The Cross Country Effect of Patent Applications on Ease of Doing Business

Georgia Institute of Technology

ECON 3161 Econometrics Fall 2022

Prof. Shatakshee Dhongde

Dylan Both, Ashwin Mahadevan, Sreya Yadlapalli

Abstract

Innovation and technological progress have long played a critical role in the economic growth of countries. While this relationship has been studied many times, it is not clear exactly how patents play into this equation. We believe that patents have a direct relationship with innovation, and thus, business growth and progress. Therefore, we hypothesized that countries with a greater number of patent applications will have better ease of business scores as measured by the World Bank. To analyze this relationship, we have used World Bank data and created economical and statistical models to understand this crucial relationship to get a better sense of the role patents play in a country's economy. Contrary to our original hypothesis, no evidence could be found to support a statistically significant effect of the number of patent applications in a country's economy on that country's ease of business score.

I. Introduction

The ability and desire to conduct business under certain conditions is a driving factor of economic growth. In addition to reducing transaction costs and incentivizing innovation, good business practices can attract foreign investors. Therefore, a country's quality of business practices, which can be inferred from how easily business is conducted in that country, directly contribute to the amount of business conducted in that economy. While ease of doing business is not the only determining factor of the amount of business conducted in an economy, it is certainly a foundational component. Logically extending from this assumption is the fact that greater ease in conducting business in a certain country leads to higher chances of economic stability in that country, which is important for many reasons including its implications on that economy's financial system. Since sound financial systems are far more likely to exist in a stable economy than an unstable economy, the importance of having a stable economy is highlighted by that economy's ability to sustain a solid financial system. Achieving, and more importantly, maintaining a stable economy is crucial to any country that prioritizes transparency and reinforcement in its financial system.

In an international system, any rational actors who prioritize a stable, growing economy are incentivized to bolster their ease of doing business. A primary consideration in the conditions for ease of doing business is intellectual property rights. It is commonly understood that without the recognition and protection of innovation, incentives to create and develop new technology are mitigated. Trademarks, trade secrets, and patents are a few of the several ways to recognize innovation by formalizing intellectual property rights. Out of these, the most popular and controversial method is patents. Not only do patents require public disclosure of how exactly an innovation functions, but they also expire after a definite date. Although standard operating procedures differ from country to country, expiration and public disclosure of how the innovation functions are universal basic functionalities of patents. While their potential to encourage innovation can be clearly seen, it is possible that there is more to the story. The question of whether or not patents promote intellectual property in a way that is truly beneficial to long term economic stability must be considered. Given that the owner of a patent can exclude anyone else from freely producing that same good or service, the chance of excess market power must also be acknowledged. Considering these competing factors at play, it is possible that the relationship between patent applications and EOB score may not be as significant as initially hypothesized.

If patents do promote sustainable economic growth, analyzing the factors that surround the institutional design of a patent system can allow a country to stimulate economic growth by optimizing their patent system. Understanding the number of patents applied for on an annual basis provides insight into the

institutional design of the intellectual property rights system within a country. After all, actors will only apply for patents in any given country if they know there is a chance it will be granted and that it will be beneficial to them. Alternatively, countries with fewer patent applications can be seen as having weaker institutional patent systems. Therefore, the number of patent applications in an economy is a direct reflection of the institutional design of that country's patent system. It follows that analyzing the effect of the number of patents applied for on the ease of doing business can provide a better understanding of how the materialization of intellectual property rights affects economic strength. In this study, we will be measuring the effect the number of patent applications has on the ease of doing business. In other words, our primary independent variable will be the number of patent applications for a given year and the dependent variable will be the ease of doing business for that year, on a scale of 1-100 (1= least business friendly, 100 = most business friendly). Our hypothesis is that an increase in the number of patents applied for in a given year will improve the ease of business rankings, signaling a better ability to conduct business and a higher ease of business score.

II. Literature Review

Recognizing the important role that technological advances play in economic growth and long-run productivity improvements, Josheski and Cane (2011) studied the dynamic link between patent growth and GDP growth in G7 economies (Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States). They used quarterly data of patent and GDP growth from 1963 to 1993 to determine if a long run relationship exists between the number of patents in an economy and that economy's growth. Although an initial scatterplot of quarterly GDP growth and patent growth revealed no clear correlation between the two variables, they utilized an autoregressive distributed lag model (ARDL) to take a deeper look. Their model assumed a form similar to a simple regression model but with adjustments made to account for the time lags between patents being granted and products making it through production, contractual obligations, and other factors such as similar products or patents. Their model concluded that there exists a positive relationship between patent growth and GDP growth in the long run. Furthermore, by accounting for error corrections and performing causality tests among other statistical safeguards, Josheski and Cane ensured that this relationship is in fact causal.

Chen (2008) uses cross-country panels to determine if patent laws have a significant positive effect on invention rates in the United States and fourteen Western European countries. The usefulness of patent laws have been debated by economists since their invention. Although few would deny that patent laws incentivize innovation, the fact that they provide a temporary monopoly for the patentee is of concern to some economists. With this debate in mind, Chen aims to provide some new literature to the mostly

empty bookshelf of the effect of patent laws on innovation. For his analysis, Chen pulls data from Clarence Strait's 1957 book, *Freedom against itself,* where Strait provides a compilation of over one-thousand major scientific discoveries and inventions from around the world between 1750 and 1950. Furthermore, Chen supplements this invention data with data from *Funk & Wagnalls New Encyclopedia* (2007), which has a much smaller but more stringent list of one-hundred-fifteen inventions that are used primarily to test robustness. Chen gets his patent law data from a few different sources for each of the fifteen countries, and he gets his GDP per capita and population data from Maddison (2001). Before regressing the data, Chen used a Quandt Likelihood Ratio test to determine that innovation is dependent on patent laws but patent laws are not dependent on innovation. Chen also makes a couple of important clarifications before discussing the results of his model. First, Chen discusses the fact that certain invention related variables such as education and copyright laws are omitted from his model, but since the fifteen countries selected for the study are fairly homogeneous with respect to these laws, these omitted variables should not affect the results of the study. Second, because sample variances in his data sets are significantly greater than their sample mean, Chen uses a negative binomial regression before using a likelihood test to determine if the model can be reduced to a simple Poisson regression. The results of his likelihood ratio test rejected Poisson regressions in favor of negative binomial regression for both the encyclopedia and Strait data. Overall, Chen found a significant positive effect of patent laws on invention rates after controlling for the sizes of each country's economy. While this study focuses on the effect of patents on invention rates, it is well established in economic theory that innovation is a strong driver of long-run economic growth; therefore, this study informs our analysis of the effect of patents on economic strength via proxy with invention rates serving as an indicator of economic strength.

The relationship between innovation and societal progress, and more specifically, the role that patents and patent rights play in driving research investments and innovation, was a strong motivation behind Williams (2011) research. Patents are a strong tool in economic growth since, by giving owners small pieces of market power, they incentivize innovating new technologies that promote societal growth. It has been found in the past that patent rights tend to have a positive relationship with innovation. Williams (2011) attempted to build upon this relationship to see if there was also a positive relationship between increasing patent rights and research innovation. While factors, such as the difficulty to determine social costs and benefits of the patent system, made it difficult for Williams (2011) to conclude whether or not stronger patent rights would be better for technological innovation and progress, it is clear that there is a correlation between the two. He advised that further research needs to be conducted to determine the best strategy for modifying the patent system to better suit the needs of developers, and encourage increased innovation.

Hall and Harhoff (2012) examine economic growth, policy issues, and business and software methods to analyze the effects of patents on economies. Through this approach, the authors discuss the tradeoffs between acquiring patents and securing returns to innovation. The findings indicate that patents do not necessarily incentivize innovations and, at times, serve as a deterrent of innovation. Through the latter, new developments are either kept as trade secrets or are not even introduced because of standard operating procedures and lack of internal support. It is important to note that certain industries develop immensely through patents. Specifically, patents can be seen as a requirement for advancement in the pharmaceutical industry due to high R&D costs which lead to innovations that may not last too long. However, industries such as this are not extremely prevalent and do not dictate general market trends.

Furthermore, the necessity of patents to further industry growth and attract investment depends on the current market atmosphere. Although they are not always a necessity for innovation incentives, a company whose competitors have patents needs patents as well. Essentially, patents can be a necessity for certain firms, although this will be a zero sum game for the comprehensive economy. The allocation of benefits from patents must also be taken into account. Hall and Haroff argue that although they do not directly add value to social welfare, they can generate large returns for a handful of private investors. Because of this, many investors and economic actors have an inherent interest in acquiring patents, which inadvertently increases the demand for research done by specific entities.

When comparing the number of patents and the ease of doing business in different countries, the institutional design of the patent system itself must also be considered. The extent to which patents impact social welfare depends on the institutional system in which the patent is applied. International comparison of patents and economic factors, such as the ease of doing business, could help bridge gaps when measuring how useful patents are between different institutional patent systems.

We attempt to take a slightly different direction in our research on patents. While the relationship between patents and innovation may be slightly muddled, we wish to dissect the relationship between patents and ease of business in countries. The ease or difficulty of doing business is sure to have an impact on innovation as well as the technological and societal progress of a country, so this relationship can be crucial in discovering what impact patents truly have on an economy. Although ease of doing business is likely strongly correlated to the dependent variables in the literature discussed above, to our knowledge, there is no current literature directly examining how patents affect ease of doing business. By examining this crucial economic indicator, we hope this paper helps complete the picture of the effect patents have on an economy as a whole.

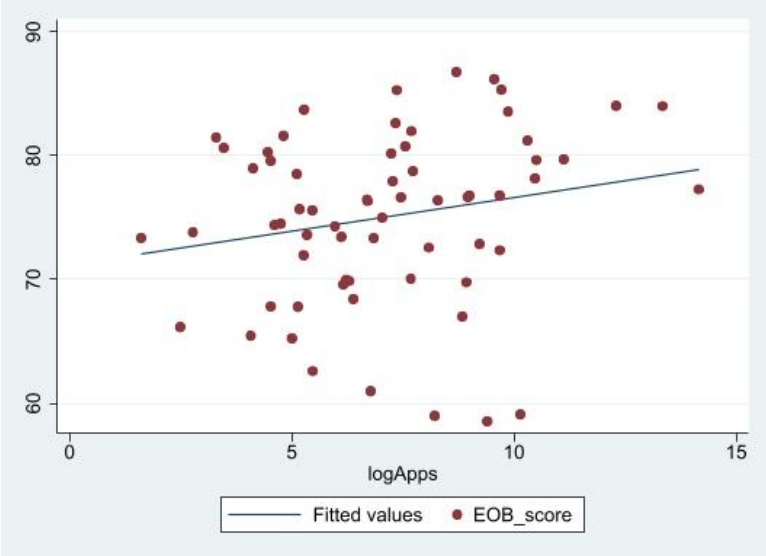
III. Data

To analyze the effect of patent applications on the ease of doing business, the independent variable must be the number of patent applications. Our data shows that China had the maximum value for the total number of patent applications with over 1.4 million applications. With approximately 600,000 annual patent applications, the United States is a distant second place to China. The dependent variable is the ease of doing business, which is denoted as the EOB score. The EOB score quantifies regulatory best practices and is calculated by taking an unweighted average of ten separate scores of proximity to the best possible performance, according to the World Bank. The ten indicators are the ease of starting a business, getting electricity, registering property, getting credit or loans, dealing with construction permits, protecting minority investors, paying taxes, trading across borders, enforcing contracts, and resolving insolvency. As mentioned previously, the EOB score can range from 0 to 100. In this case, 0 represents the worst regulatory performance and 100 reflects the best possible regulatory practice in all ten areas. Three different control variables were chosen for this study: GDP per capita in US dollars, gross domestic expenditures on research and development (R&D) as a percent of total GDP, and gross tertiary school enrollment ratio per country. The motivation behind choosing GDP per capita as a control variable involves the possibility that it should be easier to do business where the average person is more likely to have spending money. It is extremely important to note how intertwined this control variable can be in relation to the dependent variable (EOB score). Since Josheski and Cane (2011) found a significant long-term relationship between patent applications and GDP growth, we controlled for GDP per capita. The second control variable accounted for was R&D expenditure as a percent of total GDP. As Hall and Harhoff (2012) note, the institutional system design of a state's innovation environment plays a key role in deciding whether patents are pursued or not. This is susceptible to cause the EOB score to be skewed, when considering higher R&D can lead to better business or a corporate environment where it can be easier to do business. Analyzing this as a percentage of GDP levels the playing field and accounts for countries who have relatively high or low GDPs. Lastly, we included gross tertiary school enrollment as a control variable. This is the ratio of total enrollment in tertiary education to the age group's population that officially corresponds to tertiary education. In this case, tertiary education is synonymous with college education, so the variable indicates the ratio of total college enrollment to the population of the 18-22 age group. The tertiary school enrollment ratio helps analyze the proportion of citizens that pursue higher education. Our data reflected three countries having a tertiary school enrollment ratio over 100%: Turkey, Australia, and Greece. This means that their total college enrolment was higher than the national 18-22 age group population. Therefore, we conclude that these countries have a large influx of foreign students. Education is another factor that determines the institutional design of a country's innovation

environment. Post-secondary education often involves innovation and leads to later innovation through the development of necessary skills and knowledge.

All five variables, the number of patent applications, EOB score, GDP per capita in US dollars, gross domestic expenditures on research and development (R&D) as a percent of total GDP, and gross tertiary school enrollment ratio per country, were sourced from the World Bank Group which used data from the UNESCO Institute for Statistics (UIS) dataset. All data sets were aggregated through a separate weighted average. The data from resident patent applications and non-resident patent applications were combined to create the patent application variable, since these were the only two datasets involving patent applications and they were accounted for separately by the World Bank. The base year of 2019 was used to mitigate extreme impacts from the COVID-19 pandemic. To create a clean dataset, a few adjustments had to be made. In order to ensure no blank entries, countries without data for any variable were deleted. For the final datasets, the logs of GDP per capita and patent applications were used to ensure that each percent change in EOB score was accounted for proportionally rather than through raw inputs. Using the log for both of these variables also resulted in a higher R^2 value. Hence the final independent and dependent variables were logApps and EOBScore, respectively. The control variables were logGDP_perCap, RDGDP, and EnrollmentTertiaryGross.

Figure 1: EOB_score vs. logApps



The scatterplot above shows the relationship between our independent variable, logApps, and our dependent variable EOB_score. The figure shows that there is not much of a relationship between logApps and EOB_score, as the data is scattered around and there is no obvious linear relationship. While

this shows that our hypothesis might not be true, we must look at all of the variables, including the controls, to accurately determine the relationship between logApps and EOB_score. The descriptions and summary statistics for the variables are also shown below.

Table 1: Variable Descriptions

Variable	Description	Year	Units	Source
TotalApps	Total of number of patent applications	2019	Number	World Bank
EOB_score	Ease of business score	2019	Number	World Bank
RDGDP	Expenditures on research and development, as a percentage of total GDP	2019	Percentage	World Bank
GDP_perCap	GDP per capita	2019	Current US Dollars	World Bank
EnrollmentTertiaryGross	Gross tertiary school enrollment ratio per country	2019	Percentage	World Bank

Table 2: Summary Statistics

Variable	Num Observations	Mean	Standard Deviation	Min Value	Max Value
TotalApps	65	40202.7	189636.2	5	1400661
EOB_score	65	75.0	6.9	58.5	86.8
RDGDP	65	1.3	1.1	0.03	5.1
GDP_perCap	65	27429.8	24873.5	1374.0	113218.7
EnrollmentTertiaryGross	65	66.4	25.5	12.2	148.5

The data must be checked to ensure it meets the CLM/Gauss Markov Assumptions. These checks verify that the data can be used for the models and that it does not have any strong biases.

1. **Linear in Parameters**: The multiple regression models use the following equation. $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta\ x\ + u$. Since our β coefficients are constant and they are the parameters, this assumption is satisfied.
2. **Random Sampling in Data Selection**: The data was all collected from the World Bank Center. Other than omitting data with missing variables, no other changes were made. Data from the World Bank Center is drawn from populations from around the world, so the model meets this condition.
3. **No Perfect Collinearity in Variables**: The collinearity table below shows that there is no perfect collinearity among the variables used in the model. This means that this assumption is also met. This assumption is important because when it is violated, it is impossible to hold the perfectly collinear variable constant. Therefore, it is impossible to determine the effect of just one variable.

Table 3: Collinearity

	logApps	**EOB_score**	**RDGDP**	**logGDP_perCap**	**EnrollmentTertiaryGross**
logApps	1				
EOB_score	0.21	1			
RDGDP	0.49	0.49	1		
logGDP_perCap	0.25	0.61	0.69	1	
EnrollmentTertiaryGross	0.21	0.42	0.39	0.49	1

4. **Zero Conditional Mean**: The error term is included in all of the regression equations used on the model to account for the factors that affect EOB_score that we did not explicitly include in our model. However, since we do not know what these variables are, we cannot ensure that our explanatory variables do not contain information about their mean. We use multiple regression models to make this assumption more likely to be met, but for the reason above, we cannot assure that it is met.
5. **Homoscedasticity**: This assumption means that the variance in u, the error term, must be constant. Looking at Figure 1, the data points are not uniformly scattered and there is variance in the distance between the points and the line of best fit. However, this variance is clearly larger when logApps is between 5 and 10 than it is for higher or lower values of logApps, so the variance is not constant. Therefore, our model does not satisfy the homoskedasticity assumption.

6. **Normality**: Normality assumes that the unobserved error is independent of the explanatory variables and is normally distributed with zero mean. Therefore, in order for normality to be satisfied, assumptions 4 and 5 must be satisfied as well. If normality holds true in a model, that model's OLS estimators can be shown to be the minimum variance unbiased estimators. Since we cannot be confident that assumptions 4 and 5 are met, it is unlikely that normality is met. Therefore, our model's estimators likely are not the minimum variance estimators.

IV. Results

Our first model is a simple regression between the EOB_Score and logApps variables. This model predicts the linear effect that logApps has on EOB_Score. This represents the change in a country's ease of business score that results from a one percent change in total patent applications in that country.

Model 1: $EOB\_score = \beta_0 + \beta_1 logApps + u$

Running this regression in Stata gives us the values $\beta_0 = 71.14$ and $\beta_1 = 0.55$, giving us equation 1.

Equation 1: $EOB\_score = 71.14 + 0.55 logApps$

According to this model, there is a positive correlation between logApps and EOB_score. Since this is a level-log regression model, the interpretation is that a 1 percent increase in patent applications increases a country's ease of business score by 0.55. With a two-tailed P-value of 0.102, this coefficient is statistically significant at a 10.2 percent level, meaning 89.8 percent of the time, a change in logApps will have a non-zero effect on EOB_score when not controlling for other variables. 95 percent of the time, the value of this coefficient will be between -0.11 and 1.21. Since this 95% confidence interval contains 0, this result does not align with our hypothesis that an increase in patent applications will lead to a higher ease of business score, at the 5 percent significance level. This model has an R^2 value of 0.042, indicating that only about 4 percent of the variance in EOB_score can be explained by logApps.

Taking into account our control variables will give us a more realistic view of the impact that patent applications have on ease of business scores. Therefore, our second model is a multiple regression model that also considers the effects that a country's research and development spending, GDP per capita, and college enrollment can have on its ease of business score.

Model 2: $EOB\_score = \beta_0 + \beta_1 logApps + \beta_2 RDGDP + \beta_3 logGDP\_perCap + \beta_4 Enrollmenttertiarygross + u$

The regression output yields β_0, β_1, β_2, β_3, *and* β_4 values of 43.12, 0.008, 0.74, 2.90, and 0.04 respectively, giving us equation 2.

Equation 2: *EOB_score = 43.12 + 0.008logApps + 0.74RDGDP + 2.90logGDP_perCap + 0.04Enrollmenttertiarygross*

Table 4:

Independent Variable	Coefficient (standard error)
logApps	0.01 (0.31)
RDGDP	0.74 (0.94)
logGDP_perCap	2.90 (0.94)
EnrollmentTertiaryGross	0.04 (0.03)

Like our simple regression model, the multiple regression model shows a positive correlation between patent applications and a country's ease of doing business. However, the correlation is much weaker in the multiple regression model. This is no surprise since the mean independence assumption was unlikely to be satisfied in the simple regression model, so the unobserved error term likely skewed the results of our first model more than it did in the second. According to Model 2, a 1 percent increase in patent applications in a country increases that country's ease of business score by only 0.008. This means that, holding other factors constant, it would take a 125 percent increase in patent applications to increase a country's ease of business score by 1 percent. More important than the decrease in the magnitude of logApps' coefficient is the increase in the two-tailed P-value of logApps' coefficient. With a P-value of 0.979, a change in logApps only has a non-zero effect on EOB-score about 2 percent of the time. The 95 percent confidence interval shows that the value of logApps' coefficient will be between -0.61 and 0.63 95 percent of the time. Increasing the percentage of GDP spent on R&D by 1 percent leads to an increase of 0.74 in their ease of business score. If GDP per capita increases by 1 percent in a country, their ease of business will increase by 2.9. Finally, a 1 percent increase in tertiary enrollment leads to an increase of 0.04 in ease of business scores. Notice that we do not take the natural log of tertiary enrollment because EnrollmenttertiaryGross measures a percentage of those enrolled in tertiary school compared to the relevant population, so taking the log would make the interpretation how 1 percent of a 1 percent change in tertiary enrollment affects EOB_score. This is the same reason why RDGDP is not a log variable. RDGDP, logGDP_perCap, and Enrollmenttertiarygross have two-tailed P-values of 0.433, 0.003, and 0.200 respectively, indicating that logGDP_perCap is the only statistically significant variable in this

model. 99.7 percent of the time, a change in logGDP_perCap will have a non-zero effect on EOB_score. This effect is likely always positive since the 95 percent confidence interval is from 1.02 to 4.78. This model has an R^2 value of 0.402, meaning that these four variables account for about 40 percent of the variance in EOB_score. This makes it a much more accurate model than model 1.

Since logGDP_perCap was the only statistically significant variable in model 2, we decided to take a closer look at its effect on EOB_score.

Model 3: $EOB\_score = \beta_0 + \beta_1 logApps + \beta_2 logGDP\_perCap + u$

The regression output yields β_0, β_1, and β_2 values of 36.80, 0.16, and 3.81, giving us equation 3.

Equation 3: $EOB\_score = 36.80 + 0.16 logApps + 3.81 logGDP\_perCap$

Table 5:

Independent Variable	**Coefficient (standard error)**
logApps	0.16 (0.28)
logGDP_perCap	3.81 (0.66)

According to this model, a one percent change in the number of patent applications in a country leads to a change of 0.16 in EOB_score. However, the P-value for this coefficient is 0.575, indicating that the effect of logApps on EOB_score is only non-zero 42.5 percent of the time. The 95 percent confidence interval for this coefficient is (-0.397, 0.708). Like in model 2, the coefficient for logGDP_perCap, which states that a 1 percent increase in a country's GDP per capita leads to a 3.8 point increase in their EOB_score, is statistically significant at the 1 percent level with a p-value of 0.000. We can be 95 percent sure that the value of this coefficient is between 2.497 and 5.133.

Clearly the effect of logApps on EOB_score is very small in magnitude in all of our models. Furthermore, a null hypothesis that the effect of logApps on EOB_score equals 0 cannot be rejected at a 1, 5, or 10 percent level in any of the models. As seen below in table 6, which contains all our models, logGDP_perCap is the only of our variables that is statistically significant in any of our models.

Table 6:

Independent Variable	Model 1 Coefficient	Model 2 Coefficient	Model 3 Coefficient
logApps	0.55 (0.33)	0.008 (0.31)	0.16 (0.28)
RDGDP	N/A	0.74 (0.94)	N/A
logGDP_perCap	N/A	2.90*** (0.94)	3.81*** (0.66)
EnrollmentTertiaryGross	N/A	0.04 (0.03)	N/A
Intercept	71.14	43.12	36.80
No. of Observations	65	65	65
R^2	0.042	0.40	0.38

Significant at 10% 5%** 1%****

V. Extensions

After creating and analyzing our models, it was clear that the only significant variable was logGDP_perCap. The other two control variables we picked, RDGDP and EnrollmentTertiaryGross were not significant. While they were not independently significant, an F-test can be used to determine if RDGDP and EnrollmentTertiaryGross are jointly significant. We conducted an F-test using Model 2 as our unrestricted model and Model 3 as our restricted model.

Model 2: $EOB\_score = \beta_0 + \beta_1 logApps + \beta_2 RDGDP + \beta_3 logGDP\_perCap + \beta_4 Enrollmenttertiarygross + u$

Model 3: $EOB\_score = \beta_0 + \beta_1 logApps + \beta_2 logGDP\_perCap + u$

$H_0 : \beta_2 = 0, \beta_3 = 0$, $H_1 : H_0$ is False

The number of observations used for both models were 65, with the unrestricted model having four variables, and the restricted model having 2. Using the R^2 values from Model 2 and Model 3, we can calculate the F statistic as the following:

$F = [(R_{UR}^2 - R_R^2)/q] / [(1 - R_{UR}^2)/(n - k - 1)] = [(.4020 - .3780) / 2] / [(1 - .4020) / (65 - 4 - 1)] = 1.204$

Based on the number of observations and degrees of freedom, we will be using the critical value for $F_{2,60}$. At the 10, 5, and 1 percent significance levels, our critical values are 2.39, 3.15, and 4.98 respectively. Our test statistic, 1.204, is less than the critical values at all of the significance levels so we fail to reject

the null hypothesis. Therefore we can conclude that RDGDP and EnrollmentTertiaryGross are not jointly significant at the 10, 5, or 1 percent significance levels.

As our final test, we elected to see how EOB_score differs from high-income countries to low-income countries. According to the World Bank Data Team (2019), a country with a GNI (Gross National Income) per capita greater than $12,375 was considered a high income country in 2019. Using World Bank GNI data for our 65 countries, we created a dummy variable, HighIncome, assuming the value 1 if that country's GNI per capita exceeded $12,375 and 0 otherwise. From there, we ran model 4.

Model 4: $EOB\_score = \beta_0 + \delta_0 HighIncome + u$

The regression output yields β_0, *and* δ_0 *values of* 73.39 and 2.11 respectively, giving us equation 4.

Equation 4: $EOB\_score = 73.39 + 2.11 HighIncome$

According to model 4, without controlling for other variables, EOB_scores for high income countries were 2.11 points higher than EOB_scores for non-high income countries for any given number of patent applications in 2019. With a two-tailed p-value of 0.32, this coefficient is only significant at the 32 percent level. Including our variable of interest, logApps, and our statistically significant variable from models 2 and 3, logGDP_perCap, gives us model 5.

Model 5: $EOB\_score = \beta_0 + \delta_0 HighIncome + \beta_1 logApps + \beta_2 logGDP\_perCap + u$

The regression output yields β_0, δ_0, β_1, and β_2 *values of* 36.12, -1.49, 0.18, and 3.99 respectively, giving us equation 5.

Equation 5: $EOB\_score = 36.12 - 1.49 HighIncome + 0.18 logApps + 3.99 logGDP\_perCap$

Surprisingly, the sign of the coefficient for our dummy variable switched in model 5. This means that when controlling for logApps and logGDP_perCap, high income countries have EOB_scores 1.49 points lower than non-high income countries given any specific number of patent applications. With a p-value of 0.41, this result is not statistically significant at the 10, 5 or 1 percent level, so we cannot derive any strong conclusions about how EOB_scores vary from high income to non-high income countries based on our data. Like in models 2 and 3, logGDP_perCap remains significant at the 10, 5, and 1 percent levels with a STATA reported p-value of 0.00, but logApps is not significant at the 10, 5 or 1 percent level with a p-value of 0.53. Therefore, we can conclude that logGDP_perCap remains a significant indicator of EOB_score but logApps fails to be a significant indicator of EOB_score when only high income countries are studied.

VI. Conclusions

The hypothesis of a positive correlation between patent applications and a country's EOB score was ultimately not confirmed, as our models indicate. Naturally, the patent applications variable had the

highest coefficient with the simple regression in Model 1 at 0.55. Even considering this, we failed to reject the null hypothesis at the 10 percent significance level, since the two-tailed p-value was .102.

Following a similar trend, Models 2 and 3 did not prove there was a correlation between patent applications and EOB score. When including three control variables in Model 2, the logApps coefficient fell to 0.008. Furthermore, with a p-value of 0.97, logApps was not statistically significant at the 10, 5, or 1 percent significance levels. In Model 2, logGPD_PerCap emerged as the only statistically significant variable, having a p-value that proved statistical significance at the 10, 5, and 1 percent significance levels. Since the other variables were not statistically significant, they were dropped for our restricted Model 3. Therefore, to test the sole effect of GDP per capita with patent applications, Model 3 included only logApps and logGDP_perCapita. In terms of logApps' effect, the results were consistent with the rest of the models: logApps was not statistically significant at the 10, 5, or 1 percent significance levels. On the other hand, logGDP_perCapita was statistically significant at all three significance levels. When accounting for joint significance, an F-test was used to determine statistical significance between RDGDP and EnrollmentTertiaryGross variables. The F-value was lower than the critical value using 2 and 60 as the numerator and denominator degrees of freedom, respectively. Therefore, we concluded that the two variables are not jointly significant at the 10, 5, or 1 percent significance levels.

Future research examining the bridge between intellectual property and economic prosperity must take other factors into account. One way to extend this study could be to include trends over a period of time instead of using a cross-sectional dataset, which included only one given point in time. For example, analyzing how EOB scores have moved over a decade compared to the number of patent applications could lead to a better understanding of potential correlation or new effects such as having a trailing period or lead-lag effect. Furthermore, we recommend using additional control variables to test the strength of logGDP_perCapita, which dominated the statistical significance compared to the other variables in this study. Testing new control variables for statistical significance could lead to a better understanding of how patent applications can affect the ease of doing business and allow for less skews in the data.

Additionally, there is the possibility that patent applications do not truly reflect innovations levels. Since we found no statistically significant relationship between patent applications and ease of doing business, analyzing the direct relationship between innovation and ease of doing business might produce more meaningful results. We recommend that future researchers first test the relationship between patent applications and innovations levels and then try to find a variable that better captures the essence of innovation to produce a better model. Establishing a basic understanding of how innovation levels are influenced is pivotal to analyzing the effect on ease of doing business.

References

Josheski, D., Koteski, C. (2011) *The Causal Relationship between Patent Growth and Growth of GDP with Quarterly Data in the G7 Countries: Cointegration, ARDL and Error Correction Models.* Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1921908

Qiang Chen, (2008) *The effect of patent laws on invention rates: Evidence from cross-country panels* Retrieved from the Journal of Comparative Economics, Volume 36, Issue 4, 2008 https://reader.elsevier.com/reader/sd/pii/S0147596708000383?token=C60C92A791E2101397654 0AE7FE378E47054F65A2813ED76B746F1D66D40539A2C0D5681E74FA4D9DAC3041393C 831FD&originRegion=us-east-1&originCreation=20221005193503

Williams H. (2017) *How do patents affect research investments?* Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5664960/

Hall, B., Harhoff, D. (2012) *Recent Research on the Economics of Patents*, Retrieved from the National Bureau of Economic Research https://www.nber.org/system/files/working_papers/w17773/w17773.pdf

World Bank Data Team (2019) *New country classifications by income level: 2019-2020*, Retrieved from https://blogs.worldbank.org/opendata/new-country-classifications-income-level-2019-2020

Appendix

Countries:

Argentina	China	Greece	Latvia	Pakistan	Turkey
Armenia	Columbia	Guatemala	Lithuania	Poland	United Arab Emirates
Australia	Croatia	Hong Kong	Luxembourg	Portugal	United Kingdom
Austria	Cyprus	Hungary	Malta	Romania	United States
Azerbaijan	Czechia	Iceland	Mauritius	Russian Federation	Uzbekistan
Belarus	Denmark	Iran	Mexico	Serbia	Vietnam
Belgium	El Salvador	Ireland	Moldova	Singapore	Chile
Bosnia and Herzegovina	Estonia	Iceland	Mongolia	Slovak Republic	Georgia
Brazil	Finland	Israel	Montenegro	South Africa	Kyrgyz Republic
Bulgaria	France	Italy	New Zealand	Spain	Oman
Canada	Germany	Korea, Rep.	Norway	Sweden	Switzerland

Model 1 Stata output

```
. regress EOB_score logApps
```

Source	SS	df	MS	Number of obs	=	65
				F(1, 63)	=	2.76
Model	128.739426	1	128.739426	Prob > F	=	0.1016
Residual	2937.60898	63	46.628714	R-squared	=	0.0420
				Adj R-squared	=	0.0268
Total	3066.34841	64	47.9116939	Root MSE	=	6.8285

EOB_score	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
logApps	.5481457	.3298883	1.66	0.102	-.1110835	1.207375
_cons	71.13844	2.501459	28.44	0.000	66.13967	76.13721

Model 2 Stata output

```
. regress EOB_score logApps RDGDP logGDP_perCap Enrollmenttertiarygross
```

Source	SS	df	MS	Number of obs	=	65
				F(4, 60)	=	10.08
Model	1232.76567	4	308.191417	Prob > F	=	0.0000
Residual	1833.58274	60	30.5597123	R-squared	=	0.4020
				Adj R-squared	=	0.3622
Total	3066.34841	64	47.9116939	Root MSE	=	5.5281

EOB_score	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
logApps	.0081009	.3096803	0.03	0.979	-.6113519	.6275537
RDGDP	.7437539	.9426741	0.79	0.433	-1.141875	2.629383
logGDP_perCap	2.896881	.9407476	3.08	0.003	1.015106	4.778657
Enrollmenttertiarygross	.0404981	.0312756	1.29	0.200	-.0220624	.1030587
_cons	43.11764	8.451864	5.10	0.000	26.21139	60.02389

Model 3 Stata output

```
. regress EOB_score logApps logGDP_perCap
```

Source	SS	df	MS	Number of obs	=	65
				F(2, 62)	=	18.84
Model	1158.9478	2	579.473901	Prob > F	=	0.0000
Residual	1907.40061	62	30.7645259	R-squared	=	0.3780
				Adj R-squared	=	0.3579
Total	3066.34841	64	47.9116939	Root MSE	=	5.5466

EOB_score	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
logApps	.1559193	.2763967	0.56	0.575	-.3965897	.7084283
logGDP_perCap	3.814838	.6592327	5.79	0.000	2.497051	5.132626
_cons	36.79523	6.272944	5.87	0.000	24.2558	49.33467

Model 4 Stata output

```
. reg EOB_score HighIncome
```

Source	SS	df	MS	Number of obs	=	65
				F(1, 63)	=	1.02
Model	48.9008628	1	48.9008628	Prob > F	=	0.3162
Residual	3017.44754	63	47.8959928	R-squared	=	0.0159
				Adj R-squared	=	0.0003
Total	3066.34841	64	47.9116939	Root MSE	=	6.9207

EOB_score	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
HighIncome	2.109919	2.088128	1.01	0.316	-2.062872	6.282711
_cons	73.3939	1.849633	39.68	0.000	69.69771	77.0901

Model 5 Stata output

```
. reg EOB_score HighIncome logApps logGDP_perCap
```

Source	SS	df	MS	Number of obs	=	65
				F(3, 61)	=	12.73
Model	1180.42619	3	393.475396	Prob > F	=	0.0000
Residual	1885.92222	61	30.9167577	R-squared	=	0.3850
				Adj R-squared	=	0.3547
Total	3066.34841	64	47.9116939	Root MSE	=	5.5603

EOB_score	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
HighIncome	-1.486137	1.783015	-0.83	0.408	-5.051496	2.079222
logApps	.1778157	.2783223	0.64	0.525	-.378724	.7343555
logGDP_perCap	3.98744	.6925465	5.76	0.000	2.602607	5.372272
_cons	36.12459	6.339713	5.70	0.000	23.44755	48.80163