THE PERILS OF RELYING ON STATE ECONOMIC PERFORMANCE RANKINGS WITHOUT ADJUSTING FOR HETEROGENEITY

A.E. Rodriguez & Brian Marks[1]

Abstract

State performance rankings are ubiquitous. But most rankings fail to recognize the heterogeneity inherent in the seemingly "objective" variables utilized to structure the ordering. A more parsimonious representation can be accomplished by adjusting the ordering variable by its most important attributes.

To demonstrate the procedure, we utilize a state ranking based on Cumulative GDP Growth. We identify the relative importance and sensitivity of several popular variables used in explaining the variation in cumulative gdp growth performance among the states. Once identified, important variables can enhance the effectiveness of legislators and administrators' policy-making efforts. State performance rankings are recast after adjusting cumulative gdp growth for the important drivers identified.

The period examined is 2004-2014. To identify the importance and sensitivity of predictors we utilize random forests via the R packages relaimpo, Boruta, and random forests. Partial dependence depictions of the critical variables identified enable policy inferences.

Specifically, we find that the top marginal personal tax rate and the number of state employees exert and uncommonly high influence in explaining variation in state performance rankings based on cumulative gdp-growth.

The method proposed here is of general applicability and can be deployed to extract robust policy prescriptions based on a more accurate treatment of data given the limitations of traditional econometric models.

DRAFT: January 8, 2017

JEL: C43, M21, H70, R15

[1] Department of Economics, University of New Haven, Email: arodriguez@newhaven.edu, bmarks@newhaven.edu. Rodriguez is corresponding author.

"Who comes out on top, in any ranking system,
is really about who is doing the ranking."

Malcolm Gladwell (2011)

Relative performance among states is routinely gauged across any number of indicators. These indicators could be individual, conventional indicators such as gross domestic product per capita or employment growth or they could be composite indicators assembled algorithmically from elementary variables.[2] The ordering provided by performance measures simplify often complex debates by reducing the dimensions involved thereby facilitating comparisons and benchmarking. Rankings compel the questioning of individual standings. The ranks invite attention from the media and policymakers and, increasingly, from the general public – leading invariably to finger-pointing or high-fives.[3] And thus, almost inevitably rankings constitute invitations to look more closely at the explanations that underlie them (Saltelli 2007).

Despite their seeming intuitive simplicity, the construction and usefulness of composite indicators has been severely criticized - if not compromised (Artz, et al. 2016, Kolko, Neumark and Cuellar-Mejia 2013, Paruolo, Saisana and Saltelli 2012, Gladwell 2011). Less scrutiny however, has been given to the capability of individual economic indicators used for purposes of state performance rankings. Unfortunately, individual indicators – such as State GDP Growth, Absolute Net Migration or Employment Growth, *inter alia* – may betray an unwanted heterogeneity that distorts the resulting ordering. For instance, an ordering based on State Gross Domestic Product would most likely rank more populous states higher and thereby convey a misleading sense of a state's relative

2 There are many examples. For one instance of the comparative use of economic indictors see "New England Economic Indicators," Federal Reserve Bank of Boston [https://www.bostonfed.org/publications/new-england-economic-indicators.aspx]. For an example of a composite indicator-based ranking see the Freedom in the 50 States ranking published by the Cato Institute [https://www.freedominthe50states.org/].

3 See, e.g., U.S. News & World Report's 2016 Best Places to Live rankings [http://realestate.usnews.com/places/rankings-best-places-to-live], and Forbes' Best Places to Retire in 2016 [http://www.forbes.com/sites/williampbarrett/2016/04/04/the-best-places-to-retire-in-2016/#669dc952703e]. (viewed December 7, 2016).

performance. Some of the disparity in a GDP-based ranking is isolated by normalizing: the setting forth of a common denominator. Thus, and perhaps obviously, GDP normalized by a state's population into a new metric altogether is considered a more informative alternative than a standalone measure based on GDP.

However, the reason for reducing GDP by a state's population – isolating the desired dimension to curtail the likelihood of biased performance metrics - applies with equal force to any number of variables. Put differently, should we not adjust GDP by other variables to fully extract a more representative economic performance signal? Because it is not clear which other variables should be considered the issue of subjectivity in variable selection and domain relevance re-enters the debate (Sitglitz, Sen and Fitoussi 2009).

Yet, it is possible to examine and identify the most important variables in explaining the difference in seemingly objective performance metrics. In this paper we rely on the important-feature selection capabilities of random forests to identify the most important predictors of performance metric variability. With key drivers identified, state relative performance metrics are recast. Importantly, identifying the most important variables and their sensitivity constitutes a roadmap for policymakers intent of addressing shortcomings.

The methodology, references and sources of data are provided in this paper. The associated code is available upon request. The paper is arranged as follows. The next section provides a graph of state rankings based on cumulative growth in gdp. Cumulative growth in gdp stands as our archetypical example. The third section contains a discussion of random forests and feature selection. The fourth section contains our results. The last section concludes.

Relative Performance Among 50 States

The period examined is 2004-2014 and encompasses the fifty states of the United States. Relative performance is gauged from three perspectives: Cumulative Growth in State Gross Domestic Product, Absolute Domestic Migration, and Cumulative Employment

Growth. All three performance variables are responsive to state policymakers' prescriptions. And all three variables are recurring in policy and political debates.

Figure 1

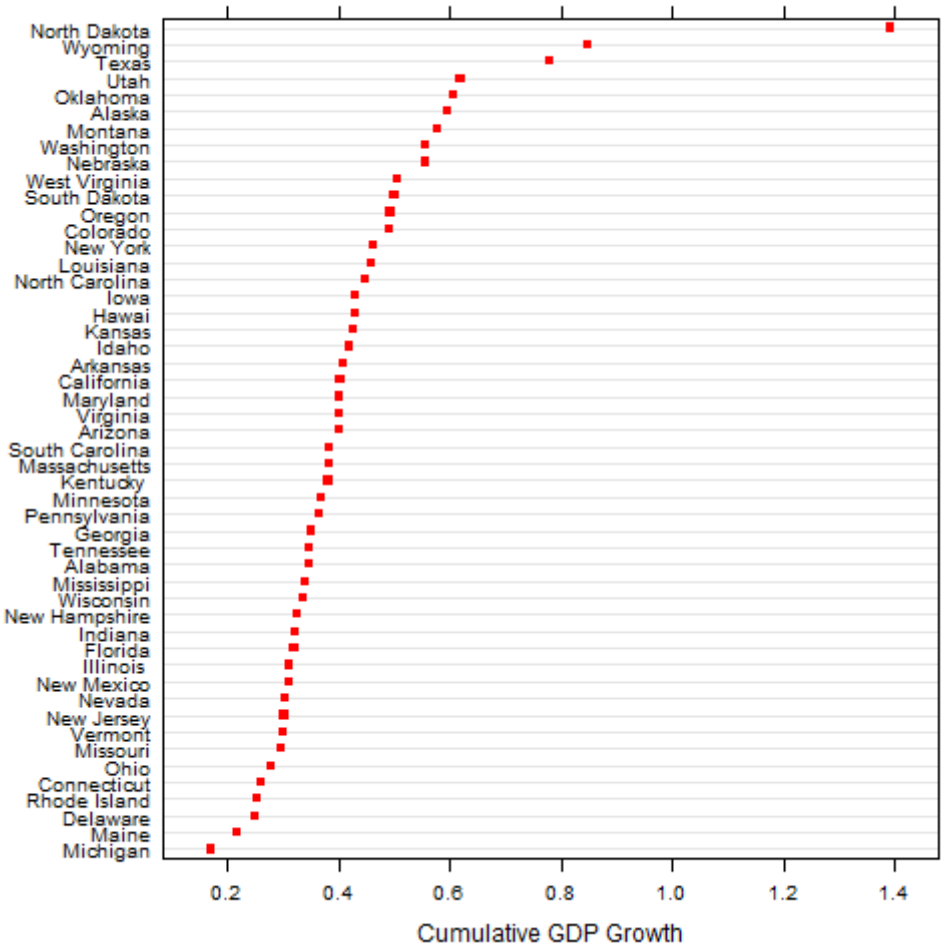


Figure 1 provides a visual depictions of the associated ordering induced by a specific performance variable. Specifically, Figure 1 displays states rank in terms of cumulative GDP growth across the period examined. To conserve space, Figure 2 & Figure 3 are displayed in the appendix to this paper. They display how states rank in terms of cumulative employment growth and absolute net migration.

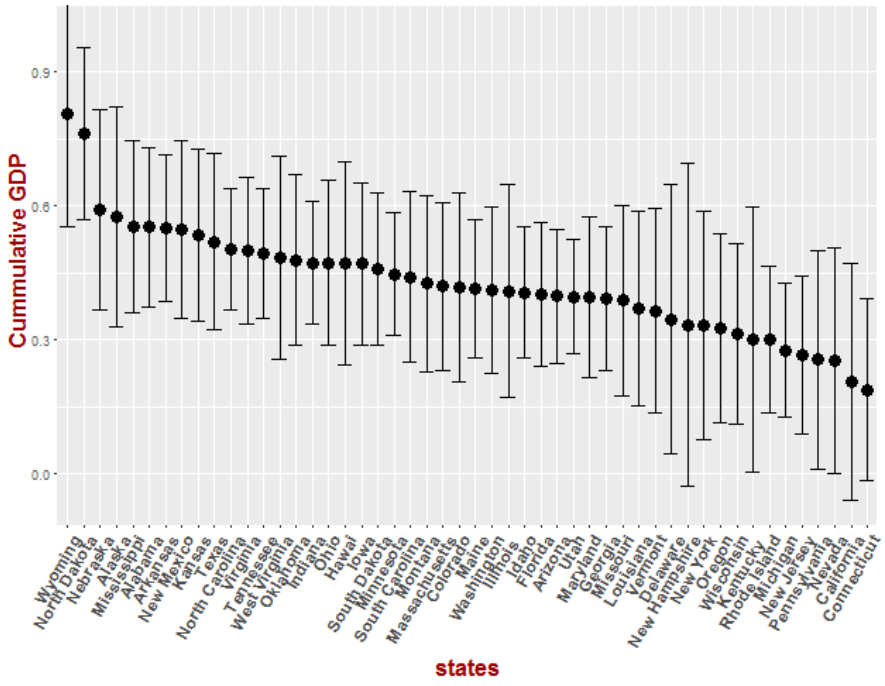
We resort to a machine learning algorithm because is empirically difficult to isolate the key explanatory variables responsible for GDP growth using conventional econometric methods. In fact, researchers have tested numerous explanatory variables and resorted to a dizzying array of econometric approaches in attempting to explain factors

underscoring the realized difference in the relative performance of states based on GDP growth. A sampling of recent work in this area include differences in tax policy (McBride 2012, Gale, Krupkin and Rueben 2015, Segura III 2016), on the composition of clusters (Delgado, Porter and Stern 2012), on historical industry structure (Higgins, Levy, and Young, 2006), and knowledge and technology (Moretti 2012, Glaeser 2011, Florida 2002).

There are other technical glitches that diminish the power of individual metrics in ordered unadjusted rankings. Few studies can reach the entire population. The performance variable is assembled by surveying samples of a population; as such it admits sampling error. The magnitude of resulting margin of error may be sufficiently wide to vitiate any meaningful difference between any two consecutive positions in the resulting orderings. Thus, for example and *arguyendo*, despite a seeming difference in gdp growth between two states ranked 20th and 21st the observed difference cannot be established as being significantly different from zero – thus rendering the ordering meaningless. Generally, the closer the resulting values are to each other the higher is the probability of error in the ranking. To illustrate this statistical artifact, Figure 2 below displays the state ranking based on cumulative gdp growth.[4] The graph includes estimated confidence intervals around each of the results obtained for each of the sampling units. At the very least, the overlapping confidence intervals suggest caution in the interpretation of the resulting ordering.

4 The confidence intervals are those associated with the forecast dependent variable of a linear model fitting cumulative gdp growth to the fifteen original explanatory variables.

Figure 2



The procedure for adjusting for heterogeneity proposed here is unlikely to eliminate altogether the confidence interval associated with the predicted dependent variable. But it is likely to reduce the margin of error.

Predictor Importance

There are numerous qualitative and quantitative approaches to determine predictor importance. Analysis of variance, principal components analysis, factor analysis, discriminant analysis, multivariate regression, various machine learning algorithms, and logistic regression *inter alia* are commonly used to relate attributes to dependent variables (LeBreton, Ployhart and Ladd 2004). All these procedures are capable of ascribing predictor influence on the variability of the response variable.

Driver importance is a difficult task across the board but especially vexing in the performance indicator literature. The difficulty lies in the fact that many economic variables reflect overlapping concepts thereby jointly contributing to the variability of the de-

pendent variable. Moreover, various steps in the construction of an index can be subjective. The process tends to artfully reflect their author's remit, or – as the case might be, their ideological leanings, theoretical preconceptions, their political identity or agenda. The variables selected, variable construction, the aggregation procedure, the time period encompassed, and the weights utilized provide considerable leeway to an index builder to shape or assist a narrative (Nardo, et al. 2005, Sitglitz, Sen and Fitoussi 2009). In addition, because of the interactions among attributes it is difficult to isolate the net effect or individual contribution of an attribute to the dependent variable.

The Relative Importance of the Determinants of Performance

We scrutinize fifteen variables as potentially relevant explanatory variables for the realized variance in state rankings. The data is obtained from Laffer, et al (Laffer, Moore and Williams 2016). Each of the variables are policy variables – presumably in control of State elected officials and administrators.

We use Random Forests to extract the importance of variables in explaining the performance metric. Operationally, a large number of unpruned trees is constructed. A random sample of predictors is taken before each node is split and classification turns on the majority vote of the full set of trees (Kuhn and Johnson 2013).

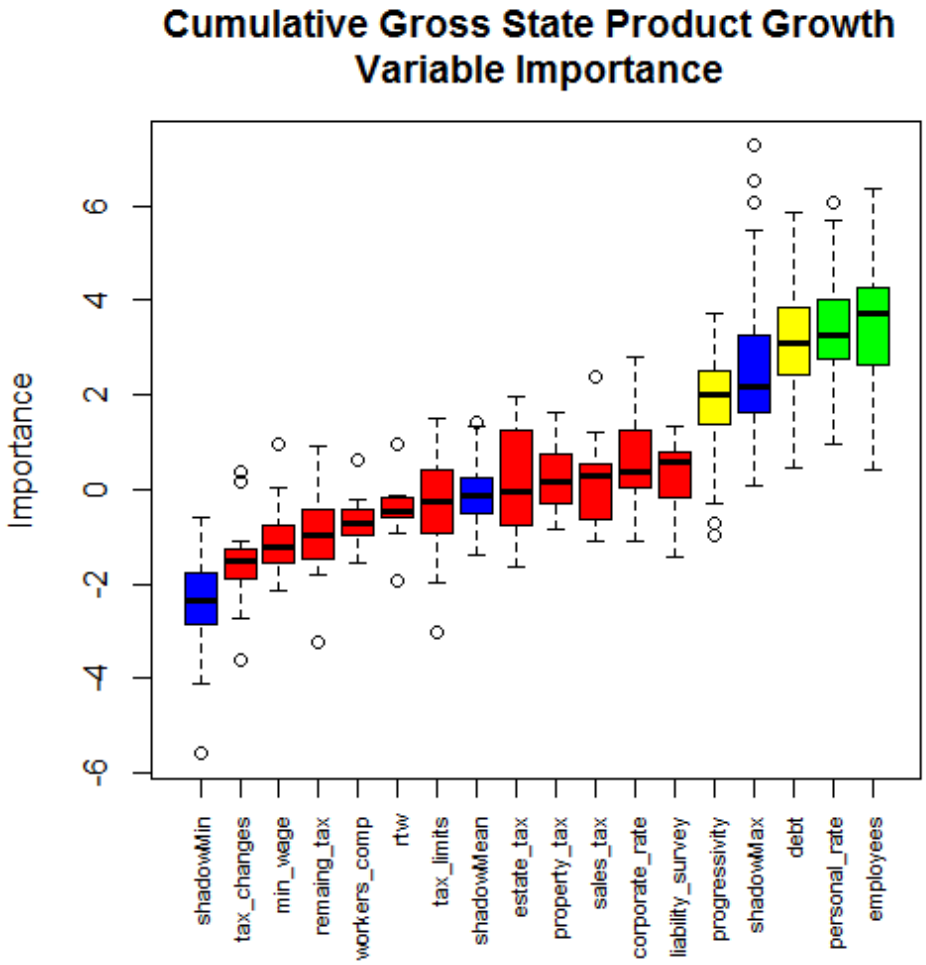
Random Forests is a machine learning technique ideally suited for the type of data assembled here. First the variables are mixed-types: binary and numeric. Second, the number of explanatory variable is large compared to the number of observations. Third, the variables are correlated and some highly correlated. Non-independence can affect standard error estimates used to determine statistical significance.

Random Forests is a non-parametric algorithm which requires no distributional assumptions and no explicit model; rather, it infers nonlinearities and interactions from the data. RF's ability to approximate arbitrary functional forms and thus its ability to identify the presence of complex nonlinear relationships accounts for its enhanced performance over conventional models in econometrics. The latter require an explicit specification of the relationship between explanatory and outcome variables.

We use the R packages *random forest, party, boruta*, and *relaimpo* to extract predictor importance and sensitivities. The importance of variables is assessed by their impact on the accuracy of predictions. These packages deploy several procedures to assign to each predictor it's percent contribution to the total variance explained. This allows for a ready assessment of a predictor for the outcome of interest. The results presented below are based on the averaging of the sequential sum-of-squares obtained from all the possible orderings of the predictors (Gromping, 2006). This procedure of identifying the relative contribution to a joint outcome is conceptually a Shapley Value consideration and an application of Shapley Value Regression (Lipovetsky & Conklin, 2001).[5]

5 Lloyd Shapley was concerned with the fair allocation of profits gained collectively by several actors. In effect, he studied how to fairly estimate the importance of each actor to an overall result where each actor varied in their contribution of effort. Shapley passed earlier this year.

Figure 3



The level of government employees accounts for a significant portion of the variation in cumulative GDP growth. The top marginal personal income tax rate contributes significantly as well. Table below presents the two variables deemed important-predictors.

Table 1

Cumulative GDP per Capita

Public Employees per 10,000 of Population
Top Marginal Personal Income Tax Rate

Sensitivity Analysis

How are the identified "most important" variables related to performance? It is possible to obtain a visual display of the marginal effect on the performance variable of the variables identified as most important. Here, for instance, we see the relationship between the top marginal income rate and cumulative gdp growth rate.

Figure 5

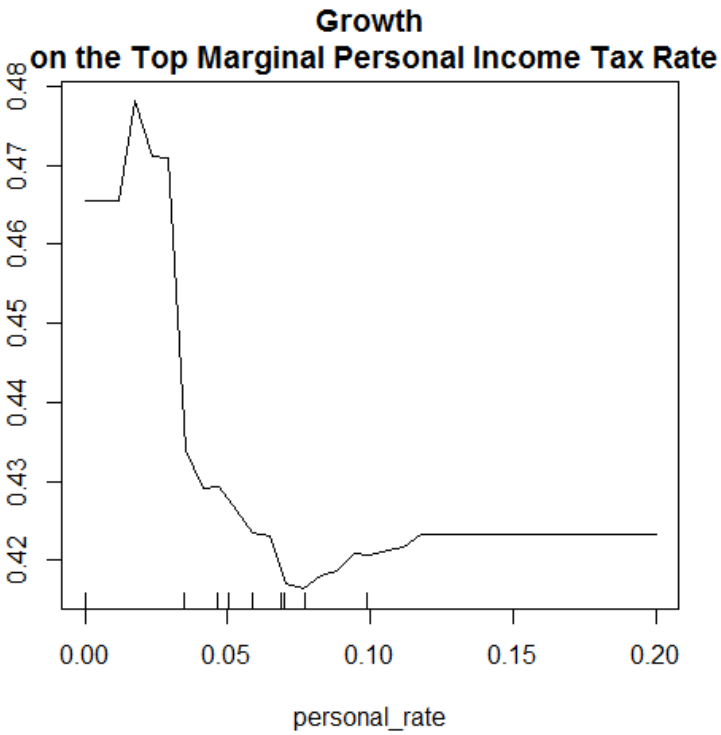
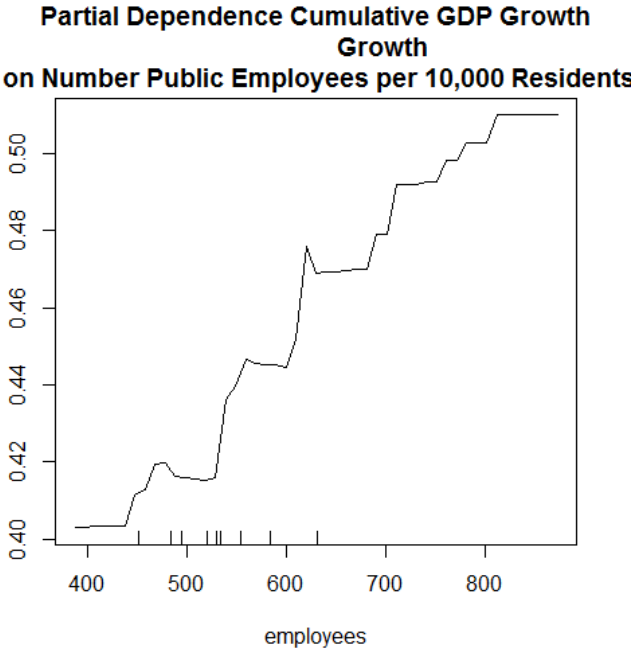
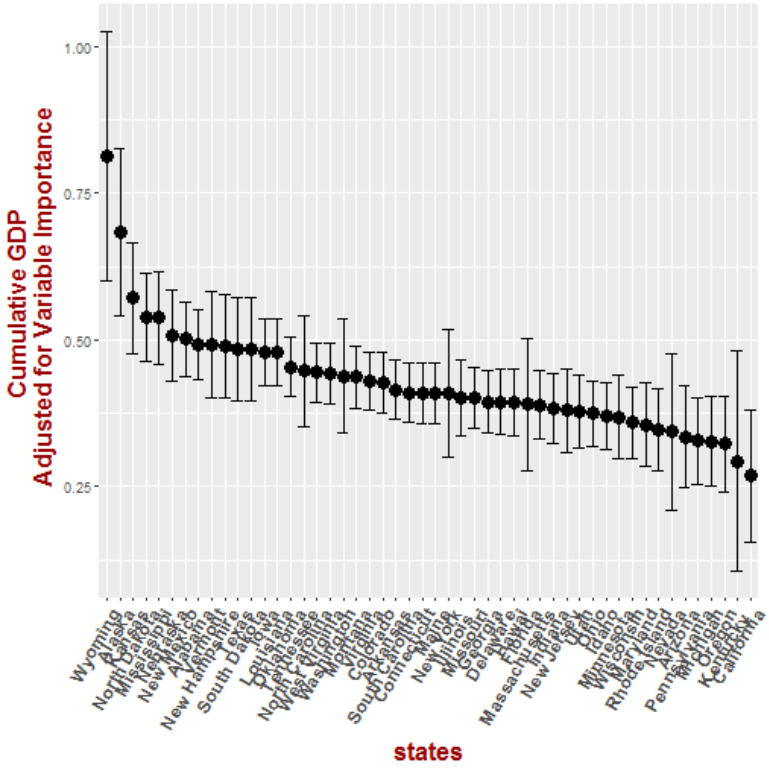


Figure 4



In order to aggregate the identified important drivers into a composite index of performance, we calculate the predicted values of a linear model relating cumulative gdp growth and the two identified drives. The resulting predicted values are used to calculate the relative state rankings. The results are presented below.

Figure 6



Concluding Comments

We rely on random forest R packages to dissemble Cumulative GDP Growth in search of the most influential policy variables. Largely due to the significant multicollinearity of the predictive variables, and the small-sample multiple-variable character of the problem at hand random forests outperforms conventional multiple regression methodology.

At least two important drivers of state economic performance based on cumulative gdp growth are identified. The number of employees in government service and the top marginal income tax rate exert considerable influence on the realized outcomes. The revised rankings based on the predicted scores of the performance variable are likely to be more reliable than the one based on the unadjusted metric.

References

Artz, Georgeanne, Kevin Duncan, Arthur Hall, and Peter F. Orazem. 2016. "Do State Business Climate Indicators Expalin Relative Economic Growth at State Borders?" *Journal of Regional Scienc* 56 (3): 395-419.

Cebula, Richard J. 2014. "The Impact of Economic Freedom and Personal Freedom on Net In-Migration in the U.S.: A State-Level Empirical Analysis, 2000 to 2010." *Journal of Labor Research* 35 (1): 88-103.

Cebula, Richard J., Vikas Agrawal, Robert Boylan, and Robert Horner. 2016. "The Impact of the Net Percentage Growth Rate in the Number of Small Firms on Diferential State-Level Employment Growth In The U.S.: An Explanatory Empirical Note." *Applied Economic Letters* 23 (3).

Delgado, Mercedes, Michael Porter , and Scott Stern. 2012. "Clusters, Convergence, and Economic Performance." *NBER Working Paper No. 18250.*

Florida, Richard. 2002. *The Rise of the Creative Class: And How It's Transforming Work, Leisure , Community and Everyday Life.* New York, NY: Basic.

Gale, William G., Aaron Krupkin, and Kim Rueben. 2015. "The Relationship Between Taxes and Growth At The State Level: New Evidence." *National Tax Journal* 68 (4): 919-942.

Gladwell, Malcolm. 2011. "The Order of Things." *The New Yorker.*

Glaeser, Edward J. 2011. *Triumph of the City: How Our Greatest Invention Makes Us Richer Smarter, Greener, Healthier, and Happier.* New York, NY: Penguin.

Gromping, Ulrike. 2006. "Relative Importance for Linear Regression in R: The Package relaimpo." *Journal of Statistical Software* 17 (1): 1-27.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. 2013. "Who Creates Jobs? Small Versus Large Versus Young." *The Review of Economic and Statistics* 95 (2): 347-361.

Kolko, Jed, David Neumark, and Marisol Cuellar-Mejia. 2013. "What Do Business Climate Indexes Teach Us About State Policy and Economic Growth." *Journal of Regional Science* 53 (2): 220-255.

Kuhn, Max, and Kjell Johnson. 2013. *Applied Predictive Modeling.* New York, NY: Springer.

Laffer, Arthur B., Stephen Moore, and Jonathan Williams. 2016. *Rich Sates, Poor States: ALEC-Laffer State Economic Competitiveness Index.* Arlington, VA: American Legislative Exchange Council.

LeBreton, James M., Robert E. Ployhart, and Robert T. Ladd. 2004. "A Monte Carlo Comparison of Relative Importance Methodologies." *Organizational Research Methods* 7 (3): 258-282.

Lipovetsky, Stan, and Michael Conklin. 2001. "Analysis of regression in game theory approach." *Applied Stochastic Models in Business and Industry* 319-330.

McBride, William. 2012. "What Is the Evidence on Taxes and Growth?" Tax Foundation Special Report No. 207.

Moretti, Enrico. 2012. *The New Geography of Jobs.* New York, NY: Houghton Mifflin Harcourt.

Nardo, Michela, Michaela Saisana, Andrea Saltelli, Stefano Tarantola, Anders Hoffman, and Enrico Giovannini. 2005. *Handbook on Constructing Composite Indicators.* OECD Publishing.

Paruolo, Paolo, Michaela Saisana, and Andrea Saltelli. 2012. "Ratings and Rankings: Voodoo or Science?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176 (3): 609-634.

Saltelli, Andrea. 2007. "Composite Indicators Between Analysis and Advocacy." *Social Indicators Research* 81: 65-77.

Segura III, Jerome. 2016. "The Effect of State and Local Taxes on Economic Growth: A Spatial Dynamic Panel Approach." *Papers in Regional Science.*

Sitglitz, Joseph, Amartrya Sen, and Jean-Paul Fitoussi. 2009. "Report by the Commission on the Measurement of Economic Performance and Social Progress." http://graphics8.nytimes.com/packages/pdf/business/Stiglitzreport.pdf.

Appendix

Table 2
Variable Definitions

Top Marginal Personal Income Tax Rate	The marginal tax rate is the percentage taken from your next dollar of taxable income above a pre-defined income threshold. The marginal tax rate includes federal, state and local income taxes, as well as federal payroll and self-employment taxes.
Top Marginal Corporate Income Tax Rate	The amount of state tax – as a percent - paid by Corporations on the additional dollar of income earned; includes local taxes if any.
Personal Income Tax Progressivity	This measures the difference between the average tax liability per $1000 at incomes of $50,0000 and $150,000. The average tax rate is the total tax paid as a percentage of total income earned.
Property Tax Burden	Tax revenues from property taxes per $1,000 of personal income.
Sales Tax Burden	Tax revenues from sales taxes per $1,000 of personal income.
Remaining Tax Burden	Tax revenues from all taxes per $1,000 of personal income. It excludes personal income, corporate income, property, sales and severance taxes.
Estate/Inheritance Tax Levied?	Yes or No.
Recently Legislated Tax Changes	Relative change in tax burden over the 2014-2015 legislative session.
Debt Service as a Share of Tax Revenue	Interest paid on debt as a percentage of total tax revenue.
Public Employees per 10,000 of Population	Full-time equivalent public employees per 1,000 population.
State Liability System Survey	Quality of state legal system. A ranking of tort systems by state.
State Minimum Wage	State minimum wage, if applicable. Otherwise the federal rate is used.
Average Workers' Compensation Costs	Worker's Compensation Index Rate per $100 of payroll.
Right to Work State?	Yes or No. Whether a state requires union memberships for its employees.
Number of Tax Expenditure Limits	Whether the state has a (i) a state expenditure limit; (ii) mandatory voter approval of tax increases; and (iii) a supermajority requirement for tax increases.

Source: (Laffer, Moore, & Williams, 2016).

Cumulative GDP per Capita	Net Migration	Cumulative Employment Growth
Public Employees per 10,000 of Population	Top Marginal Personal Income Tax Rate	State Liability System Survey
Top Marginal Personal Income Tax Rate	Average Workers' Compensation Costs	Top Marginal Personal Income Tax Rate

Figure 7

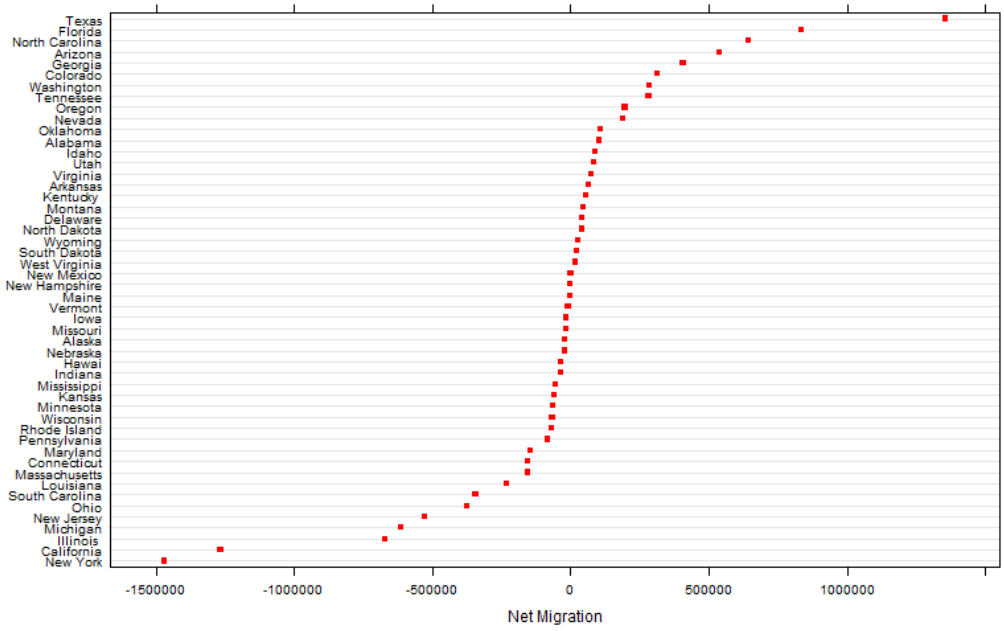


Figure 9

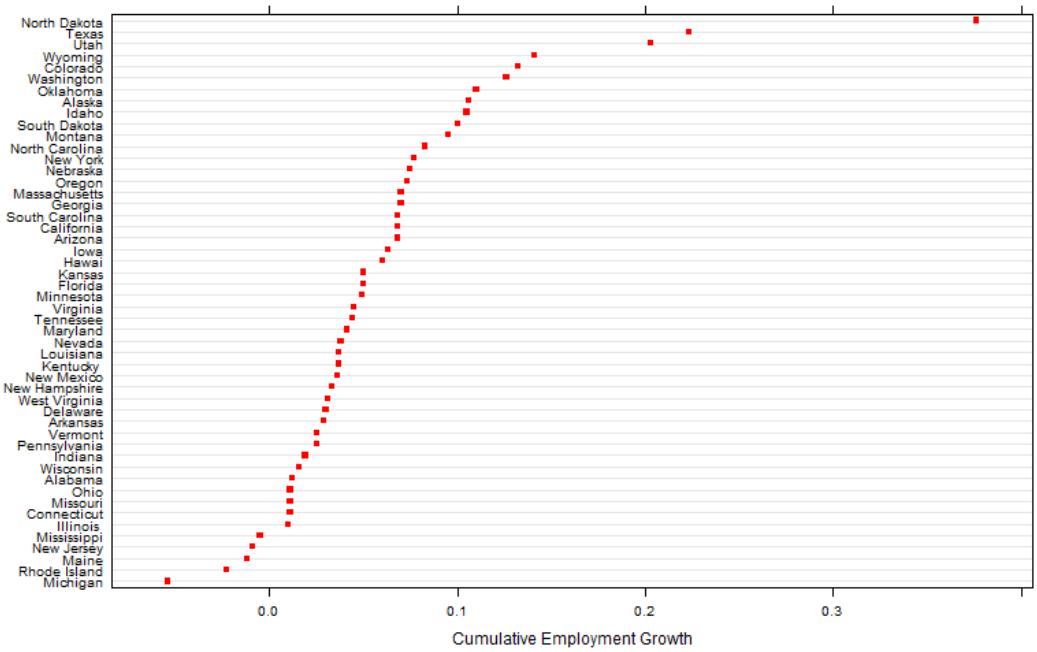


Figure 8

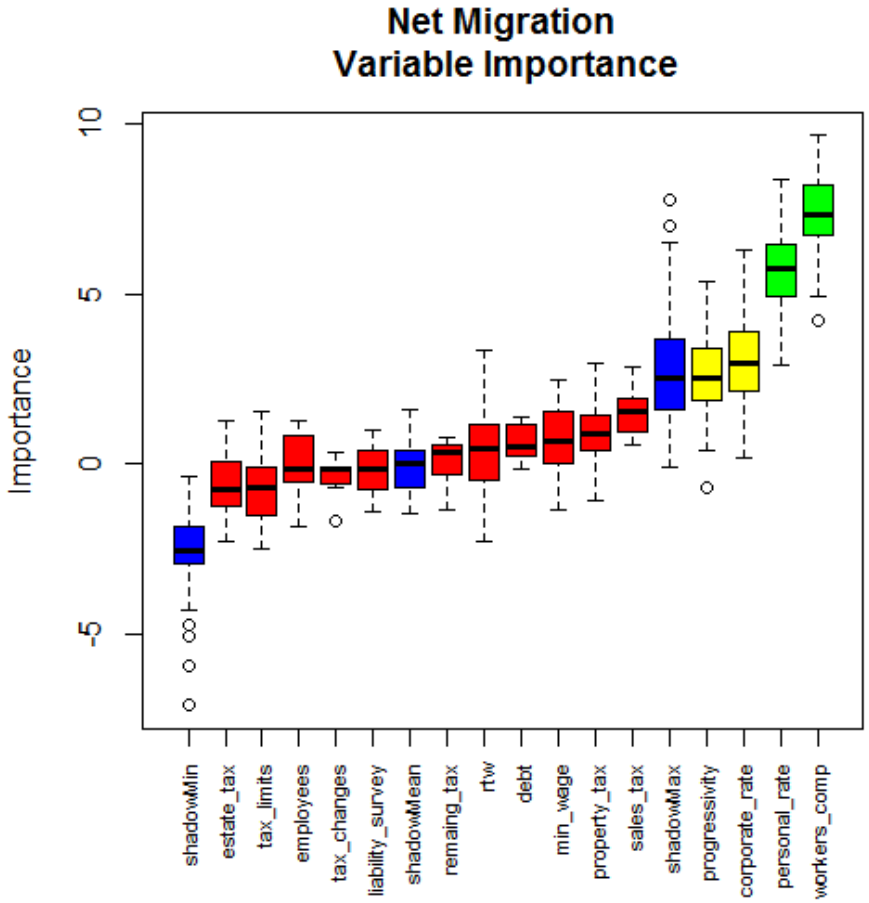


Figure 10

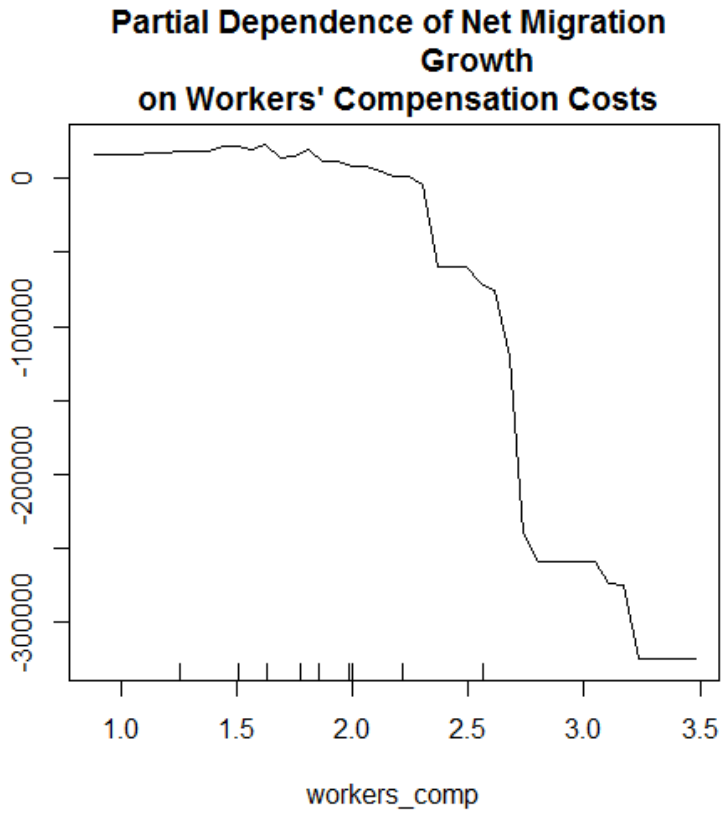
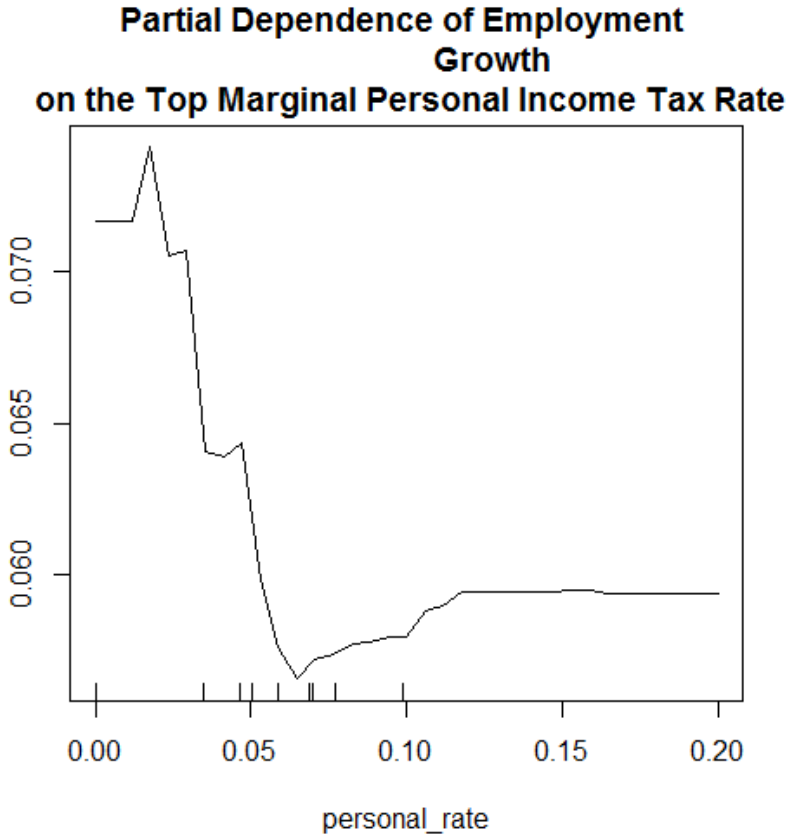
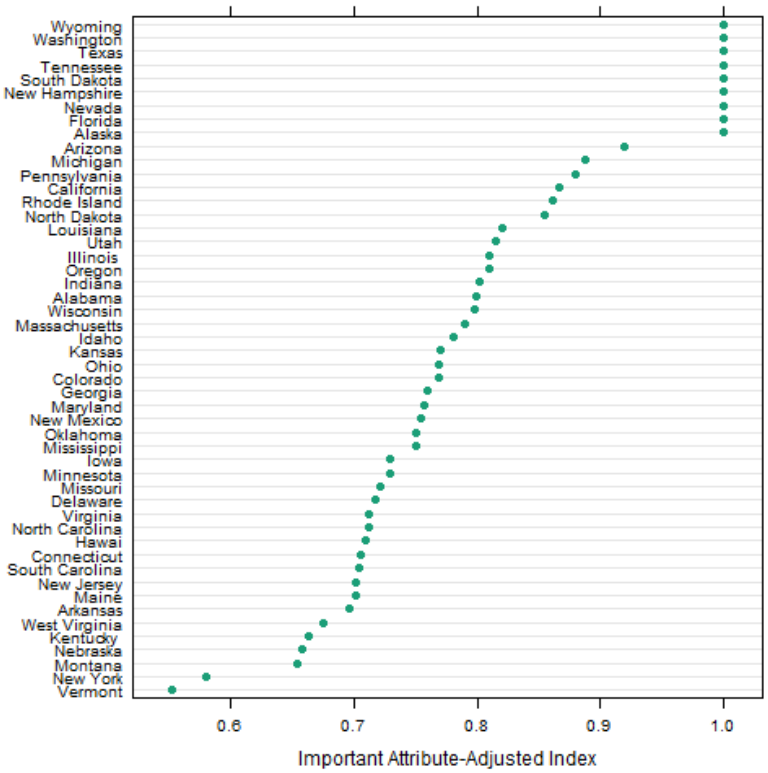
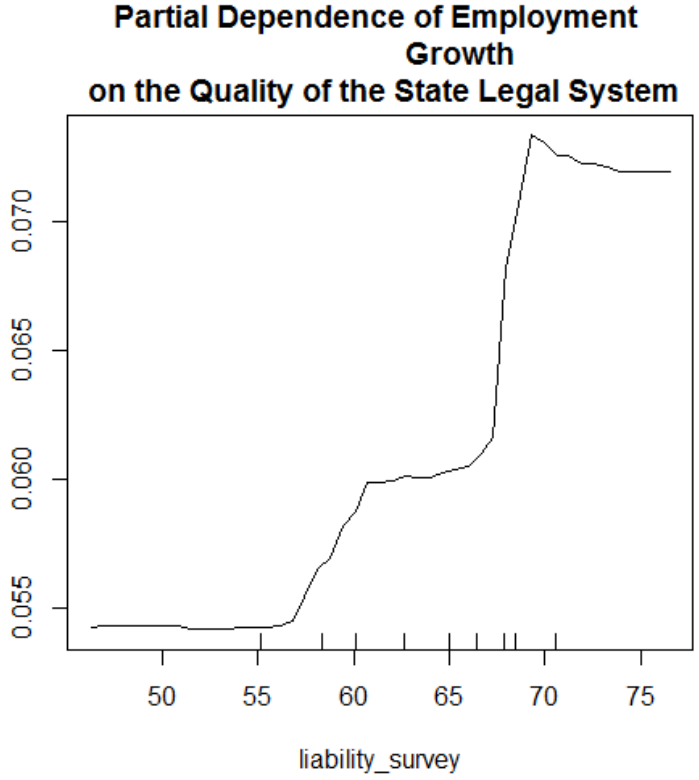

Partial Dependence of Net Migration Growth
on Workers' Compensation Costs
0
-100000
-200000
-300000
1.0
1.5
2.0
2.5
3.0
3.5
workers_comp

Figure 11


Partial Dependence of Employment Growth
on the Top Marginal Personal Income Tax Rate
0.070
0.065
0.060
0.00
0.05
0.10
0.15
0.20
personal_rate


Wyoming
Washington
Texas
Tennessee
South Dakota
New Hampshire
Nevada
Florida
Alaska
Arizona
Michigan
Pennsylvania
California
Rhode Island
North Dakota
Louisiana
Utah
Illinois
Oregon
Indiana
Alabama
Wisconsin
Massachusetts
Idaho
Kansas
Ohio
Colorado
Georgia
Maryland
New Mexico
Oklahoma
Mississippi
Iowa
Minnesota
Missouri
Delaware
Virginia
North Carolina
Hawai
Connecticut
South Carolina
New Jersey
Maine
Arkansas
West Virginia
Kentucky
Nebraska
Montana
New York
Vermont
0.6
0.7
0.8
0.9
1.0
Important Attribute-Adjusted Index


Partial Dependence of Employment Growth
on the Quality of the State Legal System
0.055
0.060
0.065
0.070
50
55
60
65
70
75
liability_survey