Volume 3 Issue 1

University of Toronto Economic Review

Spring 2017

University of Toronto Economic Review

Volume 3, Issue 1

Spring 2017

Economics Students' Association

University of Toronto Economic Review

The University of Toronto Economic Review gratefully acknowledges support from the Department of Economics and the Economics Students' Association at the University of Toronto. We are likewise indebted to the teaching staff who assigned and graded these written works.

The Department of Economics at the University of Toronto is located at 150 St. George Street, Toronto, Ontario M5S 3G7, Canada.

All correspondence with the Review can be sent to uter@esauoft.ca.

Contents

Message from the Economics Students' Association

Dear Readers,

The Economics Students' Association, like its name expresses, is a club by the students and for the students. We continually strive to enrich the undergraduate experience by fostering creative engagement.

This mission has, in the past, been realized only through academic events and seminars. However, it is through initiatives like this journal that we see the stellar work students are capable of beyond the realms of a lecture hall. We hope you enjoy this curated collection that showcases the work of our outstanding economics program, as well as others in North America.

Saarah Sheikh
President of the Economics Students' Association

Dear Readers,

In its first two years, working on the *University of Toronto Economic Review* often felt like working on an experiment. Every step of the editing and publication process was new, and we were unsure where our work would lead, or even if we would end up with anything of value. Would students submit their papers? Would this work even be worth publishing?

In its third year, however, the *Review* has largely transitioned from experiment to mainstay of the undergraduate Economics program. The editing and production are still at times unpredictable and challenging, but the fundamental questions of the *Journal* and its value are mostly answered.

This year, we have selected four papers from students at three universities in North America. We believe these papers represent the breadth and depth of topics that can be investigated using economic tools and theory, as well as the quality of work that can be produced by motivated undergraduate students. We start with Ujjwal Dahuja's work on the relation between foreign direct investment and economic growth. All seniors at Princeton University are required to complete a senior thesis or significant personal project. For the last two years – those years that we have accepted papers from outside of the University of Toronto – we have been fortunate to receive a number of senior theses from students at Princeton. Ujjwal's paper is an excellent example of the quality of research this program can produce. Next, Matthew Hong and Ayush Gupta analyze the factors influencing operating costs and ridership of public transit systems in the US. We then present Parker Griffin's work investigating the impact of the West African Economic and Monetary Union on its employment-population ratio, and finally En Hua Hu and Hon Chiu Yuan's paper on player choice in the popular online game *League of Legends*.

As far as we know, we are still the only undergraduate economics publication in Canada, and one of only a few in North America. This is a responsibility we do not take lightly, and also is something we could never fulfill on our own. This volume could not have been completed without generous help from the Economics Students' Association, the Arts and Science Students' Union, and several professors and instructors who helped us with specific questions and problems during the editing process.

As Quan Le noted in the inaugural volume of the *Journal*, economics grows by the sharing of ideas. We are honoured to be a part of this process, and are proud to present Volume 3 of the *University of Toronto Economic Review.*

Matthew Hardy
Editor-in-Chief

What is the relationship between foreign direct investment and economic growth in India and China?

Ujjwal Dahuja
Princeton University

1 Introduction

The relationship between foreign direct investment (FDI) and economic growth has been a subject of considerable research for the past few decades. In theory, FDI can have a substantial impact on economic growth – it can create jobs, increase production capacity, enable access to new technology, compensate the lack of domestic investment and create a multiplier effect on the overall economy. Additionally, FDI can also help augment the existing stock of knowledge and update the skill of the labor force in the recipient country. However, the causality between FDI and economic growth could also run the opposite way, i.e. rapid economic growth could also induce a greater inflow of FDI as often expressed by the "market-size hypothesis". As a country grows, its requirement of capital increases, thus raising the demand for capital from foreigners to close potential gaps. Lim (1983), for instance, argued that a higher economic growth rate, other things being equal, leads to a higher level of aggregate demand, leading to greater opportunities for making profits and, hence, increasing the incentive for foreigners to invest. In addition, rapid economic growth also creates opportunities for cost efficiencies of production and realization of economies of scale which makes growing markets very attractive to foreign investors (Iamsiraroj, 2015).

In this paper, I investigate the relationship between FDI and economic growth and assess whether there is a two-way causality between these two variables. If a two-way causality does exist, I also investigate whether the causality from FDI to growth is greater than that from growth to FDI. Specifically, I focus on India and China in this paper because both these countries have experienced high growth and high FDI flows. While China is the fastest growing nation in real terms among the G20 economies, India is second highest. Similarly, while China has historically attracted a huge amount of FDI, India has grabbed the top destination for FDI spot this year. On the other hand, despite being two similar emerging economies, India and China differ in some ways too. While FDI seems to be more restricted in India with quotas in certain sectors that prevent FDI more than a certain percentage, India hasn't used FDI as a macroeconomic tool in the way China many times has. In addition, China's recent focus on opening up its capital markets does not find similarities in the Indian context. Finally, the composition of both economies with China being more investment-driven and India more consumption-driven highlights an important difference between the two economies. Thus, the many similarities and yet some structural differences make India and China rather unique for the purpose of this study.

2 Literature Review

Previous literature focusing on the relationship between economic growth and foreign direct investment (FDI) has fallen largely into two main groups. The first group have implicitly assumed a one-way causality running from FDI to GDP growth and estimated the impact of FDI based on such causality. Broadman and Sun (1997), Zhang (1995), Chen (1996), Kaiser et al. (1996), Liu et al.(1997), Shan et al. (1999), Sun (1998), Sun and Chai (1998), Wang and Swain (1995, 1997) and Young and Lan (1997) belong to this first group who have used a single equation model to study the impact of FDI on growth in the Chinese case. However, as Khlody (1995) pointed out, using a single equation approach leads to simultaneity bias since growth can also influence the level of FDI. Additionally, these studies also used cross-sectional data when discussing the long-run impacts of FDI on the Chinese economy. Cross-sectional data does not allow for any causal interpretation in these studies that time series data would have otherwise enabled. As Shan (2002) pointed out, "it is not possible to infer anything, in cross-sectional context, more than a contemporaneous correlation."

These shortcomings led to the creation of a second group of empirical studies which improved upon the previous literature by not only considering two-equation models, but by also using the more advanced time series econometric techniques that developed at the time. Using Granger causality, innovation accounting, co-integration and vector error correction models, this second group showed that the causality between FDI and growth runs both ways. In the case of China, for instance, Shan, et al. (1999), Liu et al (2002) and Tang et al (2008) found that there exists a two-way causality between FDI and growth, suggesting that the previous single-equation model had some biases in estimations. In a landmark paper, Shan (2002) even investigated the strength of this causal relationship and concluded that not only was there a two-way causality between FDI and output growth, but that the strength of the causality was greater for the impact of growth on FDI. In the Indian context, Chakraborty and Basu (2002), Ray (2012) and Khan and Mitra (2014) also concluded similar results, while outside the Chinese and Indian context, Ludosean (2012), Ilgun, et al. (2010) and Frimpong, et. al (2006) also concluded with similar results in the context of Romania, Turkey and Ghana.

Yet, despite the advances made in the recent literature by this second group of empirical studies, a few important shortcomings still persist. First, in the Chinese case, most studies on this topic are rather dated and there are no studies to my knowledge that have captured the most recent data. Second, in the Chinese and Indian cases, most of the literature focusing on the relationship between economic growth and FDI has used annual data instead of quarterly data. For instance, Basu and Chakraborty (2002), Kalirajan et al. (2009), Ray (2012), Khan and Mitra (2014) have all used annual data which has led to fewer observations and thus less robust results. Additionally, quarterly data is also helpful in that it reduces the two-way causal effects. This is because firms generally cannot react as quickly with respect to their decisions on FDI based on quarterly figures of output, exports and domestic investment. FDI tends to be a long term investment and one would not expect it to be heavily influenced by quarterly macroeconomic figures. Recognizing these advantages of quarterly data, Shan (2002) did try to use quarterly estimates but he ended up using proxies for his variables that were not very accurate. For instance, he used industrial output as proxy for Chinese GDP and treated the sum of 'total capital construction investment' and 'total technical updating and transformation investment' as a measure of domestic investment. Both these proxies are inaccurate in that agricultural output also constitutes a significant portion of the Chinese economy and that capital construction investment does not account for overall fixed capital formation.

In context of these shortcomings, this paper strives to add to the existing literature in the following three ways in order of importance. First, it seeks to provide better estimates of the relevant variables with quarterly frequency which will allow for robust estimates. Second, it endeavors to update previous work done in the Chinese context by Shan (2002) whose dataset only examined the data up till 1998. Third, it attempts to compare the results found in the Chinese and Indian cases, which to my knowledge, is a first of its kind study.

3 Data

3.1 China

Data was collected on GDP, foreign direct investment (FDI), domestic investment (DI) and exports for China from Q4 1994 to Q2 2013. Since there was no common dataset with figures on all these four variables on a quarterly basis, a new dataset was created using inputs from different sources. These variables were chosen because they influence both GDP and FDI and have been widely used in empirical studies on this topic. The date range was determined by the availability of reliable estimates of quarterly FDI for China. However, it must be mentioned that the data only misses about 2 years of relevant information. This is because even though 1978 marks the year when China opened its doors to FDI, it was only after Deng Xiaoping's 'southern tour' in 1992 that led to a new phase of FDI for China. In addition, China's joining the WTO in and the consequent abolition of FDI project approval requirement, happened only by 2002. This further shows that despite being constrained by availability, the data captures the most relevant time periods in the case of China.

The description below details how the data was obtained and normalized:

- GDP: Quarterly data was obtained from the National Bureau of Statistics China website.The data was initially in current prices in RMB terms so it was first adjusted for inflation using CPI data (index 2010 = 100) obtained from Federal Reserve Economic Data (FRED). It was then converted to units of billions of dollars using IHS Global Insight data on USD-CNY market exchange rate period averages. A possible limitation of this procedure is that a GDP deflator offers a more precise way of obtaining real GDP than CPI because CPI only accounts for consumer prices and also considers imported goods. However, in the absence of quarterly data on GDP deflator, using CPI was the best possible option to account for inflation. Additionally the measure of GDP, despite the slightly erroneous use of CPI instead of the GDP deflator, is still a vastly improved estimate in comparison to the total industrial output estimate used by Shan (2002).

- Domestic investment: Quarterly data on gross fixed capital investment from DataStream was used as a proxy for domestic investment. The data was already in constant prices (2010) in units of billions of dollars. Hence, no changes were made to the data. A possible limitation of this proxy, however, is that it excludes inventory numbers. The absence of quality quarterly data on inventory measures caused this limitation.

- Exports: Monthly data was obtained from the National Bureau of Statistics China website. The data was first adjusted for inflation using CPI data (index 2010 = 100) obtained from Federal Reserve Economic Data (FRED). The monthly numbers were then added to obtain quarterly figures. Since the data was originally expressed in dollar terms, no further changes were made. Monthly data was only available from October 1994 to December 2014 which limited the start date of the main dataset to Q4 1994.

- FDI: Quarterly data on FDI was obtained from China's Ministry of Commerce website. The proxy used for FDI was "Total value of foreign direct investment actually utilized" which is a flow variable in constant prices (2010) and in US Dollar terms. Since the source figures were in year to date format, first quarterly figures were extracted from those numbers and then were expressed in units of billions of dollars. The proxy is a very good one since it measures how much FDI was actually utilized instead of overall FDI, some of which may not have been yet put to productive use.

While collecting data, emphasis was laid on both quality and quantity. All the data sources mentioned above are credible sources but a combination of them had to be used to maximize the number of data points of the variables being considered. This is why no single dataset was used in this paper.

3.2 India

Data was collected on GDP, foreign direct investment (FDI), domestic investment (DI) and exports for India from Q2 1997 to Q3 2014. Once again there was no common dataset with figures on all these four variables on a quarterly basis, so a new dataset was created using inputs from different sources. The date range was determined by the availability of reliable estimates of quarterly FDI for India. However, it must be mentioned that the data does not miss a lot of relevant information because even though the Indian economy set itself on a path of economic liberalization in 1991, it was only in 1997 that India allowed FDI in cash and carry wholesale.

The description below details how the data was obtained and normalized:

- GDP: Quarterly data was obtained from IHS Global Insight Database. The data was initially in current prices in INR terms so it was first adjusted for inflation using a GDP deflator (index 2010 = 100) obtained from the same data source. It was then converted to units of billions of dollars using IHS Global Insight data on USD-INR market exchange rate period averages.

- Domestic investment: Quarterly data on "Gross Fixed Capital Formation" was obtained from Federal Reserve Economic Data (FRED). The data was already in constant prices (2010) but it was in units of billions of INR. Hence, USD-INR market exchange period averages obtained from IHS Global Insight were used to convert the figures to billions of USD. A possible limitation of this proxy, however, is that it excludes inventory numbers. The absence of quality quarterly data on inventory measures caused this limitation.

- Exports: Quarterly data was obtained from Federal Reserve Economic

Data (FRED). The data was already in constant prices (2010) but it was in units of billions of INR. Hence, USD-INR market exchange period averages obtained from IHS Global Insight were used to convert the figures to billions of USD.

- FDI: Monthly data on FDI was obtained from Reserve Bank of India. The data was already in constant prices (2010) and dollar terms. So the only action taken was to convert the monthly data into quarterly data.

3.3 Summary Statistics

Given below are some summary statistics of the data. Note that the data has been transformed into percentage changes from previous quarter and table 1 shows the summary statistics of the transformed data.

Table 1: Detailed summary statistics expressed in percentage change over previous quarter

China (Percentage Change from Previous Quarter) – N=74											
Variables	Mean	Std. Dev.	Min	Max	p1	p5	p25	p50	p75	p95	p99
GDP	4.6%	12.7%	-22.0%	28.9%	-22.0%	-20.1%	-8.2%	10.0%	13.4%	17.3%	28.9%
FDI	7.8%	37.5%	-51.1%	112.7%	-51.1%	-43.3%	-18.6%	-2.4%	39.0%	76.3%	112.7%
Exports	4.1%	15.2%	-34.5%	27.3%	-34.5%	-25.7%	-9.8%	8.1%	16.5%	23.6%	27.3%
Dom. Inv.	2.8%	0.8%	1.5%	5.7%	1.5%	1.6%	2.3%	2.6%	3.0%	4.9%	5.7%
India (Percentage Change from Previous Quarter) – N=69											
GDP	1.5%	79.0%	-14.4%	23.7%	-14.4%	-10.%	-5.7%	-0.3%	6.3%	20.2%	23.7%
FDI	48.3%	166.7%	-294.5%	736.0%	-294.5%	-83.6%	-38.4%	-8.2%	67.6%	463.7%	736.0%
Exports	3.8%	7.8%	-19.6%	25.7%	-19.6%	-11.1%	1.3%	3.5%	8.8%	13.4%	25.7%
Dom. Inv.	2.8%	5.8%	-19.4%	17.7%	-19.4%	-7.1%	-0.7%	3.0%	6.3%	11.4%	17.7%

On average all variables have been increasing over time for both countries which is what one would expect from India and China. Additionally, on average China's GDP and exports have been growing at a faster rate in every quarter than India's which is also what one would expect given China's consistently high growth rate and focus on exports. However, China's output and exports per quarter are highly fluctuating as seen in the large standard deviations of both variables. An estimate that particularly leaps off the page from these summary statistics is that of the estimated standard deviation of FDI flows for India which is about 166.7%. While at first this seems quite large, this too is expected given that India has overtaken China in terms of FDI even though the FDI liberalization process started much later in India as compared to China. It must be noted, however, that a few results from summary statistics still do seem to defy intuition. For instance, the average quarterly growth for China at 4.6% seems to be rather high; the minimum and maximum quarterly growth for China at -22.0% and 28.9% also seem to be rather large. However, it can be argued that these statistics appear larger because of changes in CPI over time. As noted before, in the absence of a quarterly GDP deflator, CPI had to be used to convert nominal GDP to real GDP and the seemingly large statistics are only a consequence of the fact that they have been displayed in 2010 current prices terms using CPI to deflate the nominal GDP figures. Regardless, our analysis will be later repeated for the middle 90% of the sample to check for robustness to outliers. Finally, it is of note that the small sample size for both India and China is rather small which can make inference on the estimated coefficients more difficult.

4 Methodology

In this paper a Vector Auto Regressive (VAR) model will be used to determine the relationship between economic growth and FDI. This is because a VAR model can accurately capture the two way causal effects between FDI and growth and it has also been effectively employed in a number of previous studies on this topic. According to Enders (1995), VARs are useful tools in examining the relationships among economic variables in a dynamic context. Furthermore, VAR models have also been proven to generate very reliable estimates in endogenous contexts. This is because they require less a priori information and have an advantage of treating each variable under the study as an endogenous variable (Gujarati, 1995). This is crucial because macroeconomists over the years have proposed contrasting theories and as such it would not be reasonable to assume the validity of one theory over the other in understanding the relationship among different variables. Particularly in the growth-FDI debate, as mentioned earlier, it would not be reasonable to assume one theory as valid and to only make model predictions on the basis of that theory. VARs in this context thus stand out because they use very little prior information and allow great flexibility in constructing models without using much a-priori economic reasoning. Due to these advantages, this study will use a VAR model to capture the linear interdependencies among GDP, FDI, exports and domestic investment. The equation below showcases the VAR model to be used in this paper:

$$y_t = c + A_1 y_{t-1} + A_2 yt - 2 + ... + A_p y_{t-p} + \varepsilon_t \tag{1}$$

where y_t is a 4 x 1 vector (GDP_t, FDI_t,$EXPORTS_t$, DI_t), ε_t is a 4 x 1 vector of error terms for the variables included, c is a 4 x 1 vector of constants, A_1 through A_p are 4 x4 matrices of coefficients and p represents the number of lags.

4.1 Lag Selection

There are a number of different approaches that are used in economic literature to select the appropriate the number of lags (p). The F-statistic approach, for example, starts with a model with many lags and involves performing hypothesis tests on the final lag. However, the drawback of this model is that it will produce too large a model, at least some of the time. Even if the true coefficient of the final lag is 0, a 5% test using the t-statistic will incorrectly reject the null hypothesis. Thus it is not an appropriate method of model selection for the purpose of this study. Other approaches include minimizing information criteria. Generally, the Bayes or Schwarz Information Criterion (known as BIC) is preferred because it is a consistent estimator of p. However, for the purposes of this paper, the Akaike Information Criterion (AIC) will be used since consistency is not the main concern given the small sample and that AIC has been shown to be better suited as a lag selection method for VARs using small sample quarterly data. In addition, studies of the ADF statistic suggest that it is better to have too many lags than too few, so it is recommended to use AIC instead of the BIC to estimate p for the ADF statistic.[1] AIC and BIC are computed using the following formulas:

$$BIC(p) = ln(\frac{SSR(p)}{T}) + (p-1)\frac{ln(T)}{T} \tag{2}$$

[1] See Stock (1994) and Haldrup and Jannson (2006) for reviews of simulation studies of the small-sample properties of the Dickey-Fuller and other unit root statistics.

$$AIC(p) = ln(\frac{SSR(p)}{T}) + (p-1)\frac{2}{T} \tag{3}$$

where T is no. of observations, p is no. of lags.

Note that a corrected version of AIC, known as AICc will also considered later to test for robustness of lag length selection. AICc corrects the AIC for small sample sizes and can be computed in the following way:

$$AICc(p) = AIC(p) + \frac{2p(p+1)}{T-p-1} \tag{4}$$

4.2 Stationarity of Data

Once the number of lags is determined, the data will be tested for stationarity by performing a generalized least squares modified Dickey Fuller (DF-GLS) test also known as the Elliot-Rothenberg-Stock unit root test. This test is preferred to other unit root tests used in the literature because Davidson and MacKinnon (2004) have shown that the PhillipsPerron test performs worse in small samples than the Augmented Dickey Fuller (ADF) test and the ADF test itself has been shown to be less powerful than the DF-GLS test for small samples (Stock and Watson, 2002). The DF-GLS test locally detrends data series to efficiently estimate the deterministic parameters of the series, and uses the transformed data to perform a usual ADF unit root test. This procedure helps to remove the means and linear trends for series that are not far from the non-stationary region. The augmented DickeyFuller test involves fitting a regression of the form

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \gamma_3 Y_{t-3} + \gamma_4 Y_{t-4} + u_t \tag{5}$$

and then testing the null hypothesis $H_0 : \delta = 0$ against the one-sided alternative $H_1 : \delta < 0$.The DF-GLS test is performed analogously but on GLS-detrended data. The null hypothesis of the test is that Y_t is a random walk, possibly with a drift. In theory, there are two possible alternative hypotheses: Y_t is stationary about a linear time trend or Y_t is stationary with a possibly nonzero mean but with no linear time trend. This paper uses the former option.

If a regressor has a stochastic trend (has a unit root), then there could be three main problems due to non-stationarity: (1) The estimator of the autoregressive coefficient is biased toward 0 if its true value is 1; (2) The t-statistic on a regressor with a stochastic trend can have a nonnormal distribution, even in large samples; and (3) there could be a case of a spurious regression. (Stock and Watson, 2002). In order to avoid these, it is necessary to work with data that is stationary. The variables being studied in this paper exhibit gradual growth over time, that is over the long run, the series tends to grow by a certain percentage per year on average as in Figure A1. Thus, in levels, the data is not expected to pass this test. In order to then obtain stationarity, transformations to percentage changes will be used.[2] This will be done iteratively until stationarity is achieved. Later, the robustness of our results to the choice of the unit root test will be evaluated to gauge whether our choice significantly affects our results.

4.3 Granger Causality

The next step after achieving stationarity would be to estimate the VAR model and test for Granger causality. The Granger causality statistic is the F-statistic testing the hypothesis that the coefficients on all the values of one of the variables are zero. This null hypothesis implies that regressors have no predictive content beyond that is contained in the other regressors. Thus Granger-causality only measures predictive power and not necessarily causality as the name might suggest. Nevertheless, results from this test in the paper are reported even though they are not particularly meaningful in context of our research question.

4.4 Impulse Response functions (IRFs)

It is rather difficult to infer linkages between variables merely from the coefficients of the VAR model. This is because effects of each variable need to be compounded over time to see how variables are really interdependent on the current and past values of other variables. Impulse response functions offer a convenient graphical and/or tabular outlet that allows for inference. Impulse response functions can be used to trace the time path of variables over time to a 1 standard deviation shock in the error term of each estimated

[2] This is equivalent to a scaled version of differencing which is a commonly used transformation to obtain stationarity.

equation in the VAR. These shocks thus allow us to relatively compare how responses to exogenous shocks differ for each variable in our VAR.

4.5 Forecast Error Variance Decompositions (FEVDs)

Forecast error variance decomposition techniques help break down the variance of the forecast error of each variable into components that can be attributed to each of the endogenous variables in the VAR system. Note, however that FEVD's do not provide any causal information. They are merely being used as an additional tool to gauge the dynamic relationships of the variables in the VAR. FEVD's can be interpreted as partial R squares that reveal information about goodness of fit rather than causality.

5 Results

Running the DF-GLS test and using AIC for lag selection revealed that in the case of China, the first differenced series percentage change in variables quarter over quarter – with four lags, was the most appropriate model. While FDI and exports were found to be stationary at the 5% significance level, real GDP and domestic investment were found to be stationary at the 10% significance level. In the case of India, the first differenced series with 5 lags was found to be the most appropriate model at the 5% significance level for all the variables. In this study, therefore, the first differences of variables for both countries are used to allow for easier interpretation of results. Broadening the significance level to 10% instead of 5% does not degrade our estimation much while providing us a convenient way to compare the results obtained for both countries. Tables A2 and A3 in the appendix show the VAR estimates obtained using the first differences of variables for both India and China. The tables also show results of the Wald test used to determine Granger-causality. Table 2 below provides a summary of the results:

In table 2, it can be seen that both GDP and FDI Granger-cause each other in the case of India while for China the reverse is true. Both these results are in contradiction with the existing literature. While Chakraborty and Basu (2002) argue that GDP in the Indian case does not Granger-cause FDI, Shan, Tian, and Sun (1999) argue that the Granger-causality in the

Table 2: Granger-Causality Results for India and China

Causality at 5% level	China	India
GDP→ FDI	No	Yes
GDP→ Exports	Yes	Yes
GDP→ Domestic Investment	Yes	Yes
FDI→ GDP	No	Yes
FDI→ Exports	No	No
FDI→ Domestic Investment	No	Yes
Exports→ GDP	Yes	No
Exports→ FDI	Yes	No
Exports→ Domestic Investment	Yes	No
Domestic Investment→ GDP	No	No
Domestic Investment→ FDI	No	No
Domestic Investment→ Exports	No	Yes

case of China holds for both GDP and FDI. This contradiction could possibly be because this study uses data from a different time period than both Chakraborty and Basu (2002) and Shan, Tian, and Sun (1999). Additionally, Chakraborty and Basu (2002) also use annual data instead of quarterly data which can result in differences between the two studies. Another interesting result from table 2 is that in the case of India, while foreign direct investment seems to Granger-cause domestic investment, the reverse is not true. In the case of China, neither foreign direct investment Granger-causes domestic investment nor domestic investment Granger-causes FDI. It is of note, however, that Granger-causality only measures predictive power, so not much can be said about causality from this measure.

In order to make statements about causality, we turn to impulse response function analysis. Using the ordering GDP, FDI, exports and domestic investment' in the Cholesky factorization, impulse response functions have been estimated to show the effect of shocks in GDP and FDI. Note that the IRFs have been scaled to show the effect of a 1% increase in GDP growth when the impulse is GDP and a 10% increase in FDI growth when the impulse is FDI. This has been done to provide a clearer interpretation of the shocks since a shock of one standard deviation is not as easily interpretable especially since the standard deviations of each of the variables differs for both countries. In addition, we shock using a 1% increase in GDP and a 10% increase in FDI because the standard deviation of FDI is much larger than that of the standard deviation of GDP and as such a 10% increase in FDI is perfectly plausible given the data. Figure 3.1 below shows the impact of a 1% increase in GDP on FDI for China. Figure 3.2 shows the impact of a 10% increase in FDI on GDP for China. Similarly, Figures 3.3 and 3.4 showcase the relevant results for India.

Figure 3.1

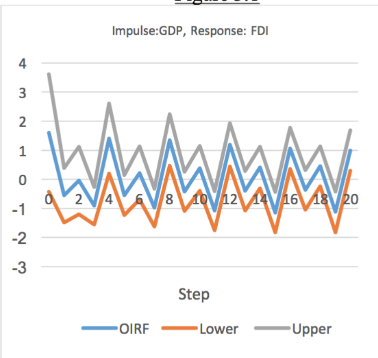


Figure 3.2

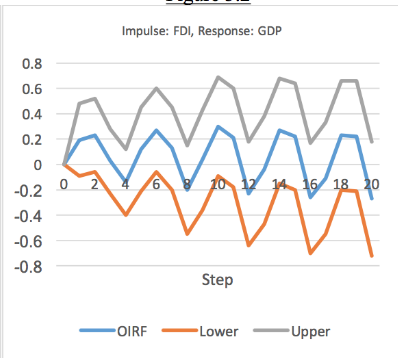


Figure 3.3

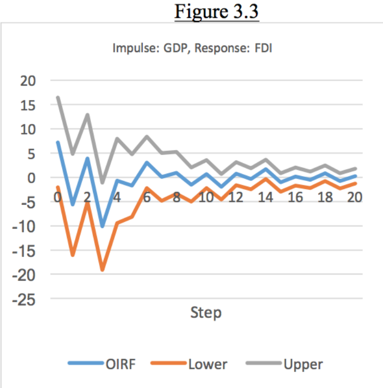
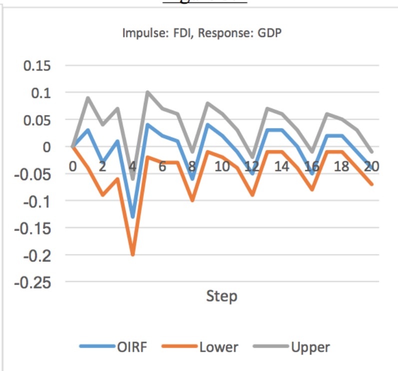


Figure 3.4



It can be seen in the figures above that a 1% increase in GDP has a larger short term effect on FDI in the case of India as compared to China. However, we also notice that the impact seems to be more persistent in the case of China even after 20 periods (or 5 years), while the impact seems to be decreasing with time in the Indian scenario. In addition, we also notice that in the case of both India and China, a 10% increase in FDI seems to have a negligible effect on GDP in both the short and the long term. The magnitude of the response is indeed very small. Thus our results seem to indicate that the causality from GDP to FDI is much stronger than the causality from FDI to GDP for both countries. Additionally, the impact of GDP shocks seems to be much larger in magnitude for India than for China. Note that this only means that the absolute magnitude of the response is larger in the case of India; the sign of the response seems to be highly fluctuating for both countries.

A possible reason as to why the causality from GDP to FDI is stronger for both India and China is that the market size of both India and China has grown substantially in the period being considered in this study. A growing market size creates a large requirement of capital thus increasing the demand for foreigners to close potential gaps. Additionally, high levels of aggregate demand in both economies have created large opportunities for making profits thus increasing the incentive for foreigners to invest. Furthermore, a possible explanation for why the impact of GDP on FDI has been greater for India than for China could be due to India's democratic political system. GDP growth rates equal, foreign investors may find investing in a democratic country preferable (Busse, 2003). Additionally, a possible explanation could be the difficulty in repatriation of profits from China. For the greatest part of the sample period considered in this study, it was much more difficult to remit profits earned in China to other countries as compared to India. All else equal, investors would prefer to add capital from where they can expect to easily remit profits.

Analyzing other IRFs shown in figures A5.1 and A5.2 for China and A8.1 and A8.2 for India reveal some other interesting patterns. For instance, a 1% increase in GDP seems to have persistent effects on exports in the case of China while the effects seem to decay over time in the case of India. In addition, the effects of a 1% increase in GDP on domestic investment are almost negligible in the case of China while for India the effect is more volatile in the immediate short term. Finally, a 10% increase in FDI seems to be have rather small effects on exports and domestic investment for both

India and China.

Moving away from causality, we analyze some results of the forecast error variance decomposition (FEVD). We find that a significant fraction of the total forecast error variability in FDI and GDP can be attributed to shocks in FDI and GDP themselves respectively for both India and China, as expected. However, this fraction reduces gradually over the longer term and we notice, for instance, that shocks in other variables begin to each contribute about 10% of the total forecast error variability in GDP at the end of 5 years (20 quarters) for China. We also notice that over time the fraction of variability in exports for both India and China is increasingly explained by shocks in GDP. For instance, at the end of 5 years, a shock in GDP in the case of China seems to explain 31% of the forecast error variability and in the case of India 22% of the forecast error variability. Finally, we also notice that in the case of India, a shock in GDP seems to explain 44% of the forecast error variability in domestic investment at the end of 5 years (20 quarters). Note once again that these numbers can only be interpreted as partial which is a measure of predictive strength and not causality.

5.1 Robustness of Results

5.1.1 Robustness to Lag Length Selection

In our estimation so far, we have been using 4 lags in the case of China and 5 lags in the case of India because AIC had estimated that those to be the optimal number of lags given our data. However, BIC suggests that we use only 1 lag in the case of China but still 5 lags in the case of India. Additionally, AICc suggests that we use 5 lags for China and India. Therefore, to test the robustness of our results, we re-run our analysis with these new lag lengths. Notice that AIC, AICc and BIC all estimate lag length to be 5 for India. The differences are only in the case of China. When we re-run our methodology with 5 lags for China, our results, as shown in Figure A11, are very similar to the ones we had obtained with 4 lags. This is expected since we have only changed the lag length by 1 and that the additional lag is using information from 5 quarters prior. When the analysis is re-run with 1 lag for China as shown in A12, the impact of a 1% increase in GDP or a 10% increase in FDI are now no longer persistent. Additionally, the impact of FDI on GDP, though still small in magnitude, is larger compared to our main results. Overall, however, our central message that the causality from GDP to FDI is much stronger than the causality from FDI to GDP still remains the same. Thus our results are robust to various lag selection criteria.

5.1.2 Robustness to Outliers

It was noted earlier that some of the summary statistics for both India and China seemed incongruous with our intuition and that the analysis would be redone without these outliers. Using the middle 90% of the data for both China and India, we see in figures A13 and A14 that there are slight differences in the output when we exclude outliers.[3] For instance, in the case of China we find a comparatively larger impact of FDI on GDP than before and we also find less persistence in the responses of both FDI and GDP. In the case of India, we find greater short run fluctuation in impact of GDP on FDI while other results remain similar. Overall, however, our central message that the causality from GDP to FDI is much stronger than the causality from FDI to GDP still remains the same. Thus our results are robust to large outliers.

5.1.3 Robustness to Choice of Unit Root Test

Our results are highly sensitive to the choice of the unit root test in that stationarity for China was not achieved in the first differenced series using the Augmented Dickey Fuller Test or the Phillips-Perron test. In both those cases, second differencing was required which would not have allowed for comparisons with results pertinent to India since the data for India passed all three stationarity tests in the first differenced form.

5.2 Comparison with Other Results from Existing Literature

Compared to others who have studied this topic using annual data in the context of China, our results show a more dominant effect of GDP on FDI in

[3] Sorted by values of GDP for China and FDI for India, since summary statistics seemed rather large for these two variables.

comparison to the effect of FDI on GDP. This difference can arise for multiple reasons different time periods of study, different VAR specifications and use of quarterly data.

Compared to others who have studied this topic using quarterly data such as Shan (2002), our results show a more dominant one-way causality rather than a two- way one. Once again this difference could arise from the factors mentioned above. For instance, Shan (2002) uses a 9 variable VAR with data spanning only about 50 quarters while this study uses a 4 variable VAR with data spanning 75 quarters in the case of China. However, an additional plausible explanation could be that Shan (2002) uses exchange rates in his VAR model which has also been used in other studies but not considered here. This key difference could cause results of this study to vary substantially from that of Shan (2002).

5.3 Limitations and Further Steps

5.3.1 Presence of Unknown Breaks in Data

As shown in figure A10, a number of breaks were identified in the middle 70% of the data for India using the QLR test on the four different variables used in the study. The presence of structural breaks indicates that alternative methods of modeling might be more suited to this dataset since OLS regression estimates a relationship that holds "on average". Note that in figure A10 only data pertinent to India has been shown. Similar breaks were found in the case of China.

5.3.2 Small Sample Size

This is by far the most relevant concern with respect to the conclusions made in this paper. In a VAR setting in particular, small sample size can hamper inference by introducing bias. However, it must be noted that the sample size included in this paper is still larger than in the literature reviewed on this topic. Shan (2002), for instance, used a sample size spanning less than 50 quarters. In comparison, this study uses data spanning 75 quarters.

5.4 Measurement Error

Measurement error could arise in results from two different sources: (i) Domestic Investment figures for China and India do not include inventory which underestimates the measure of domestic investment (ii) The data used in this paper comes from multiple different sources and is thus subject to measurement error because different sources may have slight differences in their data collection techniques. It must be noted, however, that the above mentioned limitations were in some way unavoidable because of the constraints imposed in getting a large enough sample with quarterly data.

5.4.1 Omitted Variables

Selection of variables in this study such as exports and domestic investment was based merely on economic intuition. However, a statistically rigorous way of choosing which variables to include would be to use an FAVAR (Factor Augmented VAR). The FAVAR enables us to use the most important linear combinations of different series for estimation which can be quite useful in forecasting. However, there are 2 possible drawbacks to this approach: (1) Forecasting is not the main aim of this paper so to go through the exercise of finding data on multiple macroeconomic indicators and running an FAVAR on them is outside the scope of this paper (2) Since FAVARs use a linear combination of different series, they do not have a natural economic interpretation rendering their use to be less practical for this study.

5.4.2 Co-integration Among the Different Macroeconomic Variables

The "trace statistic" of the Johansen co-integration test revealed that in the cases of both India and China there were 2 co-integrating series among the four variables. The Johansen test can be seen as a multivariate generalization of the augmented Dickey-Fuller test described earlier in this paper. The generalization is the examination of linear combinations of variables for unit roots. Thus, even though individually the time series used in this study are by themselves stationary, there are two series such that linear combinations of them may contain unit roots. This implies that the standard asymptotic distributions cannot be applied for inference in this study and that a vector error correction model may have been a more suitable specification. This also explains why a large majority of literature focusing on the relationship between foreign direct investment and economic growth uses vector error correction models to better understand the interactions between the two series.

6 Conclusion

Using a VAR model, this study finds that the causal effect of GDP on FDI is much more dominant than the causal effect of FDI on GDP in the context of India and China. In particular, the causality of GDP on FDI seems to be more pronounced in the data for India than in the data for China. These results are robust to lag length selection criteria and large outliers and contradict much of the recent literature written in the context of China which has either concluded a two-way causality or no obvious causality between the two variables. These results, however, are consistent with the market size hypothesis often associated with emerging countries such as India and China. As both these economies have grown at a fast pace in the past three decades, they have created a huge demand for capital that has been fulfilled by foreign investors who have found great profit making opportunities in these two economies.

7 Appendix

Figure A1: Top left panel: GDP over time, Top right panel: FDI over time, Bottom left panel: Exports over time, Bottom right panel: Domestic investment over time (in millions USD), China

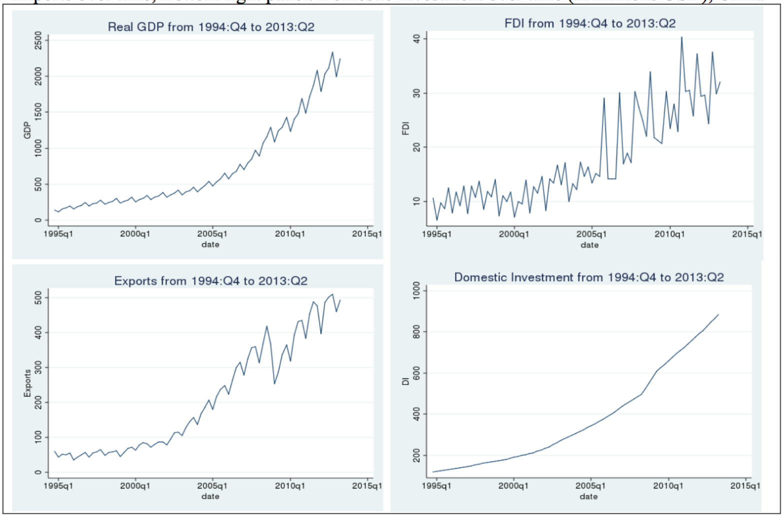


Table A2: VAR estimation for China

VARIABLES	(1) GDP	(2) FDI	(3) Exports	(4) Dom. Inv.
GDP_{t-1}	-0.0566	-0.234	-0.760***	0.00816
	(0.0638)	(0.554)	(0.178)	(0.0182)
GDP_{t-2}	0.0186	0.843	-0.644***	-0.00581
	(0.0669)	(0.581)	(0.187)	(0.0191)
GDP_{t-3}	-0.128**	-0.793	-0.282*	0.0281*
	(0.0598)	(0.519)	(0.167)	(0.0171)
GDP_{t-4}	0.766***	1.209**	0.319*	0.0327*
	(0.0616)	(0.535)	(0.172)	(0.0176)
FDI_{t-1}	0.0247*	-0.219*	0.0524	0.00242
	(0.0137)	(0.119)	(0.0384)	(0.00393)
FDI_{t-2}	0.0300**	-0.508***	0.0633*	0.00328
	(0.0137)	(0.119)	(0.0385)	(0.00393)
FDI_{t-3}	0.0260*	-0.168	0.118***	0.00407
	(0.0142)	(0.123)	(0.0397)	(0.00406)
FDI_{t-4}	0.0128	0.236*	0.0453	0.00115
	(0.0139)	(0.121)	(0.0390)	(0.00398)
$Exports_{t-1}$	0.0969**	0.137	-0.000903	0.0113
	(0.0428)	(0.372)	(0.120)	(0.0122)
$Exports_{t-2}$	0.0191	0.199	0.0847	0.0234**
	(0.0396)	(0.344)	(0.111)	(0.0113)
$Exports_{t-3}$	0.0473	0.444	-0.0155	0.000985
	(0.0374)	(0.325)	(0.105)	(0.0107)
$Exports_{t-4}$	0.157***	0.0992	0.228**	-0.00319
	(0.0350)	(0.304)	(0.0978)	(0.01000)
Dom. Inv_{t-1}	-0.963**	-1.366	-2.382**	0.734***
	(0.411)	(3.572)	(1.151)	(0.118)
Dom. Inv_{t-2}	-0.281	-2.832	0.0812	0.0401
	(0.532)	(4.619)	(1.488)	(0.152)
Dom. Inv_{t-3}	0.278	-0.269	0.484	0.101
	(0.532)	(4.618)	(1.487)	(0.152)
Dom. Inv_{t-4}	0.718*	4.244	2.299**	-0.238**
	(0.413)	(3.588)	(1.156)	(0.118)
Constant	-2.40e-05	0.0493	0.0567*	0.00467
	(0.0120)	(0.104)	(0.0336)	(0.00344)
Observations	70	70	70	70
Excluded variables	Granger Causality F statistics (Prob >F)			
GDP		0.056*	0.000***	0.169
FDI	0.126		0.038**	0.827
Exports	0.000***	0.608		0.295
DI	0.001***	0.515	0.005**	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3: VAR estimation for India

VARIABLES	(1) GDP	(2) FDI	(3) Exports	(4) Dom. Inv.
GDP_{t-1}	-0.280	-25.01***	0.330	-0.0654
	(0.179)	(6.218)	(0.302)	(0.259)
GDP_{t-2}	-0.0588	-12.67**	0.767**	0.458*
	(0.179)	(6.208)	(0.302)	(0.258)
GDP_{t-3}	-0.0916	-23.38***	0.683**	0.319
	(0.181)	(6.294)	(0.306)	(0.262)
GDP_{t-4}	0.768***	-13.63**	0.613**	0.277
	(0.177)	(6.148)	(0.299)	(0.256)
GDP_{t-5}	0.183	4.004	0.349	0.425
	(0.196)	(6.817)	(0.332)	(0.284)
FDI_{t-1}	0.00353	-0.292**	0.00816	0.000811
	(0.00327)	(0.114)	(0.00552)	(0.00472)
FDI_{t-2}	0.00170	-0.199*	0.000835	-0.000320
	(0.00331)	(0.115)	(0.00560)	(0.00479)
FDI_{t-3}	0.00259	0.151	0.00154	-0.00440
	(0.00301)	(0.104)	(0.00508)	(0.00434)
FDI_{t-4}	-0.00996***	0.0653	-0.00421	-0.0104**
	(0.00318)	(0.110)	(0.00536)	(0.00459)
FDI_{t-5}	-0.00304	0.366***	-0.00217	-0.00843*
	(0.00326)	(0.113)	(0.00551)	(0.00471)
$Exports_{t-1}$	-0.204**	5.553**	-0.555***	-0.107
	(0.0797)	(2.769)	(0.135)	(0.115)
$Exports_{t-2}$	-0.180**	2.193	-0.564***	-0.0792
	(0.0918)	(3.188)	(0.155)	(0.133)
$Exports_{t-3}$	-0.175**	3.365	-0.574***	-0.368***
	(0.0861)	(2.990)	(0.145)	(0.124)
$Exports_{t-4}$	-0.198***	3.477	-0.322**	0.0185
	(0.0764)	(2.654)	(0.129)	(0.110)
$Exports_{t-5}$	-0.0503	-0.349	-0.367**	-0.199
	(0.0849)	(2.947)	(0.143)	(0.123)
Dom. Inv_{t-1}	0.383***	19.12***	0.404**	0.318*
	(0.120)	(4.156)	(0.202)	(0.173)
Dom. Inv_{t-2}	0.272**	5.611	0.113	-0.152
	(0.123)	(4.256)	(0.207)	(0.177)
Dom. Inv_{t-3}	0.0843	1.668	0.352	0.0659
	(0.134)	(4.665)	(0.227)	(0.194)
Dom. Inv_{t-4}	-0.0391	-0.231	0.131	-0.199
	(0.123)	(4.270)	(0.208)	(0.178)
Dom. Inv_{t-5}	-0.0408	-1.068	0.346*	0.160
	(0.116)	(4.018)	(0.195)	(0.167)
Constant	0.0229***	0.163	0.0570***	0.0452***
	(0.00875)	(0.304)	(0.0148)	(0.0126)
Observations	64	64	64	64
Excluded variables	Granger Causality F statistics (Prob >F)			
GDP		0.000***	0.143	0.130
FDI	0.017**		0.710	0.190
Exports	0.036**	0.323		0.048**
DI	0.005***	0.000***	0.298	
All	0.000***	0.000***	0.000***	0.010**

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure A5.1: Orthogonalized Impulse Response to a 1% increase in GDP growth for China

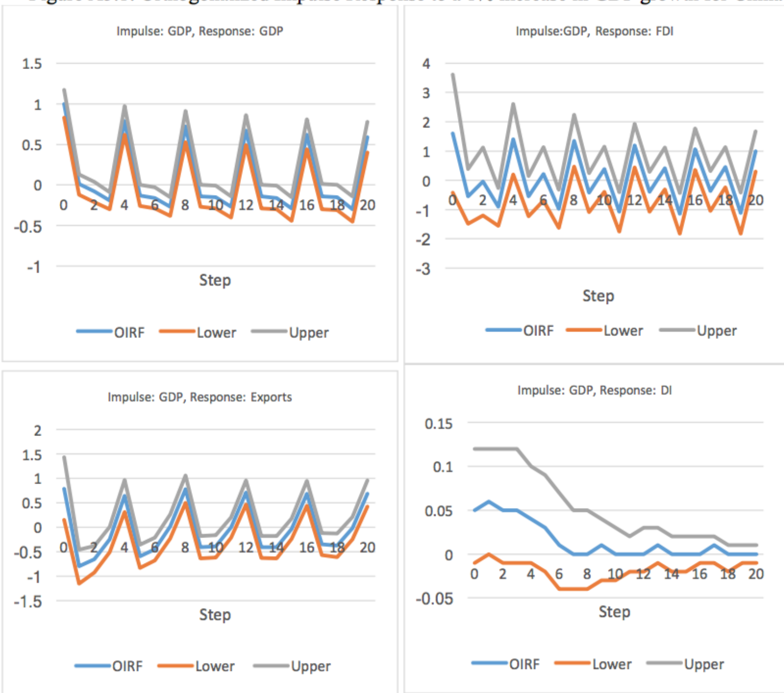


Figure A5.2: Orthogonalized Impulse Response to a 10% increase in FDI growth for China

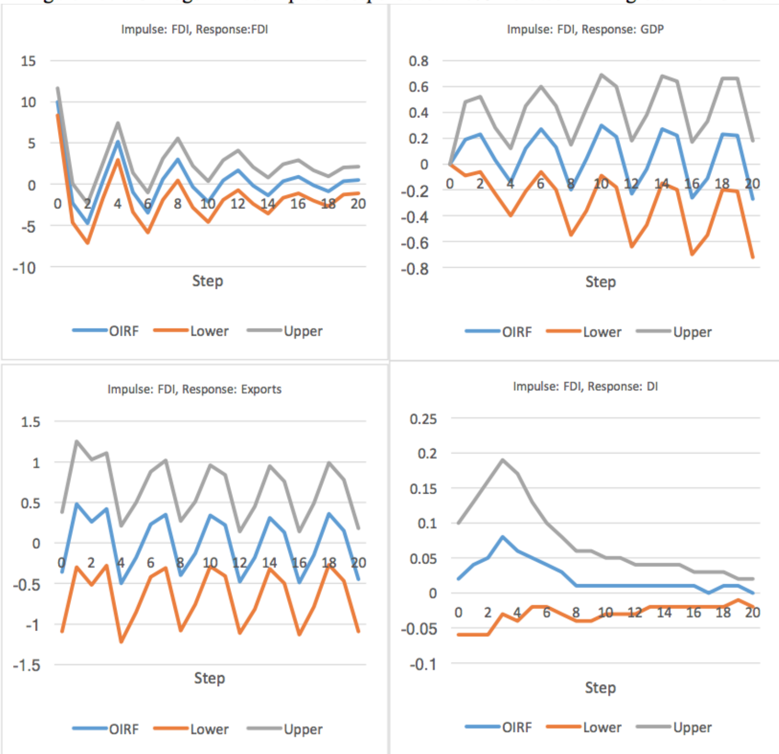


Figure A8.1: Orthogonalized Impulse Response to a 1% increase in GDP growth for India

Figure A8.2: Orthogonalized Impulse Response to a 10% increase in FDI growth for India

Table A10: QLR Test to determine breaks in data for India

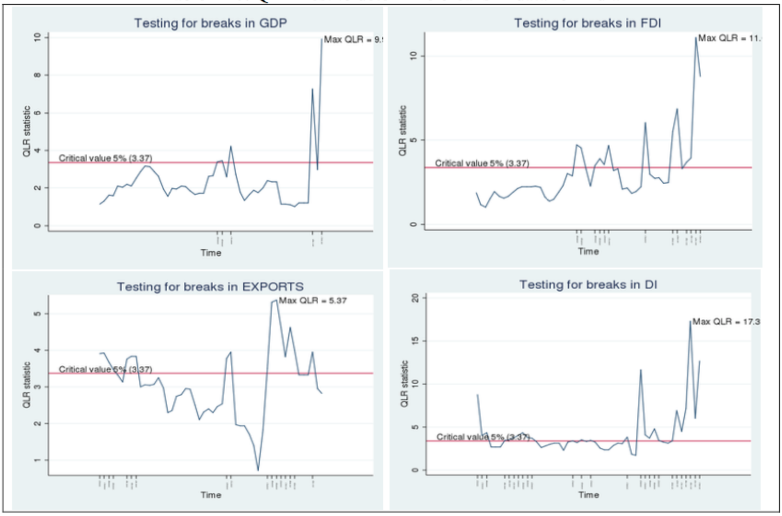


Figure A11: Estimation of IRFs using 5 lags in the first differenced series for China

Figure A12: Estimation of IRFs using 1 lag in the first differenced series for China

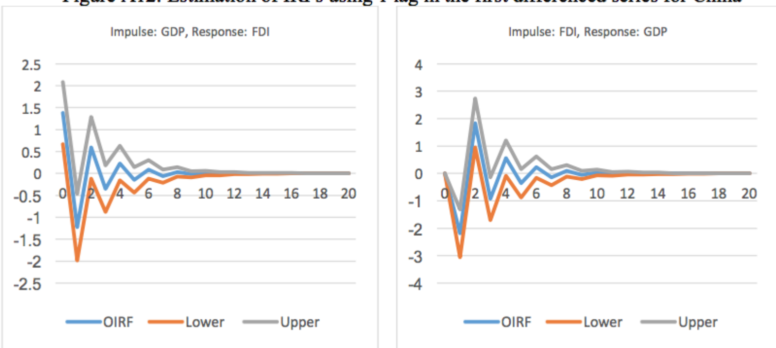


Figure A13: Estimation of IRFs without outliers for China

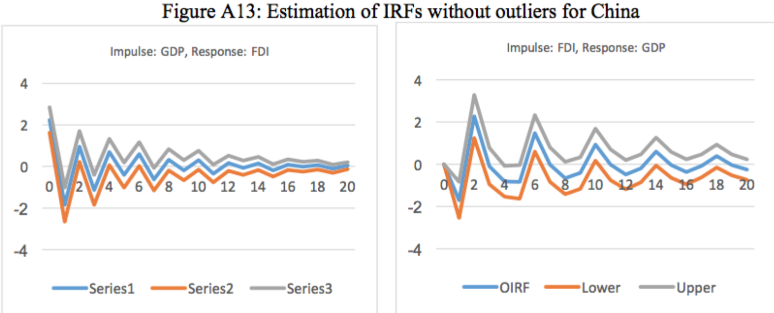
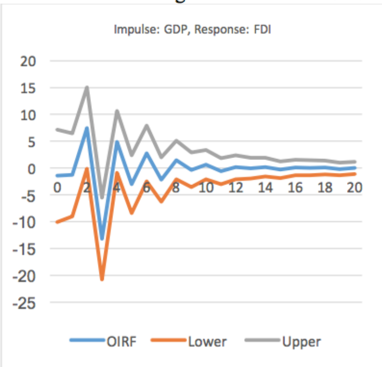
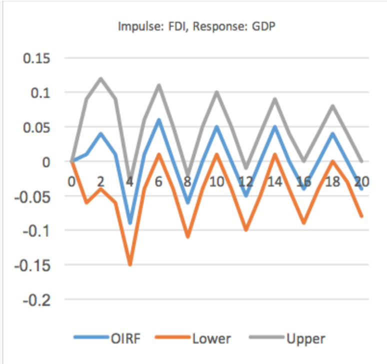


Figure A14: Estimation of IRFs without outliers for India





Broadman, Harry G. and Xiaolun Sun, "The distribution of foreign direct investment in China," *The World Economy*, 1997, *20* (3), 339–361.

Busse, Matthias, "Democray and FDI," 2003.

Chakraborty, Chandana and Parantap Basu, "Foreign direct investment and growth in India: a cointegration approach," *Applied Economics*, 2002, *34* (9), 1061–1073.

Chen, ChienHsun, "Regional determinants of foreign direct investment in mainland China," *Journal of Economic Studies*, 2002, *23* (2), 18–30.

Duttaray, Mousumi, Amitava K. Dutt, and Kajal Mukhopadhyay, "Foreign direct investment and economic growth in less developed countries: An empirical study of causality and mechanisms," *Applied Economics*, 2008, *40* (15), 1927–1939.

Frimpong, Joseph Magnus and Eric Fosu Oteng-Abayie, "Bivariate causality analysis between FDI inflows and economic growth in Ghana," 2006.

Gujarati, Damodar N. and Dawn C. Porter, *Basic Econometrics*, New York: McGraw-Hill Education, 2008.

Iamsiraroj, Sasi and Hristos Doucouliagos, "Does Growth Attract FDI?," 2015.

Ilgun, Erkan, Karl-Josef Koch, and Mehmet Orhan, "How Do Foreign Direct Investment and Growth Interact in Turkey?," *Eurasian Journal of Business and Economics*, 2010, *3* (6), 41–55.

Kaiser, Stefan, David A. Kirby, and Ying Fan, "Foreign Direct Investment in China: An Examination of the Literature," *Asia Pacific Business Review*, 2006, *2* (3), 44–65.

Kalirajan, Kaliappa, Adil Miankhel, and Shandre Thangavelu, "Foreign Direct Investment, Exports, and Economic Growth in Selected Emerging Countries: Multivariate VAR Analysis," 2009.

Khan, Gholam Syedain and Papia Mitra, "A Causal Linkage between FDI Inflows with Select Macroeconomic Variables in India An Econometric Analysis," *Journal of Economics and Finance*, 2014, *5* (5), 124–133.

Kholdy, Shady, "Causality between Foreign Investment and Spillover Efficiency," *Applied Economics*, 1995, *27* (8), 745–749.

Lim, David, "Fiscal incentives and direct foreign investment in less developed countries," *The Journal of Development Studies*, 1983, *19* (2), 207–212.

Liu, Xiaming, Haiyan Song, Yingqi Wei, and Peter Romilly, "Country characteristics and foreign direct investment in China: A panel data analysis," *Review of World Economics*, 1997, *133* (2), 313–329.

Liu, Xiaohui, Peter Burridge, and P.J.N. Sinclair, "Relationships between economic growth, foreign direct investment and trade: evidence from China," *Applied Economics*, 2002, *34* (11), 1433–1440.

Ludoean, Bianca Maria, "A VAR analysis of the connection between FDI and economic growth in Romania," *Theoretical and Applied Economics*, 2012, *10* (575), 115–130.

Ray, Sarbapriya, "Impact of Foreign Direct Investment on Trade and Economic Growth of Pakistan: A Co-integration Analysis," *Advances in Information Technology and Management*, 2012, *2* (1), 187–201.

Shan, J., G. Tian, and F. Sun, "Causality between FDI and economic growth," in Wu Yanrui, ed., *Foreign Direct Investment and Economic Growth in China*, Cheltenham, UK: Edward Elgar, 1999, pp. 140–154.

Shan, Jordan, "A VAR approach to the economics of FDI in China," *Applied Economics*, 2002, *34* (7), 885–893.

Stock, James H. and Mark W. Watson, *Introduction to Econometrics*, Pearson, 2007.

Sun, Haishun, "Macroeconomic Impact of Direct Foreign Investment in China," *The World Economy*, 1998, *21* (5), 675–694.

_ **and Joseph Chai**, "Direct foreign investment and inter-regional economic disparity in China," *International Journal of Social Economics*, 1998, *25* (2), 424–447.

Tang, Sumei, E. A. Selvanathan, and S. Selvanathan, "Foreign Direct Investment, Domestic Investment and Economic Growth in China: A Time Series Analysis," *The World Economy*, 2008, *31* (10), 1292–1309.

Wang, Zhen Quan, "The determinants of foreign direct investment in transforming economies: Empirical evidence from Hungary and China," *Weltwirtschaftliches Archiv*, 1995, *131* (2), 359–382.

_ **and Nigel Swain**, "Determinants of Inflow of Foreign Direct Investment in Hungary and China: Time-Series Approach," *Journal of International Development*, 1997, *9* (5), 695–726.

Young, Stephen and Ping Lan, "Technology Transfer to China through Foreign Direct Investment," *Regional Studies*, 1997, *31* (7), 669–679.

Zhang, Zhaoyong, "International trade and foreign direct investment: further evidence from China," *Asian Economic Journal*, 1995, *9* (2), 153–167.

Zhao, Changwen and Jiang Du, "Causality Between FDI and Economic Growth in China," *Chinese Economy*, 2007, *40* (6), 68–82.

What Determines Operating Cost Per Rider?
A panel study into drivers of operating costs and unlinked passenger trips across U.S. public transit systems

Matthew Hong & Ayush Gupta
University of Toronto

1 Introduction

In 2014, government funding for the operating expenses of public transit systems in the United States totalled approximately $52 billion dollars, which accounted for 60% of all operating costs.[1] Given the vast sums of taxpayer money involved, it is no surprise that the topic of how to maximize agency revenues and how to model transit demand in various settings has been extensively studied. Chen, Varley and Chen (2010) studied the impact of the rise and fall of gas prices and transit fares on ridership. Thompson, Brown and Bhattacharya (2012) found that time cost is the most important variable in determining ridership from studying Broward County, Florida, which has an unusually high per capita bus ridership despite a weak central business district (CBD) and a lack of officially identified transit-oriented developments.

Others have focused on analyzing the main drivers of operating costs and suggesting minimum thresholds. For instance, Guerra and Cervero (2011) focus on establishing minimum levels of population and job densities necessary for a transit agency to be relatively more efficiently run on an operating cost per passenger mile basis. Substantial literature has also been dedicated to measuring the consumer surplus (or lack thereof) that transit delivers to society, the economy and the environment. On the other hand, Harford (2006) used a more generalized cost-benefit analysis and found that more

[1] John Neff and Matthew Dickens, "2014 Public Transportation Fact Book". See http://www.apta.com/resources/statistics/Documents/FactBook/2014-APTA-Fact-Book.pdf

than half of the 81 major systems studied failed the test i.e. the total costs were greater than the total benefits from the transit system.

In this paper, we attempt to model operating cost per rider as a function of various observable characteristics. Since it is a function of total operating cost and total ridership, one imagines that operating cost per rider will be affected by all the factors that determine operating cost and ridership. We therefore model operating cost and ridership separately and then combine the two models. After running the Hausman test, we find that the random effects model is more fitting for our data, and we run a GLS regression to analyze the panel data.

2 Data

In this paper, we analyze data from 1,042 transit agencies between 2005 and 2013 (both included). Our panel comprise of annual total operating expenses (*opex*), unlinked passenger trips (*rider*), service population (*pop*), service area (*area*), population density (*popden*), difference in commute time when taking public transit instead of personal car (*comdiff*), average fare revenue per passenger (*fare*), crude oil prices (*oil*), lag variable for capital spending (*lagcap*), job densities (*den*), and CPI (*CPI*). These variables are aggregated from the American Public Transportation Association (APTA), National Transit Database (NTD), U.S. Census American Community Survey (ACS), and Energy Information Administration (EIA).

Most of the existing literature studies the costs and benefits of public transit on a per-passenger-mile basis. However, we use unlinked passenger trips to examine the data from a new perspective and feel that this metric is more suited to our analysis due to two main reasons. First, we argue that once a transit system has been set up, the marginal cost of carrying a passenger an extra mile is very small. Furthermore, most of the extra cost is captured by additional fuel cost and service coverage area, both of which are already included in our analysis. Secondly, the per-passenger-mile metric is inherently biased towards passengers who travel longer distances. In this paper, we are using data from over 1,000 different transit corporations and we do not want to put undue emphasis on the data from systems that have large service areas; particularly because area is generally an exogenous factor based on the history of the city's development. With the exception of a few new and well-planned cities, the majority of cities have merely developed from previously existing neighbourhoods. Hence the coverage distance is often beyond the control of most transit authorities. Priority when designing transit systems, furthermore, is generally placed on servicing the greatest number of people from one concentrated region to the next, and less dependent on how far apart they are. Therefore, given that distance is generally pre-determined and transit authorities focus more on people regardless of where they are, we believe that operating cost per rider is a more accurate metric than operating cost per passenger mile when comparing operational efficiencies.

Operating expenses, capital spending, unlinked passenger trips, service area, and service population collected from the NTD are observed by transit agency name and by year as well as by reporter type.[2] Different modes of transportation such as streetcars, buses, heavy and light rail are aggregated in our panel data. Our motivation for aggregating transit modes comes from the fact that many cities have interconnected modes and often design systems with the intention of complimenting each other. Especially for transit modes that are managed by the same authority or management, the effect of co-variates will tend to be overestimated. Geographical features and demographic characteristics are also important factors, since certain transit modes can be infeasible over hilly areas, for instance. Moreover, even though the operational cost per rider may be more efficient for a given mode, the initial capital outlay may not be justifiable based on other demographic factors or budget constraints. The estimate from an analysis that attempts to separate transit modes within a given city will likely be biased and unreliable, if at all possible. This paper aims to capture a more holistic view of the transit operating spending, and aggregating different modes achieves this very objective.

Microeconomic theory suggests that ridership is determined (largely) by the time and monetary cost of public transit, relative to other means of transportation. In most of the cities that we study, the primary alterna-

[2] Note that the service area and service population are to be differentiated from regular area and population figures. The service data refer to a slightly more specific region that the transit systems aim to reach. As will be discussed further, this also does not equate to small catchment areas, which other studies have used as acceptable walking distances that individuals would be willing to take

tive to public transit is automobiles. Following this theory, we consider the difference in commute times for each region between public transit and automobiles to measure the time costs. We observe that the differences were predominantly positive numbers, which means that public transit almost always takes longer to commute with than with cars. We expect this figure to be negatively correlated with ridership, as the opportunity cost of taking transit would decrease as the time gap decreases.

The monetary cost of using public transit is the transit fare that you have to pay. However, transit fare is not the same for all riders. In most cities, students and the elderly enjoy discounted fares, and almost all cities have some sort of frequent rider discount (generally in the form of a monthly pass). To account for these, we calculate average fare revenue per unlinked passenger trip. We claim that this number is a better indicator of the fare paid by the average rider and is, therefore, a better metric than the single ride fare that most other papers use.

Another factor affecting ridership is the price of gasoline. An increase in the price of gas makes it more expensive to drive your own vehicle and should lead to an increase in ridership. Historical crude oil prices are collected from the Energy Information Administration (EIA) and used as a proxy for retail gasoline prices. Although in recent years gasoline price trends have experienced a divergence from those of crude oil prices, due to heightened volatility in the commodities markets, crude oil prices maintain a strong correlation with gasoline prices for the period we analyze. In addition to a positive correlation with ridership, we also posit that an increase in the price of crude oil may increase operating costs for transit agencies.

It is easy to see how the price of crude oil may affect the operating cost of bus systems, but for heavy and light rail transit that rely more heavily on electricity, oil prices may or may not increase their energy costs depending on the local utilities provider. It is unlikely that transit agencies would be paying at the same retail rates as individual households for electricity. We instead assume that most agencies would have special contracts with local utility companies, and the cost would vary greatly depending on their method of generation. To simplify this issue, we borrow the results from previous literature that has shown electricity costs to be generally correlated with crude oil prices.

Finally, we constructed a variable called *lagcap* ; it is the moving sum of capital expenditure in the previous 10 years. We expect capital expenditure to affect operating cost per rider in a variety of ways. Most importantly, we expect it to increase total ridership because the bulk of capital investment goes towards making the transit system larger, faster or more user friendly. We also expect capital expenditure to lead to economies of scale and thereby lower operating cost per rider.

While many of the independent variables we consider in this paper were readily available, we encountered missing observations in some of the collected data. Further investigation showed that in the vast majority of the cases, data was missing for exogenous reasons and not due to the values of the dependent and independent variables themselves. This means that operating expense data was not missing from the dataset because it was too high or too low. In fact, we found that many agencies had data missing in the NTD even though their agency website continued to report annual financial reports.

There are many reasons why such data could be missing, but the NTD specifies some main reasons- failure by agencies to submit data before reporting deadlines, weak compliance to the specific format in which the NTD requires agencies to report their data, and failure to reconcile discrepancies or inaccuracies in the submitted data by the reported deadline. As demonstrated, these reasons are exogenous to our study and therefore satisfy the missing at random (MAR) conditions. MAR conditions suggest that the observed sample is indeed an unbiased representation of the population, and we can therefore use list-wise deletion to carry out parameter estimations. A potential source of bias lies in cases where transit agencies in small cities or towns merge with neighbouring agencies to realize economies of scale and to provide a more seamless service. In such cases, it is difficult to determine whether lack of data has a correlation with the operating cost value itself. However, there could also exist political motives, or accounting policy-related technicalities behind these decisions, which again we believe are exogenous to our analysis.

3 Model

This paper uses unbalanced panel data under MAR conditions to model the operating cost per unlinked passenger trip for public transit systems in America. Random effects generalized least squares method is used to run the regressions. We divide the model into two parts: one modelling total operating expense, and the other modelling ridership. The motivation behind this separation of models is due to the ambiguous effect of some of the covariates on operating cost per rider. For instance, increases in oil prices may increase ridership, since operating a personal car would become more expensive. At the same time, an increase in oil prices may also increase operating cost for transit agencies. Such an effect, therefore, is likely to weaken the marginal effect that the covariate has on the dependent variable and may make the analysis more vague. Instead, dividing the model into two parts enables us to gain a clearer understanding of the partial effects. We specify the models as the following:

$$\begin{split} opex_{it} = \beta_0 + \beta_1 area_{it} + \beta_2 pop_{it} + \beta_3 popdens_{it} + \beta_4 totrider_{it} \\ + \beta_5 oil_t + \mu_{it} + \epsilon_{it} \end{split} \tag{1}$$

$$\begin{split} totrider_{it} = \gamma_0 + \gamma_1 pop_{it} + \gamma_2 popdens_{it} + \gamma_3 oil_{it} + \gamma_4 fare_{it} \\ + \gamma_5 lapcap_t + \gamma_6 emp_t + \gamma_7 comdiff_{it} + \upsilon_{it} + \delta_{it} \end{split} \tag{2}$$

We expect service area, service population, total ridership, and oil prices to have a positive effect on operating expense. As service area, service population and total ridership increase, so would the need for increased route lengths, engineers, and drivers to operate the systems. This would increase variables such as wage expenses, which would lead to an increase in operating expense. As oil prices increase, we would expect greater operating expense due to higher energy costs in operating the vehicles as well as machinery maintenances.

For the ridership model, we expect population density and job density to be positively correlated with transit demand. Since higher population and job densities lead to a greater likelihood of traffic jam during commute times, and therefore increase the appeal of alternate modes of transit. Oil price is also another crucial factor that affects ridership, since the higher the gas price, the more costly it is to drive cars and therefore makes public transit comparatively cheaper. One should also note that in this paper, the service population density and job density is not necessarily limited to the areas surrounding transit routes that many studies discuss. Our data is much more generalized and covers a greater area than the immediate neighbourhood of transit stations because, as exemplified by Thompson, Brown, and Bhattacharya (2012), not every agency is the beneficiary of transit oriented developments (TODs). This implies that the actual transit catchment area could exceed the immediate neighbourhood used in previous literature especially for sparsely populated regions, which would have overestimated the impact of population and population density.

Lastly, we consider the potential sources of endogeneity in our models arising from correlation of dependent variables with "between" and "within" error terms. For operating costs, we do not observe the quality of management, geographical features, urban layout, and climate patterns. Differences in these factors could lead to biased results. For instance, if the north-east region of the U.S. experiences greater snowfall during winters and therefore more frequent stoppages, the effect of our other explanatory variables would be overestimated. In RE models, we must also consider the error terms for within variables, which may account for differences for a given transit system across time periods. In our case, we do not observe changes in management and/or mandates, weather patterns, and level of security. Again, if the management style and mandates change towards providing more accessible services to the community with less regard for cost effectiveness, such factors may overestimate the partial effects of covariates with positive correlation. Bias can also come from the differences in the level of safety and security around stations and on vehicles. If crimes around public transit systems occur more frequently, ridership is more likely to decrease and operating cost more likely to increase as the need for transit enforcement increases. Lastly, public attitude and environmental awareness may also have an effect between and within agencies. A more environmentally conscious population would prefer to take public transit more often than cars, leading to increased ridership.

4 Results

We used the Hausman test to conclude that RE is a more fitting model than FE since some of the unobserved factors vary across time as well as across observations. The first regression we run is *operating cost* on *service area, service area population, population density, total number of riders* and *oil price*. We find that population, ridership and oil price are significant at the 99% significance level but population density and service area are not.[3] We propose that the insignificance of service area is due to the minimal marginal cost of servicing an extra mile. Also, since we are not considering the capital costs of setting up a larger transit system, it is expected that service area is uncorrelated with our dependent variable.

The insignificance of population density is less surprising. While researchers have observed that population density increases the initial cost of setting up a transit system, there is no reason why it should affect the operating cost once the system has been put in place. This is particularly true when you consider the fact that areas with extremely high population densities generally use modes like underground heavy rail, which are more suited to the area. The rest of the results are as expected with positive correlations with population, ridership and oil prices.

Next we run a regression of *ridership on population, population density, oil price, ticket fare, lag of capital spending, number of jobs in the area* and *difference in commute times*. We find that all of the variables are significant at the 90% level except for ticket fare. This is a very surprising result because economic intuition and past research suggest that there exists a strong correlation between ridership and ticket fare. We speculate that our findings differ because of the manner in which we gathered the ticket fare data. Most of the studies done in the past use the price of a single ride as the ticket fare. However, we believe that the average fare revenue is more representative of the overall riders. Furthermore, our finding can imply that the demand for public transit is more inelastic than was previously assumed. This might be true because a large part of our data comes from small towns, where the only people who take public transit are the ones who do not have many alternatives and the cost difference between alternatives that exist can

[3] Adding the robust option leads to no appreciable change in estimates

be very large.

Another notable result is that population has a negative coefficient. At first glance this seems counter intuitive, but one must consider two things. First, the absolute value for the is so small that one can argue that it is not economically significant. Secondly, population density has a positive and very large coefficient. Taken together, they suggest that, if one additional person were to move into an area, the total ridership of the transit system will increase significantly.

Finally we run the regression of *operating expense per rider* on all the variables used in the above two regressions except for *total ridership*. We omit this variable because we already have all the regressors that we used to model ridership. We find the R^2 value of the regression is lower than the other two regressions but still significant. The only regressors that are significant are fare, population, population density and the lag of capital spending. As expected, we find that the increase in capital spending or population density decreases per rider operating costs whereas an increase in population increases per rider operating costs. However, it must be noted that all three coefficients are extremely small.

This regression demonstrates that an increase in fare is correlated with an increase in operating cost per rider. Needless to say, the direction of causation in this case is unclear because systems with high costs will probably charge their customers a higher fare. To solve this problem, we used inflation as an instrument for fare. We also argue that the price of public transit is not determined by the forces of demand and supply. Instead, it is a public good and its price is often motivated by political considerations. As noted earlier in this paper, more than half of the operating expenses of public transit systems are paid for by the government in the form of funding and subsidies.

5 Conclusion

In this paper we have identified some of the factors that determine operating cost, ridership and operating cost per rider. We have modelled operating cost per rider through a two-step approach, and in the process, discovered the partial effects that various factors have on operating expense and unlinked passenger trips respectively. By identifying statistically significant correlations between operating expense, ridership and measurable metrics such as population density, service area, capital expenditure, and crude oil price, we have created a holistic model to compare the relative efficiency of transit systems and to evaluate the feasibility of new transit projects. We have shown that certain 'obvious' factors have no correlation with our dependent variables and we have also used some hitherto underutilised data to provide a fresh perspective.

This paper is set apart from the existing literature due to its scope both in terms of the variables being considered and the sources of the data. By including data from over 1,000 transit systems, we ensure that our results are applicable, not just to big cities but also to small towns. Furthermore, our results provide justification for further research into some of the new data we used.

As noted previously, a wide array of unobserved factors could have affected our results. A great number of researchers have devoted their research to modelling demand as well as cost models. Some areas that could warrant further research is differences in management style between agencies, and perhaps public opinion towards public transit in various regions.

6 Appendix

Table 1: Random Effects GLS Results

	opex	totrider	opride(per '000s riders)
area	-42.082		-0.114
	(86.844)		(0.076)
popdens	-334.668	374.081***	-0.264*
	(456.094)	(104.892)	(0.103)
oil	259721.4***	12436.44***	-1.286
	(36320.0)	(4609.29)	(8.910)
totrider	2.175***		
	(0.025)		
pop	12.554***	-1.246***	0.001***
	(1.219)	(0.162)	(0.0002)
fare		-30590.74	648.047***
		(60643.16)	(97.755)
lagcap		0.012***	0.000***
		(0.0006939)	(0.000)
emp		79.320***	-0.010
		(25.992)	(0.022)
comdiff		-26607.46*	16.786
		(14491.72)	(18.530)
Constant	-0.000	9066469	7853.954
	(0.127)	(0.128)	(996.020)
Observations	4983	2125	2125
"Within" R^2	0.5313	0.0872	0.0001
"Between" R^2	0.9283	0.7764	0.2631

Standard errors are displayed in parentheses. Significantly different from zero at 99 (***), 95 (**), 90 (*) % confidence.

References

American Public Transportation Association, "2014 Public Transportation Fact Book," Technical Report, American Public Transportation Association, Washington, DC 2014.

Anderson, Michael L., "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion," 2013.

Chen, Cynthia, Don Varley, and Jason Chen, "What Affects Transit Ridership? A Dynamic Analysis involving Multiple Factors, Lags and Asymmetric Behaviour," *Urban Studies*, 2010, *48* (9), 1893–1908.

Guerra, Erick and Robert Cervero, "Cost of a Ride: The Effects of Densities on Fixed-Guideway Transit Ridership and Costs," *Journal of the American Planning Association*, 2011, *77* (3), 267–290.

Hassan, Morsheda and Raja Nassar, "Empirical investigation and modelling of the relationship between gas price and crude oil and electricity prices," *Journal of Economics and Economic Education Research*, 2013, *14* (3), 119–129.

Liu, Li and Guofeng Ma, "Cross-correlation between crude oil and refined product prices," *Physica A: Statistical Mechanics and its Applications*, 2014, *413*, 284–293.

Thompson, Gregory, Jeffrey Brown, and Torsha Bhattacharya, "What Really Matters for Increasing Transit Ridership: Understanding the Determinants of Transit Ridership Demand in Broward County, Florida," *Urban Studies*, 2012, *49* (15), 3327–3345.

U.S. Census Bureau, "Means of Transportation to Work by Selected Characteristics: 2010-2014 American Community Survey 5-Year Estimates," Technical Report 2014.

The Promise of a Better Job:
A Study of the Employment-Population Ratio
of the West African Economic and Monetary Union

Parker Griffin
The George Washington University

1 Introduction

Recent calls for the disintegration of the West African Economic and Monetary Union (WAEMU) have called into question the economic benefits of the union for member countries. Indeed, since the 2008 global financial crisis, many of these countries have experienced slow export-led growth. This has been partly a result of the union's artificially strong pegged currency, the CFA Franc, established during French colonial control of the region. Heads of state and leaders of the WAEMU have called for a reform of the union and its currency. As a result, many are examining economic output of member and nonmember nations in order to determine the benefits of membership. One important measure of economic performance is the rate at which the economy can produce employment for the working age population. This employment-population ratio gives valuable insight into both the labor market and overall economic performance. This paper will examine the employment-population ratios of member states of the WAEMU in relation to other African countries.

2 Literature Review and Overview

Current studies of the WAEMU have focused on financial integration, measurement of successful cohesion of member states, and econometric analysis of the integration of the WAEMU. The majority of these studies conclude that there has not been much significant economic convergence of the member states. In fact, some conclude that the multitude of macroeconomic policies undertaken by the union members may have led to divergence. Evidence suggests that the instability of the private financial sector and procyclic nature of government expenditure is driving member states apart.

Sy (2006) has investigated the financial integration of various institutions in the WAEMU since its inception. Sy points to efforts made by member states to eliminate intra-union financial barriers and to harmonize regulation, evidenced by the 1995 PARMEC Law. This law achieved a basic framework for consistent savings and standardized financial legislation across member states. Sy acknowledges, however, that the regional banking system is small and managed by a handful of European banks. For example, 56% of total bank capital and assets are foreign owned, predominantly by French firms. Indeed, because colonial banking frameworks still persists in most countries, the financial sector is somewhat fragile. Sy concludes that if further integration were to occur it would need to come in the form of an overhaul of the interstate financial system.

Dessus et al. (2014) have studied the procyclicality of public investments in the WAEMU and offer numerous policy suggestions in order to correct the issue. Increases in social services and uptakes in infrastructure projects are a result of both higher economic prosperity and increases in foreign direct investment. They suggest that member countries should adopt cohesive and union-wide counter-cyclical fiscal rules by way of risk sharing in order to mitigate shocks. A common fear is that these countries engage in procyclical public consumption that drive up public debt to levels that the governments can no longer service during periods of stagnation or recession. Accordingly, Dessus et al. determine that the WAEMU needs further regulation on government spending in order for meaningful integration to occur.

Bah (2015), by contrast, uses an econometric model to investigate the real convergence of the WAEMU. By using a CM model and taking possible structural breaks into account, Bah measures country specific linear time trends, GDP, and the ratio of the GDP per capita per year compared to the average GDP per capita. Bah concludes that some WAEMU states – such as Burkina Faso, Mali, and Togo – are experiencing convergence. However, his results also suggest that divergence within the union is occurring in both Niger and Senegal.

This study will build off of Bah's work and approach, using difference-in-differences modeling to examine the impact of WAEMU membership on the employment-population ratio. Accordingly, this paper will attempt identify a causal link between WAEMU membership and the employment-population rate in these member states. It seems reasonable that the integration and harmonization of WAEMU macroeconomic policy since 1994 has led to a higher employment-population ratio among union members as compared to other African states.

Economic theory and ex-ante study both suggest that the union should have a positive and significant impact on the employment-population ratio of the member countries. This would be true not just of WAEMU but of other economic and political unions. Indeed, recent studies of Poland have showed that its membership in the European union have increased its employment rate. Rae (2008) has analyzed these effects, finding that policies focused on economic harmonization have benefited the labor market of the country. I will examine whether this case holds for the WAEMU countries as well. The discovery of the union having a significant and positive effect on the employment levels for member states could bolster support for its continuation against the current skepticism. Further, if the WAEMU has a positive effect on the employment-population ratio on its member states, it would imply that further economic and monetary union across the African continent could be pursued in order to promote higher employment rates.

The structure of the paper is as follows: First, I describe the methodology and regression model that have been selected for this paper. Then, I discuss regression results and relevant graphs. Finally, I draw conclusions and policy suggestions for the WAEMU based on my results.

3 Data and Methodology

3.1 Data

Data for my analysis are drawn from the Penn World Table and the United Nations database for the years 1980 to 2014. This dataset includes panel data of hundreds of countries and gives annual metrics for each. The data was regressed in five different iterations. Each iteration of the regression measured the WAEMU states against various groupings of African countries.

Additionally, the UN's Population Division of the Department of Economic and Social Affairs compiles data from countries every five years to measure the net migration flow from each country – that is, the number of immigrants minus the number of emigrants. This figure is expressed in thousands, with the corresponding graphs found in the Appendix.

Employment-population ratios of the eight WAEMU countries (Benin, Burkina Faso, Côte D'Ivoire, Guinea Bissau, Mali, Niger, Senegal, and Togo) were compared against (i) all other African countries, (ii) all other African countries without South Africa, (iii) African countries formerly under British colonial rule, (iv) African countries formerly under French colonial rule, and (v) the Central African Economic and Monetary Union Community (CEMAC) member states. Iterations will serve to examine a difference-in-differences explanation to see if the effect of the union on the employment-population ratio changes depending on the sample of countries the WAEMU is measured against. The following is an explanation of the groupings of countries chosen.

The first iteration measures the impact of the WAEMU on the employment-population ratio in relation to the entire continent. Another way to look at this model is to remove South Africa. This country is often excluded from samples in other papers on African economics due to the unusual prosperity of that country. This study continues with convention by employing this practice in the second iteration. The third iteration compares the relative variations in the employment-population ratio between the WAEMU countries and former British colonies as some studies suggest that former British colonies have a natural economic advantage against former French colonies[1]. Since Guinea-Bissau is a former Portuguese colony, the fourth iteration of the model will test the WAEMU against all other French colonies to ensure there is no significant effect of Guinea-Bissau on the predominately Francophone union. For the purposes of this study, former French colonies include the former Belgian colonies of the Democratic Republic of the Congo, Rwanda, and Burundi because of the similarity of the methods of colonial rule with the French. Lastly, the final iteration of the regression will test WAEMU states against the CEMAC, another economic and monetary union created

[1] Seven of the eight WAEMU states are former French Colonies, and Guinea-Bissau a former Portuguese colony

by the French during the colonial period under the same pretenses. This sister union also has a common currency, allowing us to see which union has stronger employment rates.

3.2 Regression and Variables

The dependent variable – the employment-population ratio – was generated and calculated for each country yearly by dividing the number of people employed by the total working age population in millions. Looking at the independent variables, the $WAEMU$ variable is a dummy variable where $WAEMU = 1$ when the country is a member of the union and $WAEMU = 0$ when the country is not in the union. GDP and $Population$ are included as variables in order to control for these major macroeconomic factors. The $GDPPC$ variable represents the GDP per capita of a given country, measured by placing the economic output over the population of the country. This variable succinctly demonstrates the economic prosperity of a country and is included in this model in order to account for variations in standard of living across the groups of countries included. This also helps control for countries that may have extraordinarily large populations that would otherwise skew the data in the model. $CapitalStock$ and $PLCapital$ represent the current capital stock at current purchasing power parity (priced in U.S. dollars) and price level of capital formation, respectfully. The resources invested in capital and the price of capital impact the ability of countries to employ citizens, therefore necessitating the inclusion of these variables in the model.

The regression model is designed to account for a variety of factors that may influence the employment-population ratio independent of WAEMU effects. The model is as follows:

$$\begin{aligned} EmpPopRatio_{it} = \beta_0 + &\beta_1 WAEMU_{it} + \beta_2 GDP_{it} + \beta_3 Population_{it} \\ &+ \beta_4 GDPPC_{it} + \beta_5 CapitalStock_{it} + \beta_6 PLCapital_{it} + \varepsilon_{it} \end{aligned} \tag{1}$$

This regression provides an estimate of the effect of WAEMU membership on the employment-population ratio, controlling for the other independent variables.

4 Graphical and Regression Analysis

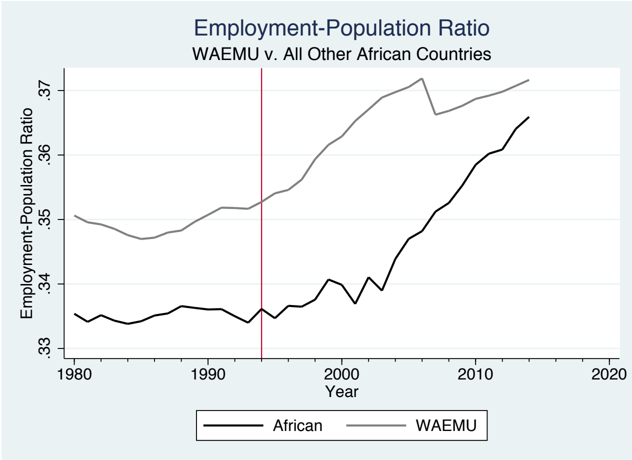
4.1 Graphical Analysis

Plotting the average annual employment-population ratio for the WAEMU states against the five alternative groupings gives a visual representation of its yearly variations. The union formation year of 1994 is demarcated in each graph by the red line. The graphs discussed here can be found in the Appendix.

Looking to trends in WAEMU members, from 1980 to 1994, the level of employment first dipped and then slowly recovered. Historians and economists often designate the 1980s as Africa's "lost decade". Standards of living fell to disparagingly low levels, wages stagnated, and inflation ate up the salaries of many African workers (Shepard, 1992). Across the board, the graphs make the lost decade abundantly clear as all trend lines fail to grow further than pre-1980 levels. Looking at Figure 1.1, the employment-population ratio of the WAEMU countries climbs steadily after 1994, then undergoing a sharp decline in the recession years of 2007-2008 followed by a recovery. Over the last decade alone millions of working age adults found employment as the ratio jumped from 0.35 to 0.37.

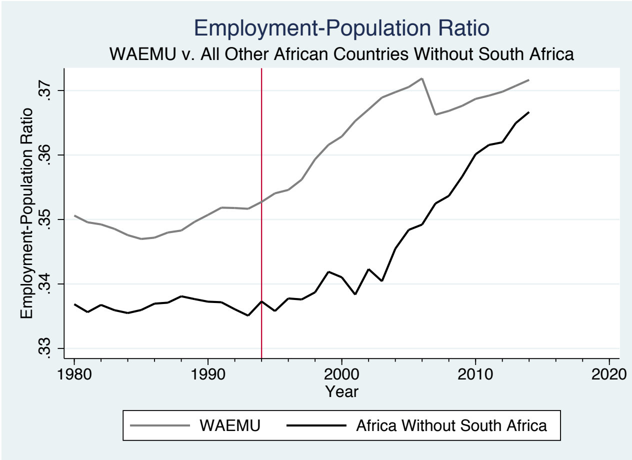
Although employment rose as a result of the macroeconomic policies pursued by WAEMU countries during the two decades following the mid-1980s, the global financial crisis was a major step back for member states. Over the course of one year the average employment-population ratio dove .05 points. After 2009, however, we can see that the line trends positive once again after this dip. However, current data shows that to date the growth of the employment population ratio has been growing at a slower rate than it had been before the crisis. Over the last 5 years the WAEMU has recovered its employment population ratio to over 0.37. The following is an in-depth look at each iteration of the graphical representation of the data.

Figure 1.1



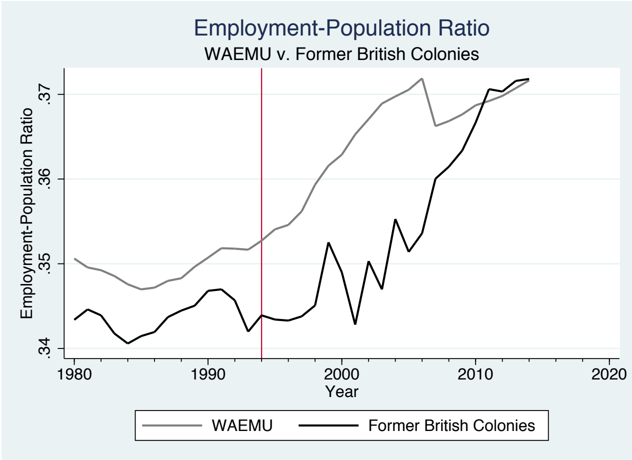
(i) Comparing the WAEMU to all other African countries. Figure 1.1 depicts the trend lines for both groups of countries, demonstrating that the whole of Africa was susceptible to a poor labor market as a result of the lost decade. However, from 1994 to 2014 the gap closes significantly with other African countries climbing from a ratio of 0.335 to 0.365. This dramatic catch-up by states outside of the WAEMU comes within 0.5 percentage points of the two groups' ratios, despite starting out a full 1.0-point behind in 1994. While the two groups appear to follow parallel trends from 1994 to the recession years of 2008-2009, it appears that the rest of Africa is performing better in this metric in recent years. Indeed, the WAEMU may now be a hindrance to growth of the employment-population ratio.

Figure 1.2



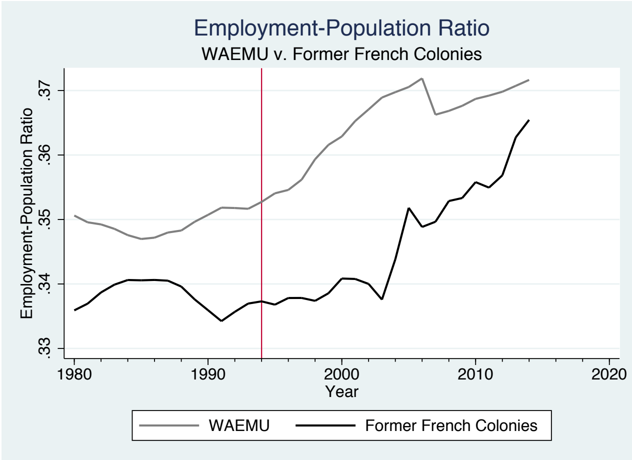
(ii) Comparing the WAEMU to All African Countries except South Africa. South Africa, the most economically prosperous country on the continent, is often an outlier and tends to skew the data in some regressions. Therefore, South Africa was removed from an iteration of analysis to determine if it had a significant impact on the results. However, the group of all other African countries performs slightly better with South Africa excluded, as show in Figure 1.2. Thus, it can be concluded that the superior performance of other African countries in the recovery period is not the result of South Africa's strength alone.

Figure 1.3



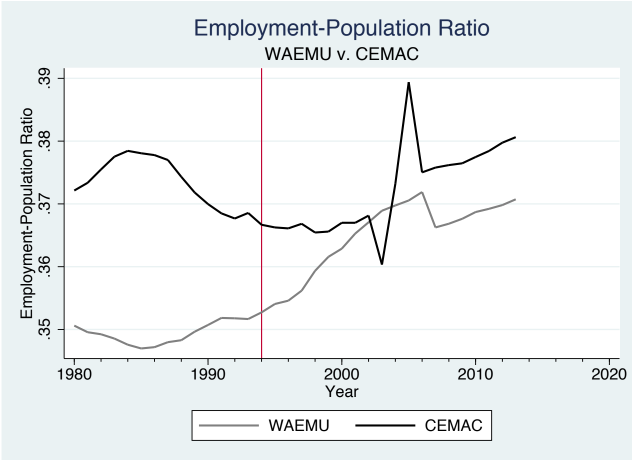
(iii) Comparing the WAEMU to Former British Colonies. Economic development has often been linked to colonial legacies, and so it is natural to compare the performance of the WAEMU with other colonial groupings. Starting with former British colonies, Figure 1.3 shows that African countries formerly under British colonial rule have equal employment-population ratios as those in the union. In fact, former British colonies overtook the union for a short time during the recent recession. Generally, however, the two groups appear to perform quite similarly. This is particularly intriguing because previous studies have suggested that former British colonies outperform former French colonies. However, here we see that these groupings of Francophone countries (with one former Portuguese colony) are largely equal measure with their English-speaking counterparts. While this doesn't upend any previous research, it does recognize the particular strength of WEAMU states in their employment-population ratios.

Figure 1.4



(iv) Comparing the WAEMU to Other Former French Colonies. Figure 1.4 investigates any variation between these two groupings of French colonies. Although the WAEMU profited from a growing labor market from 1994 until 2003, the former French colonies have caught up and have seen their employment-population ratios rise by nearly 2.5 points. It is evident that all former French colonies experienced a dip as a result of the global financial crisis in 2008. But, as noted with both Figure 1.1 and Figure 1.2, the gap between the WEAMU states and this group of African countries has been closing significantly in recent years.

Figure 1.5



(v) Comparing the WAEMU to CEMAC. Finally the WAEMU was tested against the CEMAC to determine the relative success of both of these economic and monetary unions, as shown in Figure 1.5.[2] CEMAC countries all use the Central African CFA franc, a sister currency of the WEAMU's West African CFA franc. Comparing these two economic institutions should yield very similar results since they were use the same currency and similar models. Although the two unions had identical employment-population ratios by 2002, for the past six years the CEMAC has consistently had higher rates than the WAEMU. Although these economic and monetary unions have a shared history and have undergone many of the same steps in their pursuit of economic harmonization, the CEMAC has been able to employ more workers than its West African counterpart.

Across the board, the WEAMU has struggled in recent years against reference groups. Although in three out of the five cases WEAMU countries have a higher employment-population ratio, in all cases they have a lower growth rate. It appears that the WEAMU is struggling to provide employment for its population relative to other African countries, especially since 2007.

[2] CEMAC is the French acronym for the Central African Economic and Monetary Community, in French *Communauté Économique et Montaire de l'Afrique Centrale*

4.2 Regression Results

Table 1: Regressions on Employee Population Ratio

Grouping	(1)	(2)	(3)	(4)	(5)
$WAEMU$	0.0171	0.0171	0.00877	0.0163	-0.0205
	(0.61)	(0.62)	(0.28)	(0.41)	(-0.73)
$rgdpe$	2.05e-08*	5.67e-08***	3.89e-08***	3.04e-7***	-6.59e-09
	(2.48)	(6.64)	(3.39)	(14.32)	(-0.05)
pl_i	-0.00460*	-0.00194	-0.00522**	-0.00611*	0.00101
	(4.83)	(4.62)	(1.86)	(0.54)	(-0.45)
csh_i	0.0378***	0.0348***	0.0159	0.00457	-0.00565
	(-2.36)	(-1.03)	(-2.60)	(-2.36)	(-0.26)
$GDPPC$	-6.04e-08	-3.91e-08	-1.31e-7	-3.1e-7	-4.49e-07
	(-0.42)	(-0.28)	(-0.95)	(-1.26)	(-1.43)
$Cons$	0.339***	0.337***	0.348***	0.339***	0.379***
	(28.68)	(27.98)	(25.06)	(15.42)	(18.07)
N	2170	2105	1705	1019	576

T-statistics are displayed in parentheses. Significantly different from zero at 99 (***), 95 (**), 90 (*) % confidence.

Regressions results confirm the graphical evidence. For each of the comparison groups, the WEAMU dummy variable does not have any significant effect on the employment-population ratio. The table of results given in Table 1 further establishes the evidence that the formation of the union has not had any significant effect on the employment-population ratio of member countries. This is explained by the lack of significance of the WEAMU dummy variable.

Contrary to the initial hypothesis, the legal and structural progress that the union has made towards economic harmonization has not appeared to benefit its labor market in recent years. The regression also indicates that none of the chosen variables are significant when it comes to measuring the impacts to the employment-population ratio between the WAEMU and the CEMAC countries. This result seems to suggest that there are other macroeconomic situations that the model hasn't – or can't – account for determining the employment-population ratio.

The WAEMU does not have any significant impact on the employment-population ratio of the member countries. After testing the members against various other groups of African countries, it is clear that the union has not benefited from significantly stronger employment levels in these countries, especially in the recent recovery. It is possible that the member states are not economically or politically developed enough to have proper methods for enforcing and securing the benefits from the harmonization of macroeconomic policy union-wide. As a result, the union at present has not seen a benefit in terms of its employment-population ratio after decades of attempted convergence.

4.3 The Curse of Small Countries

I now further analyze the possible effects of the WEAMU on the employment-population ratio of member states by examining a scenario in which the strongest countries of the union are removed. Côte D'Ivoire and Senegal are the union's two most economically prosperous members; they alone account for 54% of the GDP of the WAEMU. When these two countries are removed from the data set, the WEAMU countries have both a higher overall average employment-population ratio as well as a smoother trend line with softer peaks and troughs.

The overall employment-population ratio of WAEMU countries is on average is 0.01 points higher when Senegal and Côte D'Ivoire are excluded from the data.[3]. These two economic powers may be dragging down the overall WEAMU employment-population ratio because their population is growing in a way that cannot support the job market. Because workers in the member states of the WEAMU benefit from relaxed borders, many economic migrants from countries such as Mali and Niger move south and west looking for more work. This migration could inflate the population of Senegal and Côte D'Ivoire and cause them to have more people than they can employ while simultaneously opening the labor market for those who choose to stay in the more economically disadvantaged countries. Figure 2.1 and Figure 2.2 support this assertion.[4] The graphs are drawn using data from the United Nations Population Division of the Department of Social

[3]See figure 2.1 in the Appendix
[4]The figures are available in the appendix

and Economic Affairs. Although Senegal has been increasingly challenged with high levels of out-migration, their increasing population prior to 2008 suggests that their employment-population ratio would be inherently lower due to their growing immigrant demographic. By the same token, countries that are both the most disadvantaged and have larger out-migration rates – Mali, Niger, and Burkina Faso – all have weaker job markets as a result of such exodus.

In order to further verify the presence of economic migration within the union, remittance data was crosschecked to look for patterns suggesting that there is a high volume of remittances flowing out of Côte D'Ivoire and Senegal and into other WEAMU countries. An in-depth study by the Central Bank of West African States (BCEAO) found that 62.9% of Côte D'Ivoire's remittance outflows are sent to other countries within the WEAMU whereas only 8.3% of remittances received by Côte D'Ivoire are from other WEAMU countries. A study on the remittance markets in Africa by The World Bank found countries like Burkina Faso receive more that 72% of their remittance inflows come from Côte D'Ivoire and Senegal. Coupled with the migration rate data, the remittance numbers conclude that Côte D'Ivoire, and to a lesser degree Senegal, have positive net migration rates, and therefore have lower employment-population ratios than their fellow union members with positive net remittance. This data confirms that the stronger countries of the WAEMU will inherently have lower employment-population ratios due to their intake of economic migrants from other countries in the union.

It is important to note, however, that excluding both Côte D'Ivoire and Senegal from the WAEMU increases its employment-population ratio in 2007 to nearly twice its rate with both countries included. However, although these two countries may be depressing the overall employment-population ratio of the WAEMU, they also act as a countercyclical buffer in the event of an adverse economic event.

5 Conclusion and Suggestions

The result of the graphs and regression models overwhelmingly suggests that the WEAMU has not performed significantly better than other African countries in terms of the employment-population ratio. The rate at which the metric has increased since 1994 has slowed in the post-recession recovery. More recently, in fact, the WAEMU member states have lower employment-population ratios than CEMAC countries and are on par with former British colonies. The regression model estimates confirm the graphics. Despite the attempts of member states to harmonize their economic and legal structures to promote stronger labor markets, the union has not had a significant impact on the employment levels of its member states. Regardless of the iteration of the regression and the group of African countries the WEAMU states were tested against, the employment-population ratio is not affected by union membership.

Moving forward, major changes may be necessary for the union, which could include taking steps to further integrate the economies of the member states. In considering new economic policies, the WEAMU should perhaps first look towards its Central African counterpart, the CEAMC. Indeed, the sister union has been able to maintain a higher employment-population ratio than the WAEMU since 2003. Although the member states that comprise this group may have economic factors that make them a better economic and monetary union, their recent harmonization policies, such as the easing restrictions on working visas, could be useful next steps for the WEAMU. Although total disintegration seems hasty, the union appears to have no impact on employment rates and therefore should reevaluate its plan to strengthen its labor market in the coming years.

A more radical suggestion is to remove Senegal and Côte D'Ivoire from the union. These countries succeed economically on their own and have macroeconomic and structural needs that are different from the majority of economically disadvantaged countries in the union. Furthermore, the free movement of labor has become a hindrance to both the growth of the employment-population ratio and labor market of these countries. Considering this and the additional evidence discussed in this paper, it seems that change in the WAEMU may be needed.

The author would like to extend his thanks to Remi Jedwab and Mohamed Siry Bah for their support and advice, which undoubtedly added to the value of this paper. Additionally the author would like to thank Professor Donald Parsons for his comments and guidance throughout the study.

6 Appendix: Supplementary Figures

Figure 2.1

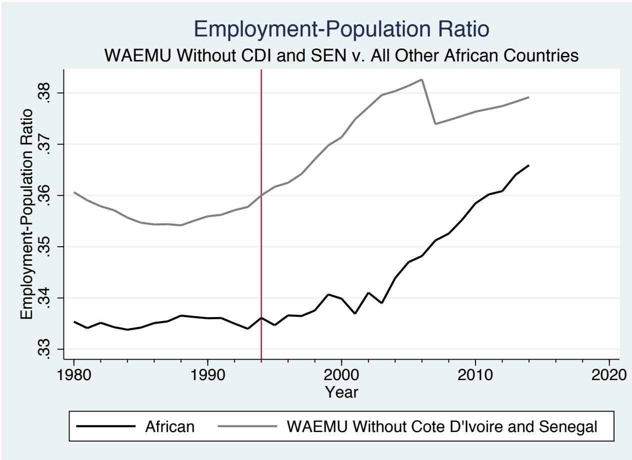


Figure 2.2

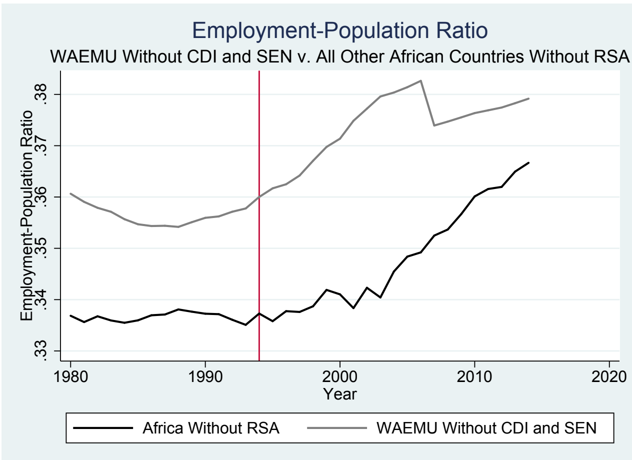


Figure 2.3

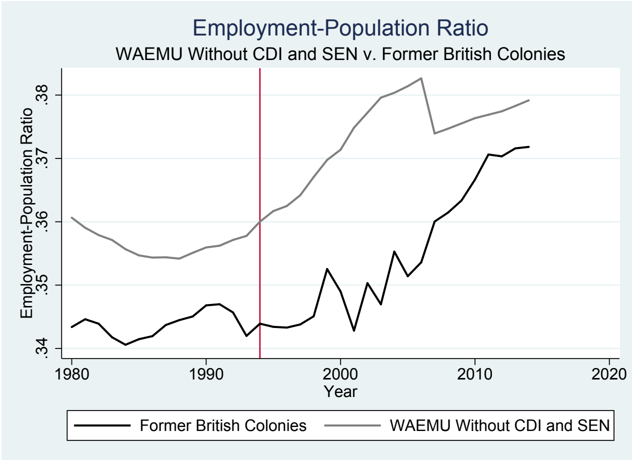


Figure 2.4

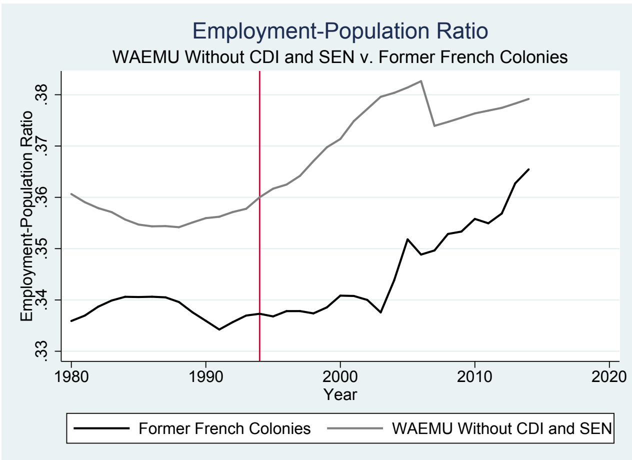


Figure 2.5

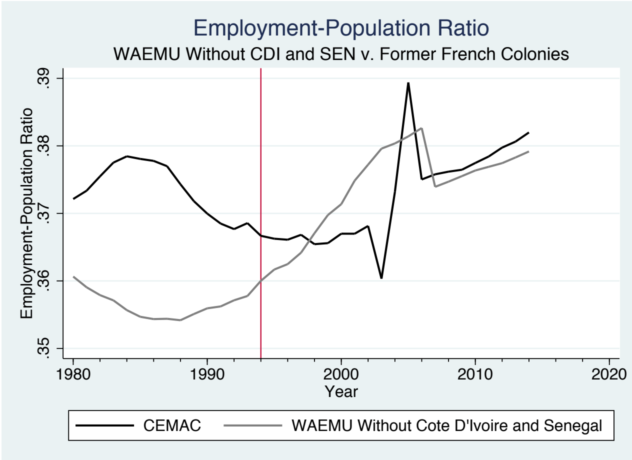


Figure 3.1

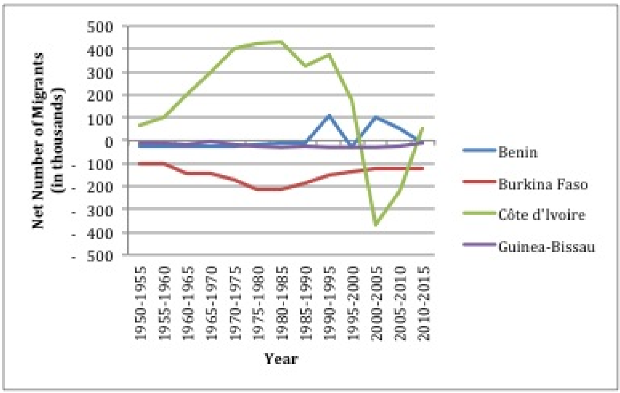
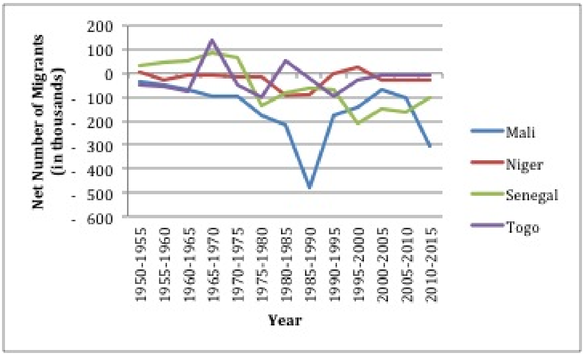


Figure 3.2


Net Number of Migrants
(in thousands)
200
100
0
- 100
- 200
- 300
- 400
- 500
- 600
1950-1955
1955-1960
1960-1965
1965-1970
1970-1975
1975-1980
1980-1985
1985-1990
1990-1995
1995-2000
2000-2005
2005-2010
2010-2015
Year
Mali
Niger
Senegal
Togo

Player Choice in League of Legends

En Hua Hu & Hon Chiu Yuan
University of Toronto

1 Introduction

The online video game *League of Legends* is hugely popular with over 100 million monthly active players. In the game, a player plays one-shot matches with 9 other players, with 5 on each team.[1] Before the game begins, players are assigned, based on preferences, to one of five roles. Once ten players are matched together, players choose from 133 "champions" to play for the duration of the match. Champions are characters with different abilities, looks, and personalities that players are then in control of during the game.

In this paper we investigate the factors determining these champion choices. In particular, we are interested in how much winning matters, how much professional play influences casual and amateur play, and how important the "fun" factor is for different champions and its relation to win-rate. We predict that this cross derivative is negative. That is, the more "fun" a champion is, the less important its win-rate becomes for players – players don't only play to win.

We use data from games played over several months from the website *op.gg* to study these questions. The data is divided into five periods. At the beginning of each period, the game owners apply a patch. A patch is an update to the game where slight modifications are done in order to weaken or strengthen champions. With different patches the win-rate changes, allowing us to study its impact on popularity.

Our analysis has two stages. First we observe that the market for 'champion' is divided into nests. Those nests correspond to different roles champions belong to in the game. This needs to be taken into account and we use

[1] A one-shot match is one that is not impacted by other matches in the tournament, similar to soccer or most other sports.

a simplified nested logit model to study this. To be more precise, the standard nested logit model does not apply since nests here do not compete and we do not have nest-specific variables. Second, we wish to investigate the fun-factor of a champion, which is unobservable. We isolate four variables which we believe are highly correlated with this fun factor which we call "attractiveness". Following that, we use these variables in factor analysis which to generate a factor which explains the common covariance between our variables.

We finally run our model and to test for an interaction between win-rate and attractiveness. Namely, our hypothesis is that both are positively related with pick rate but that their cross derivative is negative. That is, a champion that is highly attractive should suffer less from having a low win-rate and vice versa. We test this and find it to hold for four out of five roles.

The paper proceeds as follows. In Section II we review some of the existing literature. In Section III we briefly discuss the game and relevant facts and rules. In Section IV, we address potential methodological issues in our analysis and our approaches to mitigating them. Lastly, in Sections V, VI and VII we present our model, the datasets, and the results of our study.

2 Literature Review

Although there is an abundance of data on *League of Legends* matches available through public APIs, there has not been much research on this field or using this data. This could be because the access of these datasets is not simple, and requires a degree of familiarity with coding and programming. Additionally, the data available online concern only in-game and player information – that is, no information on sales or the company are released. Therefore, the economic value and interest of the data are rather low. However, the market itself is still an interesting one. The decision-making process of the players contain patterns that are not common in other markets, and it will thus be necessary to develop new techniques to analyze player choices.

Most of the related literature that uses similar datasets comes from research in computer science and machine learning. Conley and Perry (2013) use machine learning algorithms on a similar dataset to predict outcomes of games in Dota2, a game comparable to *League of Legends.* Further work, also in machine learning, is summarized in Semenov et. al (2016).

Player incentive, however, has been studied in freemium models, and there is a growing literature as this business model has also grown in popularity. Kumar (2014) has recently introduced this model and argued for its significance and effectiveness.

We use factor analysis to study player choice. The goal of factor analysis is to break down the correlations between variables, identifying them as factors. This can be done, for example, in the natural sciences to isolate key variables and their relationships. In social sciences, this is usually done to analyze a more conceptual – and often unobservable – variable. For instance, a student's performance might be related to how much he enjoys studying. Thus, researchers can use variables such as class attendance, grades, and others to estimate this unobservable enjoyment. Yong et. al (2013), provides a sound introduction to the topic and its applications.

3 Game Information

In this section, we summarize the process of choosing a champion that a player undergoes for readers who are not familiar with the game.

Players queue to be matched with other players of similar skill level. However, before a player enters a queue, she selects from one of five roles. She is then matched with players so that two players of each role are present in the game. Roles determine the position that a player has on the team, but for our purposes a role is a set of champion. Indeed, certain champions fit certain roles and in picking a role, a player has also picked a set of champions.

After being matched with her teammates and the opposing team, a player picks her champion from those of her role. It is very rare that complementary preferences occur. That is, it is rare that a player's choice of champion depends on the choice of her team. It is possible, however, that the champion the player wishes to play is banned or picked by the enemy team, in which case it is no longer available.

There are many factors which influence player choice. Our dataset uses the top 10% of all ranked players, and we therefore expect players to pick champion which have high win-rates, which are more likely to lead to wins.

Furthermore, we expect that players also play for fun, and so champions which are more fun to play will be picked more often. Lastly, players tend to follow professional play so we expect the rate of appearance of a champion in professional games to be reflected in player decision.

4 Methodological Issues and Assumptions

To estimate a structural model of player behavior, it is important to address some issues unique to this framework. There are three aspects which our model needs to take into account:

1. Bans and exclusivity of picks
2. Role of champions and independence of irrelevant alternatives (IIA)
3. Innate unobserved attractiveness

We now briefly summarize the issues and our method in dealing with them.

4.1 Bans and exclusivity of picks

During the champion pick phase, a player's intended champion may be picked away by the enemy team or banned. The multinomial logit model does not account for the fact that decisions may have different choice sets. This means that we do not observe the intended pick. However, given that the ban-rates are available, we would like to find the true pick-rate.

Fixing some period, let the true pick-rate be tp_i and the ban-rate be b_i. Also let sp_i be the probability that a champion is picked as a second most preferred option. Finally let the observed pick rate be p_i. We then have:

$$p_i = tp_i(1 - b_i)(1 - tp_i) + sp_i$$

Namely, the observed pick-rate is the probability that it was the intended pick and not banned or picked by another player added to the probability it was picked as second most preferred champion.

We then compute sp_i as:

$$sp_i = tp_i * \sum_{j=1}^{N} [\frac{tp_j}{\sum_{k=1}^{N} tp_k - tp_j}] - \frac{tp_i}{\sum_{k=1}^{N} tp_k - tp_i}$$

Here, tp_i is no longer just a function of b_i and p_i, but is still solvable since we have N unknowns and N equations. However, it is likely highly computationally complex and the probability of being second picked is likely small.

Instead, we will assume that it is uniformly distributed proportionately to observed pick-rate. We now suppose that $sp_i = 0$, and we will solve for $tp_i^{'}$, and then take the difference between $tp_i^{'}$ and p_i to obtain tp_i.

From the discussion above, we have: $p_i^{'2}(1 - b_i) + tp_i(1 - b_i) = 0$

It turns out that there is always only one solution that is positive:

$$tp_i^{'} = \frac{(b_i - 1) + \sqrt{(1 - b_i)^2 + 4p_i(1 - b_i)}}{2(1 - b_i)}$$

We now assume that that sp_i is distributed uniformly per true pick rate:

$$sp_i = \frac{\sum_{i=1}^{N} tp_i^{'} - \sum_{i}^{N} p_i}{\sum_{i=1}^{N} p_i} * p_i$$

Finally, this gives us $tp_i = tp_i^{'} + sp_i$

4.2 Roles and IIA

The IIA assumption of the logit model is problematic given that champions have specific roles and often only compete within their roles. The multinomial logit model predicts that if a champion's win-rate decreases, then all other champion's pick-rates increase by the same amount. We cannot, however, assume that a top champion's pick-rate or win-rate change affects that of a mid champion.

In the standard literature, this problem is addressed with a nested logit model, where products are further categorized into nests. However, this is also not appropriate for our analysis, since the nests compete against each other and we require nest-specific variables. In our setting the roles have fixed pick-rates. Every game will have the same number of top, mid, and other roles. Therefore we use a fixed nested model. In short, this means that the player has equal probability of choosing any role. This player's probability over champion is thus given by a multinomial model within the champion of that role.

Our model gives the following:

$$tp_{it} = 0.2 * P(i\ is\ chosen \mid i's\ role\ is\ chosen)$$

$$U_{ij} = \phi(w_{it}, \chi_i) + \beta X_{it} + \epsilon_{ij} \;\; if\ j\ is\ assigned\ the\ role\ of\ j,\ 0\ if\ not.$$

Therefore what we do is simply find the marginal probabilities for each role and run five distinct regressions. We note here that roles do not have the same number of champions. In particular, the support and adc roles have far less than other roles as we will see in the empirical results section. Therefore, it is important to accommodate for this since champions of those two roles likely have inflated pick-rates which do not depend on their win-rate or attractiveness but merely because the competition is weaker.

Our model makes two assumptions: One, that players do not change roles, and two, that champions only have one role. Neither of these are true. However, the first is not a problem because in a setting of large numbers, such as this game, the model is consistent since two players that play half top and half mid is the same as two players, one of whom playing top and the other playing mid. The second problem is harder to work with. However, in our data we have the pick-rates of champion organized by their roles, so the issue can be resolved within the model.

In the above model we use the term $\phi(w_{it}, \chi_i)$. This is to capture two implicit assumptions we make about win-rate and χ_i, attractiveness:[2]

1. *Ceteris paribus*, players prefer high attractiveness champions.

2. The more attractive a champion is, the less win-rate impacts pick-rate.

Thus, without including attractiveness we are sure to run into endogeneity issue if our hypothesis is true. However, we do not directly observe attractiveness.

To address this, we use techniques such as fixed effect models to capture certain scores which a-priori should be highly correlated with attractiveness. After that we will run a factor analysis which takes all the similarity between those scores that are supposedly caused by attractiveness. We then use this

[2]Recall that our model is: $U_{ij} = \phi(w_{it}, \chi_i) + \beta X_{it} + \epsilon_{it}$

attractiveness score and approximate the best fit of $\phi(w_{it}, \chi_i)$. To do so, we will use the following scores:

1. Fixed effects of champion over all periods.
2. Fixed effects of champion among those of same roles.
3. A variable coefficient on champion specified change of win-rate. That is, we find the change of win-rate from last period and run the regression with it as a variable, noting the coefficient.
4. Proportion of champion pick-rate and win-rate over average pick and win. To do so, we use $\frac{\frac{p_{it}}{\mu(p)}}{\frac{w_{it}}{\mu(w)}}$ and take the average for each champion.

5 Model

We run five separate regressions on the marginal intended pick-rates. Let mp_{it} be the marginal probability of a champion being the intended pick given a certain role. Our model for that role, is:

$$U_{ij} = \phi(w_{it}, \chi_i) + \beta X_{it} \;\; if\ j\ is\ assigned\ the\ role\ of\ j,\ 0\ if\ not.$$

Then, we have:

$$ln(mp_{it}) - ln(mp_0) = \phi(w_{it} - w_0, \chi_i - \chi_0) + \beta(X_{it} - X_0) + \epsilon_{it}$$

Where:

$$\phi(w, \chi) = \alpha_1 w + \alpha_2 w\chi + \alpha_3 w^2\chi + \alpha_4 w^2\chi^2 + \alpha_5 w\chi^2 + \alpha_6\chi$$

Although it is possible this may not be enough to approximate the true function, we still choose this form because we can only interpret appropriately coefficients for those degrees.

We note that in our actual regression, we include only professional play in X_{it}. We do not want to include ban-rate because that it implicitly used to derive mp_{it}. Furthermore, we do not want to use other game-specific champion characteristics since those are related to attractiveness which may take away explanatory power. Additionally, the coefficient on professional play is the only one that is statistically significant for all roles other than win-rate.

6 Data

Table 1: Summary Statistics on Champions

Variable	Mean	Std. Dev.	Min	Max
Win-rate	0.500	0.024	0.403	0.616
Pick-rate	0.071	0.063	0.0007	0.379
Ban-rate	0.045	0.087	0.0005	0.708
Popularity	0.117	0.224	0	1
N = 662				

We collected data manually from the popular website *op.gg* and *leaguepedia.com*. Our data measures the win-rate, pick-rate, and professional play-rate of all champions over five patches (i.e., periods). We then add dummy variables for champion roles and other variables based on in game information.

Our data also only consists of the results of all games played by the top 10% of all ranked players. This has both drawbacks and advantages. One advantage is that highly ranked players usually possess or own all the champions they wish to play, whereas beginners often do not own and therefore cannot pick a champion that is not their first pick. This problem does not arise in our data. Furthermore, given their rank, we expect high-ranked players to optimize well and place a higher importance on win-rate. This is good in the sense that win-rate will likely be statistically and economically significant, but this will not be representative of the whole population. Lastly, in highly ranked games, champions are likely to be played to their full potential. Therefore, win-rate is more likely to be an accurate representation of how strong a champion is.

Looking at the data, the pick-rate and win-rate differs wildly among champion which gives us good differences to measure differential impacts. Another important aspect of our data is that we do not expect any endogeneity issues. Our win-rate variable, which is roughly the equivalent of price in the standard literature, is not correlated with the error. Price is related to unobserved quality of a product but here we assume that the fun-factor (i.e., quality) is not related to win-rate, although their impacts on pick-rate is certainly intertwined. Therefore, problems due to endogeneity will not arise. To verify this we collected data on in-game information and found that the probability of being larger than χ^2 is 0.000 for all five roles. This gives us further reason to believe that endogeneity will not confound our results.

7 Empirical Results

We now report the empirical results we obtain from the three specifications and also report and interpret the results we obtain from our model.

It seems the ban-rate and exclusivity-of-pick modifications do not change our data by much since ban-rates and overlapping picks rarely happen. The role specification, however, is very significant as some roles have far fewer champions and so champions have higher pick-rates. Separating roles allows us to isolate the heterogeneity between them. Lastly we obtain very reasonable attractiveness scores. Indeed, these scores relate strongly with our four hypothesis variables and also satisfy the intuitive nature of attractiveness of champions.

7.1 Specification Results

Table 2: Intended & Observed Pick Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Intended	0.071	0.058	0.0008	0.344
Observed	0.071	0.063	0.0007	0.379
Difference	2.06e-09	0.010	-0.007	0.102
N = 662				

We first observe that when adjusting for ban-rates, the intended pick differs slightly from the observed pick. Namely we see the skew is not as “nicely” distributed. Going over our data more carefully, we see that champions with high ban-rates and observed pick-rates tend to have higher intended pick-rates which is reasonable given the context of the game.

Looking at the table, the results do not change too significantly, but this is as expected. The chances of a favored champion being taken away is quite rare and most champions have a very low ban-rate. Therefore only champions with high ban-rates and high pick-rates are changed.

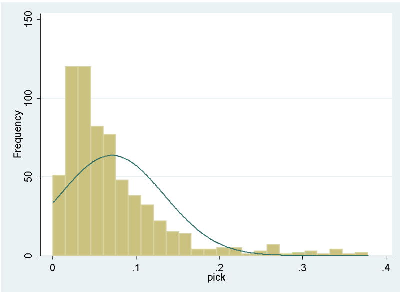


Figure 1: Observed Pick-Rate

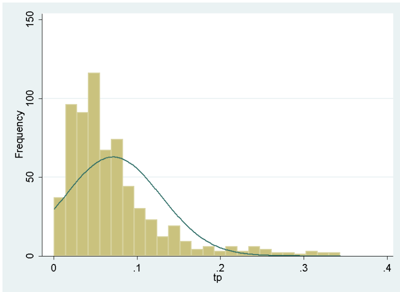


Figure 2: Intended Pick-Rate

We now discuss the resulting estimates of marginal probabilities. We used our intended pick-rates to generate estimates of them instead of the observed pick-rate. Note here that the jungle pick-rate is not a multiple of 5 despite us having 5 periods because a new jungle champion was introduced during the fourth period. Here, the marginal probability is defined as the probability of a champion being the intended pick of a player whose role is the same as the champion's, with both in the same period.

Table 3: Marginal Pick-Rates by Champion Type

Variable	Mean	Std. Dev.	Min	Max	Num. Obs.
Top	0.028	0.017	0.002	0.099	180
Jungle	0.033	0.024	0.005	0.163	152
Mid	0.030	0.018	0.003	0.0869	165
ADC	0.063	0.053	0.0004	0.186	80
Support	0.059	0.034	0.018	0.143	85

It is clear that it is necessary to separate the champions by their roles.

For example, as the ADC and Support roles have far fewer champions, we can expect each champion to have higher pick-rate for those roles. One could try to capture this with a dummy variable, but doing so did not yield reasonable results in our regression. This will nevertheless help us account for heterogeneity of the champions which stems from role differences.

Finally, we examine attractiveness results. We generate four variables:

1. *fixeall*: fixed effect of champion over all periods.

2. *fixedrole*: fixed effect of champions among their roles over all periods.

3. *cwstat*: Coefficient of change of win-rate for a champion in a regression on $ln(tp)$

4. *impdiff* The average elasticity of pick-rates as win-rate changes.

Using the command *alpha* in Stata, we find that the Cronbach's alpha for those variables is 0.8914 which indicates that there should be an underlying variable explaining the covariance relationships between our four variables. Furthermore, we will see that *cwstat* is not as strongly correlated with the other three variables as they are with each other. We then run the $factor$ command in Stata to find this variable (i.e., attractiveness).

Table 4: Factor Results

	Eigenvalue	Proportion
Factor1	2.836	0.991

There are other factors, but this one explains 99.1% of the variance in our four variables. Furthermore, the eigenvalue indicates the percentage of variance explained. Our factor now relates to our four variables in the following way:

Here, uniqueness represents the proportion of common variance not associated with the factor. This makes sense given that we observed earlier that *cwstat* is not strongly correlated with the other three.

We take Factor1 to represent attractiveness and we see that champions that are usually considered unattractive are given a low scores while those

Table 5: Factor Results

Variable	Uniqueness
fixedall	0.0244
fixedrole	0.062
cwstat	0.4643
impdiff	0.1543

that are generally attractive are given a high scores. This confirms our initial expectations.

7.2 Model Results

We now analyze the impact of the generated attractiveness score. We first note that this value seems to correspond with the general perceptions of most players. That is, champions which are typically considered to be fun are given a high score. Before we move on to the results, we would like to note that running robust versions only lead to modest changes, so outliers do not appear to be very significant. Furthermore, running an IV regression with champion characteristics as dummy variables shows that attractiveness is not related with the error term. We therefore conclude that endogeneity is also not an issue in our model.

We first run a naive regression on each role using the following model and obtaining the results that follow:

$$ln(mp_{it}) - ln(mp_0) = \alpha_1(w_{it} - w_0) + \alpha_2(\chi_i - \chi_0) + \beta(X_i - X_0)$$

Table 6: Role-wise Naive Regression

Variable	Top	Jungle	Mid	ADC	Support
Winrate	3.01***	3.28***	3.31***	7.55***	3.34***
	(0.68)	(1.08)	(0.64)	(1.82)	(0.94)
Attractiveness	6.07***	5.81***	5.978**	5.73***	5.04***
	(0.16)	(0.22)	(0.13)	(0.25)	(0.17)
Pro-Play	0.62***	0.40***	0.29***	0.49**	0.49***
	(0.09)	(0.09)	(0.07)	(0.19)	(0.07)
Constant	-2.93***	-2.79***	-2.88***	-2.74***	-2.53***
	(.12)	(.09)	(.08)	(.13)	(.09)
Observations	180	152	165	80	85
R-Squared	0.9062	0.8430	0.9300	0.9327	0.9203

Standard errors are displayed in parentheses. Significantly different from zero at 99 (***), 95 (**), 90 (*) % confidence.

Here the results are more or less similar for each role. We see that win-rate and attractiveness are both economically and statistically significant in each regression. The major differences appear to be the ADC role's high coefficient for win-rate and the relatively low R-Squared for the jungle role.

Before analyzing the potential interaction between attractiveness and win-rate, we first present the variables so that interpretation is clear.

Table 7: Win & Attractiveness Summary

Variable	Mean	Std. Dev.	Min	Max
Attractiveness	0.587	0.173	0	1
$w_{it} - w_0$	-0.001	0.035	-0.121	0.136

Here, we take the differences of attractiveness from the outside option and normalize it for easier interpretation. For win-rates, however, we do not normalize.

We would like to run a regression with $w_{it}\chi_i$, $w_{it}^2\chi_i$, $w_{it}\chi_i^2$, and $w_{it}^2\chi_i^2$ added for interaction. This choice is motivated by interpretation since cubed terms have no real interpretation here. However, running the full regression yields many insignificant terms. Therefore, it is likely not the best specification for our model. We run instead a stepwise regression, removing all terms with $p > 0.20$.

Table 8: Stepwise Interaction Regression

Variable	Top	Jungle	Mid	ADC	Support
Winrate	-41.79*	34.93	-7.36		
	(24.08)	(24.69)	(5.27)		
Attractiveness	6.25***	6.03***	6.71***	6.16***	4.94***
	(0.23)	(0.26)	(0.38)	(0.29)	(0.17)
a1w1	130.190**	-136.85		41.26***	6.84***
	(62.98)	(96.86)		(10.83)	(1.85)
a1w2		129.17	-382.69**		
		(79.69)	(191.8)		
a2w1	-90.42**	146.15	37.28*	-48.58**	
	(41.24)	(91.75)	(17.30)	(16.95)	
a2w2	-43.75**		727.61**		
	(19.93)		(363.55)		
Pro-Play	0.60***	0.36***	0.27***	0.50**	0.50***
	(0.10)	(0.099)	(0.07)	(0.19)	(0.07)
Constant	-3.06***	-2.89***	-3.27***	-2.94	-2.49***
	(0.17)	(0.11)	(0.09)	(0.20)	(0.09)
Observations	180	152	165	80	85
R-Squared	0.9113	0.8496	0.9320	0.9350	0.9211

Standard errors are displayed in parentheses. Significantly different from zero at 99 (***), 95 (**), 90 (*) % confidence.

There are several points that we must clarify here. First, there are differing thoughts on the effectiveness of using a stepwise regression for model construction. Sometimes the results it produces are unintuitive, and in our case, one may believe this since the model drops the win-rate variable in two roles. However, we will show that the effect of the win-rate is still very well captured in the model.

We have two main reasons for using stepwise regression. First, we need to test our hypothesis that the interaction terms for the variables are not significant. Therefore, it was necessary to take some of them out of the model. Secondly, it is not clear which terms to remove as the interaction terms may depend on specific roles. Thus we chose to use a stepwise regression as it is an unbiased way of choosing an appropriate model given that roles are different.

Another issue is robustness. To investigate this, we ran a stepwise regression with robustness as well and only ADC and Mid have non-trivial changes.

In our earlier hypothesis, we assume that win-rate and attractiveness both have a positive impact on pick-rates. Furthermore, we assume that their cross derivative is negative. We find the partials of ϕ for each role and plug in the observed numbers to obtain the following.

Table 9: Stepwise Interaction Regression

Variable	Top	Jungle	Mid	ADC	Support
Mean of ϕ_w	2.93	3.89	3.05	4.77	3.39
Mean of ϕ_χ	5.99	6.31	5.95	5.39	5.04
Mean of $\phi_{w\chi}$	-10.65	-7.66	5.03	-17.5	6.85
Observations	180	152	165	80	85

We see that our hypothesis regarding variables' impact on pick-rate are reasonably confirmed by the data. Further, notice that the partials are similar to the coefficients we obtained in the naive regression. Thus, our predictions are similar to the naive regression, except this model also explains the interaction between win-rate and attractiveness.

The main outliers in the model above are $\phi_{w\chi}$ for Mid and $\phi_{w\chi}$ for support. In the case of support, its difference likely stems from the model we are using. Indeed, all coefficients are positive and we have only non-squared terms that made it past the stepwise regression. For the mid role, the problem is more difficult to explain. Currently we do not have a reasonable explanation for this difference.

8 Further Issues

There are a couple of modifications we could make to investigate this player choice further. Furthermore, there is also some data which we would have liked to use but could not obtain.

First, separating the market into roles allowed us to gain insight into variation between roles. We did not, however, include cross nested champions by separating them into different objects. Instead, we simply included them in their most played role. This may have led to small inaccuracies. However, we do expect these to be small as the majority of champions only have a single role and even multi-role champions are usually played in one particular role.[3]

There are a few issues we did not account for during our three stage specification. First, we assume that roles are picked equally often. While this is true in terms of games, it is not true in terms of queues. Each game requires two of each role, but there are unequal number of players who wish to play each role. This often results in longer queue times for more popular roles. However, there is no way to tell which are more popular using our data. Second, although we could have computed the true intended pick by solving the corresponding equations, we did not do this which could lead to small inaccuracies.

If we had more data on the players' queue statistics, we likely would have a more accurate representation of pick-rates. Furthermore, it would have been good to have data on the entire playerbase, not just the top 10% of ranked players. However, due to both the lack of availability of this data online and the time it would take to collect it, for the purposes of this paper this was infeasible.

One point of interest is whether we can determine the factors of the attractiveness score. To investigate this, we ran a regression with 7 variables on gameplay data measuring champions, such as their abilities and in-game statistics. Although some coefficients are statistically significant, we did not obtain a high R-Squared and thus do not report the results here.

[3] This occurs in about 70% of these cases.

9 Conclusion

In this paper we investigate the effect of the win-factor and fun-factor of a champion on its pick-rate. The intuitive assumptions are that champions which are more fun and win more tend to be picked more. Further, it is likely that a champion which is more likely to be picked for fun has its pick-rate less influenced by the win-factor.

We find that players do value both the fun-factor and win-factor positively. We also see that the impact of both the win-rate and attractiveness are highly significant. Furthermore, players generally value the win factor less when the fun factor is higher. Therefore, most of our assumptions are verified. Lastly, we see that although professional play impacts the pick-rate positively and is statistically significant, its impact is rather negligible compared to win-rate and attractiveness. The main differences from our approach to the traditional literature in IO are the use of factor analysis to find the unobserved fun factor and the use of fixed-nests model.

In conclusion, our analysis suggests that players generally choose champions in a rational manner. As indicated by a high R-Squared, champions' utility depends mostly on their effect on the win-rate and their fun-factor. However, our model also suggests that this relationship is complicated, and intricate tradeoffs exist between the two.

References

Conley, Kevin, and Daniel Perry. 2013. "How does he saw me? A Recommendation Engine for Picking Heroes in Dota 2."

Kumar, Vineet. 2014. "Making 'Freemium' Work: Many Start-ups Fail to Recognize the Challenges of This Popular Business Model." *Harvard Business Review*, 27–29.

Semenov, Aleksandr, Peter Romov, Kirill Neklyudov, Daniil Yashkov, and Daniil Kireev. 2016. "Applications of Machine Learning in Dota 2: Literature Review and Practical Knowledge Sharing."

Yong, An Gie, and Sean Pearce. 2013. "A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis." *Tutorials in Quantitative Methods for Psychology*, 9(2): 79–94.

Contributors' Spotlight

Ujjwal Dahuja graduated with highest honors in Economics from Princeton University in 2017. He is interested in financial research using econometric tools.

Matthew Hong is a graduating economics student with minors in math and statistics. He is interested in investigating returns to education, incentives of students to pursue further education and structural/behavioural estimation of educational attainment.

Ayush Gupta is a graduating student completing a specialist in Mathematics and Economics. He will be starting his Masters in Economics at the University of Toronto in the fall of 2017. His research interests include finance and public policy.

Parker Griffin is a graduating student with a double major in International Affairs and Economics with a minor in French. He is interested in the economic potential in West Africa and the continued growth of the regional economy.

En Hua Hu is a graduating student who studied Economics, Mathematics and Philosophy. His interests lie in decision theory and applied logic.

Hon Chiu Yuan is a graduating student who studied Economics and Finance. He is interested in macroeconomics and business cycles.

Editors' Spotlight

Rida Aamer is a second year student at the university. She is a double major in economics and statistics, with interests in development and international trade.

Matthew Hardy is a graduating student in statistics, economics, and mathematics. He is interested in machine learning and econometrics.

Anthony Piruzza is a third year student specializing in Financial Economics. He is interested in macroeconomic policy.

David Rhodes is a second year student, double majoring in economics and public policy. His interests include microeconomics and political economy, especially relating to media and elections.

Ce Shang is a third year student studying Economics and Neuroscience.

Qasim Sheikh is a fourth year student with majors in economics and public policy and a minor in statistics. He hopes to pursue a Master's degree in economics after this year.

Alex Smith is a fourth year student pursuing a double major in Economics and International Relations. He is interested in macroeconomic policy and US foreign policy in the Middle East.

Dario Toman is a fourth year Economics Specialist, focusing on Applied Microeconomics. His research focuses on the Economics of Conflict, in particular the effects of armed violence on human capital accumulation.

Jiada (Kadachi) Ye is a third year student in Economics and Mathematics Specialist. He is interested in microeconomic theory and applied microeconomics.