

Vermont Economic Stimulus Draft Analysis

Abstract:

In order to increase economic activity in Vermont, I propose linking an industry that is failing in the state, while thriving in the nation as a whole, to a larger industry that is seeing strong growth in both the state and the nation. By fostering a relationship between the industries, I hope that the success of one will lead to the recovery of the other.

Paul Brennan
Professor John Coles
CIS 512 – Intro. To Data Science and Analytics
November 24th, 2016

<https://github.com/brennap/cis-512>

In this analysis, I hope to better formulate a plan for economic stimulus of a state, in particular Vermont. Admittedly, this state was chosen rather arbitrarily, although it does possess some notable advantages, such as its small size and proximity to large economic centers like New York and Boston. The size should mean that an investment will go further, and the proximity to those centers equates to a potential source of capital, market, and labor. Regardless of these advantages, the approach for this project began by browsing through economic data for the State of Vermont. It was noticed that one industry, with the description “Data processing, internet publishing, and other information services” had seen a drastic decline over the past decade, with its portion of the GDP falling from \$103M to \$46M. This seemed strange, as the information fields are generally seen as high growth. Later analysis would show that this hunch was correct, as the sector had seen positive growth nationwide. Considering that information services are ubiquitous across disciplines, it occurred to me that the industry could be stimulated by fostering synergy with a larger, more successful industry. Thus, the analysis focuses on finding a candidate industry to target, with some focus on understanding the overall economic climate of Vermont and of the identified information sector.

Most of the data for this project was obtained from the Bureau of Economic Analysis, through their Data API. It was necessary to obtain an API key to issue queries. Data was gathered using a Python script, the chosen language due to its proficiency in text analysis and data structure processing. Several tables of interest were identified, queried through the REST API as JSON, and formatted into CSVs. The CSVs were then read into R for further processing and analysis. R was chosen due to its ability to quickly analyze and graph data. The primary data

used was GDP breakdown by Industry for both state and national, as well as value added by industry nationwide, income per county, and employment rates.

The analysis in R began by creating linear regressions for GDP portion by Year per Industry, for both the state and the nation. Additional data cleaning was needed at this step as the BEA data had used industry codes in the state data that did not conform to the NIACS codes used in the national data. By looking at the regressions for each industry, we could get an idea of which were generally increasing their production, and which were decreasing. The top ten highest growth industries were selected from both the state and national set. It was important that our targeted industry saw strong growth in both sets, as that not only ensures high growth opportunity from our project, but also because, as we have seen with the “Information” sector we wish to recover, strong national growth is not necessarily connected to strong growth in state. Therefore, from the two top ten lists, I further narrowed the scope to those industries appearing in both lists. That left me with three industry sectors: “Real estate”, “Management of companies and enterprises”, and “Health care and social assistance”. Out of those sectors, real estate was by far the largest sector by GDP portion. Health care saw the highest growth nationwide, although it lagged behind real estate’s growth in Vermont. However, of when looking at the models, health care had the best fit, with an R value over 0.9. This suggests to me that the growth in health care may be more stable and predictable than that of real estate. There is a question of whether growth should be measured in percentage change or in dollar amounts. Although percentage change is arguably a better method for modeling growth, I believe change in dollars is more apt for this application. Another potential source is value added by industry, however this data is lacking for the state level.

Some analysis was attempted on population and income, but conclusions were lacking due to the type of data I was able to obtain. Change in population from 2006 to 2015 was negligible, with no more than a 5% change per county. Overall Vermont saw less than 1% total population growth, all due to four out of the fourteen counties. All counties saw income per person increase, by around \$900 - \$1500 per county over the decade. At this point, it is unclear if this is from wage increase, employment increase, or some combination. It is unclear to me if this information can be used to geographically target our revitalization plan. Further analysis is required in understanding any changes in the labor force for our paired industries.

From this analysis, I believe we should use our resources to foster synergy between our identified information sector, and the healthcare sector. The healthcare sector seems to have strong growth both in Vermont and nationwide. By investing in cross disciplinary activity, we hope to leverage the success of healthcare to aid in the recovery of our information sector. Potential methods include hosting conferences to expose practitioners of each industry to the other, investing in shared incubation space to foster more daily interactions, and grants or competitions to encourage startup growth and company expansion. However, several questions still remain. A better understanding of the labor force, as well as the economic breakdown by county, may help us better focus our areas of investment. Case studies of similar investments may also help guide our approach, as would consultation with an economist.