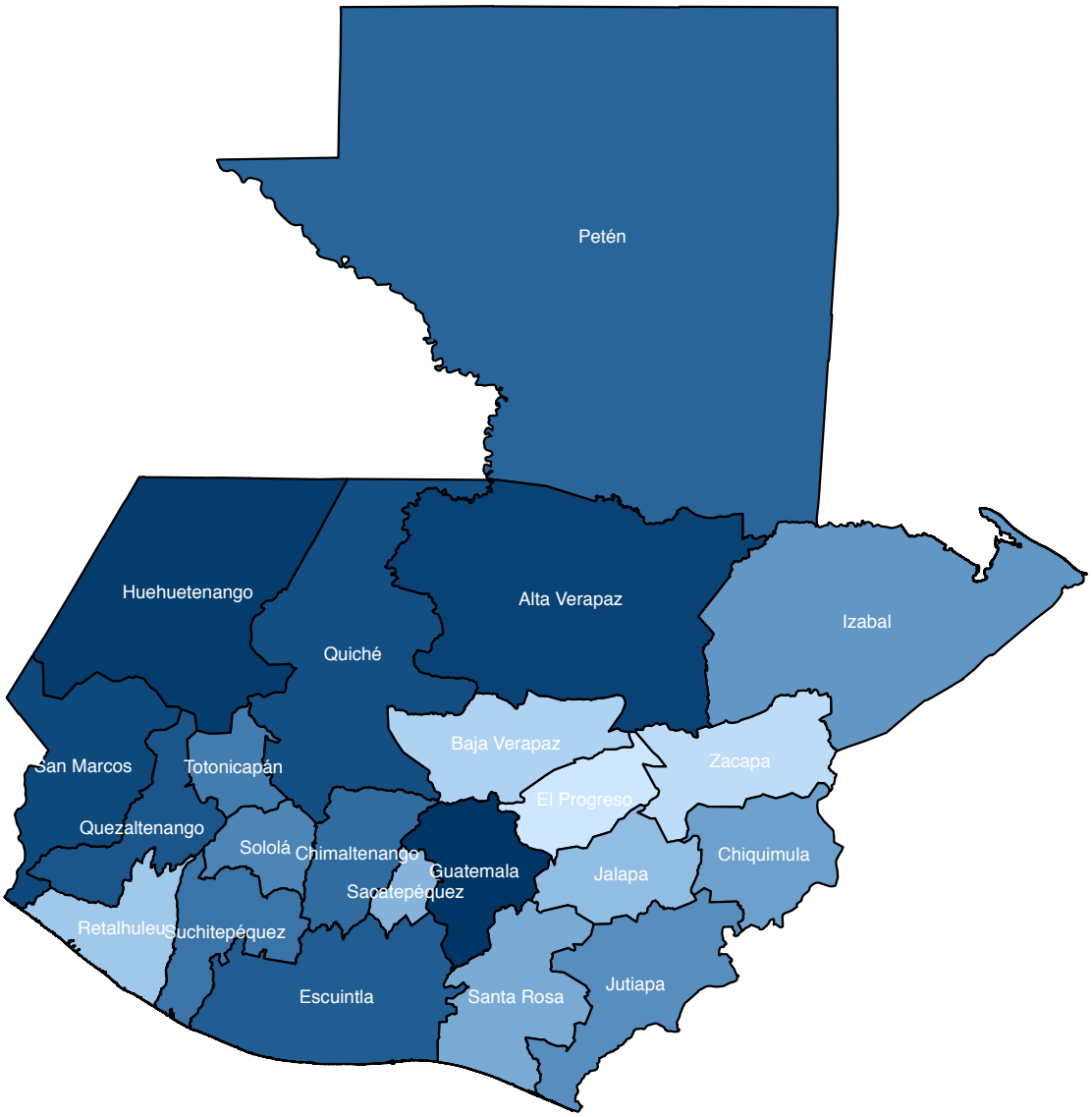
Is World Bank Investment in Guatemala Responsive to Departmental Needs?

Will Sims
May, 2019



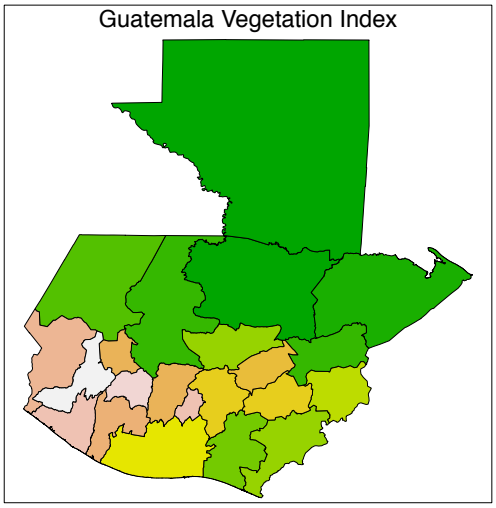
Departments of Guatemala by Population (2015)

Introduction

This paper will examine the distribution of World Bank investment amongst Guatemala's 22 regional Departments. Guatemala is an economically disenfranchised country, with especially high poverty rates amongst rural and indigenous Mayan communities. In response to these acute needs, between 1995 and 2014, the World Bank has invested $5.7 billion in development projects in Guatemala, an average of $300 million per year. This paper will seek to examine the allocation of these funds between Guatemala's 22 departments, examining the degree to which the investment allocation is responsive to observable factors like poverty rates and population. Section 1 will examine descriptive statistics of Guatemala's departments, Section 2 will use these characteristics to build a model of investment distribution, and Section 3 will critique the model and suggest areas for further examination.

Section 1: Guatemala in Context

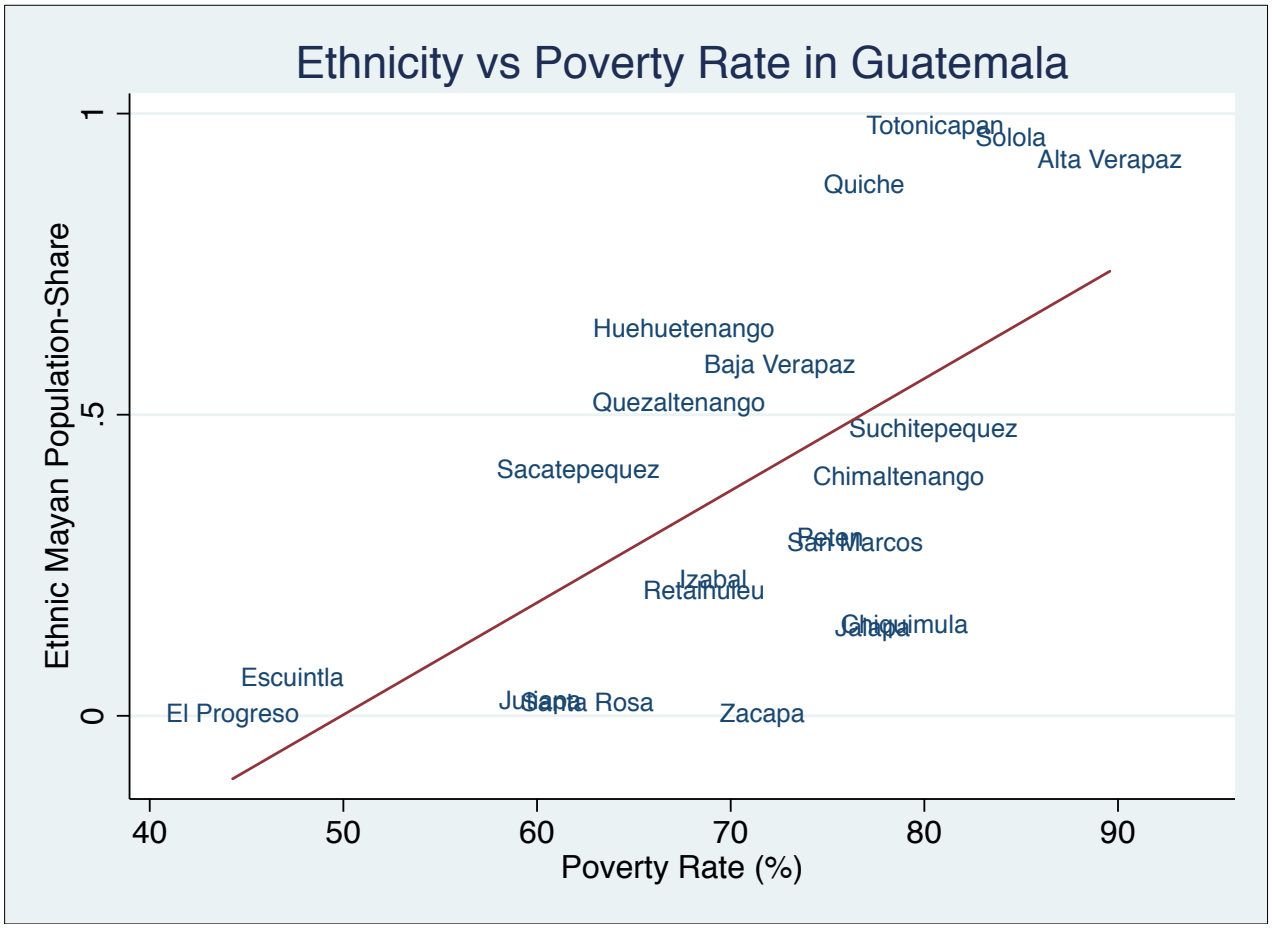
Guatemala is comprised of 22 Departments of varying demographic and geographic magnitudes. Its geography is fairly heterogenous – much of the country is mountainous highlands, with hot dry lowlands along the Pacific coast, and an extensive jungle in the department of Petén to the north. This ecological diversity is reflected in the 2016 vegetation coverage map pictured on the right.



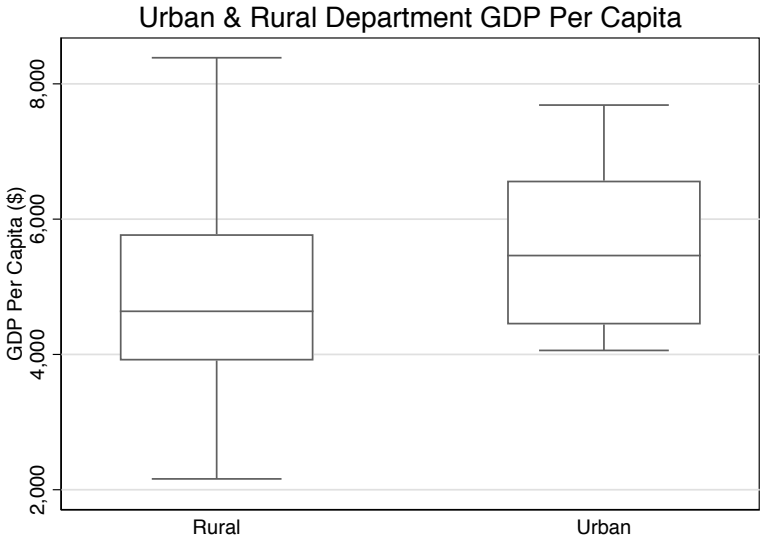
The two most significant dimensions of social and economic fractionalization in Guatemala are ethnicity and urbanization. The major urban centers are the cities of Guatemala City and Quetzaltenango, and much of Guatemala's population is concentrated in the surrounding states of Sacatepéquez and Totonicapán. These cities control much of the economic and political power within the country, while rural areas are often most affected by issues like childhood malnutrition, inadequate sanitation, and lack of economic and educational opportunities. This has created geographic stratification within Guatemala, as many young Guatemalans migrate to the cities, and those left in rural areas face increased hardship due to human capital flight and growing economic isolation.

In terms of ethnicity, the vast majority of Guatemalans are of indigenous Mayan descent, but much of the economic and political power in the country is concentrated in the hands of *ladinos* – Guatemalans of Spanish or mixed-Spanish ancestry. This division was a significant factor in the decades-long Guatemalan Civil War, which pitted a primarily ladino government and army against indigenous guerrilla forces in the rural highlands. The effects of this conflict reverberate to this day, as indigenous Guatemalans continue to face both overt and systemic discrimination throughout society.

As a result, both urbanization and ethnicity are highly correlated with poverty. Guatemala is one of the poorest countries in the western hemisphere, with poverty rates reaching 90% in some departments. As we can see in the scatterplot, there is a strong relationship between poverty rates and Mayan share of population. A simple regression on these variables shows that Mayan ethnicity can explain 40% of the variation in poverty rates in Guatemala. A 5% increase in Mayan population share is associated with a more than 1% increase in the poverty rate, and we would expect departments like Sololá or Totonicapán that are almost entirely Mayan (96% and 98%, respectively) to have approximately 21% higher shares of their population living in poverty compared to departments like Zacapa or El Progreso that are almost entirely ladino (both <1% indigenous Mayan).

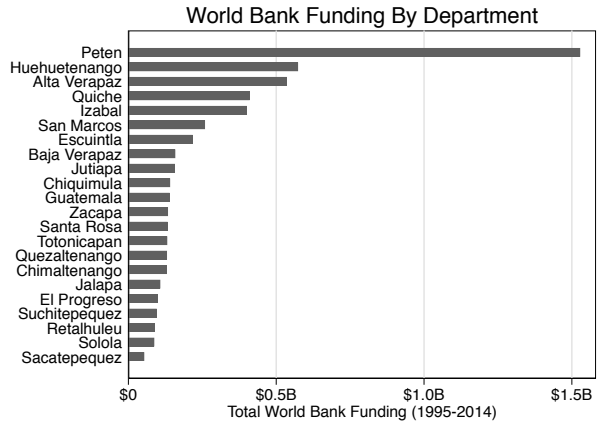


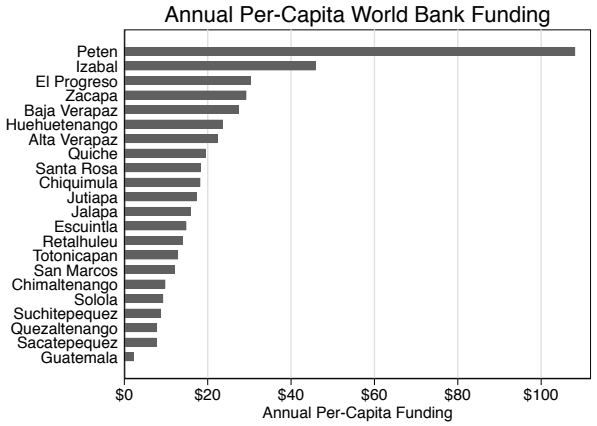
In terms of urbanization, we can compare departments that have more than one thousand people per square mile (a US-census benchmark for urbanized areas), with more rural departments. Comparing the average per-capita GDP of these regions, we find that urban areas have 15% higher levels of per-capita economic activity in rural and urban areas, however this difference does not achieve statistical significance, likely due to the small sample size. Looking at the box and whisker plot of this result, one can note both the noticeable difference in means, as well as the significantly higher range for the rural departments. This is somewhat unsurprising, as we would expect a higher degree of variation in less populous areas (see Howard Wainer on De Moivre's equation). Therefore, even though there is no statistically significant difference in the means of the two groups, it may be useful to include measures of urbanization in our model. In the next section, I will examine whether these factors of ethnicity and urbanization have measurable impacts on the allocation of World Bank investment.



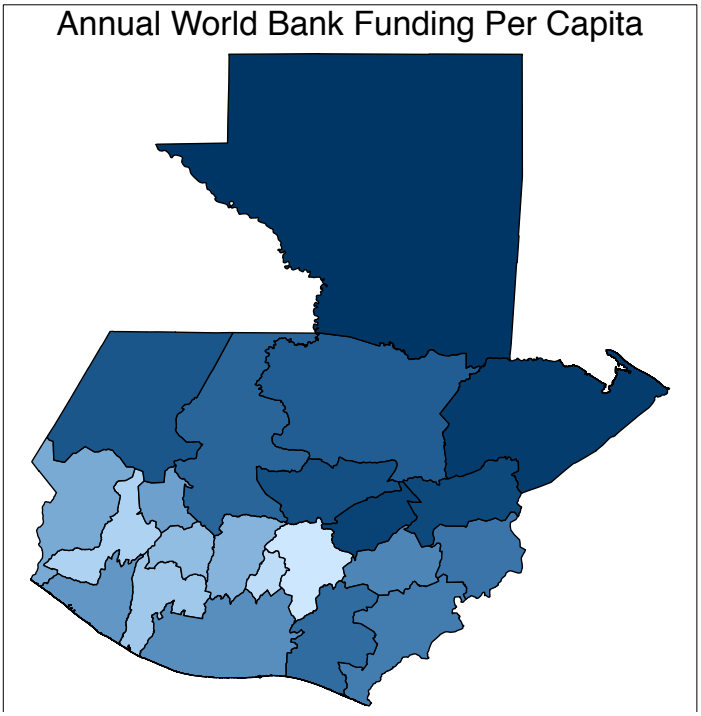
Section 2: Modeling World Bank Aid Allocation

In order to combat this widespread poverty the World Bank has invested significantly in Guatemala, but this investment has not been evenly distributed amongst departments. For instance, if we look at the bar graphs below, we can see wide divergence in World Bank investment in both aggregate and per-capita terms:





A few differences stand out between the two plots. First, Guatemala City (simply 'Guatemala' in the graphs) drops from middle-of-the-pack in total funding to the single lowest recipient in terms of per-capita funds due to its large population. Other departments, such as El Progreso and Zacapa, receive relatively more funding on a per-capita basis. Also notable are the similarities – Petén wildly outstrips all other departments in terms of both total and per-capita funds. The likely explanation is that Petén contains significant archeological and ecological treasures, including the Tikal ruins and the Maya Biosphere Reserve, which are direct drivers of World Bank investment (e.g. Community Management of the Bio-Itza Reserve). The map below shows the geographic distribution of per-capita investment, including its relative concentration in the northeastern area of the country.



Integrating the previously-discussed department-level characteristics, we can create a regression model that attempts to explain the distribution of World Bank investment. First, in order to get a sense of the fundamental relationships, I looked at a series of two-way regressions without controlling for other confounding variables:

World Bank Funding: Two-Way Regression Results				
Variable	Coefficient	Standard Error	P-Value	Covariance
Department GDP	-3,942	14,071	0.78	-0.0625
Population Density (2015)	-288,742	213,580	0.19	-0.2894
Indigenous Percent	1.18E+08	2.15E+08	0.59	0.1216
Population (2015)	64.9	105.3	0.54	0.1366

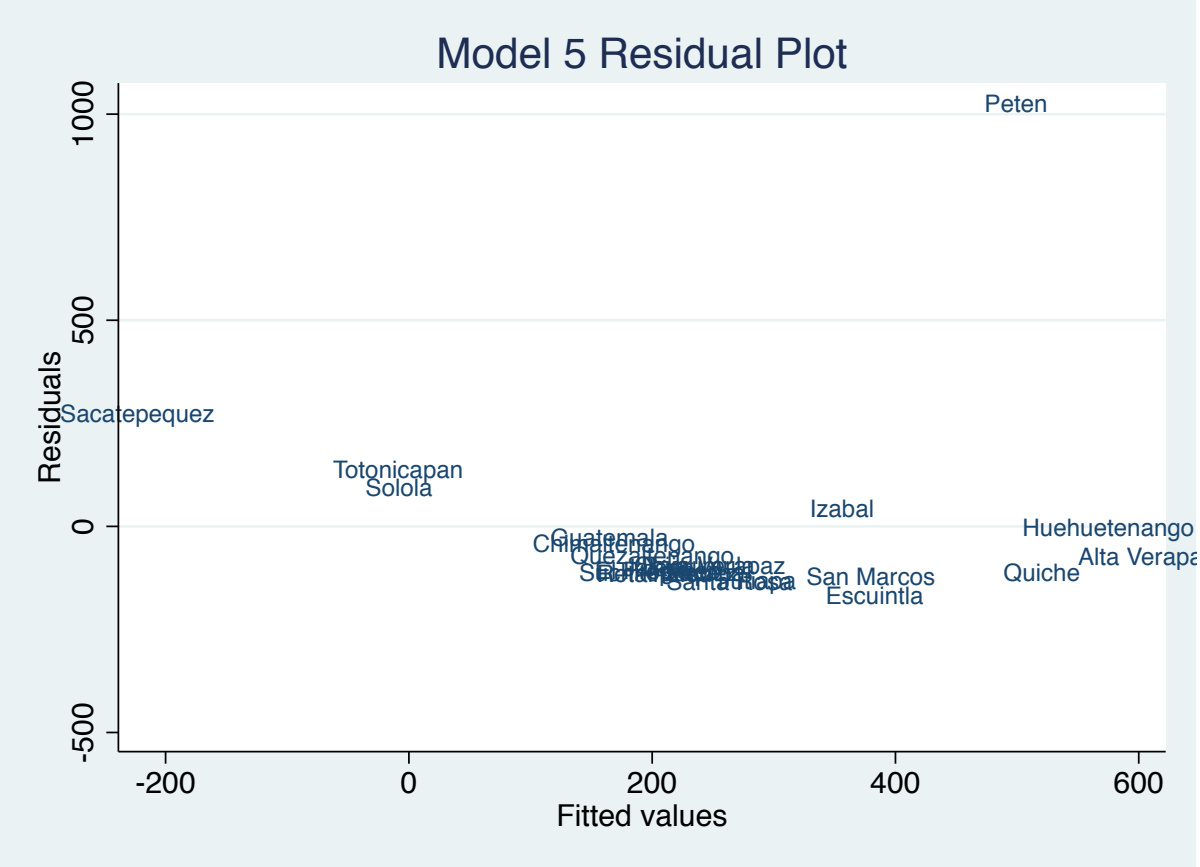
While the p-values are not significant, these results suggest that World Bank invests more in departments that have less economic activity, are more rural, have more indigenous population, and have higher population. All these are promising results, as we would like the World Bank to target investment towards departments that have these characteristics. That said, if we look at the R-squared values the correlations are all extremely weak. To attempt to address this issue, we can run a series of regressions that control for these various independent variables:

World Bank Funding: Regression Model Comparison						
	Model 1		**Model 2**		**Model 3**	
	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Department GDP	0.0627*	.029	-0.0491	.544	0.0568	.354
Population Density (2015)	-1.162*	.011	-0.612	.275	-1.176*	.019
Indigenous Percent	346.4	.115	-85.13	.812	266.6	.370
Population (2015)			0.00063	.155		
Poverty Percent					4.020	.624
Intercept	223.4	.059	184.5	.114	-12.10	.984
R-Squared (Adjusted)	**.318 (.205)**		**.397 (.255)**		**.327 (.159)**	

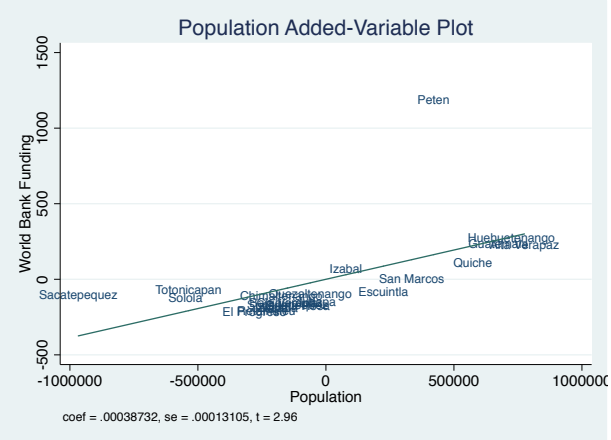
	Model 4		**Model 5**		**Combined Model**	
	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Department GDP					-0.1116	.391
Population Density (2015)	-0.899**	.005	-0.900**	.004	-0.6001	.317
Indigenous Percent	101.9	.579			-114.12	.766
Population (2015)	0.000382*	.010	0.000387**	.008	0.0008	.156
Poverty Percent					-1.145	.894
Intercept	194.2	.088	229.7*	.015	365.36	.563
R-Squared (Adjusted)	**.383 (.280)**		**.372 (.306)**		**.327 (.159)**	

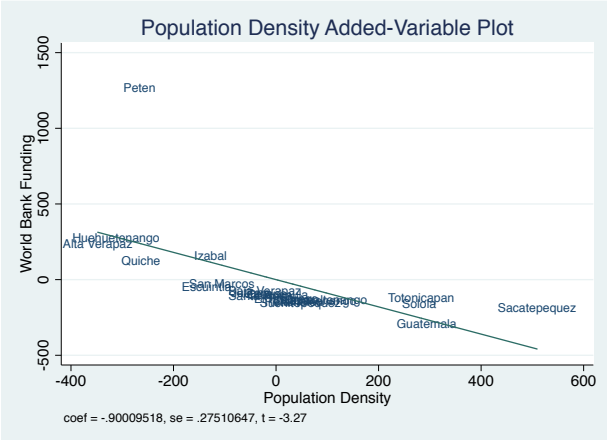
Given the low number of observations, we cannot afford to be overly-strict in terms of p-values. Even so, only Model 1, Model 4, and Model 5 have multiple significant coefficients. Of these, Model 5 has the highest adjusted R-squared value, and the most highly-significant coefficients. The independent variables in Model 5 are also the only coefficients that don't switch signs based on the model specifications, suggesting that they are the most consistent relationships. While GDP is statistically significant in Model 1, it is likely directly correlated with population, which is a better fit according to the results of Model 5. Thus we can feel comfortable using Model 5 as the most parsimonious and accurate description of World Bank Aid allocation in Guatemala.

Focusing on Model 5, we can check the post-estimation graphs to ensure that the Model is a good fit for the data. First, we can check the residual plot to see if there is heteroskedasticity in the values or any concerning outliers:

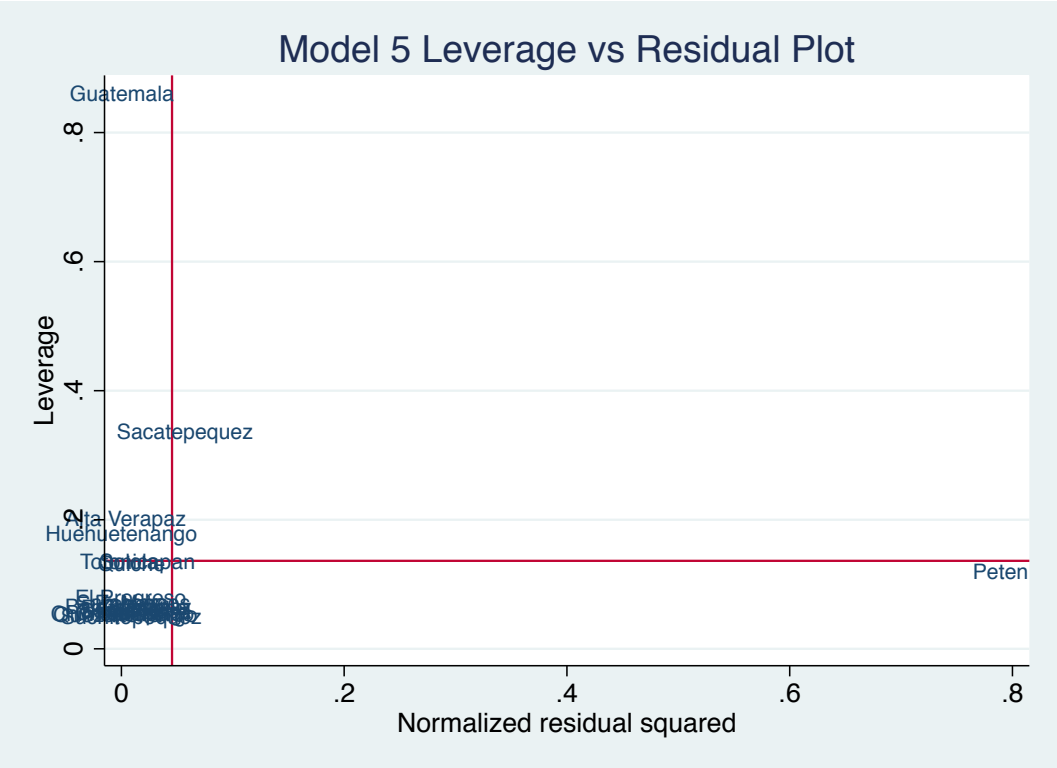


There doesn't appear to be significant widespread heteroskedasticity outside of Petén, which is a significant outlier that appears to shift the overall high end of the model. If we break it down by specific variables by looking at their added-variable plots, it looks like Petén is a strong outlier across both of the independent variables. Sacatepéquez also appears to be a moderate outlier, but otherwise all departments seem to be a good fit across the other covariates.





To check if these outliers are concerning, we look a the Leverage vs Residual plot:



The Leverage vs Residual Plot is encouraging, as it appears that despite Petén's large residual value, it has relatively low leverage. Sacatepéquez has moderate leverage, but its normalized residual value is not excessive. The department of Guatemala has huge leverage, because it has a massive population and is also very dense. Overall, while the size of the dataset is limited and there are definitely some departments that are worth keeping an eye on due to their high leverage or residuals, we can feel confident that the post-estimation has not revealed any deeply disqualifying issues with Model 5, and we can use this model to make claims about World Bank funding behavior.

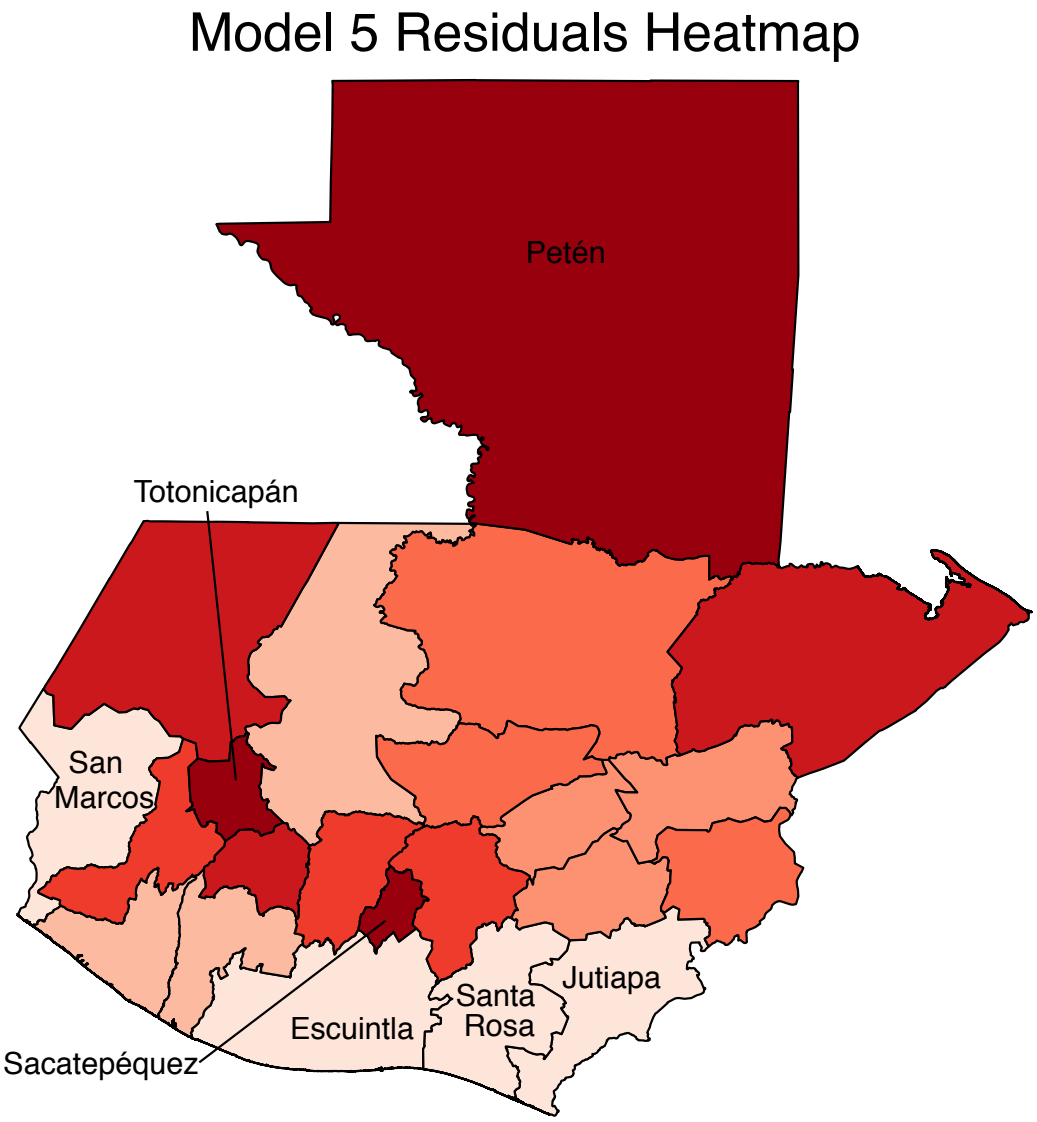
Section 3: Interpretation and Discussion

According to Model 5, there is a positive relationship between World Bank funding and population, and a negative relationship with population density. This is encouraging, as it suggests that the World Bank targets investment towards departments with larger populations to serve, and that it also invests more in departments that have lower population densities. In general, it makes sense that if the World Bank wants to maximize their impact on humanitarian outcomes, they should focus on departments with larger numbers of people. Similarly, given the widespread economic disenfranchisement of large areas of rural Guatemala, it makes sense that they would target departments with lower average population density. At least in terms of these measures, we can say that World Bank investment is appropriately targeted based on need.

In terms of the magnitudes, our model predicts that above a baseline funding level of $230MM per department, each additional person is associated with a $382 increase in World Bank investment over the period of interest, or approximately $20.11 per year. This is very close to the aggregate funding per capita of $21.60, suggesting that population is an important but not disproportionate factor in funding allocation. In terms of population density, the coefficient suggests that for each density increase of 0.9 persons per square mile, departments receive $1MM less in funding.

However there is also reason to be critical of the variables that did not appear significant in our model. Specifically, we would hope that the World Bank would also be responsive to both poverty and historically-marginalized populations like indigenous Mayans. The fact that neither of these variables proved significant in our analysis suggests that the World Bank is not adequately targeting investment towards departments with high rates of poverty and significant Mayan populations. While this is certainly better than targeting investment away from these areas, the lack of a relationship is nevertheless troubling.

We can use a heatmap to look specifically at the departments that have high residuals in our model, either because they have unexpectedly high or low levels of World Bank investment. In the graph below, shades of darker red indicate areas with funding levels that exceed their predicted values, while light areas indicate under-funding relative to our model.



Looking at this data geographically can illuminate some interesting patterns. In terms of underfunded areas, the southern coast seems to jump off the map – particularly the southernmost departments of Escuintla, Santa Rosa, and Jutiapa, as well as San Marcos on the Mexican border. One possible explanation is that these areas along the coast include some of Guatemala's most common

drug-trafficking routes, and are consequently subject to elevated levels of violence and gang activity[1]. It is understandable that given these risks to personal safety and security, World Bank staff would avoid extensive work in these areas. While I was unable to get Department-level crime statistics, including this data in the model would be a clear next-step for analysis that could allow a better understanding of World Bank funding decisions.

In terms of the high-funding areas, there are three distinct departments that stand out: Petén in the north, as well as Totonicapán and Sacatepéquez in the central highlands. Petén's unique characteristics have been discussed earlier in the paper – specifically their extensive rainforests and ancient Mayan ruins. To verify this hypothesis, it would be necessary to look at World Bank funds on a project-level to see if projects coded as related to environmental sustainability or cultural heritage are concentrated in Petén.

In terms of the other two departments, while such analysis is admittedly speculative, I can't help but notice that they are both smaller departments close to major urban hubs: Sacatepéquez is adjacent to Guatemala City, and Totonicapán is an exurb of Quetzaltenango/Xela, Guatemala's second largest city. One possibility is that the staff of the World Bank and associated partner organizations are centered in these large urban areas, so there is a natural inertia towards evaluating projects in nearby areas. The high populations of Quetzaltenango and Guatemala City could offset this effect in our model, but our model could underestimate the spillover effects into nearby yet less populated departments. In order to test this theory, we would want to look at where the offices of World Bank staff and affiliates are located in order to see if there are indeed concentrations in urban areas that could help explain these "spillover" effects.

Another possible theory is that projects are focused on developing Guatemala's tourism industry, and so they are concentrated around tourist hotspots like the Tikal ruins in Petén and the town of Antigua in Sacatepéquez. While this theory could also help explain the slightly elevated residual values in other touristic departments like Sololá and Izabal, it cannot account for the high funding levels in Totonicapán, which lacks any notable tourist infrastructure. In either case, it would be worthwhile to look at tourism-related projects as part of the project-category related analysis mentioned above.

Conclusion

While there is some reason to be encouraged by the correlation between World Bank investment and general demographic characteristics, there are some unfortunate gaps that point to an unwillingness to target investment towards historically or systematically disenfranchised and poverty-stricken populations. I do not wish to claim that these inequities are the result of conscious decisions amongst World Bank staff, but given the inherent and understandable biases towards working in regions that are easier to access and have more developed infrastructure, it is necessary to consciously combat these biases in order to ensure that World Bank investment is directed towards the areas with the highest need – those living in poverty and those whose ethnicity precludes their full participation in Guatemala's ruling power structures. Clear next steps in this analysis involve a deep-dive into project-specific data in order to identify if the types of projects – focused on tourism, the environment, or cultural heritage – can explain the remaining variation in the model.

[1] UNODOC: Transnational Organized Crime in Central America and the Caribbean: A Threat Assessment
http://www.unodc.org/documents/data-and-analysis/Studies/TOC_Central_America_and_the_Caribbean_english.pdf