Simon Miller
Row: L2

A Cross Sectional Analysis of the Effects of State Income Taxation on Innovation

By: Simon Miller, ECN 4050 Fall 2019

Simon Miller
Row: L2

Introduction

Given the societal benefits of innovation, including reduced costs, increased productivity, and higher standard of living, if one is comfortable with using policy to create an economic environment which promotes innovation, it might be beneficial to determine the optimal policy decisions, which balance efforts to maximize tax revenue while still fostering an environment where innovation can occur frequently and successfully. It would be particularly interesting to see if there exists a relationship between the amount of income taxation levied in a states and amount of innovation which occurs in it, as taxing income could alter the incentives to innovate. As a proxy for innovation, one could look at how many patents are awarded per million residents in a given state in a given year. Using this as the dependent variable, an econometric model may help determine if there exist relationships between either corporate or personal income taxation and the amount of innovation which occurs in a state.

The proposed econometric model would be unique in that no economists have previously developed models to examine how innovation activity responds at the aggregate level to changes in amounts of both corporate and personal income taxation simultaneously. While it can easily be argued that most patenting activity occurs through corporations, inclusion of personal income taxation in the model may become increasingly important. With the rise of open-source software, it is becoming easier to develop original ideas without the large research and development (R&D) funds available to larger corporations. Still, there are other avenues though which individuals can innovate and challenge the status quo. If these avenues are fruitful, then one might expect an interesting relationship between levels of personal income taxation and amount of innovation in a state. Hence, one should only remove personal income taxation from a model of this nature with caution and careful consideration.

There are multiple combinations of relationships that may exist between a state's number of patents awarded per million residents and its tax revenue from corporate and personal income taxation. By increasing corporate income taxes, a corporation may be forced to relegate funds that would have been previously used for R&D to pay for fixed costs, thus reducing their efforts to innovation, in turn reducing the number of patents that are generated by that firm. This would result in a less innovative state. Alternatively, firms may respond in a somewhat unexpected manner, where their supply of R&D funds could be considered 'inelastic.' That is to say that an increase in corporate income taxation may have a negligible impact on innovation productivity in a state.

By increasing personal income taxation, potential individual innovators could also respond in a variety of manners. If personal income taxation increases, people may have less access to expendable income, thus reducing their ability to explore new ideas and innovate. This would translate to a total decrease in the number of patents awarded to individuals in a state. Alternatively, lower personal income taxation could create an incentive to continue working or work even more, thus decreasing total free time for individuals. In turn, this would reduce the number of individuals attempting to acquire patents because they are more focused on working in salaried positions or ones with hourly wages.

The regression model will be developed by obtaining a cross sectional data set of the 50 states in the United States with relevant 2010 variables. After reviewing the exiting literature, the formal model will be introduced. Once parameter estimates are obtained, t-tests and F-tests will be used to determine the individual and joint significances of the control and testing independent variables. The necessary statistical tests will be used to test the assumptions about the error term and the model. Specifically, we will be testing for an endogenous variable with the Hausman test, determining if the functional form has been correctly specified with the Ramsey RESET test, testing for heteroskedasticity in the error term with the Breusch-Pagan test, and testing for a non-normal distribution of the error term with the Jarque-Bera test. If any of these issues are present, we will attempt to resolve them and estimate the corrected model, as we seek to preserve as many beneficial statistical properties as possible. This final set of results will aid us in answering the motivating question. Additionally, we will conduct a Chow test for stability to verify that the model developed is stable. A discussion of all results will follow. References used to construct the model will be included after this. Finally, Appendices A, B, and C, which contain the data used to construct the model, the relevant computer output, and the related hypothesis tests, are included

Review of the Literature

There is a rich body of literature that looks at different factors impacting innovation and entrepreneurship. This section is dedicated to observing the results of existing research and gaining insights into helpful models, variables, and functional forms of these variables to inform the construction of the model in this paper.

The authors of "Can Corporate Income Tax Cuts Stimulate Innovation?", Atanassov and Liu, were interested in examining how changes in taxes paid could change the incentive to invest in R&D. They test this idea by comparing two groups of firms, one of which experienced a change in taxes paid and other did not. Through their testing, they were able to confirm that tax cuts resulted in increased patenting and increased quality of patents, measured by citation per patent. They also conclude that tax increases have an opposite effect. They studied over 8000 firms from 1988 to 2016 to make their conclusions. Their dependent variables included both patents and citations per patent. The selected control independent variables include R&D expenditures, profitability, several variables related to gross state product, unemployment rate, and population. They used change in corporate tax rate as their testing independent variable. After exploring the theory and mechanisms that cause the changes observed, they compared their findings to Mukherjee et al, claiming that tax increases and decreases have effects of similar magnitude.

While the motivating question of "Can Corporate Income Tax Cuts Stimulate Innovation?" is similar to the motivating question of this paper, their use of panel data means that the model developed in this paper will not be a replication of the study completed. While they analyzed firm-level data to draw conclusions about the titular question, they included a variety of which will be valuable as control independent variables when developing the econometric model in this paper. They include macroeconomic variables related to GSP and unemployment rate, as well as firm-level variables that may be difficult to incorporate into a state-level model.

Cullen and Gordon examine how changes in taxes affect an individual's decision to divide their time between salaried work and entrepreneurial activities in "Taxes and Entrepreneurial Risk-Taking: Theory and Evidence for the U.S." They study a cross section of single individual's tax returns in 22 years between 1964 and 1993. By studying groups of individuals within tax brackets, they are able to make conclusions about the marginal incentive to engage in entrepreneurial activity. Based on their data set and the resulting model, the authors were able to conclude that level of taxation is correlated with entrepreneurial activity. They conclude that cuts in personal taxes decrease entrepreneurial activity. They also conclude that owners assume more risk when corporate taxes are cut, so they predict either a small change or no change at all in risk taking in the face of decreased corporate taxes.

If one allows "participating in entrepreneurial activity" to be interpreted similarly to "focusing on innovating", then "Taxes and Entrepreneurial Risk-Taking: Theory and Evidence for the U.S." provides valuable insight into developing the aspect of this model which deals with personal income taxation. The results of "Taxes and Entrepreneurial Risk-Taking: Theory and Evidence for the U.S." lead to a conclusion regarding taxation and individuals participating in entrepreneurial activity which were unexpected. These results then allow for the construction of more informed hypothesis testing once the model in this paper is developed. While the authors studied individuals, it may be interesting to see if their results are consistent with results of the model developed in this paper, which uses aggregate data.

"The Effect of Corporate Taxes on Investment and Entrepreneurship", Djankov examines how corporate taxation affects investment, FDI, and entrepreneurial activity. The data was collected by filing taxes as a fictional business in 85 countries and by collecting other relevant data for the cross-section of countries. PricewaterhouseCooper staff in each respective country then completed the taxes. The results were verified for accuracy. The authors consider how a basket of tax-related independent variables, including statutory corporate tax rate and labor tax, among others, are related to dependent variables such as investment and FDI. Control independent variables include GDP per capita, average inflation over a 10-year period, and an index for ease of starting a business in the given country. Interestingly, statutory tax rate did not prove to significantly affect investment, but it was significant when FDI was the dependent variable. Both statutory and effective tax rates were shown to have large impacts on entrepreneurship.

The work completed in "The Effect of Corporate Taxes on Investment and Entrepreneurship" is valuable to the construction of the model in this paper because studies both investment and entrepreneurship, which should be correlated to the number of patents that an area produces. Interestingly, the authors found that statutory corporate taxes did not affect investment, but did affect entrepreneurship. This result almost seems contradictory, so it drives further interest in the issues that this paper hopes to examine. Additionally, there does not exist a large body of relevant literature which aims to analyze cross-sectional data, so it is beneficial to observe how the analysis was conducted. If states were substituted for countries, then the same techniques could be applied when developing the model in this paper.

The article "'Success Taxes,' Entrepreneurial Entry, and Innovation" focuses on how tax policy, among other factors, can affect an individual's decision to assume the role of an entrepreneur. After examining the existing literature, the , Gentry, Hubbard, and Glenn propose that higher tax rates encourage self-employment due to opportunities to easily underreport earnings and legally claim business-related expenses when completing taxes, an idea referred to as the tax sheltering hypothesis. Additionally, taxes that affect the returns of innovating will also change the number of entrepreneurs. The authors present other avenues through which taxation may affect the number of entrepreneurs, though they are less relevant to this paper. The authors use household-level panel data to construct a probit model to address the question of interest. Independent variables used in the model include educational attainment, earnings potential, several demographic variables, and the marginal tax rate paid. There is an additional tax variable which represents how the marginal tax rate changes in response to entrepreneurial success or failure. The authors found that high marginal tax rates discouraged entry into entrepreneurship. The marginal tax spread had a similar effect. They also found that increasing an individual's level of education increased their propensity to enter an entrepreneurial role. Tax convexity discouraged entrepreneurship across all levels of education.

"'Success Taxes,' Entrepreneurial Entry, and Innovation" is another valuable article for studying how an individual's choice to engage in entrepreneurial activity is affected by the taxes they face. Some of the independent variables proposed in the model would be difficult to include in a state-level study, but many of the assumptions used by the authors provide insight into the problem that this paper focuses on. Namely, the authors assume that level of education is a good proxy for ability to innovate and they show that an individual's level of education is positively correlated with their likelihood of engaging in entrepreneurship. Their consideration of tax convexity and progressive income taxes may also be a variable of interest that could enhance the model developed in this paper.

Mukherjee, Singh, and Zaldokas, the authors of "Do corporate taxes hinder innovation," are motivated by a similar question as Atanassov and Liu. They use a differences-in-differences approach of modeling the relationship between corporate taxation and the innovative competitiveness of an economy. Their empirical findings support the notion that increasing taxation on corporations will cause a decrease in innovation in a given economy. The first measure of total innovativeness considered was patenting activity and its reaction to tax changes overtime, controlling for increased patenting seen overtime and firm-level variables that would likely cause innovation to change. They also look at how R&D expenditures, new product announcements, and quality of patents change with respect to these variables. They find that firms with the highest marginal tax rate are most responsive to changes in tax rates, and firms that utilize combined reporting suffer in patenting for a longer period of time than firms that do not use this strategy of tax sheltering. They also test some classical hypotheses surrounding innovation activity. Generally speaking, the authors found that increasing corporate taxes cause a decrease in innovation activity, while decreasing corporate taxes have an opposite, albeit weaker, effect.

Due to the similar nature of the question being asked in "Do corporate taxes hinder innovation" and this paper, "Do corporate taxes hinder innovation" is a valuable resource in increasing the level of robustness in the model developed later in this paper. The authors looked at a firm-level view of the issue though and utilized changes in tax rate over time to help answer their question of whether or not taxation influences innovation. Because this paper aims to utilize cross sectional data, the differences in tax rate are present in different observational units, i.e. each state, not observations over time. The authors' consideration of quality of patents presents another facet of the problem being considered. Ideally, this could be incorporated into the model developed in this paper, but may present difficulty due to the fact that state level data is used and the model can only consider one dependent variable. Some of the independent variables the authors included in their models may be hard to gather data on, or may be simply irrelevant, at the aggregate level, such as product announcements and tax sheltering actions. Still, the paper was valuable in confirming that the use of several independent variables was logically founded.

The Model

As the model is developed, it is important that one specifies the functional form. Principles of multiple linear regression will be used to develop the most robust model possible. Observe the functional form of the model below:

$$pat_i = \beta_0 + \beta_1 bach_i + \beta_2 bus_i + \beta_3 ed_i + \beta_4 GSP_i + \beta_5 pIncRev_i + \beta_6 cIncRev_i + \beta_7 pIncTax_i + \beta_8 cIncTax_i + u$$

The Disturbance Term

It is important that the disturbance term, u, is first addressed before the rest of the model is interpreted. In specifying this model, one should assume that

$$E(u_i \mid x_i) = 0,$$
$$E(u_i{}^2) = \sigma^2,$$
$$E(u_i u_j) = 0 \ \forall\ i \neq j,$$

and

$$u \sim N(0, \sigma^2).$$

This set of assumptions prepares us for the appropriate use of techniques associated with multiple regression analysis. In turn, one can develop a powerful model and be confident in the estimates of the relationships between variables which are obtained through regression. In fact, under these assumptions, the estimates of the parameters presented in the model will be the minimum-variance unbiased estimators. The estimators will be unbiased, consistent, BLUE, asymptotically efficient, and asymptotically normal.

The Dependent Variable

In this model, the dependent variable, pat, is the number of patents awarded per million residents in a given state in 2010. It is used as a proxy for the amount of innovation occurring in a state. States that enact policies to foster more innovation should be expected to have a larger patent count per million residents, while states enacting policies that discourage innovation should be expected to have a lower patent count per million residents, when controlling for all other factors. We consider patents per million residents due to the expected increase in patent production given a larger population. This enables us to increase the degrees of freedom in the model, while still controlling for population.

The Testing Independent Variables

To decide whether increased income taxes discourage innovation, thus answering the driving question of this paper, the variables $pIncRev$, $cIncRev$, $pIncTax$, and $cIncTax$ have been included in the model. These variables stand for personal income tax revenue a corporate income tax revenue per million residents collected in a given state in 2010, and binary variables which assume a value of 1 if a state collects any personal or corporate income tax, respectively. Additionally, the variables $pIncTax$ and $cIncTax$ have been included. These are binary variables which indicate whether or not a state collects personal or corporate income tax at all, encoded as a 0 if the state does not collect this tax and 1 otherwise. These are included because some states do not have any personal or corporate income tax and these effects would not be adequately captured with the income tax revenue variables.

Gentry et al. found that higher marginal income taxes discouraged individuals from engaging in entrepreneurial activity, while Cullen et al. found that decreasing personal taxes decreased the amount of entrepreneurial activity individuals engaged in. If $pIncRev$ increased, individuals could innovate less frequently due to limitations in access to expendable income, or more frequently due to the increased incentive to leave a typical occupational role. Thus, the value of β_5 could assume either a positive or negative value in the model.

Based on the review of the literature, the consensus is that one can expect the coefficient attached to $cIncRev$, β_6, to take a negative value (Atanassov et al. 2019, Mukherjee et al. 2017). The justification of this is essentially that by increasing $cIncRev$ while controlling for all other factors, a firm's R&D expenditures will be forced to decrease to funds being allocated to areas with a more urgent need for funding. This would cause pat to decrease.

While states that do not collect personal or corporate income taxes have other avenues through which tax may be collected, for the purposes of this model, one can hypothesize that the coefficients of both $pIncTax$ and $cIncTax$, β_7 and β_8, will take negative values. This hypothesis is supported by the idea that without collecting taxes on personal or corporate income, individuals and firms will have a larger amount of money to spend on innovation. This would likely translate to an increase in pat if the number of patents awarded in a state is an appropriate proxy for the amount of innovation occurring in a state. Just as changes in the above testing independent variables might cause changes in the incentive to innovate, these variables could as well.

The Control Independent Variables

To control for the most important factors that would cause a change in pat, the following variables have been included in the model: $bach, corp, ed$, GSP, and $r\&d$. Respectively, these variables stand for the percentage of the state's population over 25 possessing bachelor's degrees in 2010, the number of corporations per million residents in a state in 2010, the number of accredited higher education institutions per million residents in a state in 2010, and a state's Gross State Product per million residents in 2010.

The variable $bach$ has been included as a measure of how educated a state's adult population is. It is the percentage of state's population above the age of 25 that possess a bachelor's degree or higher. Gentry et al. assumed that education could be used to approximate an individual's ability to innovate, so, at an aggregate level, if a state's population is more educated, there will be more innovators in the state. Then pat would increase. Thus, we can expect β_1 to be positive.

The variables bus and ed have been included because businesses and higher education institutions are often hotspots for research and innovation. Specifically, these variables measure the number of businesses that are associated with a state by the NAICS Association and the number of public or private universities in a state. By increasing either of these variables, the R&D efforts in a state will increase, as businesses and research universities are largely responsible for pushing the boundaries of our economy and promoting new products and ideas. Hence, it would be reasonable to expect that β_2 and β_3 will take positive values in the model.

GSP has been included as a control independent variable because states with more robust economies can be expected to have larger amounts of innovation efforts occurring. Specifically, GSP measures the total production of a state in US dollars in 2010. As economic activity and innovation increase, pat would also increase. Hence, β_4 should be positive. Atanassov et al. also included GSP in their model because it is associated with the amount of innovation in a state.

The functional forms of the variables have been justified by the use of scatter plots. Each plot displays the independent variable on the x axis and the dependent variable on the y axis. The best type of fit was selected and informed the decision for a variable's form in the model. Each of these plots can be found in Section 1 of Appendix B.

Use of the Model
If the data allows for the rejection of the null hypothesis test related to β_6, β_7, β_8, or β_9, then one can conclude that $pIncRev$, $cIncRev$, $pIncTax$, and $cIncTax$ are individually significant in determining pat, and their effect corresponds to our expectations. Additionally, it would be interesting to test

$$H_0: \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0 \qquad H_1: Otherwise$$

with an F test to determine if $pIncRev$, $cIncRev$, $pIncTax$, and $cIncTax$ are jointly significant in determining pat. After obtaining the coefficient estimates and performing the necessary hypothesis tests, one will be able to determine if personal and consumer income taxes affect the amount of innovation occurring in a given state.

Initial Estimation and Hypothesis Testing

In the previous section, the model and relevant parameters were introduced. This section is dedicated to thorough discussion and testing of the hypotheses presented about each parameter estimate. The following table will summarize the results of estimation of the model and provide an analysis of the statistical significance of the results. These results will then help answer the question of whether or not increasing income taxation negatively impacts patent production per million residents.

The data used to estimate the model was collected from a variety of sources. The dependent variable data was available through US Patent and Trademark Office. The data for corporate and personal income tax revenue and percentage of state population with a bachelor's degree, as well as the population data used to transform several variables, was collected by the US Census Bureau. The data related to the number of higher education institutions was available at worldatlas.com. The data related to the number of businesses in a state was available through the North American Industry Classification System (NAICS) Association. The GSP data was available through the US Bureau of Economic Activity. This data is available in Appendix A. More thorough estimation results are available in Table 1 of Section 2 of Appendix B.

Variable Name	*Expected Sign*	*Estimate*	*Standard Error*	*t-statistic*
Intercept	> 0	-42.93838	237.9233	-0.18047
bach	> 0	28.70267***	7.960368	3.6063
bus	> 0	0.003611	0.003848	0.93841
ed	> 0	7.773414*	5.392977	1.4414
GSP	> 0	-0.011336	0.004415	-2.5674
cIncRev	< 0	0.000486	0.000262	1.8550
pIncRev	≠ 0	0.000158*	8.4E-05	1.8810
cIncTax	< 0	-555.8598**	230.3938	-2.4127
pIncTax	< 0	173.8819	192.8398	0.90169
*F statistic = 6.155491****	*R^2 = 0.545676*	*Adjusted R^2 = 0.457027*	*Std Error = 165.1164*	

** if it is significant at the 10% level; **, 5% level; and ***, 1 % level.*

Based on these results, we can make conclusions about the initial set of hypothesis tests. Among the testing independent variables, we are able to reject the null hypotheses related to $cIncTax$ and $pIncRev$ at an alpha level of 0.05. It appears that state's levying corporate income taxes have lower levels of patenting per capita, and increased personal income tax revenue is associated with higher levels of patenting per capita. Additionally, $cIncRev$ proved to be statistically significant, but the sign was opposite of that which was expected, so we fail to reject the related null hypothesis at an alpha level of 0.10. It might be worth collecting more data to see if corporate income tax revenue is positively correlated with patenting activity in a given state and make a stronger claim. Additionally, we fail to reject the null hypothesis related to $pIncTax$ at an alpha level of 0.10. There is not enough evidence to claim that state's levying personal income taxes are less innovative. For more information, view Section 1: Part 1 of Appendix C.

Among the control independent variables and the intercept, we are able to reject the null hypothesis related to $bach$ at an alpha level of 0.01. There is evidence that, when the population is increasingly educated, the propensity to innovate also increases. We fail to reject the null hypotheses for bus and ed at an alpha level of 0.10. The estimate obtained for the intercept is additionally not statistically significant. While the estimate of GSP is statistically significant, we fail to reject the null hypothesis because the sign was opposite of that which was expected. This could indicate that larger state economies are less productive in terms of innovation output per capita. For more information, view Section 1: Part 1 of Appendix C.

To calculate the joint-significance of the testing independent variables, a Wald test must be performed. A restricted model was estimated without the testing independent variables. This model produced an R^2 = 0.441370. Based on this initial Wald test, included in Section 2 of Appendix C, the testing independent variables are jointly significant at an alpha level of 0.10. Together, the testing independent variables can inform our understanding of the dependent variable. For more information, view Section 1: Part 2 of Appendix C.

Specification Error Testing

Before making definitive claims about the results of our initial estimation, it is critical that we test that the assumptions we have made, about the model in general and the disturbance term specifically, hold.

Endogeneity

The first test we will utilize is the Hausman test for endogeneity. By utilizing this test, we will determine if the assumption $E(u_i x_i) = 0$ is violated, as we lose all desirable statistical properties if this is the case.

We will be examining bus because it is likely that bus and pat are simultaneously determined. That is, factors affecting pat also affect bus. Additionally, bus and pat are likely independent variables of each other when they are treated as dependent variables. This is suspected because factors that promote an increased number of firms also likely increase the number of productive innovators and more firms is correlated with increased innovation, but increased innovation can also result in an increased number of firms. To address this issue, we introduce the instrument variable $retail$, which is the number of businesses in a state that fall under the retail classification per million residents. This data was taken from the Census and is available in Appendix A. It was selected because the number of retail stores clearly affects the number of businesses, but is unlikely related to patenting activity because retail stores focus on sales, not innovation. We first regress the remaining independent variables and $retail$ on bus. This model is

$$bus = \beta_0 + \beta_1 bach + \beta_3 ed + \beta_4 GSP + \beta_5 pIncRev + \beta_6 cIncRev + \beta_7 pIncTax + \beta_8 cIncTax + \delta retail + u.$$

The estimation results are found in Table 2 of Section 2 of Appendix B. The residuals are stored as a variable called $res2$. We then estimate the model

$$pat = \beta_0 + \beta_1 bach + \beta_2 bus + \beta_3 ed + \beta_4 GSP + \beta_5 pIncRev + \beta_6 cIncRev + \beta_7 pIncTax + \beta_8 cIncTax + \delta res2 + u$$

and perform the necessary hypothesis on the coefficient δ. The estimation results are found in Table 3 of Section 2 of Appendix B. The hypothesis test can be found in Section 2 of Appendix C. Based on the result of this hypothesis test, there is enough evidence to claim bus is endogenous. As a result, the model is estimated with Two-Stage Least Squares (TSLS), using $retail$ as an instrument variable for bus. The results are available in Table 4 of Section 2 of Appendix B.

RESET Test

We will additionally conduct a Ramsey RESET test. This tests if there is a violation of the assumption $E(u_i) = 0$, which could possibly make the estimates we obtain both biased and inconsistent. Violations are an indication that the functional form is incorrect, or that an important independent variable has been omitted. If the problem is present, we will focus on altering the functional form of the model.

We initially hypothesized that all variables would appear in the linear form. The general scheme for conducting a RESET test is estimating the model in some functional form, using this estimation to predict $\hat{y}$ for each observation based on the values of independent variables, then re-estimating the model while including $\hat{y}^2, \hat{y}^3$, and $\hat{y}^4$ in the model as independent variables. Specifically, the model we estimate as a part of our first RESET test is a TSLS model of the form

$$pat = \beta_0 + \beta_1 bach + \beta_2 bus + \beta_3 ed + \beta_4 GSP + \beta_5 pIncRev + \beta_6 cIncRev + \beta_7 pIncTax + \beta_8 cIncTax + \delta_0 \hat{y}^2 + \delta_1 \hat{y}^3 + \delta_2 \hat{y}^4 + u$$

The estimation results are displayed in Table 5 of Section 2 of Appendix B. We then conduct an F-test to determine if the coefficients on these new variables are jointly significant in determining pat, which would indicate a violation of our assumption about the disturbance term. Based on the result of the F-test associated with the function above, the model was specified incorrectly. It is available in Section 2 of Appendix C.

To resolve this, we estimated a variety of functional forms. Functional forms differed based on whether or not we took the log transform of the dependent variables, as well as whether or not independent variables were transformed with the log transform. Failed functional form estimation results are available upon request, but have been omitted for the sake of brevity. The related hypothesis tests for the initial functional form, in which we rejected the null hypothesis, and the first functional form in which we failed to reject the null hypothesis and are included in Section 2 of Appendix C.

This functional form is

$$\log(pat) = \beta_0 + \beta_1 \log(bach) + \beta_2 bus + \beta_3 \log(ed) + \beta_4 \log(GSP) + \beta_5 \log(pIncRev + 1) + \beta_6 \log(cIncRev + 1) + \beta_7 pIncTax + \beta_8 cIncTax + u.$$

The results of the restricted and unrestricted estimations used for the RESET test related to this model are displayed in Tables 6 and 7 of Section 2 of Appendix B. Note that adding one to the income tax revenue variables and then taking the log transform is an acceptable approach because this coefficient will have no effect on states that do not levy these taxes, as log(1) = 0, and that adding one only marginally effects states that do levy these taxes because the values assumed are significantly larger. One can verify this by examining Appendix A. It is critical that we introduce each variable to the model in its optimal functional form. With this model, it is more likely that the estimates we obtain are unbiased and consistent.

Heteroskedasticity
We conduct a Breusch-Pagan-Godfrey test to determine if heteroskedasticity is an issue. Heteroskedasticity is an issue when the error term has a non-constant variance. Specifically, it is a violation of the assumption $E(u_i{}^2) = \sigma^2$, as the variance should not be a function of any of our independent variables. If this issue is present and we do not address it, the estimates obtained are neither BLUE or asymptotically efficient.

In conducting our Breusch-Pagan test, we selected to regress population on the variable because it is a size variable which is frequently known to cause heteroskedasticity. The model estimated with the Breusch-Pagan test is

$$u^2 = \delta_0 + \delta_1 pop + \epsilon$$

where the residuals act as an estimate for u.

The related computer output is available in Table 8 of Section 2 of Appendix B. The related hypothesis test is included in Section 2 of Appendix C. We fail to reject the null hypothesis, so it is not necessary to make any changes to the model as a result.

Normality
We also conduct a Jarque-Bera test to determine if the errors are normally distributed. This tests the assumption $u \sim N(0, \sigma^2)$. If we observe a violation, our estimates will be neither efficient or normally distributed, though they will be asymptotically efficient. This is not particularly helpful in our case because we are utilizing a rather limited number of observations.

In practice, one can create a histogram of the residuals and examine the kurtosis and skewness to determine if it is similar to that of a normal distribution. This is done by comparing the Jarque-Bera statistic to a Chi-squared distribution. The histogram and relevant statistics are included in Table 9 of Section 2 of Appendix B and the related hypothesis test is included in Section 2 of Appendix C. As a result of the Jarque-Bera test, there is not enough evidence to claim that the error term is not normally distributed. There is no reason to transform the model to resolve issues related to a non-normal distribution of errors.

Final Estimation and Hypothesis Testing

We have corrected our model for violations of the assumptions made about the disturbance term, which allows us to claim that the model estimates should be unbiased, consistent, BLUE, asymptotically efficient, and asymptotically normal. We can re-estimate our model. To be clear, the model we are estimating is

$$\log(pat) = \beta_0 + \beta_1 \log(bach) + \beta_2 bus + \beta_3 \log(ed) + \beta_4 \log(GSP) + \beta_5 \log(pIncRev + 1) + \beta_6 \log(cIncRev + 1) + \beta_7 pIncTax + \beta_8 cIncTax + u.$$

The data used to estimate the model has not changed, but we have included a new variable $retail$ as an instrument variable. Again, this data is from the Census and is available alongside the other data in Appendix A.

The results of the final estimation are contained below. We will use these results to conduct individual hypothesis tests about each coefficient and a Wald test concerning the testing independent variables. This information is available in Section 3 of Appendix C. More thorough estimation results are available in Table 10 of Section 2 of Appendix B.

Variable Name	*Expected Sign*	*Estimate*	*Standard Error*	*t-statistic*
Intercept	> 0	5.603660	6.741835	0.831177
bach	> 0	4.779017***	1.82228	2.622548
bus	> 0	-3.74E-05	2.85E-05	-1.312281
ed	> 0	0.311808	0.310012	1.005793
GSP	> 0	-1.291306	0.750204	-1.721273
cIncRev	< 0	-0.021771	0.179162	-0.121516
pIncRev	< 0	0.030625	0.051534	0.594268
cIncTax	< 0	-0.568746	2.550351	-0.223007
pIncTax	< 0	-0.358326	0.995036	-0.360114
*F statistic = 5.401075****	*R^2 = 0.423576*	*Adjusted R^2 = 0.311103*	*Std Error = 0.649159*	

Of the testing independent variables included in the model, none of the individual estimates of the coefficients were statistically significant at an alpha level of 0.10. The appropriate signs for the coefficients of $cIncRev, cIncTax$, and $pIncTax$ were assigned, but the standard errors were too high to be very confident in the results. The sign of the coefficient for $pIncRev$ was opposite of what we were expecting, and also statistically insignificant. We failed to reject the null hypotheses related to each of these variables at an alpha level of 0.10. We cannot make very strong claims about their individual relationships to the dependent variable. Section 3: Part 1 of Appendix C contains these hypotheses tests.

Let's consider the control independent variables and intercept. We are able to reject the null hypothesis related to $bach$ at an alpha level of 0.01. There is evidence that, when the population is increasingly educated, the propensity to innovate also increases. We fail to reject the null hypotheses related to bus, ed, and gsp at an alpha level of 0.10. Interestingly, the null hypotheses related to bus and gsp was statistically significant at alpha levels of 0.10 and 0.05 respectively, but the signs of coefficients on these variables were opposite of those which were expected. This might be evidence that larger state economies are associated with less innovation, but further testing would need to be conducted. Section 3: Part 1 of Appendix C contains these hypotheses tests.

Despite the fact that only one coefficient estimate was statistically significant, we also tested for the joint significance of the testing independent variables. We estimated a restricted model by omitting the independent variables and observed an R^2 of 0.288647. The complete estimation results are available in Table 11 of Section 2 of Appendix B. This resulted in the determination that the testing independent variables are jointly-significant in determining the value of pat at an alpha level of 0.10. This indicates that a state's corporate and personal income tax policies may be related to its innovation productivity. The related Wald test is included in Section 3: Part 2 of Appendix C.

Testing for Stability

It would also be interesting to test our model for stability. We introduce a binary variable $coast$, which assumes a value of 1 if a state lies along either the East or West Coast of the United States, and assumes a value of 0 otherwise. We then estimated models for each group. That is, we separated the data based on whether or not $coast = 1$ and obtained new estimates. This would be equivalent to estimating the model

$$\begin{aligned}\log(pat) = {} & \beta_0 + \beta_1 \log(bach) + \beta_2 bus + \beta_3 \log(ed) + \beta_4 \log(GSP) + \beta_5 \log(pIncRev + 1) \\ & + \beta_6 \log(cIncRev + 1) + \beta_7 pIncTax + \beta_8 cIncTax + \delta_0 coas + \delta_1 coast \log(bach) \\ & + \delta_2 coast\ bus + \delta_3 coast \log(ed) + \delta_1 coast \log(GSP) + \delta_5 coast \log(pIncRev + 1) \\ & + \delta_6 coast \log(cIncRev + 1) + \delta_7 coast\ pIncTax + \delta_8 coast\ cIncTax + u\end{aligned}$$

These results are available in Tables 12 and 13 of Section 2 of Appendix B. This enables us to conduct a Chow test using these two new models and model developed in the "Final Estimation and Hypothesis Testing" section. The related Chow test is found in Section 4 of Appendix C.

Based on the results of this Chow test, we reject the null hypothesis of the Chow test. There is enough evidence to claim that our model is not stable. When modeling pat, it is likely necessary to consider whether or not a state belongs to the East or West coast, as the other independent variables appear to behave differently if this is the case. A thorough analysis would necessitate the construction of two models, one for coastal states and one for non-coastal states.

Conclusion

The work above allows us to conclude our study of the relationship between pat, the control independent variables, and the testing independent variables. Specifically, we were interested in examining the relationship between the levels of corporate and personal income taxation which occurred in a state and the level of innovation in a state. We developed a corrected model to account for an endogenous variable using TSLS, but it might be necessary to test other variables for endogeneity before making definite claims about the validity of the model. We then corrected the functional form of the model using the Ramsey RESET test, though we only narrowly failed to reject the null hypothesis of proper specification, so revising the functional form may also yield more meaningful results. It was not necessary to correct for heteroskedasticity or non-normality of the error term.

Utilizing the results of this model allow us to interpret the results more accurately. None of the testing independent variables were individually significant in determining the level of innovation in a given state. This is interesting because Atanassov et. al. and Mukherjee et. al. were able to obtain statistically significant estimates for the relationship between corporate income taxation and innovation, but the data they used was at a much lower level. As the data is aggregated, more noise is introduced, resulting in a more difficult analysis.

Despite the fact that our testing independent variables were not individually significant, together, corporate and personal income taxation information proved useful in explaining, to some extent, the reason innovation across states is different. This seems to align with the existing body of literature, which is reassuring. However, we are not able to use the results to make policy recommendations or better understand individual relationships, as we sought to do in the beginning of this paper.

If one were to move forward with this work, it would be beneficial to examine panel data across several years. Utilizing only 50 data points makes the process of modeling sensitive due to the lack of available information. However, the cross-sectional nature of the data set allows for a simpler analysis, appropriate for an expository article. Additionally, one could collect data on multiple developed countries. Aside from collecting more data, one might also want to move to firm-level data if it is available, as it seems results are more robust using this data. Additionally, information about individuals, such as family-level tax information, could enhance our understanding of how personal income tax impacts an individual's choice to innovate. It was ambitious to attempt to study income taxation at both corporate and personal levels simultaneously, so utilizing fewer testing independent variables may also result in improved analyses. Aside from changing the level, quantity, or quality of the data, it might be beneficial to further study the violations of the model so that they can be dealt with more appropriately. It could also be interesting to resolve the issue of instability with regard to $coast$ to identify which factors are most important in identifying the best explanatory variables for coastal and non-coastal states. A less ambitious goal may be to conduct the analysis while omitting the outlier mentioned in Section 1 of Appendix B. Overall, the results of this paper should be approached with skepticism, as there is still plenty of room to identify factors which affect innovation activity in a given state.

References

Atanassov, J., & Liu, X. (2019). Can Corporate Income Tax Cuts Stimulate Innovation? *Journal of Financial and Quantitative Analysis, 00* (00). 1-51. doi:10.1017/S0022109019000152

We hypothesize that corporate income taxes distort firms' incentives to innovate by reducing their pledgeable income. Using a differences-in-differences methodology, we document that large corporate income tax cuts boost corporate innovation. We find a similar but opposite effect for tax increases. Most of the change in innovation occurs 2 or more years after the tax change, and there's no effect before the tax change. Exploring the mechanisms, we show that tax cuts have a stronger impact on innovation for firms with weaker governance, greater financial constraints, fewer tangible assets, smaller patent stock, and a greater degree of tax avoidance.

Cullen, Julie Berry, and Gordon, Roger H. "Taxes and Entrepreneurial Risk-Taking: Theory and Evidence for the U.S." *Journal of Public Economics* 91.7 (2007): 1479–1505. Web.

How does the tax law affect individual incentives to engage in entrepreneurial risk taking? We first show theoretically that taxes can affect incentives due to differences in tax rates on business vs. wage income, due to differences in the marginal tax rates faced on losses vs. profits through a progressive rate structure and through the option to incorporate, and due to risk sharing with the government. We then provide empirical evidence using U.S. individual tax return data that each of these aspects of the tax law have clear effects on individual behavior, and together have had large effects on the amount of entrepreneurial risk taking.

Djankov, Simeon et al. "The Effect of Corporate Taxes on Investment and Entrepreneurship." *American Economic Journal: Macroeconomics* 2.3 (2010): 31–64. Web.

We present new data on effective corporate income tax rates in 85 countries in 2004. The data come from a survey, conducted jointly with PricewaterhouseCoopers, of all taxes imposed on "the same" standardized mid-size domestic firm. In a cross-section of countries, our estimates of the effective corporate tax rate have a large adverse impact on aggregate investment, FDI, and entrepreneurial activity. Corporate tax rates are correlated with investment in manufacturing but not services, as well as with the size of the informal economy. The results are robust to the inclusion of many controls.

Gentry, William M., and Hubbard, R. Glenn. "'Success Taxes,' Entrepreneurial Entry, and Innovation." *Innovation Policy and the Economy* 5 (2005): 87–108. Print.

Interest in the role of entrepreneurial entry in innovation raises the question about the extent to which tax policy encourages or discourages entry. We find that, while the level of the marginal tax rate has a negative effect on entrepreneurial entry, the progressivity of the tax also discourages entrepreneur ship, and significantly so for some groups of households. These effects are traceable principally to the "upside," or "success," convexity of the household tax schedule. Prospective entrants from a priori innovative industries and occupations are no less affected by the considerations we examine than are other prospective entrants. In terms of destination-based industry and occupation measures of innovative entrepreneurs, we find mixed evidence on whether innovative entrepreneurs differ from the general population. The results for entrepreneurs moving to innovative industries suggest that they may be unaffected by tax convexity, but the possible endogeneity of this measure of innovative entrepreneurs confounds interpreting this specification. Using education as a measure of potential for innovation, we find that tax convexity discourages entry into self-employment for people of all educational backgrounds. Overall, we

find little evidence that the tax effects are focused simply on the employment changes of less-skilled or less-promising potential entrants.

Mukherjee, A., Singh, M., & Zaldokas, A. (2017). Do corporate taxes hinder innovation? *Journal of Financial Economics, 124* (1). 195-221.

We exploit staggered changes in state-level corporate tax rates to show that an increase in taxes reduces future innovation. A variety of tests, including those based on policy discontinuity at contiguous counties straddling borders of politically similar states, show that local economic conditions do not drive our results. The effect we document is consistent across the innovation spectrum: taxes affect not only patenting and R&D investment but also new product introductions, which we measure using textual analysis. Our empirical results are consistent with models that highlight the role of higher corporate taxes in reducing innovator incentives and discouraging risk-taking.

Simon Miller
Row: L2

Appendix

Appendix A

Cross Section: Does increased income taxation decrease the number of patents awarded?												
state	pat	p_inc_rev	p_inc_tax	c_inc_rev	c_inc_tax	bach	ed	bus	gsp	retail	coast	pop
ALABAMA	93.6171493	546,008	1	92015.6274	1	21.7	13.16491162	41,989	36517.5841	3,805	0	4,785,448
ALASKA	42.0223391	0	0	1065133.22	1	27	7.003723179	57,188	74074.8782	3,513	0	713,906
ARIZONA	308.531481	350,138	1	84634.6953	1	26.3	3.901510883	53,216	38735.7607	2,728	0	6,407,774
ARKANSAS	49.2816852	732,429	1	118047.09	1	19.1	8.213614202	43,667	34801.7678	3,738	0	2,921,978
CALIFORNIA	732.431367	1,308,698	1	255495.211	1	30.1	6.028793033	52,738	52909.0762	2,851	1	37,320,903
COLORADO	422.718149	839,571	1	77113.9721	1	35.9	6.734965823	69,999	50540.0749	3,659	0	5,048,281
CONNECTICUT	524.150456	1,673,232	1	142084.727	1	35.2	8.102539028	60,947	66399.8603	3,520	1	3,579,125
DELAWARE	409.072972	997,493	1	218231.538	1	27.7	10.00450203	57,590	63859.1811	4,020	1	899,595
FLORIDA	158.125544	0	0	95768.8947	1	25.9	6.845031926	78,784	39148.8654	3,777	1	18,845,785
GEORGIA	196.564801	755,403	1	70376.1709	1	27.2	7.001784425	57,072	42925.4485	3,442	1	9,711,810
HAWAII	88.7120838	1,007,141	1	63874.1667	1	29.4	7.331577176	42,657	49999.0836	3,404	0	1,363,963
IDAHO	697.109003	692,160	1	71750.6603	1	24.3	7.002921492	58,187	35123.2801	3,702	0	1,570,773
ILLINOIS	281.369595	751,057	1	211224.303	1	30.3	8.177084818	43,765	51604.2428	3,111	0	12,840,762
INDIANA	229.722626	635,878	1	98395.5469	1	22.4	9.706589819	46,402	43158.9958	3,328	0	6,490,436
IOWA	250.10104	831,487	1	70629.4515	1	24.5	12.12809762	58,544	46375.8786	3,949	0	3,050,767
KANSAS	215.519277	968,335	1	111522.829	1	29.3	11.89554452	49,403	44693.9399	3,690	0	2,858,213
KENTUCKY	123.499379	741,947	1	100364.289	1	20.3	8.279287981	44,625	37856.0094	3,501	0	4,348,200
LOUISIANA	67.5537107	476,833	1	30697.9905	1	20.9	7.481518449	52,578	49659.0188	3,684	0	4,544,532
MAINE	158.929583	1,009,373	1	150562.053	1	26.5	11.29831158	52,360	38949.5734	4,784	1	1,327,632
MARYLAND	273.121053	1,044,588	1	124727.009	1	35.7	5.873571038	54,943	54667.6578	3,140	1	5,788,642
MASSACHUSETTS	749.874627	1,623,676	1	288903.059	1	38.3	14.16294483	56,071	62402.6202	3,702	1	6,566,431
MICHIGAN	387.748563	571,068	1	59176.4038	1	25	8.605385858	46,132	39141.9418	3,529	0	9,877,535
MINNESOTA	677.481899	1,255,475	1	170647.485	1	31.4	9.038113158	58,873	51182.3076	3,598	0	5,310,843
MISSISSIPPI	48.8127395	459,904	1	107432.463	1	19.5	6.059512492	38,211	32050.9161	3,903	0	2,970,536
MISSOURI	163.109392	737,342	1	32760.4714	1	25	12.17483192	48,912	43119.6022	3,578	0	5,995,976
MONTANA	105.983313	754,345	1	105082.96	1	27.9	12.11237865	71,958	38410.977	4,876	0	990,722
NEBRASKA	117.516135	834,188	1	76825.4902	1	27.7	13.11808021	58,953	50183.6531	3,979	0	1,829,536
NEVADA	200.187681	0	0	0	0	21.8	3.700326813	44,565	45735.3363	3,010	0	2,702,464
NEW HAMPSHIRE	552.105634	60,435	1	415995.267	1	32.9	13.66974059	59,332	48718.6517	4,653	1	1,316,777
NEW JERSEY	440.359724	1,169,129	1	239857.976	1	34.6	5.227496084	54,535	56269.3361	3,605	1	8,799,624
NEW MEXICO	211.180148	383,409	1	55673.5775	1	25.5	5.327939521	44,296	40680.8525	3,192	0	2,064,588
NEW YORK	365.20468	1,847,807	1	203203.647	1	32.1	11.59789032	48,103	62574.3244	3,993	1	19,400,080
NORTH CAROLINA	275.216144	963,364	1	113796.705	1	26.1	5.848995848	51,874	43366.7948	3,581	1	9,574,293
NORTH DAKOTA	157.104534	467,253	1	193079.99	1	26.3	20.74965541	71,073	52505.9655	4,721	0	674,710
OHIO	280.258979	703,040	1	17085.8318	1	24.1	9.61927849	47,196	43088.1454	3,166	0	11,539,327
OKLAHOMA	137.779442	618,260	1	62512.2352	1	22.6	8.777454815	49,736	40671.9062	3,471	0	3,759,632
OREGON	531.852243	1,323,473	1	110191.654	1	28.6	8.859860973	66,653	42690.8753	3,617	1	3,837,532
PENNSYLVANIA	263.941334	748,021	1	138352.776	1	26.4	12.35135304	49,117	47148.5603	3,458	0	12,711,158
RHODE ISLAND	261.874987	907,401	1	127818.714	1	30.3	11.385869	51,663	47034.8351	3,601	1	1,053,938
SOUTH CAROLINA	111.742545	481,819	1	36258.083	1	24	7.550171971	47,351	35347.0577	3,794	1	4,635,656
SOUTH DAKOTA	86.9922136	0	1	23665.5578	1	25.3	17.15339423	65,029	45927.4779	4,709	0	816,165
TENNESSEE	145.705137	27,580	1	166183.946	1	22.7	8.968890695	45,429	40271.2318	3,558	0	6,355,301
TEXAS	299.01739	0	0	0	0	25.8	11.84501851	55,801	49012.1037	3,101	0	25,242,679
UTAH	366.442381	781,774	1	87285.7105	1	29.4	6.485705865	52,789	42503.1726	3,277	0	2,775,334
VERMONT	1025.75574	800,818	1	168455.614	1	33.3	35.15050808	66,989	43312.4561	5,607	0	625,880
VIRGINIA	197.41565	1,112,604	1	96546.0985	1	33.8	7.228603334	54,257	52706.7381	3,417	1	8,023,680
WASHINGTON	779.782948	0	0	0	0	31	10.08467867	54,404	54212.8449	3,202	1	6,742,902
WEST VIRGINIA	63.6388249	809,873	1	208896.6	1	17.3	12.9434898	33,591	35203.1103	3,448	0	1,854,214
WISCONSIN	318.953817	1,030,762	1	156206.885	1	25.8	9.138070802	46,888	44698.188	3,387	0	5,690,479
WYOMING	147.0372	0	0	0	0	23.6	15.94379282	61,431	66257.4427	4,749	0	564,483

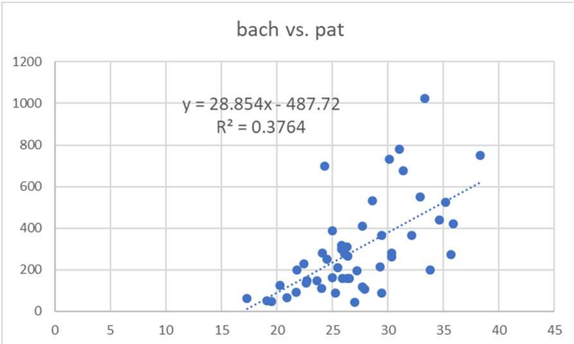
Simon Miller
Row: L2

Appendix B Table of Contents

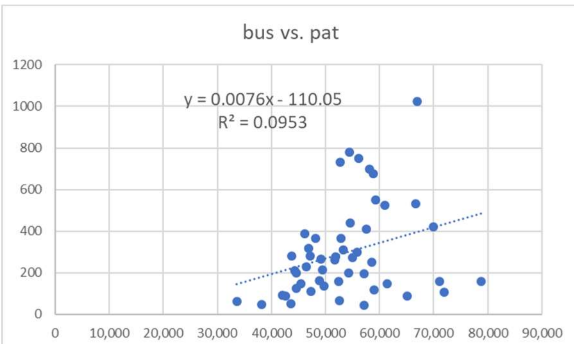
Simon Miller
Row: L2

Appendix B

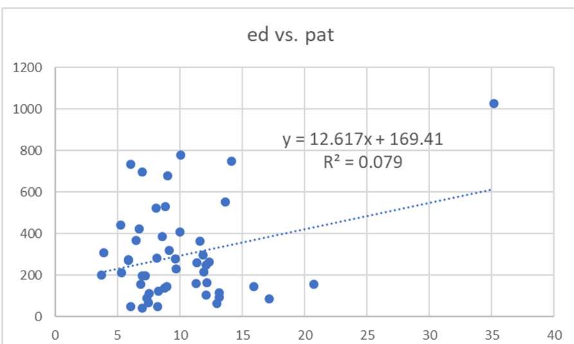
Section 1



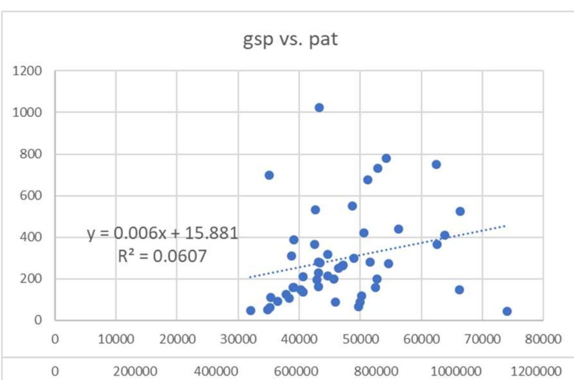
Plot 1
There is more variation about the fit line on the right-hand side, but a linear fit appears to do a good job displaying the relationship. A logarithmic fit might also be appropriate.



Plot 2
It almost appears as if there is a split, with different linear fits that would do a good job showing the relationship. There could be some underlying cause of this. Still, a linear fit seems the most logical.

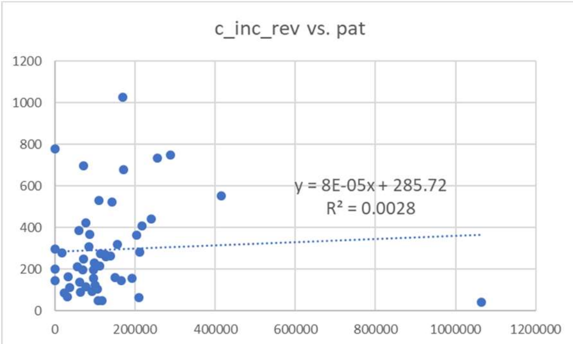


Plot 3
There is one outlier present, but there does not appear to be much correlation among these variables if it is excluded.

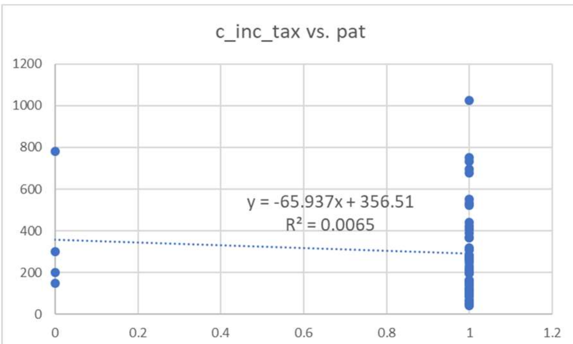


Plot 4
There is a lot of variation about the fit line, but the linear fit appears to do the best job of capturing the relationship. There is an outlier, but the linear fit appears to do the best job of capturing the relationship still.

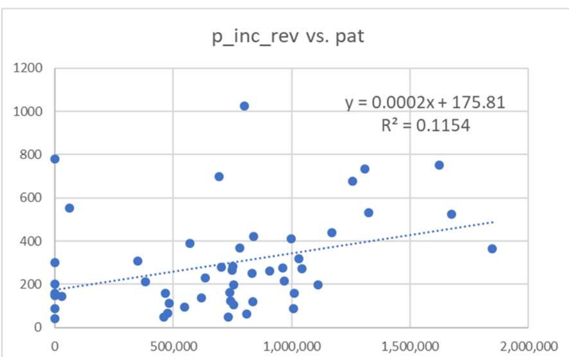
Simon Miller
Row: L2



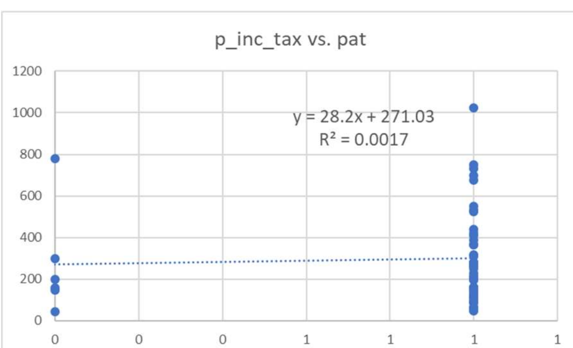
Plot 5
There is an outlier, but the linear fit appears to do the best job of capturing the relationship still. With more time, it could be interesting to conduct an analysis where the outlier is omitted.



Plot 6
It would not make sense to transform this variable, as it is a binary, but it is interesting to observe how they are correlated.



Plot 7
A linear fit appears to do the best job of capturing the relationship between these two variables, though there is only a weak correlation.



Plot 8
It would not make sense to transform this variable, as it is a binary, but it is interesting to observe how they are correlated.

Simon Miller
Row: L2

Section 2

Table 1
Initial Estimation

Dependent Variable: PAT
Method: Least Squares
Date: 11/29/19 Time: 08:41
Sample: 1 50
Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-42.93838	237.9233	-0.180472	0.8577
BACH	28.70267	7.960368	3.605697	0.0008
BUS	0.003611	0.003848	0.938500	0.3535
C_INC_REV	0.000486	0.000262	1.854140	0.0709
C_INC_TAX	-555.8598	230.3938	-2.412651	0.0204
ED	7.773414	5.392977	1.441396	0.1571
GSP	-0.011336	0.004415	-2.567759	0.0140
P_INC_REV	0.000158	8.40E-05	1.886780	0.0663
P_INC_TAX	173.8819	192.8398	0.901691	0.3725

R-squared	0.545676	Mean dependent var	295.8446
Adjusted R-squared	0.457027	S.D. dependent var	224.0790
S.E. of regression	165.1164	Akaike info criterion	13.21273
Sum squared resid	1117801.	Schwarz criterion	13.55689
Log likelihood	-321.3182	Hannan-Quinn criter.	13.34379
F-statistic	6.155491	Durbin-Watson stat	2.248572
Prob(F-statistic)	0.000033		

Table 2
Hausman Test
First Stage Model

Dependent Variable: BUS
Method: Least Squares
Date: 11/29/19 Time: 11:48
Sample: 1 50
Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8721.502	9990.822	-0.872951	0.3878
BACH	1025.110	243.1081	4.216684	0.0001
C_INC_REV	-0.027684	0.008359	-3.311951	0.0019
C_INC_TAX	27058.76	7007.995	3.861128	0.0004
ED	-143.0632	263.9165	-0.542078	0.5907
GSP	0.149681	0.153759	0.973481	0.3360
P_INC_REV	-0.003431	0.002908	-1.179773	0.2449
P_INC_TAX	-24354.21	5869.374	-4.149372	0.0002
RETAIL	8.537323	2.325428	3.671292	0.0007

R-squared	0.660583	Mean dependent var	53557.35
Adjusted R-squared	0.594355	S.D. dependent var	9127.583
S.E. of regression	5813.380	Akaike info criterion	20.33526
Sum squared resid	1.39E+09	Schwarz criterion	20.67943
Log likelihood	-499.3815	Hannan-Quinn criter.	20.46632
F-statistic	9.974417	Durbin-Watson stat	2.075822
Prob(F-statistic)	0.000000		

Table 3
Hausman Test
Artificial Regression

Dependent Variable: PAT
Method: Least Squares
Date: 11/29/19 Time: 11:51
Sample: 1 50
Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	143.6560	237.8873	0.603883	0.5493
BACH	44.49313	9.968111	4.463547	0.0001
BUS	-0.011729	0.007317	-1.603037	0.1168
C_INC_REV	4.18E-05	0.000309	0.135352	0.8930
C_INC_TAX	-74.56981	295.1835	-0.252622	0.8019
ED	16.84513	6.332626	2.660054	0.0112
GSP	-0.008644	0.004321	-2.000478	0.0523
P_INC_REV	8.40E-05	8.52E-05	0.985845	0.3301
P_INC_TAX	-213.5299	242.8079	-0.879419	0.3844
RES2	0.020383	0.008434	2.416816	0.0203

R-squared	0.603565	Mean dependent var	295.8446
Adjusted R-squared	0.514367	S.D. dependent var	224.0790
S.E. of regression	156.1547	Akaike info criterion	13.11643
Sum squared resid	975372.0	Schwarz criterion	13.49883
Log likelihood	-317.9107	Hannan-Quinn criter.	13.26205
F-statistic	6.766590	Durbin-Watson stat	2.162041
Prob(F-statistic)	0.000008		

Table 4
Two-Stage Least
Squares Estimation

Dependent Variable: PAT
Method: Two-Stage Least Squares
Date: 11/29/19 Time: 11:56
Sample: 1 50
Included observations: 50
Instrument specification: C BACH C_INC_REV C_INC_TAX ED GSP
P_INC_REV P_INC_TAX RETAIL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	143.6560	296.3043	0.484826	0.6304
BACH	44.49313	12.41594	3.583550	0.0009
C_INC_REV	4.18E-05	0.000384	0.108667	0.9140
C_INC_TAX	-74.56981	367.6704	-0.202817	0.8403
ED	16.84513	7.887701	2.135619	0.0387
GSP	-0.008644	0.005382	-1.606080	0.1159
P_INC_REV	8.40E-05	0.000106	0.791484	0.4332
P_INC_TAX	-213.5299	302.4332	-0.706040	0.4842
BUS	-0.011729	0.009113	-1.286995	0.2053

R-squared	0.369582	Mean dependent var	295.8446
Adjusted R-squared	0.246573	S.D. dependent var	224.0790
S.E. of regression	194.5010	Sum squared resid	1551056.
F-statistic	4.563783	Durbin-Watson stat	2.411910
Prob(F-statistic)	0.000497	Second-Stage SSR	1079153.
J-statistic	0.000000	Instrument rank	9

Table 5
RESET Test
Initial TSLS
Unrestricted Model

Dependent Variable: PAT
Method: Two-Stage Least Squares
Date: 11/29/19 Time: 13:42
Sample: 1 50
Included observations: 50
Instrument specification: C BACH C_INC_REV C_INC_TAX ED GSP P_INC_REV P_INC_TAX Y_HAT2 Y_HAT3 Y_HAT4 RETAIL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	244.8684	426.2099	0.574525	0.5690
BACH	278.4455	206.5714	1.347938	0.1857
C_INC_REV	-4.07E-05	0.000543	-0.074959	0.9406
C_INC_TAX	-529.4262	362.4558	-1.460664	0.1523
ED	86.11119	77.73069	1.107815	0.2749
GSP	-0.043802	0.030474	-1.437358	0.1588
P_INC_REV	0.000456	0.000304	1.500736	0.1417
P_INC_TAX	-1125.028	1180.962	-0.952637	0.3468
Y_HAT2	-0.024091	0.020793	-1.158571	0.2539
Y_HAT3	3.44E-05	3.56E-05	0.965962	0.3402
Y_HAT4	-1.41E-08	2.04E-08	-0.692293	0.4930
BUS	-0.071457	0.061611	-1.159814	0.2534

R-squared	0.064243	Mean dependent var	295.8446
Adjusted R-squared	-0.206634	S.D. dependent var	224.0790
S.E. of regression	246.1438	Sum squared resid	2302297.
F-statistic	2.487566	Durbin-Watson stat	2.161391
Prob(F-statistic)	0.018385	Second-Stage SSR	802509.5
J-statistic	0.000000	Instrument rank	12

Table 6
RESET Test
Tenth TSLS
Restricted Model

Dependent Variable: LOG(PAT)
Method: Two-Stage Least Squares
Date: 11/29/19 Time: 16:49
Sample: 1 50
Included observations: 50
Instrument specification: C LOG(BACH) LOG(C_INC_REV+1) C_INC_TAX LOG(ED) LOG(GSP) LOG(P_INC_REV+1) P_INC_TAX RETAIL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.603660	6.741835	0.831177	0.4107
LOG(BACH)	4.779017	1.182228	4.042382	0.0002
LOG(C_INC_REV+1)	-0.021771	0.179162	-0.121517	0.9039
C_INC_TAX	-0.568746	2.550351	-0.223007	0.8246
LOG(ED)	0.311808	0.310012	1.005793	0.3204
LOG(GSP)	-1.291306	0.750204	-1.721272	0.0927
LOG(P_INC_REV+1)	0.030625	0.051534	0.594277	0.5556
P_INC_TAX	-0.358326	0.995036	-0.360113	0.7206
BUS	-3.74E-05	2.85E-05	-1.312920	0.1965

R-squared	0.423576	Mean dependent var	5.411123
Adjusted R-squared	0.311103	S.D. dependent var	0.782121
S.E. of regression	0.649159	Sum squared resid	17.27772
F-statistic	5.401075	Durbin-Watson stat	2.229778
Prob(F-statistic)	0.000115	Second-Stage SSR	11.76555
J-statistic	6.86E-36	Instrument rank	9

Table 7
RESET Test
Tenth TSLS
Unrestricted Model

Dependent Variable: LOG(PAT)
Method: Two-Stage Least Squares
Date: 11/29/19 Time: 16:50
Sample: 1 50
Included observations: 50
Instrument specification: C LOG(BACH) LOG(C_INC_REV+1) C_INC_TAX LOG(ED) LOG(GSP) LOG(P_INC_REV+1) P_INC_TAX RETAIL YHAT2 YHAT3 YHAT4

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.128809	11.87382	-0.179286	0.8587
LOG(BACH)	39.41159	26.25343	1.501198	0.1416
LOG(C_INC_REV+1)	-0.177257	0.336967	-0.526038	0.6019
C_INC_TAX	-2.825519	3.470012	-0.814268	0.4206
LOG(ED)	2.544367	1.927401	1.320102	0.1947
LOG(GSP)	-9.133399	5.965669	-1.530993	0.1341
LOG(P_INC_REV+1)	0.198183	0.130470	1.518995	0.1370
P_INC_TAX	-3.282907	3.080188	-1.065814	0.2932
BUS	-0.000318	0.000242	-1.311843	0.1974
YHAT2	-0.000235	0.000173	-1.355819	0.1832
YHAT3	5.21E-07	3.91E-07	1.331309	0.1910
YHAT4	-3.24E-10	2.48E-10	-1.307682	0.1988

R-squared	-0.252910	Mean dependent var	5.411123
Adjusted R-squared	-0.615595	S.D. dependent var	0.782121
S.E. of regression	0.994124	Sum squared resid	37.55471
F-statistic	1.860696	Durbin-Watson stat	1.709386
Prob(F-statistic)	0.077194	Second-Stage SSR	9.746172
J-statistic	4.73E-31	Instrument rank	12

Table 8
Breusch-Pagan-Godfrey Test

Heteroskedasticity Test: Breusch-Pagan-Godfrey
Null hypothesis: Homoskedasticity

F-statistic	0.167933	Prob. F(1,48)	0.6838
Obs*R-squared	0.174321	Prob. Chi-Square(1)	0.6763
Scaled explained SS	0.128115	Prob. Chi-Square(1)	0.7204

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/29/19 Time: 17:06
Sample: 1 50
Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.372972	0.099476	3.749375	0.0005
POP	-4.44E-09	1.08E-08	-0.409797	0.6838

R-squared	0.003486	Mean dependent var	0.345554
Adjusted R-squared	-0.017274	S.D. dependent var	0.516095
S.E. of regression	0.520534	Akaike info criterion	1.571255
Sum squared resid	13.00587	Schwarz criterion	1.647735
Log likelihood	-37.28136	Hannan-Quinn criter.	1.600379
F-statistic	0.167933	Durbin-Watson stat	1.406664
Prob(F-statistic)	0.683778		

Table 9
Jarque-Berra Test for Normality

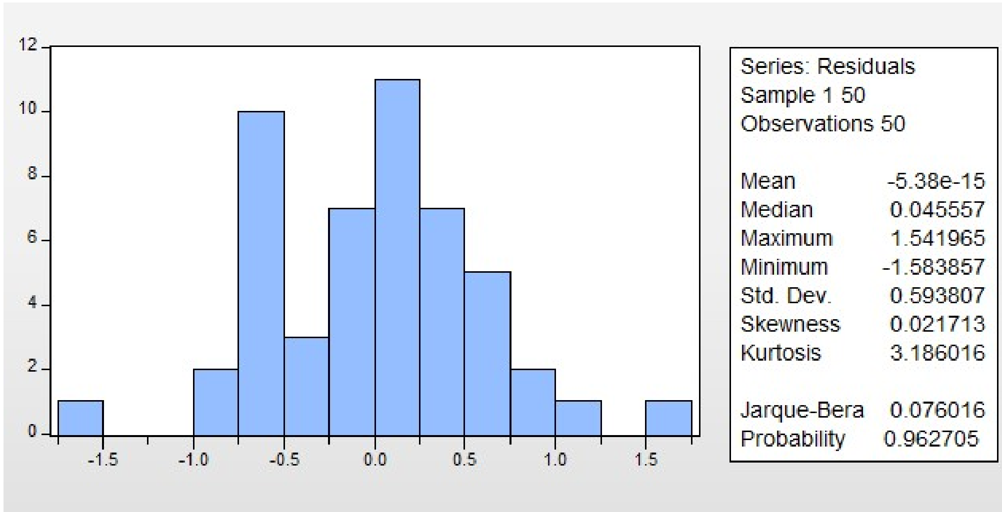


Table 10
Final Model

Dependent Variable: LOG(PAT)
Method: Two-Stage Least Squares
Date: 12/04/19 Time: 10:17
Sample: 1 50
Included observations: 50
Instrument specification: C LOG(BACH) LOG(C_INC_REV+1) C_INC_TAX LOG(ED) LOG(GSP) LOG(P_INC_REV+1) P_INC_TAX RETAIL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.603660	6.741835	0.831177	0.4107
LOG(BACH)	4.779017	1.182228	4.042382	0.0002
LOG(C_INC_REV+1)	-0.021771	0.179162	-0.121517	0.9039
C_INC_TAX	-0.568746	2.550351	-0.223007	0.8246
LOG(ED)	0.311808	0.310012	1.005793	0.3204
LOG(GSP)	-1.291306	0.750204	-1.721272	0.0927
LOG(P_INC_REV+1)	0.030625	0.051534	0.594277	0.5556
P_INC_TAX	-0.358326	0.995036	-0.360113	0.7206
BUS	-3.74E-05	2.85E-05	-1.312920	0.1965

R-squared	0.423576	Mean dependent var	5.411123
Adjusted R-squared	0.311103	S.D. dependent var	0.782121
S.E. of regression	0.649159	Sum squared resid	17.27772
F-statistic	5.401075	Durbin-Watson stat	2.229778
Prob(F-statistic)	0.000115	Second-Stage SSR	11.76555
J-statistic	6.86E-36	Instrument rank	9

Table 11
Final Restricted Model

Dependent Variable: LOG(PAT)
Method: Two-Stage Least Squares
Date: 12/04/19 Time: 10:39
Sample: 1 50
Included observations: 50
Instrument specification: C LOG(BACH) LOG(ED) LOG(GSP) RETAIL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.818534	6.039430	0.135532	0.8928
LOG(BACH)	4.584621	0.955519	4.798042	0.0000
LOG(ED)	0.311863	0.281879	1.106373	0.2744
LOG(GSP)	-0.811393	0.665074	-1.220004	0.2288
BUS	-4.58E-05	2.66E-05	-1.722577	0.0918

R-squared	0.288647	Mean dependent var	5.411123
Adjusted R-squared	0.225416	S.D. dependent var	0.782121
S.E. of regression	0.688349	Sum squared resid	21.32208
F-statistic	8.197169	Durbin-Watson stat	2.350141
Prob(F-statistic)	0.000047	Second-Stage SSR	14.43792
J-statistic	0.000000	Instrument rank	5

Table 12
Chow Test:
Not Coastal

Dependent Variable: LOG(PAT)
Method: Two-Stage Least Squares
Date: 12/04/19 Time: 12:47
Sample: 1 50 IF COAST=0
Included observations: 33
Instrument specification: C LOG(BACH) LOG(C_INC_REV+1)
C_INC_TAX LOG(ED) LOG(GSP) LOG(P_INC_REV+1) P_INC_TAX
RETAIL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.733326	10.80646	0.623083	0.5391
LOG(BACH)	5.012031	1.327721	3.774913	0.0009
LOG(C_INC_REV+1)	-0.069559	0.220243	-0.315830	0.7549
C_INC_TAX	-0.597561	3.227680	-0.185136	0.8547
LOG(ED)	0.324009	0.353983	0.915323	0.3691
LOG(GSP)	-1.477481	1.129560	-1.308015	0.2033
LOG(P_INC_REV+1)	0.033992	0.055677	0.610517	0.5473
P_INC_TAX	0.129503	1.437313	0.090101	0.9290
BUS	-3.52E-05	2.82E-05	-1.250231	0.2233

R-squared	0.488506	Mean dependent var	5.200525
Adjusted R-squared	0.318009	S.D. dependent var	0.788176
S.E. of regression	0.650898	Sum squared resid	10.16803
F-statistic	3.607759	Durbin-Watson stat	2.951294
Prob(F-statistic)	0.006907	Second-Stage SSR	7.651164
J-statistic	0.000000	Instrument rank	9

Simon Miller
Row: L2

Table 13
Chow Test: Coastal

Dependent Variable: LOG(PAT)
Method: Two-Stage Least Squares
Date: 12/04/19 Time: 12:46
Sample: 1 50 IF COAST=1
Included observations: 17
Instrument specification: C LOG(BACH) LOG(C_INC_REV+1) C_INC_TAX LOG(ED) LOG(GSP) LOG(P_INC_REV+1) P_INC_TAX RETAIL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-13.50117	113.1569	-0.119314	0.9080
LOG(BACH)	-4.604723	26.19406	-0.175793	0.8648
LOG(C_INC_REV+1)	0.284563	3.006350	0.094654	0.9269
C_INC_TAX	-20.22992	61.38334	-0.329567	0.7502
LOG(ED)	-0.228232	3.215604	-0.070976	0.9452
LOG(GSP)	0.365037	9.224750	0.039571	0.9694
LOG(P_INC_REV+1)	0.461180	2.331179	0.197831	0.8481
P_INC_TAX	9.224523	52.33135	0.176271	0.8645
BUS	0.000598	0.002899	0.206187	0.8418

R-squared	-15.875920	Mean dependent var	5.819930
Adjusted R-squared	-32.751840	S.D. dependent var	0.601571
S.E. of regression	3.494906	Sum squared resid	97.71492
F-statistic	0.050330	Durbin-Watson stat	0.380052
Prob(F-statistic)	0.999836	Second-Stage SSR	0.872172
J-statistic	3.94E-36	Instrument rank	9

Simon Miller
Row: L2

Appendix C Table of Contents

Simon Miller
Row: L2

Appendix C

Section 1: Part 1
Critical t-values:
n = 50
df = 41

	One-sided	Two-sided
$\alpha = 0.01$	2.4208	2.7012
$\alpha = 0.05$	1.6829	2.0195
$\alpha = 0.10$	1.3025	1.6829

$H_0: \beta_0 \leq 0 \qquad H_1: \beta_0 > 0$

$$t = \frac{\widehat{\beta_0} - 0}{se(\widehat{\beta_0})} = \frac{-42.93838 - 0}{237.9233} = -0.18047$$

Therefore, because -0.18047 < 1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_1 \leq 0 \qquad H_1: \beta_1 > 0$

$$t = \frac{28.70267 - 0}{7.960368} = 3.6063$$

Therefore, because 3.6063 > 2.4208, we reject H_0 at the $\alpha = 0.01$ significance level.

$H_0: \beta_2 \leq 0 \qquad H_1: \beta_2 > 0$

$$t = \frac{0.003611 - 0}{0.003848} = 0.93841$$

Therefore, because 0.93841 < 1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_3 \leq 0 \qquad H_1: \beta_3 > 0$

$$t = \frac{7.773414 - 0}{5.392977} = 1.4414$$

Therefore, because 1.4414 > 1.3025, we reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_4 \leq 0 \qquad H_1: \beta_4 > 0$

$$t = \frac{-0.011336 - 0}{0.004415} = -2.5674$$

Therefore, because -2.5674 < 1.3025 we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_5 = 0 \qquad H_1: \beta_5 \neq 0$

$$t = \frac{0.000158 - 0}{8.4\text{E} - 05} = 1.8810$$

Therefore, because 1.8810 > 1.6829, we reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_6 \geq 0 \qquad H_1: \beta_6 < 0$

$$t = \frac{0.000486 - 0}{0.000262} = 1.8550$$

Therefore, because 1.8550 > -1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_7 \geq 0 \qquad H_1: \beta_7 < 0$

$$t = \frac{173.8819 - 0}{192.8398} = 0.90169$$

Therefore, because 0.90169 > -1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_8 \geq 0 \qquad H_1: \beta_8 < 0$

$$t = \frac{-555.8598 - 0}{230.3938} = -2.4127$$

Therefore, because -2.4127 < -1.6829, we reject H_0 at the $\alpha = 0.05$ significance level.

Section 1: Part 2

Wald Test for All Independent Variables

$H_0: \beta_1 = \beta_2 = \cdots = \beta_8 = 0$

$H_1: Otherwise$

The related p-value, in Table 10 in Section 2 of Appendix B, is 0.000033. Since 0.000033 < 0.01, we reject H_0 at the $\alpha = 0.01$ significance level. The initial model is better than the null model.

Critical F-values
n = 50
df = 41
q = 4

	F-value
$\alpha = 0.01$	4.2986
$\alpha = 0.05$	2.8327
$\alpha = 0.10$	2.2225

Wald Test for Testing Independent Variables

$H_0: \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0 \qquad H_1: Otherwise$

$$F = \frac{\frac{R^2{}_{UR} - R^2{}_R}{q}}{\frac{1 - R^2{}_{UR}}{n - k - 1}} = \frac{\frac{0.545676 - 0.441370}{4}}{\frac{1 - 0.545676}{41}} = 2.3532$$

Therefore, because 2.3532 > 2.2225, we reject H_0 at the $\alpha = 0.10$ significance level.

Section 2

Hausman Test for Endogeneity

$H_0: \delta = 0 \qquad H_1: \delta \neq 0$

$$t = \frac{0.020383 - 0}{0.008434} = 2.4168$$

Refer to the table in Section 1 of Appendix C for the appropriate critical values. Because 2.4168 > 2.0195, we reject H_0 at the $\alpha = 0.05$ significance level.

RESET Test for Nonzero Mean

	F-value
$\alpha = 0.01$	4.3429
$\alpha = 0.05$	2.8517
$\alpha = 0.10$	2.2339

First Model

$H_0: \delta_0 = \delta_1 = \delta_2 = 0$

The functional form is correctly specified.

$H_1: Otherwise$

The functional form is incorrectly specified.

$$F = \frac{\frac{SSR_R - SSR_{UR}}{q}}{\frac{SSR_{UR}}{n-k-1}} = \frac{\frac{1079153 - 802509.22}{3}}{\frac{802509.22}{38}} = 4.3665$$

Since, 4.3665 > 4.3429, we reject the null hypothesis at an alpha level of 0.01. We need to correctly specify the functional form.

Tenth Model

$H_0: \delta_0 = \delta_1 = \delta_2 = 0$

The functional form is correctly specified.

$H_1: Otherwise$

The functional form is incorrectly specified.

$$F == \frac{\frac{SSR_R - SSR_{UR}}{q}}{\frac{SSR_{UR}}{n-k-1}} = \frac{\frac{11.76555 - 9.746172}{3}}{\frac{9.746172}{38}} = 2.62450$$

Since 2.62450 < 2.8517, we reject the null hypothesis at an alpha level of 0.05.

Breusch-Pagan-Godfrey Test for Heteroskedasticity

$H_0: Var(u \mid pop) \ = \sigma^2 \qquad Homoskedasticity$

$H_1: Var(u \mid pop) \ \neq \sigma^2 \qquad Heteroskedasticity$

Or equivalently:

$H_0: \delta_1 = 0 \qquad Homoskedasticity$

$H_1: \delta_1 \neq 0 \qquad Heteroskedasticity$

Refer to Appendix B for Computer Output. Since Prob. Chi Square(1) = 0.6763 > 0.10, we fail to reject the null hypothesis at an alpha level of 0.10. There is not enough evidence to claim heteroskedasticity in the error term is an issue.

Jarque-Bera Test for Normality

$H_0: The\ error\ term\ is\ normally\ distributed$

$H_1: The\ error\ term\ is\ not\ normally\ distributed$

Refer to Appendix E for Computer Output. Since the p-value from the Jarque-Bera test was 0.962705 > 0.10, we fail to reject the null hypothesis at an alpha level of 0.10. There is not enough evidence to claim a non-normally distributed error term is an issue.

Section 3: Part 1

Critical t-values:
n = 50
df = 41

	One-sided	Two-sided
$\alpha = 0.01$	2.4208	2.7012
$\alpha = 0.05$	1.6829	2.0195
$\alpha = 0.10$	1.3025	1.6829

$H_0: \beta_0 \leq 0 \qquad H_1: \beta_0 > 0$

$$t = \frac{\widehat{\beta_0} - 0}{se(\widehat{\beta_0})} = \frac{5.603660 - 0}{6.741835} = \ 0.831177$$

Therefore, because 0.831177 < 1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_1 \leq 0 \qquad H_1: \beta_1 > 0$

$$t = \frac{4.779017 - 0}{1.82228} = 2.622548$$

Therefore, because 2.622548 > 2.4208, we reject H_0 at the $\alpha = 0.01$ significance level.

$H_0: \beta_2 \leq 0 \qquad H_1: \beta_2 > 0$

$$t = \frac{-3.74\text{E} - 05 - 0}{2.85\text{E} - 05} = -1.312281$$

Therefore, because -1.312281 < 1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_3 \leq 0 \qquad H_1: \beta_3 > 0$

$$t = \frac{0.311808 - 0}{0.310012} = 1.005793$$

Therefore, because 1.005793 < 1.3025, we fail reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_4 \leq 0 \qquad H_1: \beta_4 > 0$

$$t = \frac{-1.291306 - 0}{0.750204} = -1.721273$$

Therefore, because -1.721273 < 1.3025 we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_5 = 0 \qquad H_1: \beta_5 \neq 0$

$$t = \frac{0.030625 - 0}{0.051534} = 0.594268$$

Therefore, because 0.594268 < 1.6829, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_6 \geq 0 \qquad H_1: \beta_6 < 0$

$$t = \frac{-0.021771 - 0}{0.179162} = -0.121516$$

Therefore, because -0.121516 > -1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_7 \geq 0 \qquad H_1: \beta_7 < 0$

$$t = \frac{-0.358326 - 0}{0.995036} = -0.360114$$

Therefore, because -0.360114 > -1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

$H_0: \beta_8 \geq 0 \qquad H_1: \beta_8 < 0$

$$t = \frac{-0.568746 - 0}{2.550351} = -0.223007$$

Therefore, because -0.223007 > -1.3025, we fail to reject H_0 at the $\alpha = 0.10$ significance level.

Section 3: Part 2

Wald Test for All Independent Variables

$H_0: \beta_1 = \beta_2 = \cdots = \beta_8 = 0$

$H_1: Otherwise$

The related p-value, in Table 10 in Section 2 of Appendix B, is 0.000115. Since 0.000115 < 0.01, we reject H_0 at the $\alpha = 0.01$ significance level. The corrected model is better than the null model.

Critical F-values
n = 50
df = 41
q = 4

	F-value
$\alpha = 0.01$	4.2986
$\alpha = 0.05$	2.8327
$\alpha = 0.10$	2.2225

Wald Test for Testing Independent Variables

$H_0: \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0 \qquad H_1: Otherwise$

$$F = \frac{\frac{R^2{}_{UR} - R^2{}_R}{q}}{\frac{1 - R^2{}_{UR}}{n-k-1}} = \frac{\frac{0.423576 - 0.288647}{4}}{\frac{1 - 0.423576}{41}} = 2.399314$$

Therefore, because 2.399314 > 2.2225, we reject H_0 at the $\alpha = 0.10$ significance level.

Section 4

Critical F-values

n = 50

df = 32

q = 9

	F-value
$\alpha = 0.01$	3.020818
$\alpha = 0.05$	2.188766
$\alpha = 0.10$	1.834831

Chow Test

$H_0: \delta_0 = \delta_1 = \delta_2 = \cdots = \delta_8 = 0$

$H_1: Otherwise$

$$F = \frac{\frac{SSR_P - (SSR_1 + SSR_2)}{k+1}}{\frac{SSR_1 + SSR_2}{n - 2(k+1)}} = \frac{\frac{14.43792 - (7.651164 + 0.872172)}{9}}{\frac{7.651164 + 0.872172}{32}} =$$

$$\frac{\frac{5.914584}{9}}{\frac{8.523336}{32}} = 2.467301$$

Therefore, because 2.467301 > 2.188766, we reject H_0 at the $\alpha =$ 0.05 significance level.