PRS Final Paper
Fall 2006
Jeff Diebold
Blake Skogland
Carl Turner

Introduction

The General Assembly is currently addressing the cost of regulations and their impact on the manufacturing sector in Virginia. The Joint Audit and Legislative Review Commission (JLARC) was directed by the Assembly to study this issue to determine the "cost of compliance with federal and Virginia regulation, "compare Virginia's regulations to other mid-Atlantic and southern states," and "compare Virginia's regulations for manufacturers to other business sectors."[1] As a part of this study, JLARC requested that a Public Research Seminar (PRS) team from the public policy graduate program at William and Mary asses and create regression models relevant to the study.

JLARC first charged the PRS team with assessing the regression model found in the article *The Impact of Regulatory Costs on Small Firms*. The model presented estimates the costs of domestic economic regulations be running a cross-county regression analysis bases on data from the Organization for Economic Cooperation and Development (OECD). The team first attempted to replicate the model, then addressed the specifications of the model, followed by a general comparison to other models created to measure the impact of regulation. Finally, the team manipulated the model in various ways to assess how certain variables affect the index.

The second major component of the PRS project was creating and assessing a state model based on GDP predicted by regulatory and other economic and social control

[1] JLARC. Impact of Regulations on Virginia's Manufacturing Sector. Commission Briefing, 2006. http://jlarc.state.va.us.

variables. The team created a state model with variables that related to the ones used in the international model as close as possible. The team included two regressions in the final analysis: the first with manufacturing output as the dependent variable, the second with Gross State Product (GSP) as the dependent variable. Finally the team analyzed the results of the regression and will present its ultimate conclusions on whether the international model provided and the state model created are reliable indicators of the costs and impact of regulation on Virginia's manufacturing sector.

OLS Replication of *The Impact of Regulatory Costs on Small Firm* Model

First, the team used the data provided in the article to replicate the regression results of the international regulation model.

TABLE ONE

Regression Replication

Independent Variable	**GDP per Capita, 1998-2002** Estimated Coefficients (Standard Errors)	**GDP per Capita Growth Rate, 1998-2002** Estimated Coefficients (Standard Errors)
Regulation Index	-1343	-0.01
	(2.13)**	(2.09)*
GDP in Baseline Year	1.07	0
(2001, 1988)	(23.58)*	(1.41)
Language Index	-70.0	0
	(2.97)*	(2.02)
Primary Education	-58.4	0
	(0.91)	(1.54)
% of GDP in Intl Trade	44.7	0
	(2.90)*	(3.08)*
Constant	10,487	0.13

* *= significant at .05

* = significant at .01

The results displayed in Table One show that when the team input the data provided and ran the same regression the results are the same as in the study. The team was able to replicate the regression results of the international model in the study.

International and State Model Specifications/Controls

The model under review utilized several economic and social measures as controls in order to examine the impact of regulation on economic development. This section provides theoretical justifications for those measures and their equivalent explanatory variables in the state model, in addition to the other controls included in that model. GDP per capita, 1998-2002 is the dependent variable used in the international model. Average total manufacturing share of GSP per state 1997-200 and state GSP are the dependent variables used in the state model. The regulation indexes are the key independent variables in all of the models, because they are estimations for regulation in a country or state.

Foreign Trade as Percent of GDP

Foreign trade as a contributor to economic growth has a long history in economic theory and in the empirical literature.[2] Economists think that international trade increases the welfare of all the participating countries, because each country is able to specialize in the industry or group of industries that they perform most efficiently. Costs for all tradable products are thus likely to fall, and workers' wages would presumably rise due to participation in productive activities, raising the real income of both consumers and workers on average within the trading group. In addition to comparative advantage,

[2] For discussion of the observed effects of trade on development, see Dajin Li (1996), Rivera-Batiz and Romer (1991a), and Rivera-Batiz and Xie (1993)

countries can realize greater efficiency due to economies of scale; producers can supply a larger market, the average cost of their product will drop because of the ability to increase production at a given fixed cost. Trade also facilitates a diffusion of knowledge and technology that may result in greater productivity in countries that previously lacked skills and equipment. To the degree that countries allow various levels of trade through policies such as tariffs, quotas, and subsidies that affect the integration of their economies with other countries, they are also likely to see varying levels of these benefits; this makes foreign trade as a percentage of GDP a valid proxy for level of integration.

This variable only has a statistically significant effect on GDP growth at 5% level; exclusion from model has small effect on coefficients. No good national level data is available on state trade with other states or within the national economy. The state model uses the mean of merchandise exports from 2001 to 2005[3] to an international market in dollars as a percentage of the gross state product of the state from which the manufacturing products originated as a proxy for economic integration.[4]

Primary School Enrollment

The degree to which a country's citizens are educated also has a theoretical effect on a country's development.[5] The human capital model of education holds that education, through imparting knowledge such as reading, writing, and specialized skills, makes individuals more productive once they begin work. Even the alternative to the human capital mode, the sort and select model, would presumably hold that education raises average productivity because performance in schools can be used by employers to select job candidates who would perform best in a given position.

---

[3] Data from Northeast Midwest Institute
[4] For survey on link between exports and growth, see Charles W. Hultman (1967)
[5] For discussion on the role of education and economic growth, see Maureen Woodhall (1967)

The international model uses primary school enrollment rates among children to capture this education effect. Countries with high primary school enrollment would be expected to have high per capita income. As discussed below, in the original model, primary school enrollment was shown to have no statistically significant relationship with either per capita GDP or economic growth; some potential reasons for this include the costs that come with education, both the fiscal and the opportunity costs of keeping children out of the workplace. However, because it was not significant in the original model, no equivalent was included in the state model.

Ethno-Linguistic Diversity

The ethnic diversity of a country can have an effect on economic growth.[6] Minority populations might be subject to workplace discrimination or a general stigma that makes them less productive, lowering the average GDP per capita within in a country. Ethnically fractured countries might also have higher transaction costs due to language and cultural barriers that prevent economic integration. Thus, ethnically diverse countries will likely be poorer on average than similar, but more homogenous countries.

Excluding this variable from the model makes regulation index statistically insignificant and has little effect on other variables. As a state equivalent, the percentage of the state's population that is nonwhite is used. Given that nearly every state has a white plurality, the size of the non-white minority population is likely to capture any stigma effects that might exist.

---

[6] For a discussion of the effects of ethnic diversity on economic growth in Africa, see Easterly and Levine (1997)

GDP per Capita 1988

Historical events such as wars, timing of industrialization, cultural differences and other factors that might influence economic development would be best captured through a series of dummy variables. However, given the small sample size in both the original, international model and the state model, the loss of degrees of freedom is a concern. GDP per capita in a year prior to the years measured is a good way to capture these historical factors because said effects were likely to have already been observable in that earlier period. Once controlled for, the coefficient for other explanatory variables are more likely to be consistent and unbiased. For consistency in the state model, state GSP in 1988 was used for this control.

Citizen Ideology

While levels of regulation might have an effect on economic growth, regulation policy is not randomly distributed throughout the states; it is the product of the actions of elected policy makers who presumably reflect the attitudes and thinking of their constituents. If these attitudes were reflected in areas outside of government action, perhaps through a more conservative work ethic or a more professional populace in urban, liberal areas, then both Economic Freedom and regulation indexes would be biased. To account for this, an index of citizen ideology,[7] based on the ideological rating and vote share of a state's Congressional representatives was included; the index took on a range of values from zero to one hundred, with zero being the most conservative state, and one hundred the most liberal.

Logarithmic Transformation

[7] For details on the methodology behind this index, see Berry, Ringquist, et al. (1998)

Analysts employ the use of logarithmic transformations to allow for a nonlinear relationship between the dependent and independent variables in a regression model. In this case, we have taken the log of the dependent variable measuring a country's gross domestic product. This transformation can reduce the range in a data set and make the dependent variable less sensitive to outliers. It is often used when there are a large number of possible positive values, in this case, GDP per capita. In this data set, the range in GDP per capita is to $2947 (Turkey) to $46,330 (Switzerland) and this transformation will control for heteroscedasticity, or larger variances in the dependent variable for larger values of the right-hand side variables.

The potential for outliers in this data set is relatively small, despite the small sample size because the dependent variable, GDP, is calculated per capita. Thus, output is expressed as the amount of output generated per person in the countries included in the sample. However, outliers could still exist if one or more of the countries were markedly more or less industrialized or technologically advanced relative to the other countries in the sample. Technology and the factors of production are important factors often used to determine production levels of a country and its citizens. Given the countries included in this sample, it is unlikely that outliers do exist.

While there is variation within the sample, the data set is rather homogeneous considering the variation in GDP for all countries. For the most part, all of the countries in this sample come from relatively developed nations with stable governments and economies. Over seventy-five percent have a GDP per capita over $17,500. Having more developed countries in the regression model is important for isolating the effects of regulation on GDP, but also obviates need for a logarithmic transformation. There is not variability so large between the countries in the sample that it requires a logarithmic transformation to stabilize it or to make the within-group variation more similar as the sample is already somewhat homogenous.

As Table One reveals, the model loses its ability to predict the relationships of the independent variables with GDP per capita. The estimated coefficients reported are the semi-elasticities of the variables in the model. They are the estimated percent change in GDP per capita given a unit change in the independent variables.

TABLE TWO

Log Transformation

Variable	Estimated Coefficients (Standard Errors)
Regulation Index	-.152 (.098)
GDP in 1988	.0001** (7.02e^-6)
Language Index	.005 (.004)
Primary Education	-.014 (.013)
% of GDP in Intl Trade	.004* (.002)

* *= significant at .05

* = significant at .01

After the transformation, only two variables remain significant. Most importantly, the coefficient on the regulation index is no longer statistically significant. When we allow this model to assume nonlinear relationships, it loses the significance that existed in the linear model. However, there is little theoretical reason to carryout this transformation. If heteroscedasticity were present in the original regression model, it is unlikely that there would have been any significant coefficients, given an already small sample size. Additionally, the data set is roughly homogenous thus there is little need to control for within-group variation through a logarithmic transformation. Thus, these insignificant results should raise any suspicion about the estimates in the original regression model.

GDP Changes

After verifying that the model used in the analysis was replicable, the next test of this model was for robustness of the estimates. To test for this, two additional regressions of the same model were run using updated information and data for the dependent variable GDP but not the independent variables.

Admittedly, there are some shortcomings in the analysis by not updating the right-hand side variables despite the fact that the controls were generally averages or static population measures. This concern is specific to the variable measuring the proportion of GDP resulting from international trade. It is possible that this statistic could vary from year to year, given economic and political climates within a particular country.

The averages for proportion of GDP attributable to international trade are capable of mitigating aberrational fluctuations in the data due to economic or political climates. However, not being able update this data along with GDP could compromise the estimates if the proportion of international trade in the countries used in the is analysis changed dramatically during the time periods for which the GDP variable was updated and the international trade measure was not, in this case 1988-1998 and 2001-2005.[8]

While trade may fluctuate from year to year, the other variables are relatively fixed, or change so slowly that holding them constant over the time-periods analyzed should not generate the same concerns as the variable measuring international trade. It is

[8] GDP in all updated years was adjusted for 2005 dollars.

unlikely that education levels and the ethno linguistic diversity of the populations changes so dramatically as to warrant concern over holding them constant.[9]

Updating and appending data for the years 2001-2005 and 1988-2005 illustrates that the estimates for the variables in this model were not robust over time. The regulation index remained significant in the 2001-2005 dataset, only this time, at the .10 level. All other variables, excluding the baseline control variable, lost their statistical significance. The results for the 1988-2005 indicate the effect of the regulation index over this period is not significantly different from zero.

TABLE THREE
Updated Regression with 2001-2005 Data
Compared With 1998-2001 Data

Variable	**2001-2005** Estimated Coefficients (Standard Errors)	**1998-2001** Estimated Coefficients (Standard Errors)
Regulation Index	-456.8*	-1343**
	(221)	(2.13)
GDP in Baseline Year	1.02**	1.07**
(2001, 1988)	(.023)	(.045)
Language Index	-4.33	-70.0**
	(7.91)	(23.3)
Primary Education	-36.4	-58.4
	(29.7)	(86.1)
% of GDP in Intl Trade	7.05	44.7**
	(4.95)	(14.0)

* *= significant at .05
* = significant at .01

[9] All data that was not updated was irretrievable without membership access to the World Bank, otherwise the information was adjusted for every year of the study.

TABLE FOUR

Updated Regression with 1988-2005 Data

Compared With 1998-2001 Data

Variable	**1988-2005** Estimated Coefficients (Standard Errors)	**1998-2001** Estimated Coefficients (Standard Errors)
Regulation Index	69.5	-1343**
	(334)	(2.13)
GDP in Baseline Year	.810**	1.07**
(2001)	(.034)	(.045)
Language Index	19.9	-70.0**
	(11.9)	(23.3)
Primary Education	-6.05	-58.4
	(44.8)	(86.1)
% of GDP in Intl Trade	-12.9	44.7**
	(7.46)	(14.0)

* *= significant at .05

* = significant at .01

These results could indicate that there was something particular to the time-period used in the original model, 1998-2001 that the regulation index was able to capture. During this time, the world economy was expanding and levels of productivity were high. It is therefore intuitive that during an expansion, higher regulations, as indicated by a higher regulation index, could act as an impediment to the size of a country's GDP. Conversely, during an economic downturn, when countries are generally importing and exporting less, the effect of on regulations on GDP could be less consequential. Averaging for many expansions and contractions over such a long period time, 1988-2005, would not isolate this effect were it to exist. If this process is what the original model is actually capturing, then it is reasonable that during this time, 1998-2001, regulations would have more explanatory power than in 2001-2005 or 1988-2005.

State Model

The state models we created use the average total manufacturing share of GSP per state from 1997-2005 or state GSP as the dependent variable. The independent variables used are the Economic Freedom Index *or* Regulation Index, state population in 2005, citizen ideology, percent minority population, total manufacturing share of GSP per state 1988, and international merchandise exports per state 2001-2005. Explanations for the inclusion of each variable are given in the model specifications section above, with the exception of the indexes used to measure regulation.

The Pacific Research Institute Economic Freedom Index is used as a proxy for "Economic Freedom" in each state. The index was created by compiling a list of relevant indicator variables for economic freedom. These indicators were then converted into a number of indexes using various techniques. The indexes created are fiscal, regulatory, judicial, government-size, and welfare spending. The indexes were then compared to each other in terms of their ability to explain human migration, other things the same. The best composite index was then used to rank the U.S. states in terms of economic freedom. This index is used in the first state model regression.

The second index used is the regulation index component of the Economic Freedom index. This is a useful proxy for regulation within the states. The regulation index is comprised of 53 variables based on: mandatory regulations on labor, education, and the environment and regulations that preserve and extend economic freedom such as right-to-work laws. Some examples of the variables used to make this index are licensing requirements, continuing education requirements, preference laws affecting public procurement, state purchases, right to work, minimum wage and prevailing wage laws, worker compensation and insurance laws, gun laws, and vehicle laws.

State Model – Manufacturing Share

The first state model uses the average total manufacturing share of GSP per state from 1997-2005 as the dependent variable. Using this dependent variable the Economic Freedom Index has a significant, negative effect on manufacturing output. The second model replaces the Economic Freedom Index with the component Regulation Index, which does not have significant effect on manufacturing.

TABLE FIVE

State Regression Results - Manufacturing Output		
	Economic Freedom Index	***Regulation Index***
	OLS	**OLS**
Variable	Estimated Coefficients	Estimated Coefficients
	(Standard Errors)	(Standard Errors)
Regulation Index	-584.83**	295.30
	(256)	(275)
GDP in Baseline Year	.001***	-.001
(manufacturing share)	(.000)	(.000)
Minority	-4.43	-23.0
	(70.440)	(73.3)
Citizen Ideology Index	-61.59	-156.5
	(67.35)	(61)
Merchandise Exports in 1000s	.000***	.000***
	(000)	(.000)
State Population 2005	-.001	-.001
	(.000)	(.000)
***= signif at .01		
* *= signif at .05		
* = signif at .1		

State Model – GSP

The third state model uses state GSP as the dependent variable. Using this dependent variable the Economic Freedom Index has a significant, negative effect on manufacturing output. The fourth model replaces the Economic Freedom Index with the component Regulation Index, which does not have significant effect on manufacturing.

TABLE SIX

State Regression Results - GSP		
	Economic Freedom Index	***Regulation Index***
	OLS	**OLS**
Variable	Estimated Coefficients	Estimated Coefficients
	(Standard Errors)	(Standard Errors)
Regulation Index	2205725	-811981
	(1408079)	(1482723)
GDP in Baseline Year	-.000	-.000
(manufacturing share)	(.001)	(.001)
Minority	-121469	-54095
	(387688)	(395363)
Citizen Ideology Index	537342	881828
	(370686)	(328886)
Merchandise Exports in 1000s	-.000	-.000
	(001)	(.001)
State Population 2005	44.03**	44.17**
	(2.45)	(2.54)
***= signif at .01		
* *= signif at .05		
* = signif at .1		

Two Stage Least Squares Estimation

A concern with any simple linear model is that the coefficient estimates may suffer from omitted variable bias. A variable with explanatory power that is omitted from a model due to a lack of data or perhaps to a lack of consideration could also be correlated with one or more of the included variables. If so, the resulting coefficient will be biased because the effect of the omitted variable on the dependent variable will be attributed the explanatory variable correlated with the omitted variable.

Biased estimates are a possibility in our model. The regulation index could be correlated with unaccounted factors such as geography and climate that expose a state to catastrophic risk.[10] In addition, whether a state specializes in a particularly risky industry such as coal mining might effect the state's inclination to regulate private industry. Another example might involve politically sensitive goods such as energy, which has an independent effect on economic growth.[11] For example, legislatures that respond to calls for lower energy prices might choose to regulate, but even in absence of that regulation, high-energy prices would weaken economic growth.

One solution to this is to attempt to include these variables so that our original explanatory variable coefficients are controlled for them. However, if there are a sufficiently large number of such variables, we will exhaust our limited degrees of freedom. This is especially the case if one category of determinant, such as the health risks of a state industry, can only be represented categorically, i.e. as a series of dummy variables. It is also implausible that any simple OLS estimated model would ever capture every potential explanatory variable, both because of faults within the data and the cost of finding such data.

A more viable and common technique to deal with the problems of omitted variables is to employ two-stage least squares estimation. This process involves the use of an instrument: a variable that is associated with the main independent variable, but is

---

[10] Noll, Roger G. 'Reforming Risk Regulation,' *Annals of the American Academy of Political and Social Science* (1996)

[11] Gray, Wayne B. 'The Cost of Regulation: OSHA, EPA, and the Productivity Slowdown,' *American Economic Review* (1987)

not associated with unobserved explanatory variables and has no independent effect on the dependent variable. The instrument is used, along with all included control variables, to create a new set of estimates for the relevant explanatory variable for each observation. These estimates are then used in the main model in lieu of the actual observed values to estimate coefficient for the model of the dependent variable. Because the estimates of the explanatory variable are created using independent variables not correlated with the unobserved variables, the coefficients in the main equations theoretically correct for bias and inconsistency. However, the assumptions of exogeneity on the part of the instrument, especially in regards to a lack of correlation with unobserved variables, are critical to receiving consistent estimates.

In the case of our gross domestic state product model, where we believe that the regulation coefficient might be biased due to a lack of added controls, the instrument is the political ideology of the state government. The ideology of the members of a state government can be expected to affect a range of policies, including regulation; more conservative than average governments can be expected to favor less regulation and liberal governments more, all other factors being equal. Because government officials are elected in each state, one can expect state ideology to be correlated with the ideology of a state's citizens. Citizen ideology in turn is likely correlated with a range of economic, cultural, and institutional factors that are unobserved but effect economic outcomes; this could potentially violate the violate the assumption of exogeneity by instrument that is necessary for the estimated coefficients to be consistent. To correct for this, a measure of citizen ideology is used in both the substantive and instrumental model.

As a proxy for this ideology, we use the Government and Citizen Ideology Indexes developed by William D. Berry, et al.[12] The government index rates the ideology of elected members of state government by averaging ideology scores of members of each party serving in the state's congressional delegation. The index assigns that average to each member of the state government of the same party, and then creates an average score for the entire state government, weighted according to the power of the particular position held by particular members (governor versus a simple member of the state legislature). The index ranges from zero to one hundred, with zero representing the most conservative states, and one hundred the most liberal. The Citizen Ideology Index estimates by averaging the ideology of both Congressional Incumbents and their challenger, weighted by vote share, using the same scale.

None of the resulting 2SLS estimated coefficients are statistically significant, with each coefficient dropping toward zero relative to the OLS coefficients in both the Economic Freedom Index and more narrow Regulation Index specifications. This could be interpreted as a result consistent with omitted variable bias due to a failure to take into account ideology, but it is more likely that a case of multicollinearity, probably between the citizen ideology variable and the other explanatory and control variables. Citizen ideology is expected to correspond heavily with levels of manufacturing, because of the presence of unions or unsafe work environments (and thus manufacturing exports), the diversity of the state (certain ethnic groups are consistently more likely to vote for a more liberal or conservative candidate), and the baseline GSP (which was taken from within the range of years which provided the data to calculate the ideology index). Given that

[12] Berry, Ringquist, Fording, Hanson, 'Measuring Citizen and Government Ideology in the American States, 1960-93,' *American Journal of Political Science* (1998)

the ideology variables are heavily correlated with the other explanatory variables, they essentially split their effect on state product, resulting in weak, insignificant estimated coefficients.

TABLE SEVEN
2SLS Estimated Coefficients

Variable	**Economic Freedom Index** Estimated Coefficients (Standard Errors)	**Regulation Index** Estimated Coefficients (Standard Errors)
Regulation Index	2455.78 (4621)	3841.84 (8058)
GDP in Baseline Year (manufacturing sector)	.970 (.621)	1.16 (.346)
Minority	-88.7 (193)	-50.29 (173)
Citizen Ideology Index	-486.0 (655)	-316.5 (387)
Merchandise Exports in 1000s	.004 (.000)	.0004 (.0003)
State Population 2005	-.001 (.001)	-.0015 (.002)

Conclusion

Based on regression results, we cannot make solid conclusions about the effect of state regulations on manufacturing output or aggregate GSP. The state should look to other factors for effects on manufacturing or aggregate GSP, as evidenced by inclusion of the Economic Freedom Index in the state model. We were not able to find statistical significance due to small sample sizes and potential omitted variables left in the error team. A better way to assess the effect of regulation would be through panel data and a fixed effects model. This would allow us to eliminate the state fixed effects that plague the data set available. While, fixed effects can result in the loss of degrees of freedom, it would be a better approach to measure the within variation of the states in our model.

Thus, more research is necessary to extrapolate the effects of regulation on a state's manufacturing sector.

As for the international model, when the data is updated to include recent information on GDP, the relationship becomes less significant or insignificant. This might indicate a degree of caution be considered when to extrapolate meaning from the relationships predicted in the original model. Although GDP was the only updated variable many of the other measures included in the original model this should not affect the results of the subsequent regressions because many of the independent variables in this model are static such as education and ethno linguistic diversity. The only variable that might change based on the domestic and international market conditions is the variable measuring the proportion of the GDP attributable to international trade. While this is a shortcoming in the updated model and could skew the results in the data that includes information from 1988 to 2005, it should not account for the loss of significance in the period used in the original model. Because the significance of the model does not hold over time its results should be used and interpreted with a degree of caution and an understanding of the caveats potentially plaguing this model.