The GETTYSBURG ECONOMIC REVIEW

Volume 10

Spring 2017

Department of Economics
Gettysburg College
Gettysburg, Pennsylvania 17325

The Economics Department and Omicron Delta Epsilon congratulate **Phoebe Do**, winner of the *2017 Dwight D. Eisenhower Society / R.M. Hoffman Family Memorial Prize in Economics*. The Eisenhower/Hoffman Prize is awarded to the economics student writing the best quantitative paper or project with public policy implications.

The Economics Department and Omicron Delta Epsilon congratulate **Savannah Morrissey Martin,** winner of the *2017 Best Thesis Award*.

The Economics Department and Omicron Delta Epsilon congratulate **Delia Craig**, winner of the *2017 Glatfelter Prize*, awarded to one student with junior standing possessing excellent scholarship in the social sciences.

The Economics Department and Omicron Delta Epsilon congratulate **Jack Gardner** for being selected as a *2017 Kolbe Fellow*.

The Economics Department and Omicron Delta Epsilon congratulates **Jennifer Flores** and **Polina Rozhkova**, winners of the 2017 *Dr. and Mrs. William F. Railing Fellowship for Faculty-Student Research in Economics*.

The Economics Department and Omicron Delta Epsilon congratulates **Anneliese Brown**, winner of the *2017 John Edgar Baublitz Pi Lambda Sigma Award*.

The Economics Department and Omicron Delta Epsilon congratulate **Phoebe Do** and **Nicholas Papoutsis** for their induction into Phi Beta Kappa. Phi Beta Kappa celebrates and advocates excellence in the liberal arts and sciences. Its campus chapters invite for induction the most outstanding arts and sciences students at America's leading colleges and universities.

The Economics Department and Omicron Delta Epsilon congratulate the following students for their achievements in the 2016-2017 academic year:

Economics Graduation Banner Carriers:

BA: Eric Brown
BS: Nicholas Papoutis

2017 Economics Honors Graduate:

Anna Brown
Phoebe Do
Davis Healy
Tom Segerstrom
Savannah Morrissey-Martin (F'17)

Omicron Delta Epsilon would also like to thank our outgoing officers, **Eric Brown** and **Haya Mohanna**.

CONTENTS

The Rise of Technology and its Influence on Labor Market Outcomes

Maja Thomas

Abstract: Technological progress has significantly changed the inputs and production processes utilized by firms. Such shifts have led to warnings throughout the past few decades that substantial numbers of jobs, particularly things belonging to the middle class, would be eliminated and replaced by technology. This paper examines the validity of this argument by estimating the impact of technology investment on local labor markets during that period. I find evidence for a positive, rather than negative, relationship between technology and employment. Furthermore, my estimates suggest there exists a complementary relationship between technology investment and growth in labor opportunities, rather than a substitution effect of workers moving from technology-intensive industries to non-technology intensive sectors

Introduction

The rise of technology, specifically robotics and computerization, has dramatically shifted the inputs available to businesses over the past several decades. This rapid development has transformed the production processes for many different industries. Many fear that this technological development has increased automation while not adding enough jobs to offset the drop in opportunities. If true, this decrease in the employment capacity would negatively impact the wages and incomes of many workers, namely middle skill white collar and blue collar labors performing easily codifiable tasks (Autor 2011).

The subject of automation and its expansion in recent decades has ignited fears and frustrations over its threat of making many traditional jobs obsolete. Automation has been used as anecdotal evidence to explain claims of declining productivity, employment, and the current economic slow growth. The impact of automation has discriminately hit certain industries and job types, most of which are middle-paying and moderate-skilled, while straying away from others (Autor 2011). Much of this is because automation is only viable for certain job types, most of which are middle-paying and moderate-skilled. The core tasks of these positions often require employees to follow precise methodical procedures which machines are well equipped to perform. But is the rise in technology to blame for labor market perils or has it simply provided a digestible narrative?

Per neoclassical theory, investment would actually increase labor demand due to the complementary nature of labor and capital. However, many economists consider ICT capital differently, worrying that investment would decrease the demand for labor by increasing productivity of labor. Nevertheless, with spillover effects on other industries, incomes, or aggregate demand (and thus output), the impact of ICT investment is difficult to assess per traditional theory.

Thus, this paper answers this question empirically, examining the impact of technology investment on local labor markets. In the next section, I discuss the influence of robotics and job automation on employment dynamics. In section III, I develop an econometric model to analyze the relationship between increases the level of information and communications technology investment within a commuting zone and the expected level of employment in that county. In section IV, I discuss the data collected to test my hypothesis, and in section V, I use that data to test my hypothesis and find evidence for a positive relationship between investment and employment.

I. Job Automation and Labor Market Demand

The interaction between job automation and labor market dynamics has attracted significant attention from both economists and scholars alike. With vast technological advances occurring in computing and robotics, machines have now become as or more efficient than human workers in various environments. Without a clear consensus, economists continue to question automation's bearing on the labor market while the public remains largely in fear.

A negative relationship between the levels of ICT investment and employment would hardly be surprising. Since the rise of machines and machine learning, many have feared that robots would replace human labor, leading to employment losses. Straying from traditional neoclassical framework, economists tend to view ICT investment as a substitute for labor rather than a compliment. In this case, demand for labor would decrease, thereby reducing employment. Furthermore, while decreases in job opportunities due to automation could hypothetically be made up by increases in job opportunities in other industries, labor may not be able to shift into these new opportunities due to a lack of experience or other structural problems, thus leading to structural unemployment and an overall decline in employment. These findings would uphold implications from the Solow growth model, where an increase in technological investment increases labor productivity (i.e. output per worker). Ceteris paribus, firms would need less employees and would be incentivized to cut jobs.

This negative relationship between technology investment and labor has been documented by different parts of the literature. Robots and automated systems have negatively impacted several occupations, almost entirely eliminating elevator operators, highway toll collectors, parking attendants, and other similar roles (Quereshi and Syed 2014). Qureshi and Syed found that in the health care industry, 19 Aethon TUG robots can perform $1 million in human labor each year for $350,000, saving the industry 65% in labor costs. Robots such as these, in working two shifts seven days per week, save the labor of 2.8 full time equivalent (FTE) employees while costing less than one. Ebel (1986) also noted the labor costs savings by employing robots. Robots in the automotive industry costs around $6 per hour including depreciation and maintenance costs, compared with between $23 and $24 an hour in wages and benefits for an employee.

Contradicting evidence would demonstrate either no significant relationship or a positive relationship between the commuting zone levels of ICT investment and employment. If there were no significant relationship between the dependent and independent variables, job losses would either not result from automation or losses would be made up by gains in other sectors or occupations. If there were a significant and positive relationship, commuting zone job growth would result from ICT investment due to aggregate demand effects. This would align with neoclassical theory, which states that an increase in ITC investment would increase labor demand because capital and labor are complements. Higher investment would increase production, leading to an increase in income and increase the demand for goods and services, overall employing more individuals to produce these goods and services. Additionally, if demand for output increased because of the technological investment, a decrease in employment resulting from increases in labor productivity would be offset by an increase in labor demanded to increase total output. Even if ICT investment and labor were substitutes, there could be spillover effects (i.e. increases in demand for labor in related industries, impacts of increased income or aggregate demand, etc.) which could increase employment overall.

Other parts of the literature have found ICT investment to have had a non-negative impact on the labor market, largely due to spillover effects of ICT investment. Autor (2015) found that automation had not led to significant job losses, citing that the interaction between technology and employment required ingenuity and creative thinking that cannot be adequately computerized. Autor (2011) detailed growing labor market opportunities for both high skill, high-wage and low skill, low-wage white and blue collar industries, as a result of automation-led wage-level occupational shifts. As computer and robotics technologies progressed, machines were well equipped to perform core job tasks of middle skilled industries. However, this has caused various spillover effects and led to increases in opportunities in other sectors, and likely triggered dramatic growth in service occupations as detailed by Autor and Dorn (2013). Such also appeared the case during the early 2000s, where Charles, Hurst, and Notowidigo (2016) found that the declines in the manufacturing industry were propped up by the growth in the housing sector, which benefitted from the decreases in construction costs and increases in building efficiency. Leontif and Duchin (1984) forecasted the intensive use of automation the twenty years following 1985, estimating it would conserve about 10% of the labor force required to produce the same goods. However, their models predicted an increase in the output level which would offset the effects of job displacement, finding a complementary relationship between investment and employment as would the neoclassical framework. Furthermore, they argued the impacts would involve a significant increase in professional employees and a steep decline in the relative number of clerical workers as a proportion of the labor force.

An even smaller proportion of the literature has found no relationship between ICT investment and the labor market. Doms, Dunne, and Troske (1997) found that time series results demonstrated little correlation between the adoption of technology and changes in workforce characteristics. The adoption of new technologies did not appear to impact a factory's relative share of non-production labor or high wage workers, as compared to plants which did not adopt new technologies. This relationship between factory automation technologies and employment of highly paid workers was further established by Dunne and Schmitz (1995) and Siegel (1995).

Thus, the impact of ICT investment on labor markets could reasonably be either positive or negative. This paper aims to answer the empirical question of ICT investment's impact on the change in employment, differing from the above literature which addresses similar questions utilizing historical data and qualitative methods. Further information on the model is detailed in the next section.

II. Modeling

I test whether information and communications technology investment in a commuting zone affects the level of employment in that commuting zone using methods similar to those of Autor et al (2015). Commuting zones are clusters of US counties characterized by strong within-cluster and weak between-cluster ties that have been compiled by the Economic Research Service in 1990. The average level of information and communications technology investment is computed annually over the course of two eight year periods: 1992-1999 and 2000-2007.

The benchmark regression can be written as follows:

$$EMPLOY_{i,t} = \beta_0 + \beta_1 INVEST_{\mathrm{it}} + \beta_2 YEAR_{it} + \beta_3\, REGION_{it} + \mu_{it}$$

where;

- EMPLOY measures the level of employment within each commuting zone as a percentage of total employment;
- INVEST represents the average level of information and communications technology investment over two eight year periods, 1992-1999 and 2000-2007, respectively, as a percentage of total investment;
- YEAR is a dummy variablse controlling for differences in employment growth among the two eight-year periods;
- REGION is a vector of dummy variable controlling for differences in employment among census divisions;
- μ is the error term.

From the above regression, the null hypothesis for this model can be written as follows:

H_0: $\beta_1 \geq 0$

An increase in the level of information and communications technology investment within a commuting zone does not negatively impact the level of employment in that commuting zone.

H_A: $\beta_1 < 0$

An increase in the level of information and communications technology investment within a commuting zone negatively impacts the level of employment in that commuting zone.

III. Data

The data in this study comes from the European Union level analysis of Capital, Labor, Energy, Materials, and Service (EU KLEMS) and David Autor, Daron Acemoglu, and David Dorn. The unit of analysis in this data set is commuting zone-year (e.g. commuting zone 100-2007) and the data is compiled in the years 1991-1999 and 2000-2007. The EU KLEMS data measures information and communication technology investment and is part of a larger dataset which includes other variables related to capital, labor, and output from the 1970s to 2007. The Autor et al dataset was the focus of their 2015 paper and includes commuting zone-level data on employment and import penetration in the years 1991, 1999, 2007, and 2011. The data used in this analysis includes their 722 commuting zones and encompasses the entire mainland United States for the years 1999 through 2007. These commuting zones are clusters of counties with strong internal commuting ties (Autor 2014). The data sets utilized in creation of this study are codified by industry and year. Autor employs SIC codes to signify industry type, while EU KLEMS uses broad sector categories. Thus, to combine the data sets, I recode all SIC codes into broad sector categories for ease of merging.

My dependent variable is the change in commuting zone employment. As noted above, commuting zones are clusters of US counties characterized by strong within-cluster and weak between-cluster ties that have been compiled by the Economic Research Service in 1990. Employment is defined as the number of employees who are on payroll in the pay period in

March of each year. Paid employees consist of full time employees, part time employees, employees on sick leave, holidays, or vacations. The data used to construct this variable come from David Autor and the County Business Patterns series from the United States Census. I utilize industry level employment data within each commuter zone and year and manipulate it to construct my dependent variable. I start by finding total employment within each commuting zone by coding a new variable adding each industry together within a commuting zone and removing duplicate observations, leaving only commuter zone and year. This value is then divided by number of working age individuals in each commuter zone to construct an employment-population ratio. I then construct a new variable measuring the change in the employment population ratio for my two years, 1991-1999 and 2000-2007, which will represent 1999 and 2007, respectively. The data includes 1444 observations ranging from -.093% to 2.697% of total employment across all commuting zones.

The independent variable in this study is the percentage of information and communication technology, as a share of total investment, within a commuting zone. Information and communications technology (ICT) is a broad category of technology and can be used as a proxy for robot-type capital. Calculation of ICT capital is based on the database described in Jorgenson, Ho, and Stiroh (2005) and sourced from EU KLEMS. The independent variable is constructed by taking the eight-year average of EU KLEMS' ICT as a percentage of total investment from years 1991-1999 and 2000-2007. Next, I create a variable representing employment share of each industry within each commuting zone by dividing industry employment by total employment within the commuting zone. I then multiply the average ICT investment by employment share. Finally, I sum the industries to create a weighted average of ICT investment in each commuting zone and eight-year period. The finalized variable includes

1444 observations ranging between 9.57% and 23.21% of total investment across all commuting zones. The correlation coefficient between the independent and dependent variables is -0.2718, demonstrating a negative relationship between ICT investment and employment and following the narrative that increases in automation remove jobs from the labor market without adding sufficient new opportunities.

Nine control variables are utilized in this model: one dummy variable accounting for changes in employment level due to time period and eight other dummy variables accounting for changes in employment level due to geographic region (see Figure 1). These variables are coded either '0' or ‘1’. The year dummy is coded ‘1’ for observations which take place in 1999 and ‘0’ for observations in 2007. Each regional dummy is coded ‘0’ if the commuting zone is not part of that geographic region and ‘1’ if it is. No commuting zone can belong to more than one geographic region. The Mountain region is omitted in the regression analysis, leaving a variable to compare the other regions to. A summary of all variables and their respective descriptive statistics can be seen in the appendix in Table 1.

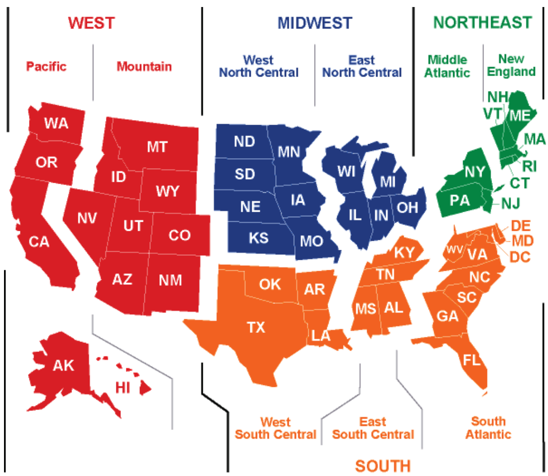


Figure 1: US Census Divisions

Figure courtesy of the US Energy Information Administration (eia.gov)

IV. **Findings**

V. Appendix Tables 3 and 4 display the results of the models constructed in this paper; that is, the impact of an increase in the level of ICT investment within a commuting zone on the expected level of employment in that county using an ordinary least squares (OLS) regression. At a first glance, there is a substantially negative relationship between the two, as seen in the scatterplot in Figure 2.

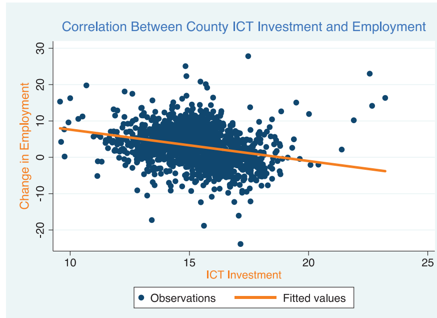
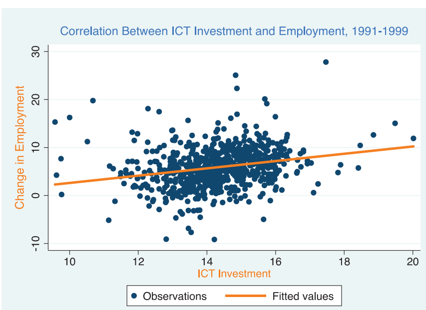


Figure 2: ICT Investment and Employment Relationship

The correlation coefficient is -.2718, again demonstrating the negative relationship between the ICT investment and the change in employment. However, as you I add in control variables such as year, there emerges, if anything, a positive relationship. This is supported by the results of the scatterplots on the left in Figures 3 and 4, where the data is separated out by year. In the period from 1991 to 1999, there exists a positive relationship (correlation coefficient of 0.2373) between the level of ICT investment and the change in employment, which is likely due to the economic boom of the 1990s.



Then, through the 2000s the relationship becomes slightly negative (correlation coefficient of -.0809) and less uniform as the market gears up for the Great Recession. Additionally, the summary statistics (see Table 2) show a higher average change in employment during the 1991-1999 period (5.90% versus 0.22%) and lower average ICT investment levels in 1991-1999 than the following eight year period (14.32% versus 16.21%).

While the first glance correlation coefficient supports my hypothesis, the first OLS model does not; I therefore fail to reject the null hypothesis and cannot conclude that there exists a negative relationship between ICT investment and employment. The results of the model (see Table 3a) indicate that there is actually a positive relationship between ICT investment and employment, although they are not significant at the 5% level. But let us not fetishize the 5% level—with a p-value of 0.063 we hold reasonably the same assurance in the coefficient as we would if it were 0.05 or under. These findings suggest that a one-percent increase in the level of ICT investment within a commuting zone, as a percentage of total investment, would lead to a 0.168% increase in the expected change in employment-population ratio in that commuting zone. These findings dispel fears of technological unemployment and the narrative of robots taking human jobs, proving consistent with the complementarities between ICT investment and human labor. However, the small size of the coefficient and borderline

Figure 3: ICT Investment and Employment Relationship, 1991-1999

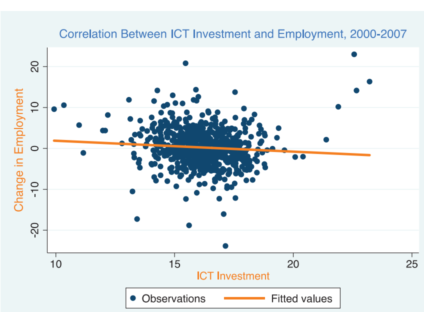


Figure 4: ICT Investment and Employment Relationship, 2000-2007

significance of its p-value may also be in accordance with Autor's (2015) findings that there exists no significant negative relationship between automation and job losses.

I implement various controls for year and region in the model. The regions are comprised of the following divisions: New England, Mid-Atlantic, south Atlantic, East North Central, West North Central, East South Central, West South Central, and Pacific. Of the nine control variables tested in this model, eight are significant below the 5% level: year, New England, Mid Atlantic, South Atlantic, East North Central, South North Central, West South Central, and Pacific. All control variables hold negative coefficients except year. This relationship between year and employment supports the results of the earlier correlation coefficients and scatter plots, suggesting that employment was expected to be 6% higher in the period from 1991-1999, regardless of region or ICT investment level.

From the results of the first model, I create a second model to include Autor's (2014) import penetration variable to account for differences arising from trade, and assess whether it was an important omitted variable in the first model (see Table 3b). Upon running the mode, I find that the change in import penetration, while significant and negative (as in Autor's findings), does not substantially change the ICT investment coefficient. The coefficient lowers slightly to 0.160 and keeps significance at the 10% level. Thus, I conclude there exists no problem of omitted variables present within the first model.

Next, I construct models which estimate the relationship between the dependent and independent variables in one of the two eight-year periods, to see if the relationships implied by the scatterplots and correlation coefficients hold true that there are differing impacts on the relationship between ICT investment and employment which are dependent on the eight-year period investigated. My first model utilized data only during the 1991-1999, and the results demonstrated a strongly significant and positive correlation between the two variables with a correlation coefficient of 0.566. The results of the 2000-2007 model, however, were negative and insignificant, even at the 10% level. Thus, the models demonstrate that the gains from ICT investment were to be made during the 1990s but did not last not through the 2000s, when the overall employment population ratio tumbled due to the 2001 recession.

Finally, I construct three models to allocate the 27 broad sector industries in each commuting zone into three categories: ICT intensive investment, moderate ICT intensive investment, and non-ICT intensive investment. From the year-commuting zone-industry stage of my data manipulation, I identify the top 9 industries by computing the simple average of the average ICT investment over the two periods, constructing one value from 1991-2007. Then, I compute the total employment in each commuter zone for each bracket, leaving 6,498 observations and three new variables corresponding to each ICT investment level. Finally, I find the change in employment for the two periods and drop the 1991 values from the data set. More information on the industry breakdown and their respective summary statistics can be found in Tables 5-7.

The results of the three ICT models (Table 4a-c) suggest that increased ICT investment positively impacts ICT intensive segments while negatively impacting non-ICT intensive industries. The ICT intensive model demonstrates a positive and strongly significant relationship between the two variables, suggesting that a one percent increase in ICT investment will increase expected employment by .42%. This result further demonstrates the complementarity of ICT investment to the labor market, particularly its addition to ICT intensive industries. On the other hand, the expected relationship between ICT investment in non-ICT intensive industries and employment is significant and negative, with a coefficient of -.21%. This disproves the idea that the increase in employment in the first model was the result of a substitution effect in non-ICT intensive industries. The moderate ICT investment model is insignificant, with a near-zero coefficient that implies no definite relationship between ICT investment and employment. This coefficient is in line with the results of the other two models because of the complementary relationship between intensiveness and employment and substitute relationship between non-intensiveness and employment.

However, the results of the three categorical models may indicate an omitted variables bias problem in the models. If an industry category—ICT intensive, for example—expands, companies may concurrently hire more employees and invest in ICT. In this case, the relationship between ICT investment and change in employment would necessarily be causal, but a response to a third variable which is driving expansion in that sector. Instituting an additional variable to control for this difference would solve this potential problem, but I could not conceive of any measurable instruments to utilize in the model. Thus, further research should attempt to correct for hypothetical bias by using an instrument correlated with ICT investment and not directly linked with employment in those industries.

I was unable to account for all possible influences on level of commuting zone employment which could misconstrue the relationship between the dependent variable and commuter zone ICT investment. Particularly, there is no control for the type of industry employment or the makeup of commuter zone employment in the first model, and the three models which consider industries only do so using intensive, moderately intensive, and non-intensive ICT brackets. However, it is unclear whether the addition of this variable would actually significantly impact the results of the model, and there would exist difficulties in coding this variable for all industries included in the initial dataset. Additionally, research conducted by

Autor et al (2015) did not find industry to have a significant impact in their model. Nevertheless, while the model demonstrates a significant relationship between the dependent and independent variables, there could exist an omitted variable or variables which impact the findings of the model.

As ICT investment is a relatively broad category of technology, further research may be needed to look specifically at the impacts of robotics and possible resulting job automation. In the creation of this model, ICT investment appears to be an adequate proxy for robotics. However, it may be that another indicator of robotics development could have been better served to estimate the model, as it would analyze the funding on specifically technologies which could be used to automate tasks. Additionally, further research should aim to include a larger number of years so as to compute both the change in employment and change in investment. This would allow the model to analyze the impacts of increasing investment in ICT technologies on employment rather than average level. Using an independent variable measuring its change, would, regardless of impact, have more straightforward policy implications.

VI. Concluding Remarks

Job automation and its growth in recent decades have awakened suspicions and frustrations over their risk of making many traditional jobs obsolete and decreasing employment opportunities for the newly jobless. Yet, according to the results of the model, this does not seem to be the case. The findings from this paper challenge my hypothesis of a negative relationship between the dependent and independent variables, instead suggesting that an increase in the level of ICT investment within a commuting zone, as a percentage of total investment, would lead to an increase in the expected employment population ratio in that commuting zone. These results are significant at the 10% level with a p-value of 0.063. Thus, the findings ultimately indicate that ICT investment leads to increased employment.

From these findings, policy recommendations are less than straightforward; the first model dictates that increasing ICT investment would push employment in commuter zones, but due to differences in the two time periods tested and the negative and insignificant coefficient in the third and fourth models, implications for the current slow growth era may be not be effective. However, the differences may be due to the 2001 recession and decrease in growth. Thus, further research is recommended to determine whether periods of slow growth can receive the employment benefits of ICT investment. This paper does not attempt to define the correct limit of spending nor does it serve to understand the optimal distribution of ICT investment by industry. What this paper does, however, is dispel fears of a negative relationship between the two variables.

The US labor market remains a major source of discussion, particularly as the economy has been plagued by slow growth. While the official unemployment rate was 4.9% as of October 2016, the labor force participation rate and employment-population ratio remain far below pre-2007 levels. A struggling labor market in the aftermath of recession and dramatic rise in technology have caused many to couple the two together, and fear that technological developments have contributed to unemployment rates. However, the use of technology appears to be a scapegoat for other issues putting downward pressure on the labor market. The rise of the service sector, as noted by Autor and Dorn (2013) has allowed another outlet for American workers. The results of the models tested in this paper, however, demonstrate a complementary relationship between ICT investment and growth in labor opportunities, rather than a substitution effect of workers moving from ICT-intensive industries to non-ICT intensive sectors. Thus, the public should embrace—rather than fear—information and communication technology investment as a way in which to spur growth and expand labor market opportunities.

VII. Appendix

Table 1: Summary of Variables

Variable	Description	Observations	Source
Employment	Employment within czone as percentage of total employment	1444 observations 1999, 2007	Autor et al.
ICT Investment	Average level of ICT investment as percentage of total investment over eight year periods	1444 observations 1992-1999, 2000-2007	EU KLEMS
Year	Dummy variable representing either 1999 ('0') or 2007 ('1')	1444 observations 1999, 2007	Autor et al/EU KLEMS
New England Division	Dummy variable representing New England czones	1444 observations 1999, 2007	Census Bureau County Business Patterns
Mid-Atlantic Division	Dummy variable representing Mid-Atlantic czones	1444 observations 1999, 2007	Census Bureau County Business Patterns
East North Central Division	Dummy variable representing East North Central czones	1444 observations 1999, 2007	Census Bureau County Business Patterns
West North Central Division	Dummy variable representing West North Central czones	1444 observations 1999, 2007	Census Bureau County Business Patterns
East South Central Division	Dummy variable representing East South Central czones	1444 observations 1999, 2007	Census Bureau County Business Patterns
West South Central Division	Dummy variable representing West South Central czones	1444 observations 1999, 2007	Census Bureau County Business Patterns

Pacific Division	Dummy variable representing Pacific czones.	1444 observations 1999, 2007	Census Bureau County Business Patterns

Table 2: Summary of Independent and Dependent Variables by Year

Variable	Mean	Standard Deviation	Minimum	Maximum
ICT investment 1991-1999	14.32083	.28453	9.5747	20.0174
ICT investment 2000-2007	16.21238	1.396903	9.9275	23.2145
Change in employment, 1991-1999	5.90318	4.114131	-9.162831	27.81029
Change in Employment, 2000-2007	.2150155	4.606847	-23.85641	22.99899

Table 3: Regression Analysis: ICT Investment Across All Levels

Variable	(a) OLS regression Change in commuting zone employment	(b) OLS regression Change in commuting zone employment	(c) OLS regression Change in commuting zone employment in 1999	(d) OLS regression Change in commuting zone employment in 2007

IT investment in commuting zone	.167 (.090)	.160 (.089)	.566*** (.124)	-.116 (.119)
Year	6.01*** (.279)	5.407*** (.287)		
Import penetration		-.971*** (.136)		
New England	-1.67* (.813)	-1.049 (.804)	.833 (1.051)	-4.251*** 1.126
Mid Atlantic	-3.201*** (.678)	-2.389*** (.676)	-3.292*** (.878)	-3.253** (.938)
South Atlantic	-3.152*** (.419)	-2.443*** (.424)	-.732 (.542)	-5.572*** (.580)
East North Central	-2.728*** (.456)	-2.017*** (.459)	.753 (.592)	-6.343*** (.629)
West North Central	-.239 (.400)	.181 (.398)	1.142* (.525)	-1.840** (.548)
East South Central	-2.756*** (.464)	-1.439** (.493)	-.384 (.600)	-5.141*** (.643)
West South Central	-1.471*** (.418)	-1.080** (.415)	-1.317* (.542)	-1.735** (.577)
Pacific	-2.487** (.541)	-2.279*** (.533)	-2.710*** (.701)	-2.307** (.749)
Constant	-0.688** (1.439)	-.213 (1.417)	-1.941 (1.732)	5.224** (1.896)
N	1444	1444	721	722
R^2	.358	.380	0.142	0.218

*p<.05; **p<.01; ***p<.001

Table 4: Regression Analysis: ICT Investment Across All Levels

Variable	(a) OLS regression Change in commuting zone employment for high ICT industries	(b) OLS regression Change in commuting zone employment for mid ICT industries	(c) OLS regression Change in commuting zone employment for low ICT industries
IT investment in commuting zone	.425*** (.060)	-.056 (.041)	-.209*** (.0459)
Year	3.773*** (.185)	1.77*** (.129)	.456*** (.142)
New England	-1.214* (.539)	-.373 (.375)	-.128 (.413)
Mid Atlantic	-1.978*** (.449)	-.965** (.313)	-.317 (.345)
South Atlantic	-.868** (.278)	-.719*** (.193)	-1.560*** (.213)
East North Central	-1.704*** (.302)	-1.106*** (.210)	.115 (.232)
West North Central	-.928*** (.265)	.171 (.185)	.558** (.203)
East South Central	-1.223*** (.308)	-.422* (.214)	-1.098*** (.236)
West South Central	-1.163*** (.277)	-.449* (.193)	.143 (.212)
Pacific	-1.087** (.359)	-.871*** (.250)	-.484 (.274)
Constant	-5.77*** (.953)	1.561* (.664)	3.459 (.731)
N	1444	1444	1444
R^2	.257	.225	0.137

*p<.05; **p<.01; ***p<.001

Table 5: ICT Level Industry Breakdown: ICT Intensive Industries

ICT-Intensive Industry Name	Broad Sector Code	Average ICT Investment, 1991-1999	Average ICT Investment, 2000-2007	Average ICT Investment, 1991-2007
Transport and storage	26	0.229	0.374	0.360
Education	35	0.300	0.349	0.344
Electrical and optical equipment	15	0.238	0.345	0.335
Machinery, nec	14	0.244	0.308	0.302
Financial intermediation	29	0.297	0.248	0.253
Wholesale trade and commission trade	22	0.226	0.246	0.244
Transport equipment	16	0.204	0.239	0.235
Construction	19	0.138	0.205	0.198
Community social and personal services	33	0.165	0.178	0.176

Table 6: ICT Level Industry Breakdown: Moderate ICT Intensive Industries

Moderate ICT Industry Name	**Broad Sector Code**	**Average ICT Investment, 1991-1999**	**Average ICT Investment, 2000-2007**	**Average ICT Investment, 1991-2007**
Pulp, paper, paper, printing and publishing	7	0.132	0.170	0.166
Manufacturing nec; recycling	17	0.163	0.166	0.166
Health and social work	36	0.149	0.153	0.152
Chemicals and chemical products	10	0.135	0.146	0.145
Retail trade, repair of household goods	23	0.124	0.132	0.131
Sale, maintenance and repair of motor vehicles and motorcycles	21	0.129	0.115	0.117
Basic metals and fabricated metal	13	0.101	0.102	0.102
Coke, refined petroleum and nuclear fuel	9	0.097	0.099	0.099
Other non-metallic mineral	12	0.089	0.094	0.093

Table 7:: ICT Level Industry Breakdown: non-ICT Intensive Industries

Non-ICT Intensive Industry Name	**Broad Sector Code**	**Average ICT Investment, 1991-1999**	**Average ICT Investment, 2000-2007**	**Average ICT Investment, 1991-2007**
Food, beverages and tobacco	4	0.076	0.091	0.090
Textiles, textile, leather and footwear	5	0.065	0.091	0.088
Real estate, renting and business activities	30	0.068	0.073	0.072
Electricity, gas and water supply	18	0.062	0.070	0.069
Wood and of wood and cork	6	0.059	0.066	0.065
Rubber and plastics	11	0.045	0.061	0.059
Hotels and restaurants	24	0.044	0.050	0.049
Mining and quarrying	2	0.061	0.040	0.042
Agriculture, hunting, forestry and fishing	1	0.014	0.018	0.018

VIII. References

Autor, David H. "The polarization of job opportunities in the US labor market: Implications for employment and earnings." *Center for American Progress and The Hamilton Project* (2010).

Autor, David H. "Why are there still so many jobs? The history and future of workplace automation." *The Journal of Economic Perspectives* 29.3 (2015): 3-30.

Autor, David H, et al. *Import competition and the Great US Employment Sag of the 2000s*. No. w20395. National Bureau of Economic Research. (2014).

Autor, David H., and David Dorn. "The growth of low-skill service jobs and the polarization of the US labor market." *The American Economic Review* 103.5 (2013): 1553-1597.

Autor, David H., Lawrence F. Katz, and Alan B. Krueger. *Computing inequality: have computers changed the labor market?* No. w5956. National Bureau of Economic Research (1997).

Berman, Eli, John Bound, and Zvi Griliches. "Changes in the demand for skilled labor within US manufacturing: Evidence from the annual survey of manufacturers." *The Quarterly Journal of Economics* (1994): 367-397.

Berndt, Ernst R., Catherine J. Morrison, and Larry S. Rosenblum. *High-tech capital formation and labor composition in US manufacturing industries: an exploratory analysis*. No. w4010. National bureau of economic research, 1992.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo. "The Masking of the Decline in Manufacturing Employment by the Housing Bubble." *The Journal of Economic Perspectives* 30.2 (2016): 179-200.

Doms, Mark, Timothy Dunne, and Kenneth R. Troske. "Workers, wages, and technology." *The Quarterly Journal of Economics* (1997): 253-290.

Dunne, Timothy, and James A. Schmitz Jr. "Wages, employment structure and employer size-wage premia: their relationship to advanced-technology usage at US manufacturing establishments." *Economica* (1995): 89-107.

Ebel, Karl H. "Impact of Industrial Robots on the World of Work, The." *Int'l Lab. Rev.* 125 (1986): 39.

Leontief, Wassily, and Faye Duchin. "The Impacts of Automation on Employment, 1963-2000. Final Report." (1984).

Qureshi, Mohammed Owais, and Rumaiya Sajjad Syed. "The impact of robotics on employment and motivation of employees in the service sector, with special reference to health care." *Safety and health at work* 5.4 (2014): 198-202.

Siegel, Donald S. "The impact of technological change on employment: evidence from a firm-level survey of Long Island manufacturers." *Economics of Innovation and New Technology* 5.2-4 (1998): 227-246.

Sullivan, Daniel, and Till Von Wachter. "Job displacement and mortality: An analysis using administrative data." *The Quarterly Journal of Economics*(2009): 1265-1306.

The Impact of Aid on the Economic Growth of Developing Countries (LDCs) in Sub-Saharan Africa

Maurice Phiri

Abstract: Least Developed Countries (LDCs) of Sub-Saharan African have been recipients of official development assistance for more than 5 decades; however they are still characterized by chronic problems of poverty, low living standards and weak economic growth. The hot question is: Is aid effective in promoting economic growth? Thus, this paper investigates the impact of aid on the economic growth of 12 least developed countries in Sub-Saharan Africa over a period of 20 years. I take a fixed effects instrumental variable approach and the results imply that aid has a statistically insignificant negative impact on economic growth. I therefore conclude that aid is ineffective in promoting growth, perhaps due to misallocation of aid or inefficient use.

1. Introduction

The fundamental role of foreign aid, given in the form of loans and grants, is to mitigate poverty and promote economic growth in developing countries. However, the results of official development assistance (foreign aid) have not universally met the fundamental objective of aid in different countries (Lohani 2004). According to Dambisa Moyo, Zambian economist and author of *Dead Aid*,

> Over the past 60 years at least $1 trillion of development-related aid has been transferred from rich countries to Africa. Yet real per-capita income today is lower than it was in the 1970s, and more than 50% of the population -- over 350 million people -- live on less than a dollar a day, a figure that has nearly doubled in two decades" (Moyo 2009).

Proponents of aid argue that aid has a positive impact on economic growth for the following reasons: 1) aid supplements domestic savings and capital formation; 2) it can close the foreign exchange gap (Fayissa and El-kaissy, 1999). 3) In Askarov and Doucouliagos' 2015 study, (cited in Morrissey 2001), "Aid can increase investment in physical and human capital. 4) Aid is also associated with technological transfer that increases capital productivity and promotes endogenous technical change."

On the other hand, opponents of aid argue that foreign aid is ineffective in Africa for several reasons including: 1) it comes at a cost and heavily in debts African governments; 2) it perpetrates corruption when aid is given to corrupt governments; 3) it increases dependency syndrome and weakens governments' efforts of collecting revenue; 4) large inflows of foreign currency can strengthen the recipients' domestic currency and raise its export prices, in turn making the country less competitive in the global market (Moyo 2009).

Furthermore, prior research on the impact of aid on economic growth is not unanimous. Hansen and Tarp (2000) found that effectiveness of aid is dependent on human capital and investment. Malik (2008) found that aid is not effective in the short run and has a negative effect on growth in the long run. Minoiu and Reddy (2009) found that effectiveness of aid is conditional on whether the aid is developmental or not. Also, there are several common challenges that face the empirical investigations of the effectiveness of aid including: 1) accounting for the lagged effect of aid on growth; 2), properly accounting for the two-way causal relationship between aid and growth and 3), properly controlling for the underlying heterogeneity of countries used in regression analysis (Askarov and Doucouliagos 2015). The study of the effectiveness of aid on economic growth is important because it can help donor countries and aid recipients understand how aid can be effectively used to alleviate poverty and attain sustainable economic growth in the least developed countries of Sub-Saharan Africa.

The results of my study support the argument that aid is ineffective for economic growth in least developed countries of Sub-Saharan Africa. For example, after correcting for problems like time fixed effects, heteroscedasticity, unit roots and endogeneity in my model, a percentage increase in net official development assistance (ODA) is associated with a 0.03% decrease in real gross domestic product (GDP); this is not statistically different from 0. However, real total factor productivity and capital accumulation have one of the largest statistically significant impacts on real GDP and therefore I argue that proper allocation of aid in the economy makes aid very effective.

The rest of the paper is organized as follows: section 2 discusses existing literature and my contribution to it. Section 3 gives an overview of the methods I have used in this study, while section 4 explains where I got my data and describes the nature of the data set used in this study. A discussion of my analysis and interpretation of my results is given in section 5 and finally, section 6 discusses my conclusion based on the empirical results of this paper.

2. Literature Review

Prior empirical economic literature on the relationship between aid and growth in developing countries is mixed. Mallik (2008) uses co-integration analysis to study the relationship of foreign aid and economic growth of the poorest six African countries. In 5 out the 6 countries, Mallik found aid has no significant effect on growth in the short run, while there is a significant negative relationship between aid and growth in the long run.

Hansen and Tarp (2000), conducted a cross country study using a growth model that captures non-linear effects between aid and growth. Their results show that when human capital and investment are not controlled for, aid increases economic growth, but with decreasing returns. Hansen and Tarp conclude that capital accumulation is the channel through which aid impacts growth. In another cross country study, Minoiu and Reddy (2009) structured their research by looking at the effect of two kinds of aid (developmental and non-developmental aid) on per capita GDP growth over long periods. Their results indicate that developmental aid has a positive, large and robust effect on economic growth, while the effect of non-developmental aid on economic growth is mostly neutral and occasionally negative.

On the other hand, Ouattara (2006) uses panel data technique to study the effect of aid on fiscal behavior given that aid is channeled through the public sector and its effect on the economy is contingent on how it is used by the public sector. Ouattara's empirical results suggest that aid has a significant positive impact on public investment and developmental expenditure, while it has a significant negative relationship with non-developmental expenditure. In addition, Tavares (2002) studied the impact of foreign aid on corruption and found that aid has a robust significant negative relationship with corruption.

I add to the existing economic literature by using an instrumental variable approach where I use percentage of population with access to improved water source as an instrumental variable for foreign aid. There are a lot of studies that have taken the instrumental variable approach: for instance Brückner (2009) used rainfall as an instrumental variable to study the impact of growth on Aid; Rahajan and Subramanian (2008) used colonial links and relative population size of the donor to recipient; and Magesan (2015) used Participation in United Nation's Human Rights Treaties. However, I am not aware of any study that uses the instrumental variable I have exploited in this paper. Some prior studies that have used the instrumental variable approach have been criticized for using weak and invalid instruments (Magesan, 2015). Some instrumental variables used in prior studies have been criticized on two to three grounds: 1) high collinearity with aid (e.g. lagged aid, lagged aid squared); 2) not truly exogenous to the economy (e.g. lagged GDP per capita, lagged arms imports) and 3) time invariance (Werker et. Al 2008).

3. Methodology

The objective of this paper is to study the impact of foreign aid on the economic growth of some least developed countries (LDCs) in Sub-Saharan Africa. In this study, I use the Solow

Growth Model's aggregate production function as a guide to structure my regression model. According to Solow Growth Model's aggregate production function, output is a function of capital accumulation (K), labor force/ Population (N) and state of technology (A) (Blanchard and Johnson, 2013). This is written out as

$$Y = F\,(K, N, A).$$

I use Total Factor Productivity (TFP) to estimate technological progress or state of technology. According to Comin, **"**Total Factor Productivity (TFP) is the portion of output not explained by the amount of inputs used in production" (Comin 2006)**.** The Solow residual defined as

$$gY - \alpha * gK - (1 - \alpha) * gL$$

is used as a measurement for TFP growth, where gY denotes the growth rate of aggregate output, gK the growth rate of aggregate capital, gL the growth rate of aggregate labor and alpha the capital share (Comin 2006). TFP is multidimensional and some of its important determinants include human capital, physical infrastructure, institutions (political and economic), financial development, geographical predicament and absorptive capacity (Issakson 2007).Cognizant that TFP accounts for both political and economic institutions, I use TFP to control for quality of government, nature of policies and corruption which appear to be determinants of aid effectiveness (Fayissa and El-Kaissy 1999).

Furthermore, I include the variable "net exports" in my model since it is argued that increasing Sub-Saharan Africa's trade share in the world can outweigh the impact of aid. According to One, "Sub-Saharan Africa's tiny share (3.5%) of global exports was worth approximately \$442 billion in 2014, around 10 times the amount of aid the region received the same year[1]." Hence my primary model in this study:

$$rgdp_{it} = \beta_0 + \beta_1 NetODA_{it} + \beta_2 NetExp_{it} + \beta_3 rtfp_{it} + \beta_4 rkstock_{it} + \beta_5 pop_{it} + u_{it}$$

Where *rgdp* is real gdp (as a measure of economic growth), *NetODA* is net official development assistance received (measure of aid), *NetExp* is trade balance, *rtfp* is total factor productivity, *rkstock* is capital stock, *pop* is population and *u* is the error term.

I use different regression methods that potentially correct for heteroscedasticity, unit roots, trending behavior, serial correlation, unobserved fixed variables and endogeneity. I then compare these regressions and make a conclusion. My main contribution to the existing literature is my instrumental variable approach where I use percentage of population with access to improved water sources (H_20_pop) as an instrumental variable for foreign aid. Human well-being indicators such as infant mortality, life expectancy, literacy etc. rather than macroeconomic indicators are the recommended determinants of aid allocation to a country (Fayissa and El-Kaissy 1999). On the other hand, real GDP only accounts for total final output in the economy. Therefore, theoretically, percentage of population with access to improved water sources is not used in the accounting of real GDP; however it is a wellbeing indicator that can potentially be used to determine aid allocation. Therefore, I suspect that H_20_pop is highly correlated with aid, but is not directly correlated with real GDP and therefore is uncorrelated with the error term of my model.

---

[1] One. "Trade and Investment" http://www.one.org/international/issues/trade-and-investment/

4. Data

My study uses panel data for 12 African countries over the span of 20 years (1995 – 2014). All the data used in this study is from Penn World Table version 9.0 and the World Bank's Database: World Development Indicators. The African countries of interest are Benin, Burkina Faso, Burundi, Mauritania, Mozambique, Rwanda, Senegal, Lesotho, Sierra Leone, Tanzania, Togo and Sudan. My key variables from Penn World Table 9.0 include real gross domestic product (GDP) at constant national prices (in million 2011US$); total factor productivity at constant national prices (2011=1); capital stock at constant national prices (in million. 2011US$); and Population (in millions). Data on the following variables are from the World Bank's Database: net official development assistance received (as percentage of gross national income (GNI); external balance on goods and services (percent of GDP), commonly referred to as trade balance or net exports; and improved water source (percent of population with access).

The summary statistics of these key variables are presented in Table 1. During 1995 to 2014, the average net official development assistance received was 13.15 % of GNI while the average real GDP of these African countries was US$ 25707.81 Million (constant 2011 US$). The mean on net exports (-19.75 % of GDP) implies that these African countries have, on average, been running trade deficits for 20 years. On the other hand, only 53.8% of the total population of these African countries, on average, has access to improved water sources.

Table 1. Summary Statistics of Key Variables

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Net ODA received (% of GNI)	240	13.15	8.61	1.22	53.48
Real GDP (Constant 2011 Million US$)	240	25707.81	37874.27	2546.94	180328.80
Net Exports (% of GDP)	239	-19.75	20.44	-118.26	6.10
Capital Stock (Constant 2011 Million US$)	240	63160.07	95285.89	6654.39	512623.80
Total Factor Productivity	240	0.95	0.15	0.56	1.28
Population (Millions)	240	13.73	12.75	1.75	50.44
Access to Water (% of Population)	240	62.06	11.84	35.70	82.10

5. Analysis and Results

Table 2: Preliminary Regression

Source	SS	df	MS
Model	2.9595e+11	5	5.9191e+10
Residual	4.6470e+10	233	199442183
Total	3.4242e+11	238	1.4388e+09

Number of obs	=	239
F(5, 233)	=	296.78
Prob > F	=	0.0000
R-squared	=	0.8643
Adj R-squared	=	0.8614
Root MSE	=	14122

rgdp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
net_oda	-503.3378	128.9794	-3.90	0.000	-757.4528	-249.2229
net_Exp	2.259717	52.0087	0.04	0.965	-100.2077	104.7271
rtfp	20967.31	7186.75	2.92	0.004	6807.996	35126.63
rkstock	.1031206	.0199218	5.18	0.000	.0638707	.1423704
pop	1892.539	148.8858	12.71	0.000	1599.205	2185.874
_cons	-19997.76	8474.099	-2.36	0.019	-36693.41	-3302.11

Preliminary regression results show that aid and real GDP has a negative relationship where a one point increase in net ODA reduces real GDP by US$ 503.34 and this coefficient is statistically significant from zero. The rest of the independent variables have statistically significant positive coefficients, except for the coefficient on net exports which has a statistically insignificant positive coefficient. However, there is evidence of heteroscedasticity, serial correlation, non-stationarity, unit roots and trending behavior in this regression output - the specific tests for these problems are included in the appendix. Thus, I potentially correct for these problems by running a first differenced as well as a de-trended regression using robust standard errors and logged variables – except for net exports because it has negative values.

Table 3: De-trended Regression

```
Linear regression                               Number of obs     =        239
                                                F(5, 233)         =    1282.46
                                                Prob > F          =     0.0000
                                                R-squared         =     0.9652
                                                Root MSE          =     .19139
```

lrgdp_dt	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnetODA_dt	-.1198737	.0306706	-3.91	0.000	-.1803007	-.0594466
netEXP_dt	.0006892	.0006669	1.03	0.303	-.0006248	.0020032
lrtfp_dt	1.200889	.1096828	10.95	0.000	.9847921	1.416986
lrkstock_dt	.4601635	.0228998	20.09	0.000	.4150463	.5052806
lpop_dt	.5660361	.0266018	21.28	0.000	.5136252	.6184469
_cons	.0008038	.0124336	0.06	0.949	-.0236929	.0253006

The results from the regression of de-trended show that there is still a negative relationship between aid and real GDP where a percentage increase in aid reduces real GDP by 0.12% and the coefficient is statistically different from zero. Surprisingly the coefficient on net exports is not practically and statistically significant from zero. The rest of the independent variables have statistically significant positive coefficients. Furthermore, the first differenced regression yields similar results to the regression of de-trended variables as far as the sign, magnitude and significance of coefficients estimates are concerned. See first differenced regression output below:

Table 4: First Differenced Regression

```
Linear regression                               Number of obs     =        237
                                                F(5, 231)         =     212.31
                                                Prob > F          =     0.0000
                                                R-squared         =     0.9567
                                                Root MSE          =     .07791
```

clrgdp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
clnetODA	-.0670772	.0402963	-1.66	0.097	-.1464724	.012318
dnetEXP	.0008234	.0010942	0.75	0.453	-.0013326	.0029793
clrtfp	1.058939	.1812455	5.84	0.000	.701833	1.416044
clrkstock	.5049199	.1021231	4.94	0.000	.3037081	.7061316
clpop	.5192678	.1409524	3.68	0.000	.2415512	.7969843
_cons	.0016393	.0052268	0.31	0.754	-.008659	.0119375

On the other hand, Cognizant that the countries in my model are heterogeneous, I also estimate my model using time and country fixed effects to net out unobserved fixed variables. The results show that all my dependent variables have a positive relationship with real GDP except for aid and net exports. Also, all the coefficient estimates of my model are statistically significant from zero. However, the negative coefficients on net exports does not make sense as a majority of the economies of LDCs in Sub-Saharan Africa are tethered to commodity prices of their exports; Rodrik (2007) asserts that there is a direct relationship between the profitability of a country's tradable commodities and economic growth. The coefficient on net official development assistance suggests that a percentage increase in net ODA reduces real GDP by 0.03%, while TFP has the largest impact on real GDP. A percentage increase of TFP increases real GDP by 0.91%. See Table below

Table 5: Fixed Effects Regression

```
i.country         _Icountry_1-12      (_Icountry_1 for country==Benin omitted)
i.year            _Iyear_1995-2014    (naturally coded; _Iyear_1995 omitted)

Linear regression                                Number of obs     =        239
                                                 F(35, 203)        =   18513.81
                                                 Prob > F          =     0.0000
                                                 R-squared         =     0.9995
                                                 Root MSE          =     .02532
```

lrgdp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnet_oda	-.0254101	.0054001	-4.71	0.000	-.0360576	-.0147626
net_Exp	-.0006843	.0002782	-2.46	0.015	-.0012327	-.0001358
lrtfp	.9099721	.018708	48.64	0.000	.8730851	.9468591
lrkstock	.3660657	.011382	32.16	0.000	.3436237	.3885077
lpop	.6720259	.0875759	7.67	0.000	.4993509	.8447009
_Icountry_2	-.0944086	.0459574	-2.05	0.041	-.1850236	-.0037935
_Icountry_3	-.3446654	.0217897	-15.82	0.000	-.3876285	-.3017024
_Icountry_4	.0502099	.1385289	0.36	0.717	-.2229301	.3233499
_Icountry_5	.3693702	.0836801	4.41	0.000	.2043766	.5343637
_Icountry_6	-.2784852	.0865143	-3.22	0.001	-.449067	-.1079034
_Icountry_7	.0663196	.0163586	4.05	0.000	.0340651	.0985741
_Icountry_8	.007316	.0301465	0.24	0.808	-.0521243	.0667564
_Icountry_9	.0124697	.0447454	0.28	0.781	-.0757556	.1006949
_Icountry_10	.5017463	.1383099	3.63	0.000	.229038	.7744546
_Icountry_11	-.079905	.1388949	-0.58	0.566	-.3537667	.1939567
_Icountry_12	-.2371432	.0356729	-6.65	0.000	-.3074801	-.1668063
_Iyear_1996	.006397	.0126093	0.51	0.612	-.018465	.031259
_Iyear_1997	-.0122268	.0169176	-0.72	0.471	-.0455837	.02113
_Iyear_1998	.0055923	.0126349	0.44	0.659	-.0193202	.0305047
_Iyear_1999	.0148259	.0138954	1.07	0.287	-.0125719	.0422237
_Iyear_2000	.0277451	.0156064	1.78	0.077	-.0030264	.0585165
_Iyear_2001	.0202424	.0170877	1.18	0.238	-.0134498	.0539346
_Iyear_2002	.0339268	.0190738	1.78	0.077	-.0036813	.0715349
_Iyear_2003	.0315997	.0215682	1.47	0.144	-.0109267	.0741261
_Iyear_2004	.0390008	.0244625	1.59	0.112	-.0092324	.087234
_Iyear_2005	.0385662	.0261948	1.47	0.142	-.0130825	.0902149
_Iyear_2006	.0546239	.0280784	1.95	0.053	-.0007388	.1099867
_Iyear_2007	.0616927	.0306953	2.01	0.046	.0011703	.1222151
_Iyear_2008	.0628272	.032922	1.91	0.058	-.0020857	.1277401
_Iyear_2009	.0609339	.0353162	1.73	0.086	-.0086996	.1305675
_Iyear_2010	.0730919	.037615	1.94	0.053	-.0010744	.1472581
_Iyear_2011	.0833452	.0396493	2.10	0.037	.005168	.1615224
_Iyear_2012	.0866613	.0418423	2.07	0.040	.0041599	.1691626
_Iyear_2013	.1004795	.0445913	2.25	0.025	.012558	.188401
_Iyear_2014	.1135807	.0480149	2.37	0.019	.0189089	.2082525
_cons	4.260529	.2091888	20.37	0.000	3.848068	4.672991

However, I suspect that foreign aid and real GDP have a spurious relationship, or there might be some underlying endogeneity in the model. This is because the economic performance of a developing country can determine if aid should be allocated to it and on the other hand foreign aid has an effect on GDP through different channels in the economic structure of the country. In order to correct for this problem I use improved water source (percent of population with access to improved water source) as an instrumental variable for aid. As a robustness check of my instrumental variable I ran a regression of log (net ODA) on log(H_2O_pop) and other dependent variables that affect aid or have been used in prior research as instrumental variables as cited in Werker et. Al 2008.

Table 6: Instrumental Variable Quality

```
Linear regression                                  Number of obs     =        239
                                                   F(5, 233)         =      49.77
                                                   Prob > F          =     0.0000
                                                   R-squared         =     0.4849
                                                   Root MSE          =     .50304
```

lnet_oda	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lh2o_pop	-1.009491	.1650581	-6.12	0.000	-1.334688	-.6842939
lrgdp	-.980642	.104675	-9.37	0.000	-1.186872	-.7744116
lrgdp_1	.1502587	.087444	1.72	0.087	-.0220233	.3225406
lpop	.6916483	.0627507	11.02	0.000	.5680171	.8152795
year	.0244286	.0068967	3.54	0.000	.0108408	.0380165
_cons	-36.11949	13.37586	-2.70	0.007	-62.47257	-9.766407

The results make intuitive sense: as percentage of people with access to improved water sources increases, net ODA decreases. The coefficient on real GDP implies that as the economic performance of the country improves the amount of aid decreases. This was the case of Botswana after it gained its independence; the role of aid decreased as revenues from diamond mining increased (Togo and Wada 2008).

Table 7: Fixed Effects IV Regression

```
. xtivreg lrgdp (lnet_oda = lh2o_pop) net_Exp lrtfp lrkstock lpop, fe vce(robust
> )

Fixed-effects (within) IV regression            Number of obs     =        239
Group variable: ccode                           Number of groups  =         12

R-sq:                                           Obs per group:
     within  = 0.9942                                         min =         19
     between = 0.9346                                         avg =       19.9
     overall = 0.9390                                         max =         20


                                                Wald chi2(5)      =   4.45e+06
corr(u_i, Xb)  = -0.3985                        Prob > chi2       =     0.0000

                                 (Std. Err. adjusted for 12 clusters in ccode)
------------------------------------------------------------------------------
             |               Robust
       lrgdp |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    lnet_oda |  -.0309319   .0542281    -0.57   0.568     -.137217    .0753533
     net_Exp |  -.0003928   .0005298    -0.74   0.458    -.0014313    .0006456
       lrtfp |   .9020897   .0392832    22.96   0.000     .8250961    .9790834
    lrkstock |   .3769718   .0249396    15.12   0.000     .3280912    .4258525
        lpop |   .8541086   .0494813    17.26   0.000      .757127    .9510901
       _cons |   3.801544   .2827884    13.44   0.000     3.247289    4.355799
-------------+----------------------------------------------------------------
     sigma_u |  .29751608
     sigma_e |   .0261815
         rho |  .99231546   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:   lnet_oda
Instruments:    net_Exp lrtfp lrkstock lpop lh2o_pop
------------------------------------------------------------------------------
```

The regression results of the fixed effect (within) IV regression show that all the dependent variables have a positive relationship with real GDP, except for net exports and net ODA. Also, all the coefficients of the variables are statistically significant, except for net exports and net ODA. The coefficient estimates are similar to the coefficient estimates of the regression with time and country fixed effects. The IV (within) fixed effects model also implies that a percentage increase in net ODA reduces real GDP by 0.03%. However, there is not enough evidence to support this relationship as the coefficient on net ODA is statistically insignificant. In contrast, the TFP, capital stock and population coefficient estimates are practically significant and support macroeconomic theory. For instance, according to macroeconomic theory a country's labor force increases as the population of the country increases and hence in the long run when a country reaches its steady state, output grows at the growth rate of technology (estimated by total factor productivity in my model) and population growth (Blanchard and Johnson, 2013).

Table 8: Fixed Effects IV Regression (Using detrended Variables)

```
. xtivreg lrgdp_dt (lnetODA_dt = lwater_dt)  netEXP_dt lrtfp_dt lrkstock_dt lpop
> _dt, fe vce(robust)

Fixed-effects (within) IV regression           Number of obs      =        239
Group variable: ccode                          Number of groups   =         12

R-sq:                                          Obs per group:
     within  = 0.9592                                        min =         19
     between = 0.9487                                        avg =       19.9
     overall = 0.9485                                        max =         20


                                               Wald chi2(5)       =    2254.47
corr(u_i, Xb)  = 0.2218                        Prob > chi2        =     0.0000

                                 (Std. Err. adjusted for 12 clusters in ccode)
------------------------------------------------------------------------------
             |               Robust
    lrgdp_dt |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  lnetODA_dt |  -.0238895   .0496233    -0.48   0.630    -.1211494    .0733703
   netEXP_dt |  -.0007491   .0003594    -2.08   0.037    -.0014535   -.0000448
    lrtfp_dt |   .9084642   .0443209    20.50   0.000     .8215969    .9953316
 lrkstock_dt |   .3712589   .0248677    14.93   0.000      .322519    .4199988
     lpop_dt |   .6691325   .0813287     8.23   0.000     .5097312    .8285339
       _cons |  -.0000394   .0007549    -0.05   0.958    -.0015191    .0014402
-------------+----------------------------------------------------------------
     sigma_u |  .24443658
     sigma_e |  .02526101
         rho |  .98943291   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:   lnetODA_dt
Instruments:    netEXP_dt lrtfp_dt lrkstock_dt lpop_dt lwater_dt
------------------------------------------------------------------------------
```

As a robustness check I also ran fixed effects within instrumental variable regression using de-trended variables since most of the variables trend with time. The coefficients are similar to the regression results in table 7, however, the coefficient on net exports is now statistically significant at the 5 % level. Again, the coefficient on net exports doesn't make sense, nevertheless its coefficient is not practically significant. A summary of my regression approaches is presented in Table 9.

Conclusion

My study investigates the impact of aid (official development assistance) using panel data for 12 least developed countries (LDCs) in Sub-Saharan Africa observed over a period of 20 years (1995 – 2014). An understanding of the historical context of aid given to Africa or developing countries in general might be helpful in interpreting the story that my data supports. According to Moyo 2009, starting from the 1980's, multilateral aid was given in order to help indebted developing countries meet their debt obligations as many countries had accumulated a lot of debt following the oil crisis of the 1970's. However, multilateral aid like budgetary support was provided on condition that developing countries implement policy reforms in order to promote free market systems and good governance. This is in contrast to aid that was given in the 1960's which primarily focused on building physical infrastructure like airports, roads, power stations, telecommunications, schools, health centers among others (Moyo 2009).

My regression results imply that that a percentage increase in net official development assistances reduces real GDP by about 0.03%. However, this is statistically not different from zero and arguably practically insignificant as well. Thus, there is not enough evidence to support this relationship; therefore this goes to show that aid that was transferred around this period (1995 – 2014) was ineffective towards achieving high levels of economic growth. My results also show that TFP, capital accumulation and population have one of the largest impacts on economic growth. For instance, in the fixed-effect (within) IV regression, a percentage increase in TFP increases GDP by 0.9% and a percentage increase in capital stock increases economic growth by 0.38%. Therefore if aid is inefficient in increasing economic growth over a long-run, it must be the case that it is being misallocated in the economy or it is practically doing little to promote robust capital accumulation, technological progress and labor force participation.

Bibliography

Anders Isaksson. "Determinants of total factor productivity: a literature review" Research and Statistics Branch, United Nations Industrial Development Organization. July 2007 accessed November 2, 2016

Arvind Magesan. "Foreign Aid and Economic Growth in Developing Countries: An Instrumental Variables Approach" Department of Economics University of Calgary. June 24, 2015 accessed November 21, 2016

B. Ouattara. "Foreign aid and government fiscal behavior in developing countries: Panel data evidence" Economic Modelling 23 (2006) 506– 514, February 1, 2006 accessed September 28, 2016.

Bichaka Fayissa and Mohammed I. El-Kaissy. "Foreign Aid and the Economic Growth of Developing Countries (LDCs): Further Evidence" Middle Tennessee State Universit. Fall 2009 accessed September 27, 2016

Camelia Minoiu and Sanjay G. Reddy. “Development Aid and Economic Growth: A Positive Long-Run Relation” International Monetary Fund May 29, accessed September 28, 2016.

Dambisa Moyo. *Dead Aid: Why Aid Is Not Working and How There Is Another Way For Africa*. New York: Farrar, Straus and Giroux, 2009

Dambisa Moyo. “Why Foreign Aid Is Hurting Africa” Wall Street Journal March 21, 2009 Accessed September 28, 2016 http://www.wsj.com/articles/SB123758895999200083

Diego Comin. “Total Factor Productivity” New York University and NBER. August 2006 accessed November 21, 2016

Eric Werker, Faisal Z. Ahmed, and Charles Cohen. “How is Foreign Aid Spent? Evidence from a Natural Experiment” Harvard Business School July 2008 accessed November 21, 2016

Girijasankar Mallik. “Foreign Aid and Economic Growth: A Cointegration Analysis of the Six Poorest African Countries” Economic Analysis & Policy, vol. 38 no. 2, September 2008 accessed September 28, 2016.

Henrik Hansen and Finn Tarp, “Aid and growth regressions” Journal of Development Economics Vol. 64 _2001. 547–570 August 1, 2000 accessed September 28, 2016.

Jose´ Tavares. “Does foreign aid corrupt?” Economics Letters 79 (2003) 99–106, August 28, 2002 accessed September 28, 2016.

Markus Brückner. “Getting the Effect of Foreign Aid on Economic Growth Right:An Issue of Addressing Endogeneity Bias” Department of Economics, Universitat Pompeu Fabra. September 2009 accessed November 21, 2016

Morrissey, O. (2001). “Does aid increase growth?” Progress in Development Studies, 1(1), 37–50.

Oliver Blanchard and David Johnson. Macroeconomics, 6th ed. Pearson Education, Inc, 2013.

One. “Trade and Investment” –n.d accessed April 30, 2015 http://www.one.org/international/issues/trade-and-investment/

Raghuram G. Rajan and Arvind Subramanian. “Aid And Growth: What Does The Cross-Country Evidence Really Show?” The Review of Economics and Statistics, November 2008, 90(4): 643–665, accessed September 27, 2016.

Rodrik Dani. “The Real Exchange Rate and Economic Growth” John F. Kennedy School of Government, Harvard University Cambridge, MA Revised, October 2008

Satish Lohani. “Effect of Foreign Aid on Development: Does More Money Bring More Development?” Department of Economics, Illinois Wesleyan University. April 22, 2004 accessed September 28, 2016

Ken Togo and Yoshio Wada. “Development Assistance and Economic Growth: A Case Study of Botswana.” Musashi University Discussion Paper No.48. February, 2008 accessed November 30, 2016

Zohid Askarov and Hristos Doucouliagos. "Development Aid and Growth in Transition Countries" World Development Vol. 66, pp. 383–399, 2015 accessed September 28, 2016

Appendix

Table 9. Summary of Regression Analysis of the effect of aid (net ODA) on real GDP

Dependent Variable: Log (Real GDP) Time Period: 1995 - 2014					
Variable	**1st Differenced**	**De-trended**	**Fixed Effects (Time and Country)**	**Fixed Effects IV Regression**	**Fixed Effects IV Regression (De-Trended)**
log (Net ODA)	- 0.0671* [0.0403]	- 0.1199*** [0.0307]	- 0.0254*** [.0054]	- 0.0309 [0.054]	- 0.024 [0.05]
Net Exports (% of GDP)	0.0008 [0.0011]	0.0007 [0.0007]	- 0.0007** [0.0003]	- 0.0004 [0.0005]	- 0.0007** [0.0004]
Log (TFP)	1.059*** [0.1812]	1.201*** [0.1097]	0.90997*** [0.0187]	0.9021*** [0.0393]	0.9085*** [0.044]
Log (Capital Stock)	0.5049*** [0.1021]	0.4601*** [0.0229]	0.3661*** [0.0114]	0.37697*** [0.0249]	0.3713*** [0.0249]
Log(Population)	0.5192*** [0.14095]	0.566*** [0.0266]	0.67203*** [0.0876]	0.8541*** [0.0495]	0.6691*** [0.0813]
Total Observations	237	239	239	239	239
R-Squared	0.9567	0.9652	0.9995	0.9390	0.9485
Prob (F-Statistic)	0.000	0.0000	0.0000	0.0000	0.0000
(*), (**), (***) represent 10%, 5%, and 1% levels of significance. Robust standard errors in brackets [ ]. The instrumental variable used in the Fixed effects IV regressions is Improved water Source (percent of population with access)					

Preliminary Regression

```
. reg rgdp net_oda net_Exp rtfp rkstock pop
```

Source	SS	df	MS
Model	2.9595e+11	5	5.9191e+10
Residual	4.6470e+10	233	199442183
Total	3.4242e+11	238	1.4388e+09

Number of obs	=	239
F(5, 233)	=	296.78
Prob > F	=	0.0000
R-squared	=	0.8643
Adj R-squared	=	0.8614
Root MSE	=	14122

rgdp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
net_oda	-503.3378	128.9794	-3.90	0.000	-757.4528	-249.2229
net_Exp	2.259717	52.0087	0.04	0.965	-100.2077	104.7271
rtfp	20967.31	7186.75	2.92	0.004	6807.996	35126.63
rkstock	.1031206	.0199218	5.18	0.000	.0638707	.1423704
pop	1892.539	148.8858	12.71	0.000	1599.205	2185.874
_cons	-19997.76	8474.099	-2.36	0.019	-36693.41	-3302.11

White's Test for Heteroscedasticity

```
White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(20)      =    217.75
         Prob > chi2   =    0.0000

Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	217.75	20	0.0000
Skewness	78.86	5	0.0000
Kurtosis	11.90	1	0.0006
Total	308.51	26	0.0000

Therefore there is evidence of heteroscedasticity.

Testing for Serial Correlation in Stata

```
predict u, resid
(1 missing value generated)
.
. gen lagu = u[_n-1]
(2 missing values generated)

. reg u lagu

      Source |       SS           df       MS      Number of obs   =       237
-------------+----------------------------------   F(1, 235)       =   1586.70
       Model |  4.0419e+10         1  4.0419e+10   Prob > F        =    0.0000
    Residual |  5.9864e+09       235  25473856.1   R-squared       =    0.8710
-------------+----------------------------------   Adj R-squared   =    0.8705
       Total |  4.6406e+10       236   196634385   Root MSE        =    5047.2

------------------------------------------------------------------------------
           u |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lagu |    .954532   .0239631    39.83   0.000      .907322    1.001742
       _cons |   154.4315   327.8851     0.47   0.638    -491.5383    800.4013
------------------------------------------------------------------------------
```

The p value for the lagged coefficient of the error term is 0.000; therefore serial correlation is a problem that needs to be corrected for.

Fisher Type Augmented Dickey Fuller Test for Unit Roots

Variable	**p-value**
rgdp	**1.0000**
net_oda	**0.0000**
net_Exp	**0.3268**
rtfp	**0.9964**
rkstock	**1.0000**
pop	**0.0000**

These results show that all the variables have unit roots except for net official development assistance (net_oda) and population (pop) and therefore I can't rule out non-stationarity.

Furthermore, I ran regressions of each variable on a time variable, year, and I found that all the variables were trending except for net exports.

Is there a path for green growth? Evidence from India

Anh Trinh

Abstract

This paper uses historical temperature fluctuations in India to idenify its effects on economic growth rates. Using a climate-adjusted form of the Solow growth model, I find that one degree Celsius increase in temperature decreases GDP per capita growth by 0.71%. This finding informs debates over the role of climate on economic development and suggests the possibility of a green path for economic growth, a policy agenda that is both sustainable and pro-growth.

I. Introduction

Climate change from greenhouse gas emission is infamously known as the "mother" of all negative externality of the market, a problem that requires international corporation to mitigate. While scientists are still debating the severity of this problem, in my opinion it is still very hard to agree with the 45th President of the United States. Climate change is not a hoax created by the Chinese government when 195 countries have already signed the Paris Agreement in March to reduce temperature by 1.5° Celsius by cutting greenhouse gas emissions. The potential repercussions of one country's pulling out from an important agreement like this are the motivation for my paper. Thus, the purpose of this paper is not to provide new insight on the science of climate change, but only to use empirical data from India to establish that temperature change negatively affect economic growth.

Often, when growth is taught in undergraduate neo-classical economics classes, there are only three factors involved: technology, labor and capital represented in the Solow growth model. At steady state, the only catalyst for economic growth according to the Solow growth model is technology. In context of a developing country where agriculture contributes mainly to annual GDP growth - the measure of economic growth in this paper – temperature change plays a role in economic growth. Technology may increase crops productivity to a certain extent, but unusual heat and drought or excessive precipitation and flooding affect the year's agricultural outcomes almost instantly, not to mention other non-economic consequences such as diseases and conflicts (Hsiang, Burke & Miguel, 2013) . These non-economic outcomes have been found to affect human capital and productivity, which is the catalyst for growth in the Solow growth model (Zivin and Neidell, 2012 & 2013). In addition to agriculture, industrial output might suffer when extreme weather affects resource productivity. If the rate of temperature change is as significant as most environmental scientists speculate, long term economic growth for a developing country like India will suffer. Thus, for economists, a relationship between temperature anomaly and economic growth contributes to the growing research on the economic consequences of "one of the biggest market failure the world has seen" (Stern, 2007). The development of a growth model that encompasses systematic changes like climate change will open new path to more creative policies with even more potentials improve people's lives especially in the more vulnerable population of the world.

While there has been significant progress towards growth in the developing world, the challenge of overcoming poverty and inequality will be greatly compounded by climate change and environmental degradation, which disproportionately hurt the poor and most vulnerable. These increasingly interlinked crises threaten development gains and prospects for continued progress. While the Paris agreement is one commitment on paper to do more, the world's collective response has fallen far short of what is needed. Unmitigated warming is expected to reshape the global economy by reducing average global incomes roughly 23% by 2100 and widening global income inequality (Burket, Hsiang and Miguel, 2015). Thus, if adequately examined, this research question poses an interesting policy outlook: if there is a relationshop between economic growth and climate change, then any investment in a sustainable economy will in turn have a positive feedback on the economy, open up opportunity for green growth. On top of that, there are great potentials for delivering a "triple bottom line" of job–creating economic growth coupled with environmental protection and social inclusion (World Resources Institute, 2012). Developing countries might benefit greatly from an investment in sustainable growth, both economically and environmentally. The economic benefits of a transition to a green economy is a question that not only policymakers would want answers to, but also every sector of the economy and are relevant to all investors and businesses. For investors, if there is consensus on how climate change negatively affects the economy, investments in "socially responsible" businesses are more attractive as these businesses are contributing more to the economy's growth than regular businesses. The benefits of being a sustainable business may outweigh the costs, which incentivizes businesses to internalize their carbon emission. Decisions made by private sector investors and financial institutions will have a major influence on how society responds to climate change.

For many developing nations, current climate policies agenda means relying heavily on financial and technical assistance from developed countries. Additionally, many developing nations are not solely concerned about climate change, but also prioritize expanding energy access to their peoples in order to move toward a better standard of living. One country that faces this dichotomy is India, for its economic status, population challenge and energy issues. It is the fourth largest greenhouse gas (GHG) emitter, accounting for 5.8 percent of global emissions. India's emissions increased by 67.1% between 1990 and 2012, and are projected to grow 85% by 2030[2].Yet, India faces a major energy issue: nearly 300 million people that do not have access to even one electric light bulb[3].This is even more challenging because the mean rate of population growth in is 1.9% (Table 2), which is relatively high when compared to developed nations[4]. How India balances expanding electricity access and economic targets while at the same time achieving its climate targets will indeed be paramount to the future of global climate change action. Thus, the answer to my research question is will provide a clear picture to achieve the twofold challenge of green economic growth. Ebinger (2016), in the Brookings policy brief even asserts that, “If India fails, Paris (Agreement) will fail”.

In the next section, I will describe what has been done in the literature surrounding the relationship between economic growth and temperature change. In section III, I will develop a regression model to answer my research question based on a climate-Solow growth model. In section IV, I will discuss the data collected to test my hypothesis, and in section V, I will use that data with my theory as evidence for my question. In section VI, I will conclude.

II. Literature review

There is a large and growing literature that examines the causal effect of temperature change on economic growth. It is not my objective to review all studies; rather, the goal is to review those studies that have some connections to my research question. The literature suggest that impact of climate change on GDP growth are found through two channels: climate direct impact on aggregate output and pollution impact on human capital.

---

[2] "India's Climate and Energy Policies." *Center for Climate and Energy Solutions*, October 2015.

[3] Ebinger, Charles K. "India’s Energy and Climate Policy: Can India Meet the Challenge of Industrialization and Climate Change?" *The Brookings Institution*, June 2012.

[4] The World Factbook, Center Intelligence Agency.

The first channel is found in studies that examine the *level* impact of climate change as an equivalent of income gain or loss in percent of GDP. Frankhauser and Tol (2005) justifies their hypothesis by arguing that the prospect of future damages (or benefits) of global warming affects capital accumulation and people's propensity to save, which in turn, affects output. In terms of capital accumulation, with a constant saving rate, if climate change has a negative impact on output, the amount of investment in an economy is reduced which lead to a lower GDP and capital stock. Lower in investment can also slowdown technical progress and/or labor productivity or human capital accumulation. The savings effect is when faced with uncertainty posed by climate change: people change their behavior to save less and consume more today. Both effects are found to be negative, and in an endogenous growth model, there is a different rate of technical progress, thus enhances the savings and capital accumulation effects. The authors examined the statistical approach in Mendelsohn's work (Mendelsohn, Morrison, Schlesinger, and Andronova, 2000; Mendelsohn, Schlesinger, and Williams, 2000). It is based on direct estimates of the welfare impacts, using observed variations (across space within a single country) in prices and expenditures to discern the effect of climate. Mendelsohn assumes that the observed variation of economic activity with climate over space holds over time as well; and uses climate models to estimate the future effect of climate change. Mendelsohn's estimates are done per sector for selected countries, extrapolated to other countries, and then added up, but physical modeling is avoided. Nordhaus (2006) and Maddison (2003) use versions of the statistical approach as well. However, Nordhaus uses empirical estimates of the aggregate climate impact on income across the world (per grid cell), while Maddison (2003) looks at patterns of aggregate household consumption (per country). Like Mendelsohn, Nordhaus and Maddison rely exclusively on observations, assuming that "climate" is reflected in incomes and expenditures—and that the spatial pattern holds over time. Rehdanz and Maddison (2005) also empirically estimate the aggregate impact, using self-reported happiness measures from dozens of countries. The problem with these research is that, even though they are able to establish and justify a clear linkage between climate and change in the level of GDP, they did not employ a clear representation of climate within their research models.

Other groups of researchers try to incorporate a clearer link between climate and output into their analysis. Hsiang and Jina (2013) are the first to provide the first global estimates of the effect of large-scale environmental disaster on long-run growth. Through an extensive examination 6,700 tropical cyclones on the planet found that national incomes decline, relative to their pre-disaster trend, and do not recover within twenty years. Income losses arise from a small but persistent suppression of annual growth rates spread across the fifteen years following disaster, generating large and significant cumulative effects: a 90th percentile event reduces per capita incomes by 7.4% two decades later, effectively undoing 3.7 years of average development. This finding substantially alters the costs global climate change, especially on developing countries. However, these are only projections, based on a theoretical derivation under the assumption that the frequencies of cyclones are certain. Similarly, Dell et al. (2012) examine temperature shock and economic growth from panel data from 125 countries from 1950 to 2005. The authors aggregate weather data to a country-year level from a gridded monthly mean temperature and precipitation dataset at 0.5x0.5 degree resolution. Economic data is the value-added agriculture and industrial as percentage of GDP from the World Bank, World Development Indicators. Using various regression models with lags, interaction between dummy variables such as poor and hot countries and political stability, Dell et al. (2012) find three main results. Poor countries, but not wealthier ones, suffer from reduction in economics growth and growth rates because of higher temperature. More specifically, a 1° Celsius increase in average temperature over a given year will decrease economic growth by 1.3%. In addition, agricultural and industrial output along with political stability decrease with increase in temperature. These findings suggest that poorer countries are the ones suffer more from the negative externality that is climate change. Hsiang (2010), using surface temperatures from National Centers for Environmental Prediction and value added aggregate income by industry data from the United Nations, shows similar findings using annual variation in a sample of 28 Caribbean-basin countries over the 1970–2006 period. National output falls 2.5 percent per 1°C temperature increase. This study further examines output effects by time of year and shows that positive temperature shocks have negative effects on income only when they occur during the hottest season. Low-income countries tend to be in tropical zones closer to the equator. They are already hotter, and their output already suffers to some extent from their higher temperatures in sectors like agriculture. Moreover, low-income countries are typically less able to adapt to climate change both because of a lack of resources and less capable institutions (Adger, 2006; Alberini, Chiabai, and and Meuhlenbachs 2006; Smit and Wandel, 2006; Tol, 2008; Tol and Yohe, 2007b; Yohe and Tol, 2002). In the papers by Dell et al (2012) and Hsiang (2010), the economic impact of climate change is assessed and valued separately – by industry output as percentage of GDP. However, this method has potential issue: it may ignore interlinkages between the sectors which could possibly affect overall growth data.

One criticism to the cross-sectional studies of temperature effect is that they are driven by country specific characteristics – meaning that the models employed have omitted variables bias. However, Dell, Jones and Olken (2009) also examine the short run effects using sub-national data from 12 countries in the Americas, and provide new evidence that the negative cross-country relationship between temperature and income also exists within countries and even within states. The fact that the cross-sectional relationship holds within countries, as well as between countries, suggests that omitted country characteristics are not wholly driving the cross-sectional relationship between temperature and income. Nonetheless, a deficiency in the 2009 paper is the lack of empirical estimates of long term GDP growth in relation to climate change. They only attempt to reconcile the long run effect through two theoretical mechanisms: convergence and adaptation. The theoretical model suggests that half of the negative short-term effects of temperature may be offset in the long run through adaptation. Thus, it is crucial to look at the empirical evidence from one country over time, to account for the interlinkages cross sectors, and to find meaningful causal effect between temperature and economic growth.

A second channel that climate and pollution can affect growth is through human capital, measured by labor supply, productivity, and cognition. Zivin and Neidell (2011 & 2013) working papers published by the National Bureau of Economic Research find both theoretical and empirical evidences of this channel. Zivin & Neidell (2013) provide a theoretically linkage through the contemporaneous and latent effects of the environment on human capital by doing a meta-analysis of multiple studies. Their justification is that pollution may lead to direct brain development which affects cognitive ability. Alternatively, decrements in lung functioning may affect one's ability to focus and thus perform a wide range of tasks. They categorize the impacts of pollution into contemporaneous latent effects. The indicators of contemporaneous effect are schooling outcomes and labor market outcomes. Currie et al. (2009) use administrative data from the 39 largest school districts in Texas to estimate schooling outcomes. When carbon monoxide (CO) levels rise, absences also rise, 10 unit increase in CO2 decreases test scores by 2.4% of a standard deviation. As for labor market outcomes, Hanna and Oliva (2011) focus on the labor supply of workers in Mexico City and find that a 1 percent increase in sulfur dioxide levels decreases hours worked by 0.72 percent. In addition, Clay et al. (2010) found that workers with higher levels of lead exposure, while lead is still believed to be safe in the 20th century to make pipes, had substantially lower wages, value added per worker and value of capital per worker.

The latent effects stem from the hypothesis that negative shocks early in life may lead to a wide range of lasting effects, which may arise even without noticeable impacts at the time of exposure (Almond and Currie, 2011). In 2011, Zivin and Neidell look at the impacts of pollution on labor market outcomes. Labor market productivity of agricultural workers is measured to examine the impact of ozone pollution on productivity. Their data on daily worker productivity is derived from an electronic payroll system used by a large farm in the Central Valley of California who pays their employees through piece rate contracts (in which the employee is paid for each unit of production at a fixed rate). Piece rates reduce shirking and increase productivity over hourly wages and relative incentive schemes, particularly in agricultural settings. To quantify for pollution, Zivin and Neidell used measures of environmental conditions come from data on ozone levels from the system of monitoring networks maintained by the California Air Resources Board. Ozone is not directly emitted but forms from complex interactions between nitrogen oxides (NOx) and volatile organic chemicals

(VOCs). They found that 10 parts per billion decrease in ozone concentrations increases worker productivity by 4.2 percent.

Considering the theoretical and empirical evidences of the two channels that link climate change and economic growth, this paper proposes to capture this dynamic effect by using a different model to assess empirical data. I want to combine effect of temperature and the effect of pollution on long run economic development, which has not been done before. I use carbon emission as an indicator of pollution as informed by Burke et al. (2015). They found that under business as usual emissions throughout the 21st century will decrease per capita GDP by 23% below what it would otherwise be. Using data from India, I am able to capture the long run effects of temperature and carbon emissions on one country's GDP growth.

III. Modeling

To answer my research question: "Is there a negative effect of climate on economic growth?" I use the simplified Solow-like growth model derived by Tsigaris and Wood (2016) as a theoretical basis. To account for the effect of climate through the direct and human capital channels discussed in section II, I consider environmental conditions as an important factor of production into my model. First, consider a simple economy:

$$Y_t = A_t L_t^{\alpha} \quad (1)$$

where Y is aggregate output, L measures population, A measures total factor productivity. A damage function $D_t = e^{\theta_1 T_t E_t}$, where T_t is temperature anomaly in year *t* from year *t-1*, E_t is the growth of carbon emission in year *t* from year *t-1*, and θ_1 is a constant less than 0. The damage function is added to the output per worker Cobb-Douglas production function $y_t = A_t L_t^{\alpha}$. The climate-Solow growth model is:

$$Y_t = D_t A_t L_t^{\alpha} \quad (2.1)$$

Ceteris paribus, output per worker is reduced with increased temperatures. Along the balanced growth path, output per worker grows at a rate dependent on growth rates of temperature and carbon emission, the growth rate of total factor productivity, g_{At} and the growth rate of the capital labor ratio weighted by the income share of capital, α. In addition to Tsigaris and Wood (2016)'s climate-Solow model, I followed Dell et al.'s (2008) idea to incorporate climate growth's effect on productivity growth:

$$g_{At} = g_{it} + \theta_1 T_t E_t \text{ (2.2)}$$

Equation (2.1) captures the *level* effect of climate on production. For example, the effect of current temperature on output per capita. Equation (2.2) captures the *growth* effect of climate; e.g. the effect of climate on features such as institutions that influence productivity growth. The growth equation in (2.2) accounts for weather shocks while allowing separate identification of level effects and growth effects. In particular, both effects influence the growth rate in the initial period of a temperature. A temperature shock may reduce agricultural yields, but once temperature returns to its average value, agricultural yields bounce back. By contrast, the growth effect appears during the climate shock and is not reversed: a failure to innovate in one period leaves the country permanently further behind. Taking the logs of equation (2.1):

$$g_t = \theta_1(T_t + E_t) + g_{At} + \alpha g_{Lt} \text{ (3)}$$

The growth effect is identified in (3) as the summation of the climate effects over time. To estimate the effects of temperature and carbon emission on economic growth, I run regression of the form:

$$g_{it} = \alpha_1 T_t + \alpha_2 E_t + \alpha_4 g_{Lt} + \varepsilon_i \quad \text{(4)}$$

where $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are estimates of the effects on GDP per capita growth of the growth rate of temperature, CO2 emission and population, respectively. From this regression model, I hypothesize that the temperature and carbon emission growth rates (the difference between the natural log of temperature and emission from year *t-1* and year *t*) negatively affect economic growth.

IV. Data

In an exhaustive review of literature on this topic, Dell et al. (2014) found that most often used in climate-economics literature are gridded datasets, which a balanced panel of weather data for every point on a grid. The most frequently used gridded datasets in the studies reviewed here are the global temperature and precipitation data produced by the Climatic Research Unit (CRU) at the University of East Anglia with spatial resolution of 0.5x0.5. In this paper, I chose to use the World Bank group's data set for three independent variables from year 1972 to 2012 to maintain the consistency of all observations. Given the complexity of data manipulation and problem with accessibility of the ideal datasets from the University of East Anglia, I averaged out monthly temperature data from the World Bank Climate Change Knowledge Portal to get annual temperature data and then find the difference between the natural log of the temperature from year to year to get temperature growth rate. I manipulated similarly CO_2 emissions as metric tons per capita data from the World Bank. I used Indian annual real GDP per capita and population growth rates data from the OECD dataset (OECD, 2016).

The descriptive statistics from Table 1. suggest that India's growth rates of temperature change, CO_2 and GDP per capita fluctuate wildly. The variation of the growth rate of GDP per capita is the most notable, from a decrease of 7.4 percent to an increase of 8.7 percent. This variation is Co_2 emission decreases by 2.4 metric tons per capita in one year and increase 4.3 metric tons per capita in another. Climate literature suggests that the average global temperature on Earth has increased by about 0.8° Celsius (1.4° Fahrenheit) since 1880 (NASA Earth Observatory, 2010). However, the mean annual temperature from 1972 to 2012 decreases by 0.001° Celsius. Its minimum and maximum values nonetheless suggest that temperature fluctuates from decreasing 0.7 degree Celsius to increasing almost 1° Celsius. The data indicate that the growth effects certainly cannot be ignored in order to answer this research question.

V. Evidence

I estimate the dependent variable which is annual growth of GDP per capita on the following independent variables: growth rates of temperature, CO2 and population. Since my empirical model uses ordinary least squares estimates on time series data, it suffers from Gauss-Markov assumptions. Table 3 in section VI. Appendix summarizes the tests used and results to evaluate the violation of these assumptions. First, the Ramsey's test was used to test for omitted variables bias, which determines whether there are neglected nonlinearities in the model. The *p*-value for this test is less than 5% for my model, meaning that the correct functional form to estimate the independent variable the model was used. Second, time series data are often subject to the correlation its past and future values. Nonetheless, my model passes the Durbin-Watson test for autocorrelation for time series data, with a test statistics equals to 2.29. To test for multicollinearity to make sure two or more predictor variables in a multiple regression model are not highly correlated, I used the variance inflation factor (VIF). The VIF statistics (Table 3) for all three of my independent variables show that the variance of the estimated regression coefficients are not inflated (values are close to 1) as compared to when the predictor variables are not linearly related. The Breusch-Pagan test for heteroscedasticity tests the null hypothesis that the variance of the error is the same for all individuals. My model did not pass the because my *p*-value is slightly higher than 0.05. This means that the variance around the regression line is not the same for all values of the predictor variables. The violation of homoscedasticity can be fixed using a robust standard error, based on the covariance matrix estimates which are consistent in the presence of arbitrary forms of heteroscedasticity. I used the ',robust' command on STATA after my original regression command to fix the problem.

After fixing for heteroscedasticity with robust standard errors estimates, I am able to obtain the best linear unbiased estimators. According to my regression results (Table 2), the coefficient on temperature is positive and statistically significant. I find that the temperature change significantly affect growth rates of GDP per capita at the 5% significance level. Holding other independent variables constant, one degree Celsius increase in temperature decrease GDP per capita growth by 0.71%. In addition, population growth significantly affect GDP per capita growth at the 1% level, with a one percentage point increase in population growth decreases GDP per capita growth by 4.4%. Given the average 1.9% current growth rate of population (Table 1), the Indian economy has to growth at approximately 8.7% to make up for its population growth. Yet, in 2015 the economy is only growing at a rate of 7.57% (World Bank). The economic growth and climate dichotomy is apparent in India.

VI. Conclusion

I find one degree Celsius increase in temperature decrease GDP per capita growth by 0.71% and 1% increase in population growth decrease growth by 4.4%. My techniques could have affected my results in several ways. First, I only used data from India with only 42 observations from 1971 to 2012. I averaged the mean annual temperature from monthly data to match the GDP per capita and the CO2 emission annual data. The results could have been improved I could find quarterly data for all independent variables. Moreover, the weather data set used in this paper is not ideal. A gridded spatial weather data might improve the accuracy of weather results.

Second, I employed a very simple version of the Solow growth model to estimate my data. As suggested in Frankhauser and Tol (2005), the Solow model's emphasis on physical capital accumulation makes it less sensitive to climate change. The authors suggested using the Mankiw-Romer-Weil and Romer models for future research, which emphasize human capital and knowledge accumulation, respectively, as they are more sensitive to climate change. A more elaborate endogenous growth model might improve the results of this paper. Third, the model used in this paper and other papers in the literature review section only examined this hypothesis in a closed economy. Globalization may exacerbate the negative impact in one place and alleviate the positive benefits in another because climate change would affect the supply of capital as well as the relative rates of return on investment (Frankhauser & Tol, 2005). Finally, the objective of answering this research question is to figure out policy recommendations and/or ways to internalize this problem to best improve social welfare. The goal of the growth model chosen is to maximize aggregate social welfare. However, there are ethical concerns with this approach to welfare, especially when it comes to climate policy (Sen, 1979).

It is important to note that, the negative relationship between growth and temperature change found in this paper implies a challenge in the reality of the Indian economy. Policymakers in India realize this challenge, and have been implementing significant actions. India has taken steps on renewable energy with increasing installed capacity[5]. The renewable energy goals require continued effort, strong implementation, and improved utilization of capacity, but there are favorable signs. In 2008, India launched its NAPCC, featuring eight national missions, ranging from R&D to sustainable agriculture, with centerpiece programs to scale up solar power

[5] Central Electricity Authority, "Executive Summary: Power Sector," January 2014,

and energy efficiency[6]. With respect to renewable energy, there are great opportunities for India and its international partners. In an Ernst & Young report, in emerging markets "renewable energy potential is attracting high levels of foreign investment, generating new jobs and creating local supply chains.... For investors, renewable energy assets are generating robust returns." [7]Thirdly, with challenges come opportunities, especially for government-government cooperation, public-private partnerships and so on. There are endless opportunities if everyone works together to combat this issue.

The solution for this negative externality is not as simple as simply creating a carbon tax, cap-and-trade, or use property rights, as most economics models typically show. As mentioned in the introduction, the idea of a green economy show great potential for delivering a "triple bottom line" of job–creating economic growth coupled with environmental protection and social inclusion. Admittedly, there are obstacles to realize this potential on a multinational level and in practice. Building a green economy that is not only sustainable but also equitable requires carefully designed policies and investments towards developing countries to benefit from this transition. As suggested by a report by the World Resources Institute report (2012), of particular importance is the need for governance and policy reforms that extend to poor people secure rights over the environmental assets that underpin their livelihoods and well-being, and that ensure a greater voice in decisions affecting how these assets are managed. At the same time, policies and measures such as green protectionism and aid conditionality that

[6] Neha Pahula et al., "GHG Mitigation in India: An Overview of the Current Policy Landscape," World Resources Institute (WRI), WRI Working Paper, March 2014,

[7] "Renewable Energy Country Attractiveness Index (RECAI)," EY, February 2014.

could adversely impact low and middle-income countries and people living in poverty must be avoided if the benefits of an inclusive green economy are to be realized.

While my paper show the benefits of having a sustainable economic growth agenda, future research might examine the costs of a green path for grow to actually suggest practical policies for countries in this climate-conscious world. Another interesting question could be to use econometrics techniques to predict the rate of output growth under the predicted rate of temperature growth and constant carbon emission. Moreover, in this paper I only examined the two channels of climate change on economic development. However, there are more indirect and interdisciplinary channels that temperature can affect long-term economic development. For example, Hsiang, Burke and Miguel (2013) conducted a meta-analysis of studies on the link between climate variability and conflicts from disciplines such as psychology, political science and economics, and found that increase from normal rainfall and temperature increase the intergroup violence by 4% and interpersonal violence by 14%. A country under conflicts is very likely to not involve in meaningful economic activities that contribute to growth. Future research can look at this intersection between disciplines to even further quantify the effects of global warming and economic growth.

In the grand scheme of things, understanding the problem of global warming is crucial in today's interrelated world because this is a problem that carries across disciplines, nations, and generations.

VII. Appendix

Table 1: Descriptive Statistics

Variable	Mean	Std. Dev	Min	Max
Temp change	-.001	.344	-.704	.989
CO_2 Growth (%)	1.593	1.275	-2.413	4.311
Population Growth (%)	1.918	.375	1.27	2.361
GDP Growth (%)	3.704	3.004	-7.383	8.755

Table 2. Regression Results

Dependent variable: GDP per capita annual growth (in %)	
Intercept	12.207
	(1.810)
Temp change	-.705
	(.263)*
CO2growth	-.003
	(0.172)
Population growth	-4.391
	1.087**
R-squared	0.358

Robust standard errors are in parentheses.
*significant at 5%, **significant at 1%

Table 3. Tests for Gauss-Markov assumptions

Assumption	Test Used	Test Statistics	Rejection Rule	Results
Omitted variables	Ramsey	0.41	p-value = 0.74 > 0.05	Passed
Heteroscedasticity	Breusch-Pagan	7.68	p-value = 0.0056 > 0.005	Did not pass
Autocorrelation	Durbin-Watson	2.29	dL = 1.098 dU = 1.518 (4-d) > dU	Passed
Multicollinearity	Variance Inflation Factors	Population: 1.05 Co2: 1.05 Temp: 1.03	< 10	Passed

Bibliography

Adger, W. Neil. “Vulnerability.” Global Environmental Change, 16(3) (2006): 268–81

Alberini, Anna, Aline Chiabai, and Lucija Muehlenbachs. “Using Expert Judgement to Assess Adaptive Capacity to Climate Change: Evidence from a Conjoint Choice Survey.” Global Environmental Change, 16(2) (2006): 123–44.

Almond, Douglas and Janet Currie. “Killing Me Softly: The Fetal Origins Hypothesis,” the Journal of Economic Perspectives, 25(3) (2011), 153-72.

Currie, Janet and Johannes Schmieder. "Fetal Exposure to Toxic Releases and Infant Health,” American Economic Association Papers and Proceedings, 99(2) (2009): 177-83.

Dell, M., B. Jones, and B. Olken. "What do we Learn from the Weather? the New Climate-Economy Literature." *Journal of Economic Literature* (2014).

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." *American Economic Journal: Macroeconomics* 4, no. 3 (2012): 66-95.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. "Temperature and Income: Reconciling New Cross-Sectional and Panel Estimates." American Economic Review 99, 2 (2009): 198-204.

Fankhauser, Samuel and Richard SJ Tol. "On Climate Change and Economic Growth." *Resource and Energy Economics* 27, no. 1 (2005): 1-17.

Hanna, Rema and Paulina Oliva “The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City.” NBER Working Paper 17302. (2011).

Hsiang, S. M. "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America." Proceedings of the National Academy of Sciences 107(35) (2010): 15367-15372

Hsiang, Solomon M., and Amir S. Jina." The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones". No. w20352. National Bureau of Economic Research, 2014.

Hsiang, S. M., M. Burke, and E. Miguel. "Quantifying the Influence of Climate on Human Conflict." *Science (New York, N.Y.)* 341, no. 6151 (Sep 13, 2013): 1235367.

Maddison, David J. “The Amenity Value of the Climate: The Household Production Function Approach.” Resource and Energy Economics, 25(2)(2003): 155–75.

Mendelsohn, Robert O., Wendy N. Morrison, Michael E. Schlesinger, and Natalia G. Andronova. . “Country-specific Market Impacts of Climate Change.” Climatic Change, 45(3– 4)(2000): 553–69.

Mendelsohn, Robert O., Michael E. Schlesinger, and Lawrence J. Williams. “Comparing Impacts across Climate Models.” Integrated Assessment, 1(1)(2000): 37–48.

NASA Earth Observatory. "Fact Sheet: Global Warming" (11, 2010). http://earthobservatory.nasa.gov/Features/WorldOfChange/decadaltemp.php

Nordhaus, William D. “Geography and Macroeconomics: New Data and New Findings.” Proceedings of the National Academy of Science, 103(10) (2006): 3510–17.

Rehdanz, Katrin, and David J. Maddison. “Climate and Happiness.” Ecological Economics, 52(1)(2005): 111–25.

Sen, Amartya. "Personal Utilities and Public Judgements: Or what's Wrong with Welfare Economics." *The Economic Journal* 89, no. 355 (1979): 537-558.

Stern, Nicholas. "What is the Economics of Climate Change?" *World Economics-Henley on Thames-* 7, no. 2 (2006): 1.

Smit, Barry, and Johanna Wandel.. “Adaptation, Adaptive Capacity and Vulnerability.” Global Environmental Change, 16(3) (2006): 282–92.

Tol, Richard S. J. “Climate, Development, and Malaria: An Application of FUND.” Climatic Change, 88(1) (2008): 21–34.

Tol, Richard S. J., and Gary W. Yohe. “The Weakest Link Hypothesis for Adaptive Capacity: An Empirical Test.” Global Environmental Change, 17 (2007): 218–27

Tsigaris, Panagiotis and Joel Wood. "A Simple Climate-Solow Model for Introducing the Economics of Climate Change to Undergraduate Students." *International Review of Economics Education* 23, (9, 2016): 65-81.

Vale, Petterson Molina. "The Changing Climate of Climate Change Economics." *Ecological Economics* 121, (1, 2016): 12-19.

Weitzman, Martin L. "On Modeling and Interpreting the Economics of Catastrophic Climate

Yohe, Gary W., and Richard S. J. Tol.. “Indicators for Social and Economic Coping Capacity—Moving Towards a Working Definition of Adaptive Capacity.” Global Environmental Change, 12(1) (2002): 25–40.

Yohe, Gary W., and Michael E. Schlesinger. “The Economic Geography of the Impacts of Climate Change.” Journal of Economic Geography, 2(2002): 311–41.

Zivin, Joshua Graff and Matthew Neidell. “The impact of pollution on worker productivity,” American Economic Review, 102(7) (2012)

Zivin, Joshua Graff & Matthew Neidell. "Environment, Health, and Human Capital," Journal of Economic Literature, American Economic Association, vol. 51(3), pages 689-730, (9, 2013)

World Resources Institute, "Building an Inclusive Green Economy for All." *Poverty-Environment Partnership*, (6, 2012). http://www.wri.org/sites/default/files/pdf/building_inclusive_green_economy_for_all.pdf.

The Effects of Airline Behavior on Aircraft Accidents

Anneliese S. Brown
Gettysburg College
Browan03@gettysburg.edu

April 17, 2017

Abstract:

The purpose of this paper is to study the effects of specific airline business decisions on aircraft accident propensity. Airline safety affects everyone and has large regulatory and policy implications. Existing research has focused largely on three areas: airline financial health, safety and the resulting effects of accidents. I use both Poisson and Negative Binomial models to study two different airline features: low-cost carriers and flight length, and how they relate to the probability of an aircraft accident. Based on results using a Generalized Negative Binomial model, I find statistically significant evidence at the 99% confidence level that a 1-unit increase in the flight length leads to a 0.11% decrease in the number of accidents. I also find statistically significant evidence at the 99% confidence level that when an airline is classified as a low-cost carrier, the number of accidents decreases by 79.16%. These results indicate that a homogenous safety regulation framework is not appropriate for the airline industry with regard to flight length and cost structure.

I. Introduction

This paper investigates the following two questions: Do budget or low-cost airlines have more aircraft accidents than their counterparts of legacy carriers? Do airlines that provide longer average flight routes have more airplane accidents than their counterparts?

Intuitively, it may be expected that budget airlines only take the minimum safety precautions in order to provide the same services as their counterparts for a lower cost. Thus, an airline classified as a budget airline may have more accidents than a non-budget airline as a result of less investment in safety. Alternately, budget airlines may spend more on safety in order to preserve their reputation and thus experience fewer accidents than their counterparts. A longer flight length may cause an increase in the number of accidents because the more time an aircraft is in the air, the more time there is for an accident to occur. Conversely, if the probability of an accident occurring is greatest during taxiing, takeoff and landing, operators who service short-haul flights may experience more accidents as they rely on quick turnaround times and incur a larger number of takeoffs and landings.

Existing research relating to these topics focuses on the subsequent effects of airplane accidents, the effect of an airline company's financial health on safety and the ways in which airlines make business decisions. The Poisson model for discrete independent variables is used consistently throughout the research related to accident rates. Using this model, existing research has found contradicting evidence on the statistical significance between financial health and safety (Wang, Hofer and Dresner, 2013; Rose, 1990; Golbe, 1986).

This paper closes the gap in existing research between business decisions and safety as I investigate the effect of business decisions, specifically whether or not the airline is a budget airline and flight routes, on accident rates. I make use of count models, specifically Poisson and

Negative Binomial, to answer my questions of interest because my dependent variable, number of accidents, is a positive count variable. While there is an abundant amount of existing research which uses the Poisson model and number of accidents as a dependent variable, no other research has combined these things with independent variables which relate specifically to deliberate business decisions such as flight length and whether or not an airline is a budget airline. Applying the Generalized Negative Binomial model closes a gap in existing research while also generalizing my conclusions by eliminating the assumption that the variance of my dependent variable is linear and equal to the mean.

This topic is important because it relates to issues of safety, transportation routes and business efficiency. Understanding the connection between a firm's decision making incentives and the frequency of accidents can help to prevent airplane accidents in the future through more effective regulation and improved business efficiency. Airlines adapt to changing economic environments while continuously aiming to maximize profits. Recognizing these decisions in relation to accident frequency may help businesses to understand the results of their actions and thus, change them accordingly to increase safety.

These research questions address issues of public policy and customer awareness, both nationally and internationally. The potential risks associated with flying are large and affect many more individuals than just those who fly. It is important for both consumers and the public to recognize the risks associated with flying, particularly if the risk is not uniform across airlines or flight routes. The results might help to determine if a universal regulatory framework for all types of airlines is the best form of safety-related policy.

The paper is organized as follows: In section II, I review related literature, important variables and common models used to answer similar questions. In section III, I outline the

Poisson and Negative Binomial models, my hypotheses and describe my research method. In section IV, I discuss the data and define each variable. In section V, I present the empirical results of my research. In section VI, I conclude my analysis with the implications and applicability of my results.

II. Literature Review

Existing research related to the effects of airline business decisions on aircraft accidents falls into two categories. A first line of this research focuses on safety as it relates to profits, financial health, investment and demand. A second line of this research studies business decisions as they relate to both topics of low cost competition and flight routes. My research provides a link between the existing yet isolated research on business decisions and safety.

First I discuss existing research relating to safety, and accidents in particular. A useful study is conducted by Golbe (1986), who examines the relationship between profits and safety precautions taken by an individual airline. She implements both cross-sectional and time-series techniques on data of U.S. airlines aggregated at the industry level from 1952 – 1972. Golbe (1986) emphasizes key variables of number of departures, load factors and net income, as a measure of profitability. Golbe (1986) uses airline accident experience as a measure of safety and models both accident experience and net income as dependent variables. Her research concludes that there is no significant relationship between profits and safety (Golbe 1986).

Bornstein and Zimmerman (1988) investigate the effect of an aircraft accident on flight demand using time series data for U.S. air carriers from 1960 - 1985, modeling revenue per passenger as a function of elapsed time since an accident, seasonal dummies, and firm and time fixed effects. They conclude that although an accident results in a significant $4.5 million loss for a firm, there is not a significant relationship between accidents and flight demand before deregulation of the industry and only weak evidence of an effect on demand after deregulation (Bornstein and Zimmerman 1988).

Rose (1990) studies the effect of an airlines' financial health on accident rates using panel data across thirty-five U.S. airlines from 1957 - 1986. She measures safety as a risk distribution, gathering data on both safety investment and physical conditions in which firms operate their aircraft. Similar to Wang, Hofer and Dresner (2013), Rose (1990) uses the Poisson probability distribution to model the dependent variable of accident rate. Using fixed effects, Rose (1990) separately models both total accidents and fatal accidents as an effect of departures (system departures in thousands), average stage length (thousands of miles), carrier type, foreign flights, size of firm, airline operating experience (billions of miles) and time variant characteristics of technology. While I use some of the same variables, all of my models use only total accidents as the dependent variable. She concludes that an increase in operator profit leads to a statistically significant decrease in accident rates (Rose 1990).

Wang, Hofer and Dresner (2013) measure the effect of safety investment on accident propensity and financial health. They use panel data on airlines from the National Transportation Safety Board (NTSB) and the U.S. Department of Transportation (DOT) from 1991 – 2008. Due to the entry and exit of airlines within the industry, they treat their panel dataset as unbalance. These authors model Poisson functions of number of accidents as I will do in this paper. Further, they create a variable for average accidents per departure, substituting this as the dependent variable in their reduced form model. They conclude that safety investment reduces accident propensity and find no relationship between financial condition and accident propensity nor financial condition and safety investment (Wang, Hofer and Dresner 2013).

Other pertinent research emphasizes airline business decisions in relation to budget airlines and flight routes. Fischer and Kamerschen (2003) examine the relationship between low-cost operator presence at airports and average airfare. They use the DOT's form 41 for Air Carrier Traffic Statistics to crease a time-series data consisting of the four quarters of 1996. They use a cross section regression model in which the dependent variable is average yield (price/distance) with independent variables including total passengers, distances (stage length) and ValueJet. They measure ValueJet as a binary variable valued at 1 if the airline ValueJet services a particular airport and 0 otherwise; this variable accounts for the presence of low cost carriers at any given airport. Fischer and Kamerschen (2003) conclude that the presence of low-cost competition for a particular route has a statistically significant negative effect on revenue.

Garrow, Holte and Mumbower (2012) study the phenomenon of product de-bundling as it relates to the emergence of low-cost carriers. They collect airline data from individual airline websites regarding baggage fees, cancelation fees, seat fees and ticket change fees. They find statistically significant evidence that low-cost carriers are the most likely carriers to charge additional fees.

Gillen and Hazledine (2015) study the effect of regional route fluctuations on firm pricing strategy. They use data from a total of six regions on various flight routes and use the Hirschman-Herfindahl index to account for airline concentration. They find no significant relationship between supply of seats and route length but find a significant difference in airfares across regions (Gillen and Hazledine 2015).

The limitation of prior research addressed in this paper is the lack of research examining the cause of accidents as related to business decisions. Although there is abundant research on airline safety and business decisions relating to budget airlines and flight routes, these topics have only been studied in isolation from each other. Current research focuses on the effects of accidents but little has been examined regarding the cause of the accidents. My research utilizes many of the same variables, models and tests as those introduced above but I investigate the link between these industry characteristics and accidents to determine the effect of both flight length and budget airlines on accidents.

III. Model and Methodology

I test whether budget airlines have more accidents than their counterparts and whether an increase in flight length leads to an increase in the number of aircraft accidents using a unified model. I hypothesize that budget airlines have more accidents than their counterparts as budget airlines may cut safety costs in order to provide cheaper fares than legacy or non-budget airlines. I expect an increase in flight length to cause a decrease in the number of accidents as I suspect that operators who provide long-haul flights invest more in safety and experience fewer takeoffs and landings, which are most damaging to the engines and aircraft, than operators who provide more frequent short-haul services.

I use a unified Negative Binomial model to answer my two questions of interest because of the similarity in potential control variables. I have included control variables which intuitively affect aircraft accidents without being directly related to flight length or whether or not an airline is a budget airline.

Similar to previous research such as that of Wang, Hofer and Dresner (2013) and Rose (1990), I begin by using the Poisson model to estimate the relationship between flight distance, budget airlines and accidents. The Poisson model is applicable to this data set because the dependent variable, aircraft accidents, is a count variable. This model requires the dependent variable to be a discrete, non-negative value including zero, which is true of aircraft accidents.

As the number of accidents may equal zero for any given year, we cannot take the log of the dependent variable. Instead, I use the following exponential function:

$$E(y|x_1, x_2, \ldots, x_k) = \exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k) = \exp(X_{it}\beta) \quad (1)$$

Where x_{it} represents various independent or control variables for airline i at time t while β represents corresponding estimated coefficients. However, with the Poisson model, equation 1 can be simplified because the distribution is determined by the mean; in fact, the mean and variance of Y are equal in the Poisson model. This is represented in the following equation:

$$P(Y_{it}) = (\exp[-\exp(x_{it}\beta)][\exp(x_{it}\beta)^{Y_{it}}] / Y! \quad (2)$$

Where $P(Y_{it})$ is the probability of Y accidents for airline i at time t, $\exp(x_{it}\beta)$ is the expected number of accidents for airline i at time t or the average accident rate per departure and $Y = 0, 1, 2, \ldots, \exp(x_{it}\beta) > 0$.

Further, in the Poisson model, the mean and the variance are equal. This is represented in the following equation:

$$E(Y_{it}) = \exp(x_{it}\beta) = Var(Y_{it}) \quad (3)$$

However due to the nature of accident rates, there may be more or less variation in the data than expected under Poisson. Thus, the Negative Binomial model may provide a better fit for the relationship of interest as the Poisson model may produce biased coefficient estimates in the presence of over- or under-dispersion (Shankar, Mannering and Barfield, 1995).[8] As shown

[8] As stated by Shankar, Mannering and Barfield (1995), "It is well known, based on the finding of many previous research efforts, that accident frequency data tend to be over-dispersed, with the variance being significantly greater than the mean" (Shankar, Mannering and Barfield, 1995).

by Shankar, Mannering and Barfield (1995), who study the effect of roadway accidents using the Negative Binomial model, equation 3 can be altered to represent the relationship with a Negative Binomial model in the following way:

$$\mathrm{Var}(Y_{it}) = E(Y_{it})[1 + \alpha E(Y_{it})] \quad (4)$$

From the above equation, the variance is no longer equal to the mean when using the Negative Binomial model due to the existence of the term $[1 + \alpha E(Y)]$, when $\alpha E(Y) \neq 0$. When α is equal to 0, $\mathrm{Var}(Y) = E(Y)$ and I am left with variance which is represented in the Poisson model. However, when α is not equal to zero, there is evidence of either over - or under-dispersion. It is important to note that the Negative Binomial model is only applicable in the presence of over-dispersion using the Poisson distribution, in which the variance is greater than the mean; when there is under-dispersion using the Poisson distribution, the Negative Binomial model is not valid (Shankar, Mannering and Barfield, 1995). As used by Shankar, Mannering and Barfield (1995), the following equation represents the probability distribution using the Negative Binomial model:

$$P(Y_{it}) = \frac{\Gamma(\theta + Y_{it})}{\Gamma(\theta) Y_{it}!} (u_{it}{}^{\theta})(Y_{it})(1 - u_{it}) \quad (5)$$

Where $u_{it} = \theta/(\theta + \exp(x_{it}\beta))$, $\theta = 1/\alpha$ and $\Gamma()$ represents a function of gamma (Shankar, Mannering and Barfield, 1995).

I will also implement the Generalized Negative Binomial model in which the form of the variance is not assumed to be linear, as it is in the Negative Binomial model. Thus, the Generalized Negative Binomial model makes my results more precise as the form of the variance is not assumed to be linear.

In my regression, I specify the following model:

$$E(\mathrm{Accidents}_{it}) = \mathrm{Departures}_{it} * \exp(\beta_0 + \beta_1 \mathrm{Budget\ Airline}_{it} + \beta_2 \mathrm{Average\ Stage} \quad (6)$$

$\text{Length}_{it} + u_{it})$

Consistent with existing research, the expected number of accidents is the number of departures multiplied by the average accident rate per departure because of the stochastic or random nature of accident data (Wang, Hofer and Dresner, 2013; Rose, 1990).

Based on equation 6, my hypothesis that budget airlines are less safe is supported when $\beta_1 > 0$. When an airline is considered to be a budget airline and β_1 is positive, there is a positive effect on the expected value of accidents and thus my hypothesis is supported. My hypothesis that an increase in average flight length leads to a decrease in accidents is supported when $\beta_2 < 0$, as an increase in the average stage length should be negatively related to the number of accidents, according to my prediction.

IV. Data

I use data from the National Transportation Safety Board (NTSB), Federal Aviation Administration (FAA) and the Bureau of Transportation Statistics (BTS) as has been used in previous research. To minimize measurement errors, I make use of a consolidated data set from the Airline Data Project at Massachusetts Institute of Technology (MIT) which contains data from the BTS form 41 which gathers quarterly billing data and monthly airline data. Using these data sources, I construct a panel data set which varies across fifteen U.S. airlines over twenty-one years, from 1995 through 2015. Data on all fifteen airlines in the MIT project is included; a list of these airlines along with the years for which data is available for each airline can be found in table 1 of section VIII.

Due to mergers and acquisitions within the industry, there is no data for all twenty-one years for all fifteen airlines. It is important to note that while this is considered “missing data” in terms of the raw data, the data is not in fact missing as the airlines simply were not in existence or operating during the years in which I do not have data. I have verified with individual airline websites that the years in which there is “missing data” align with mergers, acquisitions, entries or closings within the industry. Because of these gaps in the data, together with the fact that my panel is relatively narrow in the sense that I only include data on fifteen airlines, I continue my analysis by treating my panel data set as cross sectional data as done by Golbe (1986).[9]

I use a dependent variable of aircraft accidents as used by Golbe (1986) Borenstein and Zimmerman (1988), and Rose (1990) and Wang, Hofer and Dresner (2013). I have gathered the information from the FAA which has the NTSB’s Accident and Incident Database. According to the FAA Aviation Safety Information Analysis and Sharing (ASIAS), an aircraft accident is defined as “an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight and all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage” (ASIAS). I have included all U.S. aircraft accidents, including fatal and non-fatal, from January 1995 through December 2015 for all fifteen airlines used in my dataset. Due to the nature of aircraft accidents, this variable is a non-continuous, discrete count variable.

In order to answer my question regarding the effect of flight length on accident propensity, my primary independent variable of interest is average stage length which is used by Golbe (1986), Rose (1990) and Wang, Hofer and Dresner (2012). This variable is available in the consolidated MIT study, which pulls data from the BTS form 41, and measures the total number

[9] I report results using fixed effects in tables 9 and 10 of section VIII. While the signs of the average stage length and budget airline variable coefficients are the same when implementing cross sectional data methods, neither coefficient is statistically significant at even the 10% significance level.

of miles flown divided by the total number of departures. Thus, the average stage length represents average flight length of each departure, measured in miles.

In order to answer my question regarding the effect of being a budget airline on accident propensity, I have investigated three potential independent variables including a dummy variable, total baggage fee and total cancelation fee. Based on the research of Garrow, Holte and Mumbower (2011), who study the phenomena of product de-bundling in the airline industry, I have created a binary variable valued at 1, which is attributed to a budget or low-cost carrier and 0, which is attributed to a non-budget or legacy airline. Their research includes a total of eleven U.S. airlines, ten of which I also include in my data set. Although Garrow, Holte and Mumbower (2011) do not precisely define budget or legacy carriers, they state that the legacy carriers "participate in well-established alliances that enable them to further increase the number of destinations they can serve; these major carriers also tend to have a moderate number of other airline partners that further enhance their networks" (Garrow, Holte and Mumbower, 2011). Based on their classification of low cost carriers, I identify the following same four budget airlines: Southwest, AirTran, JetBlue and Frontier.[10] I classify the remaining eleven airlines in my data set as legacy or non-budget airlines, six of which are also considered to be legacy carriers by Garrow, Holte and Mumbower (2011). Thus, I assume that the five airlines included in my data set, but not included in the specific reference literature, are also legacy carriers.

I have also included total baggage fees and total cancelation fees as potential key independent variables to account for budget airlines. I have gathered both fee variables from the consolidated MIT study, both of which are measured in thousands of U.S. dollars. I use the

[10] In conducting further company research, I find both Allegiant Air and Sprit to be considered budget airlines. While I do not include these classifications in my primary results, tables 11 and 12 in section VIII show the results of my research with additionally categorizing both Allegiant Air and Spirt as budget airlines.

conclusion of Garrow, Holte and Mumbower's (2011) research that budget airlines are the most likely to charge additional ancillary fees. Thus, I use the fee variables, interchangeably, as proxy variables to represent an airline behaving "more like a budget airline." I assume that a 1-unit increase in either fee variable indicates an airline behaving more like a budget airline. However, due to structural breaks and variation across low-cost carriers, as mentioned by Garrow, Holte and Mumbower (2011), there is potential bias in the way these fee variables may represent budget airlines. Due to the difficulty in defining a budget airline precisely, as shown in previous research, I include all three variables (baggage fee, cancelation fee, budget airline) to interchangeably account for budget airlines.

I use the number of incidents reported for each airline in each year, from the FAA ASIAS as done by Rose (1990). An incident is defined as "an occurrence other than an accident, associated with the operation of an aircraft, which affects or could affect the safety of operations" (ASIAS). Due to the nature of aircraft incidents, this variable is a non-continuous, discrete count variable. As I have not been able to include the average age of the aircraft, I presume that incidents will work to control for age of aircraft-related characteristics, which may affect accidents as an increase in incidents intuitively leads to an increase in the probability of an accident.

The following control variables that I mention are all gathered from the MIT project and thus the BTS form 41. I control for size of aircraft by dividing average seat miles (ASM) by the total number of miles flown. ASM is an industry standard measurement of utilization and airline output and measures the total number of available seats per departure multiplied by the total number of miles traveled. However, because ASM includes mileage, there is potential for collinearity with my independent variable of interest, average stage length. Thus, I divide ASM by miles and am left with the average number of seats per departure.

I control for airline size by including the number of functioning aircraft and total operating revenue measured in billions of U.S. dollars. I include the average salary of both pilots and co-pilots, measured in U.S. dollars, to control for pilot experience and skill level. I include maintenance per aircraft in which I divide the total maintenance expenditure, measured in thousands of U.S. dollars, by the total number of aircraft in the fleet to account for maintenance cost per aircraft. Summary statistics of all variables can be found in table 2 of section VIII.

While I attempt to create a robust data set including industry standard, intuitively sound and previously used variables, I have not been able to collect data on average aircraft age and airline profitability. Aircraft incidents may serve as a proxy variable for aircraft age while total revenue may serve as a proxy variable for profitability, although neither fully capture the effect of the absent variables.

V. Results

I present my basic Poisson regressions in table 3 of section VIII. In running the most simplified regression presented in column 1, the sign of the coefficient of interest in positive and statistically significant at the 99% level. When I include control variables to the same model, as seen in columns 2 and 3 of table 3, the estimated coefficient of the average stage length variable becomes negative while remaining statistically significant. The results in column 1 indicate that a 1-unit increase in flight length leads to a 0.059% increase in the number of accidents while the results in columns 2 and 3 indicate that a 1-unit increase in average stage length leads to a 0.10% decrease in the number of accidents, which are all statistically significant at the 99% confidence level. Further, seen in the regressions in columns 2 and 3 of table 3, when an airline is a budget airline the number of accidents decreases by 61.74% and 61.41%, respectively.[11]

Based on the regression represented in column 3 of table 3 in section VIII, I run both Deviance and Pearson goodness-of-fit tests. With P-values of 0.0017 and 0.0005, respectively, I reject the null hypothesis that the Poisson model fits my relationship of interest well.

The regressions in table 5 utilize the Negative Binomial model. The basic regression in column 1 indicates that a 1-unit increase in the average flight length leads to a 0.07% increase in the number of accidents which is statistically significant at the 5% significance level. This positive sign of the coefficient is similar to that of the basic regression using the Poisson model shown in column 1 of table 3. When I implement the Negative Binomial model and run the LR test of alpha = 0, I get a P-value of 0.000. Thus, I reject the null hypothesis that alpha is equal to zero and conclude the Negative Binomial model to be a good fit for my data as I find over-dispersion and cannot assume the variance of accidents to be equal to the mean or for alpha to be equal to 0.[12]

When I include control variables to the basic Negative Binomial model, as seen in columns 2-5 of table 5, the estimated coefficient of the average stage length variable becomes negative. The difference between the regressions represented in columns 2-4 is the variable in which I use to account for budget airline. In column 2 of table 5, I include the baggage fee variable while in column 3 of table 5, I include the cancelation fee variable. Intuitively I expect an increase in baggage or cancelation fees to lead to an increase in the number of accidents, as I assume that airlines that charge higher fees behave more similarly to budget airlines. From the

[11] The output in table 4 of section VIII represents the marginal effect interpretations associated with the Poisson regressions represented in table 3.
[12] Further, because the mean of accidents is 1.28 while the variance is 3.43, I can simply identify the presence of over-dispersion within my data.

regression output seen in columns 2 and 3, neither estimated coefficient of the baggage nor cancelation fee variable is statistically significant at even the 90% confidence level. Due to the insignificance of the estimated coefficients, structural breaks and potential measurement error, I conclude that neither baggage nor cancelation fees accurately represent budget airlines.[13]

The regressions represented in columns 4 and 5 of table 5 include a binary budget airline variable as opposed to a fee variable to account for budget airlines. Both regressions show that a 1-unit increase in the average stage length leads to 0.11% decrease in the number of accidents, which is statistically significant at the 1% significance level. The estimated coefficients of the budget airline variable are large in magnitude and statistically significant at the 99% level; I find that when an airline is a budget airline, the number of accidents decreases by 71.84% and 79.16%. It is worth noting the changes in significance of the estimated coefficients of the average stage length, maintenance per aircraft, number of seats and incidents variables from column 3 to column 4.[14] The large magnitude of the budget airline coefficients in columns 4 and 5 may be explained by the measurement error in the variable and thus I am not confident in these conclusions drawn to answer my question regarding the effect of budget airlines on accident propensity.

The regressions represented in table 7 are the same as those presented in table 5, although they implement the Generalized Negative Binomial model as opposed to the Negative Binomial model. The results are almost identical to those of the Negative Binomial model but because the generalized model even further loosens the assumptions of the variance structure, I have decided

[13] The output in table 6 of section VIII represents the marginal effect interpretations associated with the Negative Binomial regressions represented in table 5.
[14] In line with previous literature, I also run these regressions with an added time trend variable in order to account for advances in technology over time which may decrease accident propensity. However, because the estimated coefficient of the time trend variable is consistently statistically insignificant, I do not include it in my final results.

to treat the regression in column 5 of table 7 as my final regression. As some of the estimated coefficients in the regression represented in column 4 of table 7 are not statistically significant at even the 90% confidence level, I run the regression in column 5 of table 7 in order to more accurately estimate the coefficients of interest.

From the regression output represented in column 5 of table 7, I have statistically significant evidence at the 1% significance level that a 1-unit increase in average stage length leads to a 0.11% decrease in the number of accidents while I have statistically significant evidence at the 1% significance level that when an airline is classified as being a budget airline, the number of accidents decreases by 79.16%. All of the signs of the estimated coefficients align with intuition.[15]

The negative and statistically significant coefficient of the average stage length variable does not align with the research of Wang, Hofer and Dresner (2013) nor Rose (1990), who both find statistically significant positive coefficient estimates.[16] However, the negative sign of the average stage length coefficient does align with the findings of Golbe (1986) though she does not find the negative average stage length coefficients to be statistically significant at any level.

VI. Conclusion

From the previous section, I conclude that there is statistically significant evidence at the 1% significance level that a 1-unit increase in average stage length leads to a 0.11% decrease in

[15] The output in table 8 of section VIII represents the marginal effect interpretations associated with the Negative Binomial regressions represented in table 7. Column 5 of table 8 in section VIII represents the marginal effects corresponding to my final regression in which a 1-unit increase in average flight length is associated with 0.00093 fewer accidents and an airline being a budget carrier is associated with 0.57 fewer accidents.

[16] Wang, Hofer and Dresner (2013) find statistically significant evidence at the 1% level that "longer stage lengths are associated with a higher accident propensity" (Wang, Hofer and Dresner, 2013).

the number of aircraft accident. I have statistically significant evidence at the 1% significance level that when an airline is a budget airline, the expected value of an accident decreases by 79.16%.

As stated in section III, I hypothesize that an increase in the average stage length leads to a decrease in the number of accidents as operators who provide short-haul services incur more takeoffs and landings, which put the engines and aircraft under the most stress. Based on the negative sign of the coefficient of the average stage length variable, this hypothesis is supported. I also hypothesize the number of accidents increases when an airline is a budget airline as budget airlines may spend less on safety in order to provide comparable services to non-budget airlines. However, due to the negative sign of the estimated coefficient of the binary budget airline variable, my hypothesis relating to budget airlines is not supported.

Intuitively the negative and statistically significant, at the 99% level, coefficients of both independent variables of interest may be explained by airline business decisions. Based on my results, an increase in average flight length leads to a decrease in the number of accidents. This may mean that carriers that provide longer flights put more resources toward flight safety as opposed to carriers which provide flights with shorter average stage lengths.[17] After further investigating the specific position of each accident during the flight, I find that 30.58% of accidents occur while the aircraft is on the ground, 16.25% of accidents occur while at cruising level, 44.35% occur during either takeoff or landing and 8.82% of accidents occur with an "other" or undefined reason. Thus, it makes sense that short-haul carriers, that experience a larger number of takeoffs and landings, have more accidents as 44.35% of accidents occur at

[17] In testing the effect of average stage length on maintenance expenditure per aircraft, I find statistically significant evidence that a 1-unit increase in flight length leads to an increase in maintenance expenditure per aircraft. Thus I conclude that longer-haul carriers have higher expenditure on maintenance per aircraft than that of their counterparts.

takeoff and landing. Conversely, it makes sense that airlines that provide longer flight lengths have fewer takeoffs and landings than their short-haul provider counterparts and thus incur a smaller number of accidents. These results indicate that airlines that provides longer-haul flights have inherently different operating methods and flight safety structures than those of shorter-haul carriers.

Based on the results, when an airline is a budget airline, the number of accidents decreases by an extremely large magnitude. Although these results may support the idea that budget airlines may be sensitive to an unsafe reputation and thus may allocate more resources toward safety than that of their counterparts in order to maintain strong reputations of safety, after further investigation I find, this is not the case.[18] While these results indicate that budget airlines have different safety structures than that of non-budget or legacy airlines I am not confident in my results regarding the budget airline variable. The unrealistically large coefficient signifies an error within the application. I suspect measurement error of the budget airline variable to be a large potential issue within my model which leaves me with little confidence in my results associated to the budget airline variable.[19]

Ultimately these results indicate that a homogenous airline regulation framework is not appropriate for budget nor long-haul airlines. With statistically significant evidence that both an increase in average flight length and an airline being a budget airline lead to a decrease in the number of aircraft accidents, it is apparent that not all airlines should be held to identical

[18] In further investigation, I find statistically significant evidence at the 1% level that budget airlines spend less on maintenance per aircraft than non-budget airlines.

[19] It is worth noting that in testing the difference between accident rates of budget and non-budget airlines, I find the mean of accidents for budget airlines to be .8133 while that of non-budget airlines is 1.45. Thus my regression results and conclusions align with the variable within my data set; thus I assume there to me measurement error within the variable and an “outside” factor affecting the large decrease in accident rate of budget airlines.

benchmarks. Airline business decisions have shown to significantly affect aircraft accident rates; thus airlines should be regulated and upheld to specific standards based on these decisions.

VII. References

Works Cited

Bornstein, Severin, and Martin Zimmerman. "Market Incentives for Safe Commercial Airline Operation." *American Economic Review* 78.5 (1988): 913-35.

Cameron, Adrian Colin, and Pravin K. Trivedi. *Microeconometrics Using Stata*. College Station, Texas: Stata, 2010.

FAA Accident and Incident Data System (AIDS). N.d. Raw data. U.S. Department of Transportation Federal Aviation Administration, Washington, DC.

Fischer, Thorsten, and David R. Kamerschen. "Measuring Competition in the U.S. Airline Industry Using the Rosse-Panzar Test and Cross-Sectional Regression Analysis." *Journal of Applied Economics* 6.1 (2003): 73-93.

Garrow, Laurie A., Susan Hotle, and Stacey Mumbower. "Assessment of Product Debundling Trends in the US Airline Industry: Customer Service and Public Policy Implications." *Transportation Research Part A: Policy and Practice* 46.2 (2012): 255-68.

Gillen, David, and Tim Hazledine. "The Economics and Geography of Regional Airline Services in Six Countries." *Journal of Transport Geography* 46 (2015): 129-36.

Golbe, Devra L. "Safety and Profits in the Airline Industry." *The Journal of Industrial Economics* 34.3 (1986): 305-18.

Greene, William H. *Econometric Analysis*. 2nd ed. New York, NY: Maxwell Macmillan International Group, 1993.

Gujarati, Damodar N. *Basic Econometrics*. New York: McGraw-Hill/Irwin, 2003. Print.

NTSB Aviation Accident and Incident Data System (NTSB). N.d. Raw data. U.S. Department of Transportation Federal Aviation Administration, Washington, DC.

Rose, Nancy L. "Profitability and Product Quality: Economic Determinants of Airline Safety Performance." *Journal of Political Economy* 98.5 (1990): 944-64.

Shankar, Venkataraman, Fred Mannering, and Woodrow Barfield. "Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies." *Accident Analysis & Prevention* 27.3 (1995): 371-389.

Swelbar, William S., and Peter P. Belobaba. Global Airline Industry Program; Airline Data Project. N.d. Raw data. Massachusetts Institute of Technology, Cambridge, MA.

Wang, Zuozheng, Christian Hofer, and Martin Dresner E. "Financial Condition, Safety Investment and Accident Propensity in the US Airline Industry: A Structural Analysis." *Transportation Research Part E: Logistics and Transportation Review* 49.1 (2013): 24-32.

Wooldridge, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 5th ed. Mason: Cengage Learning, 2013.

Wooldridge, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 6th ed. Boston: Cengage Learning, 2016.

VIII. Supporting Tables

Table 1: Airlines included in the Analysis and Available Observations for Each Airline

Airline	Available Data (Inclusive)
AirTran Airways	1995 - 2011
Alaska Airlines	1995 - 2015
Allegiant Air	2000 - 2015
America West Airlines	1995 - 2007
American Airlines	1995 - 2015
Continental Airlines	1995 - 2011
Delta Air Lines	1995 - 2015
Frontier Airlines	1995 - 2015
Hawaiian Airlines	1995 - 2015
JetBlue Airways	2000 - 2015
Northwest Airlines	1995 - 2009
Southwest Airlines	1995 - 2015
Spirit Airlines	1995 - 2015
United Airlines	1995 - 2015
US Airways	1995 - 2014

Table 2: Summary Statistics

Airline	Units	Observations	Mean	Standard Deviation	Minimum	Maximum
Number of Accidents	Count	283	1.282686	1.852146	0	9
Average Stage Length	Total Miles Flown / Aircraft Departures	282	935.0396	278.6823	256.0417	1720.326
Budget Airline	Binary	282	.2659574	.4426272	0	1
Baggage Fee	Thousand U.S. $	252	108026.3	198955.6	20.54	1125846
Cancelation Fee	Thousand U.S. $	237	267122.2	460147.5	2690.4	3117848
Number of Aircraft in Fleet	Count	282	265.3815	235.9029	.9863014	971.8904
Pilot and Co-Pilot Average Salary	U.S. $	260	130379.8	113164.9	16694.64	1859096
Maintenance Per Aircraft	Maintenance Expenditure ($1,000)/ Fleet Size	264	2420.938	996.9043	428.5007	5586.672
Total Revenue	Billion U.S. $	274	8.340641	9.127099	.0536117	41.08443
Number of Incidents	Count	287	8.355401	10.39193	0	58
Number of Seats	ASM / Miles	282	160.6285	32.88094	93.34768	265.6832

Table 3: Poisson Regressions
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)
Average Stage Length	.0005898 (.0001858)***	-.0010085 (.0003297)***	-.001009 (.0002927)***
Budget Airline		-.6174235 (.2218317)***	-.6140813 (.1752401)***
Maintenance / Aircraft		-.0002912 (.0001192)**	-.0003702 (.000085)***
Fleet Size		.0041103 (.0006356)***	.0042339 (.0003065)***
Pilot Salary		-9.72e-07 (1.41e-06)	
Number of Seats		-.0036111 (.0041183)	
Number of Incidents		.0033688 (.0055896)	
Total Revenue		.0030551 (.0154579)	
Intercept	-.3181392 (.1898117)*	1.181443 (.7035263)*	.6731913 (.2754536)**
Robust Standard Errors?	No	Yes	Yes
Pseudo R^2	0.0095	0.3019	0.3031
Chi Squared	9.96	222.40	216.34
Number of Observations	282	260	264

Notes: Standard Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level

Table 4: Poisson Regression Interpretations (Marginal Effects)
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)
Average Stage Length	.0007448 (.00023)***	-.0008845 (.00029)***	-.0008721 (.00025)***
Budget		-.4748103 (.1541)***	-.4656173 (.12154)***
Maintenance / Aircraft		-.0002553 (.0001)**	-.00032 (.00007)***
Fleet Size		.0036048 (.00055)***	.0036592 (.00023)***
Pilot Salary		-8.53e-07 (.00000)	
Number of Seats		-.003167 (.0036)	
Number of Incidents		.0029544 (.0049)	
Total Revenue		.0026794 (.01355)	

Notes: Standard Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level

Table 5: Negative Binomial Regressions
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)	(4)	(5)
Average Stage Length	.0007184 (.0003591)**	-.0005928 (.0002924)**	-.0006922 (.0003893)*	-.0011078 (.000349)***	-.0011111 (.0003296)***
Budget Airline				-.7184402 (.2333835)***	-.7916491 (.2245004)***
Baggage Fee		7.40e-07 (7.01e-07)			
Cancelation Fee			3.28e-07 (2.61e-07)		
Maintenance / Aircraft		-.0002036 (.000133)	-.0000663 (.0001418)	-.0002223 (.0001257)*	-.0002559 (.0001094)**
Fleet Size		.0044434 (.0009075)***	.0042938 (.0010027)***	.0042367 (.0006174)***	.0044162 (.0002986)***
Pilot Salary		-1.80e-06 (2.05e-06)	-3.28e-06 (2.90e-06)	-9.61e-07 (1.14e-06)	
Number of Seats		-.0018323 (.0040688)	-.0036962 (.0043345)	-.0064759 (.0043613)*	-.0060738 (.0040212)
Number of Incidents		.0132486 (.0059269)***	.018139 (.0079114)***	.0037346 (.0058818)	
Total Revenue		-.0226676 (.0324717)	-.0210113 (.0333182)	.0033395 (.0155194)	
Intercept	-.441502 (.353355)	.2237119 (.6562952)	.4707563 (.663461)	1.538596 (.7558503)*	1.450162 (.6742874)**
Robust Standard Errors?	No	Yes	Yes	Yes	Yes
Pseudo R^2	0.0046	0.1718	0.1676	0.1864	0.1865
Chi Squared	4.00	225.76	225.04	254.70	237.78
Number of Observations	282	244	230	260	264

Notes: Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level

Table 6: Negative Binomial Regression Interpretations (Marginal Effects)
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)	(4)	(5)
Average Stage Length	.0009044 (.00045)**	-.00056 (.00028)**	-.0005847 (.00033)*	-.0009521 (.0003)***	-.0009337 (.00027)***
Budget Airline				-.5315556 (.15305)***	-.5652535 (.1409)***
Baggage Fee		6.99e-07 (.00000)			
Cancelation Fee			2.77e-07 (.00000)		
Maintenance / Aircraft		-.0001924 (.00012)	-.000056 (.00012)	-.0001911 (.00011)*	-.0002151 (.00009)**
Fleet Size		.0041975 (.00086)***	.0036268 (.00083)***	.0036416 (.00054)***	.0037113 (.00024)***
Pilot Salary		-1.70e-06 (.00000)	-2.77e-06 (.00000)	-8.26e-07 (.00000)	
Number of Seats		-.0017309 (.00385)	-.0031221 (.00365)	-.0055662 (.00374)	-.0051043 (.00338)
Number of Incidents		.0125155 (.00569)**	.0153216 (.00691)**	.00321 (.00505)	
Total Revenue		-.0214132 (.03074)	-.0177477 (.02811)	.0028704 (.01333)	

Notes: Standard Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level

Table 7: Generalized Negative Binomial Regressions
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)	(4)	(5)
Average Stage Length	.0007184 (.0003591)**	-.0005928 (.0002924)**	-.0006922 (.0003893)*	-.0011078 (.000349)***	-.0011111 (.0003296)***
Budget Airline				-.7184398 (.2333834)***	-.791649 (.2245004)***
Baggage Fee		7.40e-07 (7.01e-07)			
Cancelation Fee			3.28e-07 (2.61e-07)		
Maintenance / Aircraft		-.0002036 (.000133)	-.0000663 (.0001418)	-.0002223 (.0001257)*	-.0002559 (.0001094)**
Fleet Size		.0044434 (.0009075)***	.0042938 (.0010027)***	.0042367 (.0006174)***	.0044162 (.0002986)***
Pilot Salary		-1.80e-06 (2.05e-06)	-3.28e-06 (2.90e-06)	-9.61e-07 (1.14e-06)	
Number of Seats		-.0018323 (.0040688)	-.0036962 (.0043345)	-.0064759 (.0043613)	-.0060738 (.0040212)
Number of Incidents		.0132486 (.0059269)***	.018139 (.0079114)***	.0037346 (.0058818)	
Total Revenue		-.0226676 (.0324717)	-.0210113 (.0333182)	.0033396 (.0155194)	
Intercept	-.441502 (.353355)	.2237118 (.6562952)	.4707563 (.6634611)	1.538595 (.7558502)**	1.450162 (.6742874)**
Robust Standard Errors?	No	Yes	Yes	Yes	Yes
Pseudo R^2	0.0046	0.1718	0.1676	0.1864	0.1865
Chi Squared	4.00	225.76	225.04	254.70	237.78
Number of Observations	282	244	230	260	264

Notes: Standard Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level

Table 8: Generalized Negative Binomial Regression Interpretations (Marginal Effects)
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)	(4)	(5)
Average Stage Length	.0009044 (.00045)**	-.00056 (.00028)**	-.0005847 (.00033)*	-.0009521 (.0003)***	-.0009337 (.00027)***
Budget Airline				-.5315554 (.15305)***	-.5652534 (.1409)***
Baggage Fee		6.99e-07 (.00000)			
Cancelation Fee			2.77e-07 (.00000)		
Maintenance / Aircraft		-.0001924 (.00012)	-.000056 (.00012)	-.0001911 (.00011)*	-.0002151 (.00009)**
Fleet Size		.0041975 (.00086)***	.0036268 (.00083)***	.0036416 (.00054)***	.0037113 (.00024)***
Pilot Salary		-1.70e-06 (.00000)	-2.77e-06 (.00000)	-8.26e-07 (.00000)	
Number of Seats		-.0017309 (.00385)	-.0031221 (.00365)	-.0055662 (.00374)	-.0051043 (.00338)
Number of Incidents		.0125155 (.00569)**	.0153216 (.00691)**	.00321 (.00505)	
Total Revenue		-.0214132 (.03074)	-.0177477 (.02811)	.0028704 (.01333)	

Notes: Standard Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level

Table 9: Fixed Effects Negative Binomial Regressions
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)	(4)	(5)
Average Stage Length	-.0007067 (.0003892)*	-.0003581 (.0007181)	-.0004855 (.0008489)	-.0003422 (.0007117)	-.0003761 (.0005644)
Budget Airline				-1.015469 (1.756771)	-.9040006 (1.641604)
Baggage Fee		-8.00e-08 (5.60e-07)			
Cancelation Fee			1.46e-08 (2.23e-07)		
Maintenance / Aircraft		-.0000784 (.0001749)	-.0000348 (.0001956)	-.0000821 (.0001694)	
Fleet Size		.0037966 (.0012328)***	.003085 (.0014513)**	.0037115 (.0012205)***	.0037001 (.0010353)***
Pilot Salary		-3.46e-06 (3.21e-06)	-3.85e-06 (3.42e-06)	-3.75e-06 (3.15e-06)	-3.70e-06 (2.96e-06)
Number of Seats		.0020996 (.0116373)	-.0105001 (.0135941)	-.0006037 (.0114217)	
Number of Incidents		-.0015042 (.0092532)	-.0022328 (.0127798)	-.0025023 (.0092082)	
Total Revenue		-.0198427 (.0286438)	-.007008 (.0293961)	-.0216388 (.019868)	-.0233521 (.0160176)
Intercept	2.718387 (.7503188)***	1.644222 (1.989476)	3.955269 (2.439316)	2.48019 (2.208632)	2.162544 (1.235409)*
Chi Squared	3.30	18.97	15.87	19.70	19.52
Number of Observations	282	238	230	260	260

Notes: Standard Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level

Table 10: Fixed Effects Negative Binomial Regression Interpretations (Marginal Effects)
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)	(4)	(5)
Average Stage Length	-.0007067 (.00039)*	-.0003581 (.00072)	-.0004855 (.00085)	-.0003422 (.00071)	-.0003761 (.00056)
Budget Airline				-1.015469 (1.75677)	-.9040006 (1.6416)
Baggage Fee		-8.00e-08 (.00000)			
Cancelation Fee			1.46e-08 (.00000)		
Maintenance / Aircraft		-.0000784 (.00017)	-.0000348 (.0002)	-.0000821 (.00017)	
Fleet Size		.0037966 (.00123)***	.003085 (.00145)**	.0037115 (.00122)***	.0037001 (.00104)***
Pilot Salary		-3.46e-06 (.00000)	-3.85e-06 (.00000)	-3.75e-06 (.00000)	-3.70e-06 (.00000)
Number of Seats		.0020996 (.01164)	-.0105001 (.01359)	-.0006037 (.01142)	
Number of Incidents		-.0015042 (.00925)	-.0022328 (.01278)	-.0025023 (.00921)	
Total Revenue		-.0198427 (.02864)	-.007008 (.0294)	-.0216388 (.01987)	-.0233521 (.01602)

Notes: Standard Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level

Table 11: Re-defined Budget Variable Regressions
Dependent Variable: Number of Aircraft Accidents

Regressor	(1)	(2)	(3)	(4)	(5)
Average Stage Length	-.0011291 (.0003257)***	-.0011291 (.0003257)***	-.0010107 (.0003017)***	-.0003294 (.0007111)	-.0003649 (.000564)
Budget Airline	-.9677346 (.2219163)***	-.9677344 (.2219163)***	-.955858 (.2032952)***	-.8518152 (1.811383)	-.7464848 (1.708464)
Maintenance / Aircraft	-.0003749 (.0001323)***	-.0003749 (.0001323)***	-.0003526 (.0001056)***	-.0000837 (.0001694)	
Fleet Size	.0041583 (.0006001)***	.0041583 (.0006001)***	.0041529 (.0003086)***	.0037185 (.0012219)***	.0037109 (.0010365)***
Pilot Salary	-1.23e-06 (1.47e-06)	-1.23e-06 (1.47e-06)		-3.71e-06 (3.15e-06)	-3.68e-06 (2.96e-06)
Number of Seats	-.0054923 (.0040084)	-.0054923 (.0040084)	-.0057792 (.0036146)	-.0003993 (.0114145)	
Number of Incidents	-.0027056 (.0062859)	-.0027056 (.0062859)		-.0023595 (.0092196)	
Total Revenue	.0050343 (.0151806)	.0050343 (.0151806)		-.0217331 (.0198984)	-.023557 (.0160326)
Intercept	1.987853 (.7174252)***	1.987852 (.7174251)***	1.717778 (.6175659)***	2.392866 (2.194456)	2.113079 (1.220386)*
Type of Regression	Negative Binomial	Generalized Negative Binomial	Generalized Negative Binomial	Fixed Effects Negative Binomial	Fixed Effects Negative Binomial
Robust Standard Errors?	Yes	Yes	Yes	No	No
Pseudo R^2	0.1974	0.1974	0.1984		
Chi Squared	265.39	265.39	263.31	19.53	19.39
Number of Observations	260	260	264	260	260

Notes: Standard Errors are given in parenthesis. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level