or pumps	0.1946372	0.00
Physics	1.226501	0.14
Electricity	0.9786238	0.89
Technology indicators by CPC		
Human necessities	1.610621	0.00
Performing operations; transporting	1.92963	0.00
Chemistry; metallurgy	1.533529	0.00
Textiles; paper	1.505251	0.00
Fixed constructions	1.722936	0.00
Mechanical engineering; lighting; heating; weapons; blasting engines or pumps	1.965105	0.00
Physics	1.357513	0.00
Electricity	1.482946	0.00
Logit		
Bachelors	0.999495028	0.04
Masters	0.999589084	0.43
PhD	0.993760149	0.00