



Yep, Stuart's right about GSP/Cap...VMT <i>does</i> matter, even multivariately...

Posted by [Prof. Goose](#) on January 18, 2006 - 9:58pm

Topic: [Demand/Consumption](#)

Tags: [gdp](#), [hubbert peak](#), [oil prices](#), [peak oil](#), [us states](#), [vmt](#) [list all tags]

Stuart and I have been emailing back and forth regarding [his post on modeling state gsp using vmt](#) as an independent variable.

Paula, [both here and on her blog](#), correctly suggested that education should also be considered as an independent variable. So, what I did was pull together [some data on % college educated for each state](#) and include it in with the data Stuart had already collected...and then I conducted a multivariate regression on the data, which is presented below. That regression allows us to find out what the effects are for all of these independent variables on gsp/cap *after controlling for the effects of the other variables*. This allows to get a better picture of what's going on (though we lose the visual facility that Stuart had with his bivariate graphs). Much more under the fold.

So, what's the takeaway? Stuart's right: states with higher vmts have lower gsp/cap, even after controlling for education and population density.

It should also be noted that I also agreed with much of the criticism in the comments regarding the logarithmic transforms that need to be done on the variables...however, if after taking the ln of each variable, the comparative magnitude of the coefficients remains present, then we can go back to the pre-transformed data and make easier inferences using the coefficients that are present.

So, the dependent variable is gsp/cap, the independent variables are vmt/cap, population density, and education. The unit of analysis is state, (however we are dropping DE and WY for reasons mentioned earlier.) (fyi: including them in the analysis weakens the case a bit for VMT, but the education result is still present).

Here's the multivariate regression (using robust standard errors, just because I like overkill) results (using Stata):

```
gspcap=educ+popden+vmt cap
```

```
. regress gspcap educ popden vmtcap, robust beta plus
```

```
Regression with robust standard errors  
48
```

```
38.49
```

```
0.0000
```

```
0.6664
```

```
Number of obs =
```

```
F( 3, 44) =
```

```
Prob > F =
```

```
R-squared =
```

3555.4

```

-----
--
--
Beta      |          Coef.   Robust
          |          Std. Err.   t      P>|t|
-----+-----
--
educ      |    659.0628    107.6447    6.12   0.000
.5328718
popdens   |   -1.443011    7.739639   -0.19   0.853
.0235832
vmtcap    |   -1.79709    .5305359   -3.39   0.001
.4813086
_cons     |   38455.69    7496.157    5.13   0.000
.
-----+-----
--

```

what does this gobbledygook mean? well, each coefficient is the change in the dependent variable resulting from a one unit change in that independent variable, controlling for the other variables present in the equation. So, for a one percentage change in percent education of a state, gsp/cap goes up \$659. A one unit change in vmtcap results in a -1.79 unit change in gspcap *even after controlling for education*. The effects for education and vmtcap are statistically significant at p<.001 or greater. (I can explain that more if you all want me to).

The final column out there is what we call a "standardized beta." It's not the best measure of strength around (trust me, I could explain it, but you don't want me to), but it is an indicator of "standardized explanatory strength." We cannot directly compare the regression coefficients' magnitudes, but betas can help us do that (caveat: somewhat). Because the magnitude of the betas are relatively the same, though in different directions (vmt is an inverse relationship, educ a direct relationship), we can say that these two variables have a relatively similar explanatory power.

Now, the question of the logarithmic transforms. Regression is really robust, but it has some assumptions...one of them is that these variables are normally distributed. They aren't. The solution is to attempt to make them more normal by transforming them using some standard...in this case, using natural logs of those variables that are "large," gspcap, vmtcap, and popdens, because it makes them more normal in their distribution. This is a pretty standard trick. Here's the results:

```

. regress lngspcap educ lnpopden lnvmtcap, robust beta plus

Regression with robust standard errors          Number of obs =
48                                              F( 3, 44) =
36.49                                          Prob > F    =
0.0000                                          R-squared   =
0.6753                                          Root MSE   =
.09517
-----
--

```

Robust

	Coef.	Std. Err.	t	P> t	Beta
educ	.0188446	.0030129	6.25	0.000	.5615207
lnpopden	-.0079779	.0159703	-0.50	0.620	.0678694
lnvmtcap	-.4890024	.121637	-4.02	0.000	.4718397
_cons	14.55413	1.181904	12.31	0.000	.

no changes, except in the raw coefficients. The significance levels stay the same, the betas stay relatively similar. So, we can say with some confidence that the transforms aren't all that necessary.

It should be noted that, because of the transforms we did, we would have to reinterpret (exponentiate them and the dv over e) the coefficients in order to make direct inferences about prediction (such as a rise of 1% in a state's education = \$659 in gspcap that I discussed above). However, even if we change them back, we get substantively similiar results.

So, what's the takeaway? Stuart's right. States with higher vmts have lower gsp/cap, even after controlling for education and population density.



This work is licensed under a [Creative Commons Attribution-Share Alike 3.0 United States License](https://creativecommons.org/licenses/by-sa/3.0/).