USAID
FROM THE AMERICAN PEOPLE

DETERMINANTS OF MIGRATION

AND PUBLIC POLICY IMPLICATIONS (PHASE 2 ANALYSIS)

For the Creating Economic Opportunities Project

JANUARY 2022

JAN // 2022

This report was produced for the Creating Economic Opportunities Project 0
72052018C000001 for review by the United States Agency for International Development

Prepared by: David M. Kemme & Jorge Benavides

[This Page Left Intentionally Blank]

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION

This analysis identifies determinants of migration from Guatemala at the department and municipal level and illustrates the direct links between those determinants of migration and goals of USAID's Creating Economic Opportunities project. A unique aspect of this study is that the determinants, frequently discussed in literature and by policy makers, are analyzed herein at the Department/Municipal level, particularly those Departments which are in Project geographies. The Project is designed to attract Foreign Direct Assistance (DFDI), mobilize financial resources, and generate jobs, in and beyond Guatemala's capital city, thereby lessening the likelihood of immigration of Guatemalans to the US. This analysis therefore focuses its attention on five departments in the Western Highlands: Quetzaltenango, Huehuetenango, Totonicapán, San Marcos and Quiche, and Guatemala (CEO6, hereinafter). Using an innovative technique to create a proxy for unobservable illegal migration, and a newly constructed panel database on municipal-level determinants, the data reveals, and the study concludes, that Project goals are focused on the single most important determinant of migration: enhancing economic opportunity, as measured by municipal income per capita.[1]

The analysis extends the research of Kemme (2019), which closely examined the determinants of irregular migration from Guatemala through the north, and the links with Public Policy implications. That study was unique because it utilized official Municipal and Department level data rather than surveys. This allowed us to establish a hard link between local income or economic opportunity and immigration. A host of other potential determinants were examined but the linkages were not as robust. It was clear that Project goals are important in reducing immigration because of the associated income created. Using the same innovative technique to create a proxy, a latent variable, for unobservable illegal immigration and a newly created, and now expanded data base of municipal level determinants, the analysis confirmed earlier results and found that other determinants of immigration are also significant. Even so, increasing economic opportunities[2] remains the dominant factor in the decision-making process for Guatemalans.

It is important to highlight that the phenomenon of irregular migration is of relevance to Guatemala's foreign policy and its social and economic relationships with the United States. Migration is not a recent event, but rather has evolved during the past 20 years, but at a much faster rate in the past six to eight years. The decision to migrate is complex and occurs at the intersection of "pull factors" -- predominately economic opportunity in a foreign destination -- and "push factors" -- political and socio-economic conditions in a specific home locale. While the "pull factors" in the US are relatively the same for all potential immigrants, local "push factors" vary.

There has been extensive academic and scholarly research on irregular migration; predominately case studies or survey-based research. These analyses provide important insights, but most have focused on national or international surveys, are typically not long-term longitudinal studies and research questions, countries and regions vary across studies. They are also expensive, time consuming and quickly out of

[1] Note that estimated municipal income per capita is calculated approximately the same way that GDP is calculated but using municipal level data: the sum of consumption expenditures, investment, government expenditures and net exports of goods and service from the municipality to other regions. Project objectives focus heavily on investment, income through jobs creation, increasing labor productivity and financial education. Other determinants of migration are controlled for in the models below, for example violence, drought, and nutrition, but are not the focus of the Project's objectives.

[2] We use two measures of economic opportunity, municipal income per capita as in the first study, but also a different broader measure, that includes measures of poverty, labor force, and graduates as explained in the document.

date[3]. Their conclusions are often highly qualified by the nature of the data. Nonetheless, from this research "push factors", conditions at the local level that greatly influence the decision to migrate, have been identified. These are discussed in Kemme (2019) and briefly below. Their relative importance varies by source country. This study takes these as the starting point in the analysis below.

The growth in irregular migration over the last decade has been a major concern to United States authorities and policies for limiting migration are being reconsidered. The source of migrants has rotated from predominately Mexico to the Northern Triangle of Central America, first from Guatemala and more recently from El Salvador and Honduras. A possible reason for the slowdown from Mexico is that economic development and job creation driven by foreign direct investment (predominately from the US) has increased the standard of living and made migration less attractive. Another reason may be an increase in the use of legal pathways through H2 visa temporary labor opportunities in the US. For Central America the same cost-benefit calculation still leads to higher levels of migration. While the lower standard of living is generally defined by income it is also reflected in socioeconomic conditions, such as health status, violence, housing conditions, and demographics, for the most vulnerable groups.

One issue consistent since the previous report is that socioeconomic conditions continue to evolve and are more quickly reflected in the regularly collected official data rather than irregular surveys. In the previous research, the study built a robust departmental and municipal level data base of official data rather than surveys. Since then, the data base has been updated, adding one more year of observations, and several new variables. This data allows for the use of more sophisticated statistical and econometric methods which then allows for on-going replication and extensions at relatively low cost. An additional benefit of using official municipal level data is that findings for specific determinants can then be used to focus government policy on the most important measurable determinants at the municipal level. Clear metrics for success can be developed and tracked. Comparisons of these determinants across municipalities will allow policy makers to assess the impact of public policies, and as result, local policies can be fine-tuned for more meaningful change.

Finally, this second round of findings leads to the conclusion that it is important to dedicate resources to continuously track the evolution of variables and identify the link between programs and policies to monitor their effect in reducing or slowing the growth of irregular migration. When the focus of the policy is the people that are being displaced from their homes and lands, appropriate and objective analyses and studies should guide policymakers to improve the quality of life of these families and their offspring. The hope for the future is that there will be interest in continued modelling to understand this phenomenon more fully and the impact that it has on thousands of families over time.

[3] A novel approach by Ceballos and Hernandez (2020) uses household level survey data to calculate a household propensity to migrate index. Their final twelve variables constructed from survey questions (collected in 2012, 2013 and 2014) closely parallel some of the determinants we examine below.

II. CHALLENGE: BUILDING A COUNTRY WITH OPPORTUNITIES FOR ALL

What are the costs and benefits considered by Guatemalans when they decide to migrate? Following the expression of deep concerns about the increase in unaccompanied minors attempting to migrate from Guatemala, the Government of the United States implemented an aggressive initiative to reduce the flow of Guatemalans northward, known as the "Plan of the Alliance for Prosperity in the Northern Triangle of Central America". The commitment made by the Governments of Guatemala, Honduras and El Salvador requires improvement, as soon as possible, in the economic and security conditions that lead people to migrate from their place of birth and residence.

After extensive fieldwork in those departments and municipalities in Guatemala from which most migrants originate, researchers focused on the cost-benefit analysis subconsciously made by individuals when they decide to migrate. This decision-making process involves the rough quantification of costs (e.g., loss of home property, distance from family, actual transportation and smuggling costs, dangers on the way[4] and expected insecurity at the place of arrival) and benefits (e.g., opportunities for more education, or potential acquisition of new competencies, substantial increase in income associated with more job opportunities, and overall improvement in living conditions).

When the expected benefits exceed expected costs, migration is undertaken, usually from rural and very impoverished areas to more urban and developed areas.[5] This process does not stop when people migrate to major cities in Guatemala. Instead, the decision-making process is repeated, considering a broader set of opportunities including those outside the country, and not surprisingly, they consider the US as a feasible option[6]. Also, migration is not necessarily a decision made by one isolated individual. Migration from Guatemala (and other countries in the region) appears to implicate more families, children, and youth under 18. This suggests the decision to migrate is a family life-changing decision, not necessarily an individual one-time decision, and this is the reason policy interventions should be tailored accordingly.

Under the U.S. Strategy for Engagement in Central America, the United States is working with Central American governments to promote economic prosperity, improve security, and strengthen governance in the region. The program was launched by the Obama Administration and continued with some modifications by the Trump Administration. Toward this effort, Congress appropriated $2.1 billion in FY2016-2018, doubling the annual aid for the region. The governments of El Salvador, Guatemala and Honduras collectively also allocated about $7.7 billion to the initiative in 2016-2018 (CRS (2019)).

Persistent challenges regarding economic development, health status, violence and crime, and citizen confidence in government and stability, are well known and are being addressed with measurable effectiveness by USAID programs. However, a new set of challenges due to COVID-19 and the most recent impacts from hurricanes ETA and IOTA[7] have arisen. These require not only short-term

[4] The recent tragic truck crash in the south border of Mexico, killing over 50 migrants, makes clear the dangers of irregular migration. For more information, see: https://www.washingtonpost.com/world/49-migrants-dead-dozens-hurt-in-truck-crash-in-south-mexico/2021/12/09/0729dfc2-594a-11ec-8396-5552bef55c3c_story.html

[5] The regression results below suggest that municipalities with GDP per capita in the lower quartile of the distribution may have higher than otherwise predicted immigration even controlling for all other determinants, and idiosyncratic variations in municipalities and migration is less responsive to changes in income than the average municipality.

[6] Note that a well the established family migration network may lead to direct migration to the US rather than the two-stage process as described.

[7] Two strong hurricanes which struck Central America in November 2021, causing damage in central-western Guatemala (Quiche, Huehuetenango)

humanitarian and medical assistance, but also additional interventions and longer-term support for health care systems, institutions, and policies.[8] USAID continues providing support to the communities hardest hit and most vulnerable to future shocks.

In the results below, with the additional data available for this report (pre-COVID), health status is found to be negatively associated with migration. As health improves, in part a reflection of improved economic well-being, migration is less. The effects of COVID-19 are yet to be observed in the data, but the better comprehension of the pandemic could be reflected in further versions of the analysis.

Another element policymakers should consider is the effect of the flow of capital from the US to Central American countries. Because domestic development policies include expansion of foreign direct investment, changes in the global economy, ranging from a continuation of supply chain disruptions, inflation in developed economies and the impact on exchange rates, should also be anticipated. Guatemala is a country with a high dependency on remittances to sustain moderate levels of consumption and investments at the local level. Any increment in the flow of foreign direct investment will help create more jobs and accelerate the rate of economic growth. It also will enhance the necessary investment in infrastructure, housing, health, education and training of the labor force and thereby further reduce migration. Simultaneously improvement in the quality of the labor force, may have a positive multiplier effect serving to attract additional FDI.

The analysis demonstrates that migration from Guatemala will decrease as improvements in economic opportunities and income increases, thereby changing the benefit-cost ratio. Thus, CEO Project interventions emphasizing job creation, increasing local income and wealth, moving individuals into the formal sector of the economy, and improving financial literacy, are most appropriate.

[8] See IFPRC (2021b) for recent analysis of the impact of COVID19 on food security for example.

III. CONTEXT

While irregular migration to the United States from Mexico and Central America has been a long-standing issue, in the last six to eight years the home country composition of the potential illegal migrants has changed rather dramatically. This study assumes that "pull factors" (i.e., conditions in the US that make migration attractive) are constant across home country groups. If border enforcement efforts are assumed to be independent of country of origin of those apprehended, then apprehensions provide a clear indication of how the composition has changed. Apprehensions and how they varied over time and countries are illustrated in Table 1.

Most recent and publicly available numbers show that total apprehensions of individuals attempting to cross the US border illegally was at a 45 year low in 2017. Earlier migrants apprehended tended to be single adult males from Mexico. But, in the last six to eight years, the growth in apprehensions and irregular immigration has been from the Northern Triangle of Central America rather than Mexico, and families and unaccompanied children[9] rather than single males[10]. Additionally, an increasing share are requesting asylum.

Table 1: *Border Apprehensions (source: U.S. Border Patrol)*

	GUATEMALA		MEXICO		EL SALVADOR + HONDURAS		MEX + NORTH TRIANGLE		GLOBAL
	# people	% of global	# people	% of global	# people	% of global	# people	% of global	# people
2020	**47,830**	**11.81%**	254,647	62.87%	57,103	14.10%	359,580	88.78%	**405,036**
2019	265,129	30.85%	169,536	19.72%	344,646	40.10%	779,311	90.67%	859,501
2018	116,808	28.90%	155,452	38.46%	108,764	26.91%	381,024	94.28%	404,142
2017	66,807	21.51%	130,454	42.01%	97,911	31.53%	295,172	95.05%	310,531
2016	75,246	18.10%	192,969	46.41%	125,420	30.16%	393,635	94.67%	415,816
2015	57,160	16.96%	188,122	55.80%	77,412	22.96%	322,694	95.72%	337,117
2014	81,116	16.67%	229,178	47.09%	158,113	32.49%	468,407	96.25%	486,651
2013	54,692	13.00%	267,734	63.63%	84,014	19.97%	406,440	96.59%	420,789
2012	35,204	9.65%	265,755	72.86%	53,111	14.56%	354,070	97.07%	364,768
2011	19,061	5.60%	286,154	84.10%	23,071	6.78%	328,286	96.48%	340,252
2010	18,406	3.97%	404,365	87.26%	27,303	5.89%	450,074	97.13%	463,382
2009	15,583	2.80%	503,386	90.53%	26,322	4.73%	545,291	98.07%	556,041
2008	16,396	2.27%	661,766	91.43%	32,029	4.42%	710,191	98.12%	723,825
2007	17,338	1.98%	808,688	92.24%	37,019	4.22%	863,045	98.44%	876,704

In the September 2019 report the last year of available Border Apprehensions data was for 2018. Then, as illustrated in Table 1 above, apprehensions from Guatemala showed a steady upward trend till 2017 and then a stunning increase in 2018 to 116,808. Simultaneously while the number of apprehensions from Mexico slowed over the decade to 155,452 in 2018, total apprehensions from Mexico and the Northern Triangle varied over the post-financial crisis years ending with 381,024 in 2018, or about 94% of total global apprehensions (of 404,142). The country of origin varied significantly though as apprehensions from Mexico declined and from the Northern Triangle increased, particularly from Guatemala.

Recent updated data reveals an even more stunning increase in apprehensions, including from the Northern Triangle. From 2018 to 2019 total global apprehensions more than doubled, from 404,142 to 859,501. For Mexico and the Northern Triangle apprehensions also more than doubled, from 381,024 to 779,311; but

[9] Under age 18, lack legal status and without parent or guardian in the United States.
[10] See CRS (2019).

as a share of total apprehensions, they fell from 94% to 91%[11]. Apprehensions from Mexico increased modestly, by about 15,000. However, growth from the Northern Triangle was explosive, from 225,572 in 2018 to 609,775 in 2019; or 170%. From Guatemala apprehensions more than doubled from 116,808 to 265,129, and for El Salvador and Honduras the numbers passed from 108,764 to 344,646 (El Salvador nearly tripled from 31,636 to 90,085 and Honduras more than tripled from 77,128 to 254,561).

In 2020 the Department of Homeland Security changed the reporting methodology to include deportations and the data is not comparable to earlier years[12] (see source: Southwest Border Apprehensions, U.S. and Border Protection https://cbp.gov). For 2020 reported "encounters" fell dramatically and the composition by country of origin changed significantly. Some of this change may be related to the pandemic, the related lockdowns and individuals' fear of traveling. However, media coverage suggests that apprehensions are likely much higher in 2021, but as of the date this document was released the last version of publicly available information was not updated (these years are not included in the sample for the analysis because Guatemalan data is not yet available).

The previous study noted that early Rand Corporation studies suggested that only one in three border crossing attempts were successful. Using the same likelihood of success in 2018 there would be 175,212 attempts with 116,808 apprehensions and 58,404 successes. Estimates of current migration based on these success rates should be reconsidered though. With greater border security the success rate would be expected to be lower, and many more attempts to migrate are through amnesty applications rather than illegal crossings at ports. While the actual number of successful illegal crossings per attempt may have changed, it seems clear there is a strong relationship between apprehensions and irregular migration which may deserve additional scrutiny. Further, note that there is no information, at this moment, about the exact location, department or municipality, from which the individuals apprehended originate.

What accounts for the explosive growth in 2019? In the previous report it was emphasized that "pull factors", conditions in the United States, were roughly the same for all potential immigrants and they were relatively stable since the financial crisis and great recession. Then, variations in the "push factors", conditions in the home country, were the likely cause of immigration. The decision to migrate was primarily an economic decision based on the very large relative difference in economic opportunity.

With this new version of the analysis, this analysis includes an extra year of observations at the municipal level and in the regressions below a range of domestic push factors do appear to be significant determinants of migration. However, overall, it appears that from 2018 to 2019 domestic conditions did not worsen and in many areas may have improved for Guatemala. The fact that apprehensions increased for all countries of the Northern Triangle, even more so from El Salvador and Honduras, suggests that pull factors in 2019 and later may have changed. Perhaps a stronger demand for labor in the U.S. as the economy continued to expand, pre-COVID, or a perceived likely change in U.S. immigration policy. The impact of COVID in winter2019/spring2020 is difficult to ascertain, but this is not included in our sample for the analysis below.

Also, as mentioned for the 2018-2019 period most push factors were stable or did not worsen. Table 2 presents descriptive statistics for municipal GDP per capita. We see from 2018 to 2019 there was a modest increase in median GDP per capita. However, the shape of the distribution did change as the

[11] There was a notable increase from other South American countries including Brazil, Venezuela, Cuba, and Nicaragua, but also more than 7,000 from Indonesia and over 2,000 from China.

[12] For 2018 and earlier data includes deportable aliens only, for 2019, total apprehensions and for 2020 beginning in March 2020 USBP "encounters" include both Title 8 apprehensions and Title 42 expulsions. No such categorizations were noted explicitly in the 2019 and earlier tables. Trends in data reported for 2020 is very different than that of earlier years, nonetheless our sample for Municipal level data for Guatemala is for 2018 and 2019 so this does not affect the analysis in this paper.

maximum increased substantially and the minimum decreased substantially from 2018 to 2019. The municipalities that are in the lowest quartile of the distribution are listed in Appendix 1, Table 3.

Table 2: *Municipal Income per capita, descriptive statistics (2018-2019)*

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum	Standard Deviation	Sample (# Obs)
2018	$ 1,405.76	$ 3,137.28	$ 3,809.28	$ 4,196.59	$ 4,817.09	$ 13,788.45	1,706.24	333
2019	$ 1,331.50	$ 2,834.59	$ 3,856.89	$ 4,326.26	$ 4,817.81	$ 19,546.76	2,345.37	340
All	$ 1,331.50	$ 2,950.93	$ 3,815.90	$ 4,262.10	$ 4,817.30	$ 19,546.76	2,053.64	673

Additionally, even though the median (50th percentile) and the third quartile (75th percentile) remain almost the same, the standard deviation significantly increased from 2018 to 2019, which indicates a greater range of income for the fourth (upper) quartile, from about $4,800 to $19,546.

More in depth analysis below shows how the behavior of potential migrants changes depending on the income quartile of the municipality. The next sections contain more details about the composition of the sample and implications for public policy.

IV. OBJECTIVE OF THE ANALYSIS

A natural question is what is causing the sudden upsurge in the flow of potential migrants from Guatemala. Below the general determinants of migration are discussed. It is also believed that the six Departments that are the focus of the CEO project are the origin of a relatively larger number of potential migrants, primarily because of high levels of poverty, overall low standard of living and lack of new businesses and jobs. But this analysis also finds that low graduation rates from secondary school, and poor health status, including malnutrition and exogenous shocks, such as drought, also may further contribute to migration. The CEO project is designed to enhance economic opportunity in general, thereby giving potential migrants in-country options to improve their immediate and long-term economic conditions and build community resiliency to other shocks.

As discussed in Kemme (2019) U.S. Government policy should be informed by a full understanding of the root causes of the rise in migration from Guatemala. Following the approach of the CEO Project, a primary objective is to increase economic opportunity, improving the economic conditions of potential immigrants and thereby reducing the likelihood of immigrating. The analysis aims to identify determinants of migration using municipal level data. While income or economic opportunity is always identified as a determinant, other identified determinants like violence, health status, and housing are themselves also directly related to levels of income and poverty. Median age and population characteristics are also identified as determinants and not tied directly to income. These may serve to better inform and focus policy interventions (e.g., youth employment initiatives).

In the analysis, first, because irregular immigration is not observable, this analysis constructs a proxy or latent variable based on Municipal level observations: remittances and returnees. Then, because there are many measures of determinants (e.g., for violence we have four variables: homicides, extortions, family violence against females and family violence against males), the study used principal components to reduce the dimensionality of each determinant thereby eliminating correlation among independent variables in the regression models. Third, the analysis estimates a wide array of regression models. The final empirical results are robust and statistically significant and provide clear guidance for policymakers; also affirming the approach of the CEO Project.

V. DETERMINANTS OF MIGRATION: "PUSH FACTORS"

Numerous studies have consistently identified four broad categories of determinants of migration, so called "push factors":

1. economic opportunity
2. health/nutrition
3. violence
4. climate/drought[13]

It is well established that lack of economic opportunities (i.e., formal employment, income growth and accumulation of even modest wealth) naturally leads individuals to move to areas where opportunities are greater. Health (i.e. malnutrition) is an extreme reflection of poor economic conditions and is associated with regions of high migration. Levels of violence is a social condition that is also associated with economic conditions but accentuated by lack of effective legal and judicial systems to deter criminal behavior. If these systems are dysfunctional, migration to avoid violence is certainly an option. Climate or weather conditions are most important to agricultural regions and for subsistence or near subsistence farmers. In these territories, drought, significant changes in precipitation or temperature, may entirely eliminate people's livelihoods, naturally leading to migration. The study also examined the role of social capital, civic engagement, the sense of identification with community, and population or demographic characteristics that may influence the decision to migrate (family reunification is a goal of many potential migrants and the existence of a "family migration network" may also be important). Also, weak institutions and lack of confidence in government institutions may be a "push factor" and utilize a proxy for that factor with a newly collected variable on the efficiency of local government management, not available for the earlier study[14]. All of the above push factors are detailed in Table 4 below.

The analysis proceeds in the same manner with all the variables described in the next section and documented in detail in Appendix 1, but the additional year of data provides enough observations to construct synthetic measures of each of these determinants differently than those in the earlier report.

Importantly, the decision to immigrate is made at the nexus of these push factors and the pull factor of economic opportunity in the U.S. The analysis assumes that economic opportunity in the United States remains the predominate "pull factor" and it remains constant. Thus, variation in the push factors in their home country are the primary determinants in the changes in migration. A caveat is that U.S. policies and Border Enforcement efforts have varied in the last few years. It is unclear if these changes have affected immigration though. Hiskey, et al (2018) found via survey-based research that potential immigrants in Honduras were aware of enhanced border enforcement and the dangers of the migration journey, but that did not affect their decision to immigrate.

[13] We used specific measures related to drought conditions in Guatemala – precipitation and temperature –, which are simply local weather conditions rather than climate, *per se*.

[14] These are discussed in more detail in Kemme-CEO 6 (2019), section V. Hiskey, et al (2018) also look at similar characteristics constructed from survey data for the Northern Triangle and a subsample for Honduras (their Table 1 provides details). Their results for Guatemala regarding economic determinants are broadly consistent with ours, and the decision to migrate is primarily an economic one.

Finally, given the decision to migrate, the question of "to where" must be answered. The study argues that destinations with desirable economic conditions and greater economic opportunity are the most likely destination. Immigrants see the U.S. as their best option (perhaps the only real option) until domestic economic opportunities improve.

VI. ANALYTICAL PROBLEM, DATA AND METHODOLOGY

A. ANALYTICAL PROBLEM

Total immigration from a source country to United States may be divided into two components. First, there is regular and legal migration, which is recorded via issuance of visas and temporary and permanent resident status and amnesty applications. Second, there is irregular or illegal migration, including visa overstays, irregular crossings at ports of entry, and illegal border crossings; the latter two are not clearly measured or observable.

The primary difficulty then is that the dependent variable of interest contains a large component that is not observable, either at the national level, or at the department/municipal level. The best estimates for irregular or illegal immigration are at the country level and are not contemporary. For Guatemala and other Central American countries, legal migration is trivial in magnitude vis-a-vis illegal immigration. Because there is no direct measure of irregular migration for Guatemala at the national or the municipal level, this analysis uses a proxy, a latent variable, from variables that are highly correlated with migration. This proxy is the first principal component of remittances and returnees as discussed below.

B. AVAILABLE DATA

While the analysis is constrained by availability of data, this study includes an excellent data set which may be continuously updated for the on-going study of illegal immigration from Guatemala to the United States. The database consists of **34 variables** associated with migration determinants for **22 Departments** and **340 municipalities**. These variables are defined and documented in Appendix I (codebook: the first column lists the categories, then the label or E-Views variable name, the variable itself, the metric or unit of measure, a brief description, the years available, and the data source). For this extension of the analysis, this study has added new variables and an additional year, which provides two years, 340-municipality cross sections, and a panel.

It is important to mention that part of the analysis focuses on territories covered by the programs implemented through the CEO Project. In the Appendix section, Departments and Municipalities are listed along with a code number for the municipality and a dummy variable (CEO included) to identify municipalities in the Departments of interest to the Project. A list of municipalities in the lowest quartile of municipal GDP per capita is also provided.[15]

C. METHODOLOGY

As in the previous analysis, the study proceeds in two steps. First, a latent variable is developed, or proxy for migration, and then this proxy is used as the dependent variable in standard regression models. It is clear that remittances and returnees are highly correlated with migrants in the U.S.[16] There are good measures of these variables at the municipal level that can be used to construct a proxy for migrants. The first principal component of remittances and returnees is the latent variable that will be used as the proxy

[15] Note data and variables in Tables in Appendix I are discussed in the previous report, but here the presentation includes new variables and the newly available two-year sample.

[16] Also, in our previous analysis our migration variable was also correlated with Orozco's estimates of illegal migration.

for immigration and the dependent variable in the regressions that follow. The first principal component is an equally weighted (.707) linear combination of remittances and returnees, and it alone explains 78% of the variation. The second principal component is a linear combination orthogonal to the first and it explains the remainder of the variation.

Note that the first principal component is a single variable, a linear combination of remittances and returnees, whose own variance explains the maximum share of the variance of the two variables selected[17]. The principal components methodology is also used to reduce dimensionality. For example, if there are four different measures of the same phenomenon, it would be inappropriate to include all four in a regression model because they would likely be highly collinear. In addition, one alone may not appear significant in the regression estimation. As a result, it is possible to construct a latent variable, the first principal component of the four variables explaining the same phenomenon. This is done for several of the potential explanatory variables. (E.g., we have four measures of violence which are reduced to one in the analysis below; the details are included in the next section).

The second step of the analysis was then to construct and estimate a model of migration determinants using the dependent latent variable for migration and those variables likely to be significant causes of migration, or the latent variables constructed to represent those determinates. The dataset has a second year of observations at the municipal level, so the data set has 340 municipalities and two years for a total of 680 observations. This allows the analysis to estimate a panel model with cross section fixed effects.

Tables 3-5 present available variables in the data set examined in the modelling process and how they are categorized as potential determinants.

Table 3: *Guatemala Measures of Migration – National Level*

	National Level Conditions	
MEASURES OF MIGRATION	Observed Data	
	○ Visas Issued ○ Visa Overstays ○ Apprehensions ○ Remittances ○ Returnees	
		Unobserved and Unknown
		○ Illegal Migrants

Table 4: *Guatemala Measures of Migration – Municipal Level*

MEASURES OF MIGRATION	Observed Data
	○ Remittances ○ Returnees

[17] See Rencher (1998), Chapter 10 for an excellent discussion of the principal components methodology.

Synthetic Measure of Migration to U.S.	First principal component as proxy for observed migrants

Table 5: *Potential Determinants of Migration (based on available data)*

DETERMINANTS OF MIGRATION	PUSH FACTORS – 9 groups of variables
1. Economic Opportunity - INCOME	○ Municipal Income per capita (quartile dummy variables) ○ Poverty Rate ○ Labor Force ○ Secondary School Graduates
2. Economic Opportunity - WEALTH	○ Deposit accounts: number of accounts ○ Deposit accounts: value of accounts ○ Savings accounts: number of accounts ○ Savings accounts: value of accounts ○ Housing Quantity ○ Electrification Rate
3. Health Status	○ Social Security Affiliates ○ Public Expenditure on Health ○ Chronic Malnutrition
4. Violence	○ Homicides per 100,000 people ○ Extortions per 100,000 people ○ Intra-family violence against Female ○ Intra-family violence against Male
5. Civil Engagement	○ Local Management index ○ Voters registration
6. Drought / Climate	○ Volume of Precipitation ○ Deviation from long-run average Precipitation ○ Temperature ○ Deviation from long-run average Temperature
7.A. Youth Population	○ Median Age ○ Secondary School Graduates
7.B. Population Characteristics	○ Urban Population ○ Population Density
8. Local Housing Characteristics	○ Housing: Qualitative Deficit ○ Housing: Quantitative Deficit

9. Local Infrastructure	○ Water access (households) ○ Sanitation access (households) ○ Electricity access (households)

CEO (cross section fixed effects)	Departments	CEO Project Components
	○ Guatemala ○ Quetzaltenango ○ Totonicapan ○ San Marcos ○ Huehuetenango ○ Quiche	○ Promote Trade and Investment ○ Upgrade Productive Infrastructure ○ Mobilize Financial Resources ○ Increase Businesses Competitiveness

VII. RESULTS OF THE ANALYSIS

In Appendix 2 various regression models and principal components analyses are reported. Of many estimated specifications four are reported below in Table 7. Here are described the construction of the dependent variable, the migration proxy, and the independent variables representing the determinants of migration, listed in Table 6 with further details in Appendix 2.

The first step of the analysis is to construct the migration dependent variable using principal components. The results are presented in Table 1 of Appendix 2, section II. It is possible to see that remittances and returnees explain about 78% of the variation in the latent variable (migpc1), to be used in the regressions. Also, each element of the calculated eigen vector is positive indicating that if either remittances or returnees increases the migration proxy also increases.

The independent variables in the regression are various measures of the determinates of migration discussed above in Table 5 and are presented in Table 6 below. Note that as in Kemme (2019) the analysis has three measures of economic opportunity. The first is Municipal GDP per capita, a straightforward measure of prosperity of the local community. In Table 5 there are nine other variables associated with economic opportunity or prosperity, but these variables are highly correlated with one another, and the coefficient estimates in regressions including them individually were often insignificant and not robust. As a result, to reduce the dimensionality and eliminate the potential multicollinearity, the study again employed principal components to construct two individual latent variables one based on measures of income, and the other one based on measures of wealth.

Therefore, the second measure of economic opportunity is a broader measure that includes those variables associated with income; the first principal component of municipal income per capita, poverty, labor force and graduates, as reported in Table 2a of Appendix 2. These variables explain about 58% of the variation in the latent variable econopinc-pc1 and each of the elements of the eigen vector have the proper sign. The third is a measure of economic opportunity designed to capture wealth as a measure of economic opportunity, the first principal component of savingsacc, savingsam, depositsacc, depositsam, quantihousing and electrification, as reported in Table 2b of Appendix 2. These variables explain 67% of the variation in the latent variable and all the elements of the eigen vector are of the proper sign.

Each of the three measures of economic opportunity is negatively correlated with migration individually. However, again because they are correlated with each other it is inappropriate to include them all in one individual regression equation. For the first two measures, the regression equations including other independent variables, discussed further below, are quite meaningful and statistically sound. The measures of economic opportunity remain orthogonal to all other independent variables in each specification. The migration proxy and econop_wealth_pc1 are also negatively correlated (the coefficient estimates were negative and statistically significant for econop_wealth_pc1 alone). When other explanatory variables were included with this proxy, the regression results indicated a high degree of correlation among the independent variables, the coefficient estimates were not robust, and therefore these specifications were not pursued[18].

[18] See Table 1a and in Table 1b of section I.A of Appendix 2 for complete regression output.

The analysis therefore uses municipal income (GDPpercapita) and the latent variable econopincomepc1 as measures of economic opportunity. Note also because there may be some income threshold effects with respect to the decision to migrate, the study constructs quartile dummy variables for municipal GDP per capita for municipalities in each of the four quartiles.

For categories 3 to 9 of potential determinants in Table 5 (health status, violence, civil engagement, drought, youth population, population characteristics, housing, and local infrastructure), the analysis also uses principal components to construct an individual latent variable. Results are reported in Appendix 2, section II, Tables 3 to 9. The variables included in the principal components analysis explained from about half to over 95% of the variation in the latent variable used as a proxy for each category. Also, in each case the elements of the eigen vector were of the appropriate sign. Table 6 summarizes the independent variables included in the regression models. The study uses the number of cellphone users as a proxy for the existence of a migration network, or family reunification. Note that the existence of a strong family migration network may lead individuals to migrate directly to the US rather than to a larger municipality first.

Table 6: *Definitions of Independent Variables included in Table 7 Regressions, final specifications (see Table 1 of the Appendix for additional details and original sources and see Appendix 2 for details regarding calculations of principal components)*

VARIABLE	DESCRIPTION
CEOINCLUDED	CEOINCLUDED = 1 if municipality is in a CEO Department, 0 otherwise
CELLPHONES	Number of cell phone users
GDPPERCAPITA	Municipal income per capita, US dollars a year
ECOONOPINC_PC1	First principal component of four variables associated with income. From Table 2a of Appendix 2, II.
HEALTH_PC1	First principal component of three variables associated with health status. From Table 3 of Appendix 2, II.
VIOLENCE_PC1	First principal component of four measures of violence. From Table 4, Appendix 2, II.
HOUSING_PC1	First principal component of two variables related to the quality and quantity of housing. From Table 8 of Appendix 2, II.
YOUTH_PC1	First principal components of median age and graduates. From Table 7a Appendix2, II.
DEMOG4_PC1	First principal component of urban population and population density. From Table 7b, Appendix 2, II
DGDPQ25_18	Dummy variable = 1 if municipal GDP per capita is in First quartile of distribution

The final regression results are reported in Table 6 below. The study uses panel OLS regression with cross section, municipality, fixed effects[19]. These control for idiosyncratic variations at the municipal level. There are four final specifications. Specifications 1 and 2 employ the entire sample of all municipalities in all departments. Specifications 3 and 4 employ a sample of municipalities in the departments that are the focus of the CEO project (designated CEO6). The coefficient estimates are similar and the goodness of fit, the adjusted R^2, in all cases is quite high.

For the country as a whole (all 340 municipalities), Specification 1 includes gdppercapita as the measure of economic opportunity and Specification 2 uses econopincomepc1 as the proxy. The coefficients are negative and statistically significant indicating a strong negative association between higher local economic opportunity and the level of migration. This is also true for the subset of municipalities in the six CEO Departments as indicated in specifications 3 and 4. This strongly suggests that the decision to migrate is an economic decision. Specifications 5 and 6 include the dummy variable for municipal GDP being in the lowest quartile for 2018 (DGDPQ25_18) and the interaction of that dummy with GDP per capita[20]. The coefficient for DGDPQ25_18 is positive and statistically significant, indicating that the general negative association between municipal GDP per capita and migration is less strong for municipalities in the bottom quartile of GDP per capita in 2018 than the effect for the average municipality. The coefficient of the dummy variable interacted with municipal GDP per capita is also positive and statistically significant[21]. This suggests that increasing incomes in the lower quartile does not have as large a deterrent effect on migration abroad as in the typical municipality. This may be because there are already few migrants abroad, remittances and returnees are very low to begin with, and the potential decline or slower growth is not significant or that individuals are not migrating abroad, may be migrating internally or simply not at all.

Importantly though, other determinants are also significant in these specifications. The analysis finds a negative association between youth_pc1 (median age and graduates) and migration, and also a positive association between violence_pc1 and migration. The individual associations are strong and consistent with the survey literature. Also, Clemons (2017) finds a more focused link between violence and child migration from the Northern Triangle[22]. The previous analysis did not find population characteristics to be significant, but here not only youth, but also urbanization and population density (demog4_pc1) are positively associated with migration and statistically significant. Note, importantly, that the youth principal component is a linear combination of median age and share of graduates in the population and the weights for each are positive (from Table 7a in Appendix 2, II). Thus, as median age increases, or the share of graduates increases, the youth proxy increases and that is negatively correlated with migration, so migration would fall. Policies aimed at increasing graduation (secondary school) should lower migration. The positive association between both urbanization, population density, and migration reflects the location of a large number of prospective migrants.

Our measure of health status is more comprehensive than in Kemme (2019) and here improved health status is negatively associated with migration, and it is statistically significant in every specification.

While the measure of economic opportunity based on measures of wealth (bank accounts) was less important, an alternative measure, housing_pc1, based on the quality and quantity of housing is negatively

[19] Note that the cross section fixed effects eliminates the significance of the CEO6 dummy and the individual department dummies if they are included in the regressions.

[20] The entire set of quartile-year dummies were included in other regressions but only this dummy was significant throughout. Note that most municipalities in this quartile for 2018 are in Huehuetenango, Quiche (CEO Departments) and Alta Verapaz.

[21] These specifications were also estimated for the CEO6 subsample and results were very similar.

[22] Hiskey, et al (2018) does not find a link between violence and migration for Guatemala but does for Honduras and El Salvador. They suggest that violence in Guatemala is quite different than in the other two countries which is predominately larger scale organized gang and drug smuggling.

associated with migration. The number of cell phones is positively related to migration. This may be an indication that an individual has enough wealth to consider migrating, has access to much more information in general and there is greater awareness of opportunities in the U.S., and/or that there is a strong family migration network.

The variables used for civic engagement and drought/climate are not statistically significant in the specifications reported in Appendix 2 and therefore are not included in the final specifications reported here in Table 7.

Table 7: *Regression Results, Dependent Variable MIGRATION_PC1*

Variable	Coefficient P value					
Dependent variable	MIGPC1	MIGPC1	MIGPC1	MIGPC1	MIGPC1	MIGPC1
	Spec 1 (AT1a -13)	Spec 2 (AT1b-13)	Spec 3 (BT13CEO6)	Spec 4 (B13aCEO6)	Spec 5 (AT4a-13)	Spec 6 (AT4b-13)
C	-0.117753 0.1542	-0.22910* 0.0013	-0.518407 0.1809	-0. 825563** 0.0236	-0.13157** 0.0992	-0.13889*** 0.0823
GDPPERCAPITA	-2.73E-05** 0.0154		-6.73E-05** 0.0217		-1.94E-05*** 0.0760	-1.90E-05*** 0.0825
ECONOPINC_PC1		-0.08115** 0.0465		-0.181716*** 0.0972		
VIOLENCE_PC1	0.160043** 0.0200	0.150970** 0.0279	0.442348** 0.0256	0.415994** 0.0370	0.134996 0.0760**	0.1365** 0.0399
CELLPHONES	1.71E-06** 0.0140	1.75E-06** 0.0126	3.00E-06** 0.0231	3.06E-06** 0.0251	1.71E-06 0.0109**	1.74E-06 0.0399
HEALTH_PC1	-1.414196* 0.0000	-1.489997* 0.0000	-2.382806* 0.0032	-2.555407* 0.0022	-1.385601* 0.000	-1.4009* 0.000
HOUSING_PC1	-3.704773** 0.0188	-3.50139** 0.0267	-8.598438*** 0.0575	-8.756618*** 0.0579	-2.17085** 0.1607	-2.30226 0.1355
YOUTH_PC1	-0.468548* 0.0000	-0.462875* 0.0000	-0.894493* 0.0000	-0.875144* 0.0000	-0.334985* 0.000	-0.3430* 0.000
DEMOG4_PC1	0.137379* 0.0079	0.136873* 0.0084	0.230578** 0.0361	0.229282** 0.0396	0.110748* 0.0271	0.11026** 0.0278
DGDPQ25_18					0.231429** 0.000	

Variable Coefficient P value						
Dependent variable	MIGPC1	MIGPC1	MIGPC1	MIGPC1	MIGPC1	MIGPC1
	Spec 1 (AT1a -13)	Spec 2 (AT1b-13)	Spec 3 (BT13CEO6)	Spec 4 (B13aCEO6)	Spec 5 (AT4a-13)	Spec 6 (AT4b-13)
DGDPQ25_18* GDPPerCapita						8.64E-05* 0.000
N	672	672	263	263	672	672
Adjusted R-squared	0.949	0.949	0.944	0.943	0.952	0.953

NOTE: * statistically significant at the 1% level ** at the 5% level or less, and ***10% or less

VIII. FINDINGS, IMPLICATIONS & LINKS TO CEO PROJECT

The above econometric analysis has identified the most significant root causes of migration and provides a strong foundation for policy recommendations. Given the extent and quality of the data base and the statistical significance and robustness of the results, the study provides an improved understanding of the migration decision and the determinants of migration. The qualitative results are summarized in Table 8. Clear findings and implications include:

1. Economic opportunity in the United States, the dominate "pull factor", has remained unchanged. Therefore, altering the domestic "push factors" is essential to addressing the root causes of migration.

2. The most significant determinate of migration is the level of local municipal income, or GDPpercapita, or economic opportunity. Migration from Guatemala is predominately an economic decision, but may also be affected by other factors, like violence/crime and health status, themselves influenced by level of income. These factors are also statistically significant in the results above.

3. As noted above, foreign migration is negatively associated with the level of income in the municipality. However, a new finding is that the effect of increases in income on potential foreign migration from municipalities in the lower quartile of the distribution of municipal income per capita is less strong than the average municipality after controlling for all other determinants. For the lowest quartile municipalities, foreign migration is less effected by increases in municipal GDP per capita than the average municipality. Resulting migration may be internal as well.

4. Migration is sensitive to income, which in turn, for an individual, is driven by labor productivity. Education, human capital investments, and improved health status increases labor productivity and individual income. Again, these are statistically significant in the results above.

5. Migration may be reduced by creating jobs in the formal sector, promoting financial education and bankability (savings such as wealth accumulation) and access to credit (the possibility of improving productive activity).

6. Youth, as measured by median age and graduates, of the municipality is negatively associated with migration. Policies that increase the number of graduates (secondary school) should lower foreign migration and increase expected income in the long term.

7. Urbanization and population density of the local municipality, and lower quality/quantity of housing are positively associated with migration in the analysis above.

8. The above points emphasize the weight that improved economic conditions have in reducing migration. The effect of personal security conditions was also important. However, for domestic violence, homicides and extortion, only extortions showed a consistent effect on migration, presumably from urban areas since rural areas generally have low levels of violence and criminality.

9. Improved health status and better housing conditions (water, sanitation, electricity and other infrastructure like telecommunications) are also negatively associated with migration

10. While all the above conclusions are relevant determinants of migration in Guatemala, the study also finds that climate/weather and civil engagement (social capital) are less important. Most likely, municipal income per capita indirectly captures the effects of these determinants. Table 8 below illustrates the relationship between the factors analyzed and the change in migration to U.S.:

Table 8: *Relation between Independent Variables and Migration to U.S.*

←	Income per capita, Municipal	→	←	Migration to United States
←	Bankability and Financial Education	→	←	Migration to United States
←	Health Status	→	←	Migration to United States
←	Secondary Education Graduates	→	←	Migration to United States
←	Housing, quantity and quality	→	←	Migration to United States
□	Violence (i.e., extortions)	→	←	Migration to United States
←	Cell phone users	→	←	Migration to United States
←	Urbanization and population density	→	←	Migration to United States
←	Youth (population gets older)	→	←	Migration to United States
←	Income at Lowest Quartile Municipalities	→	=	Possible Internal migration
← ←	Climate (precipitation and temperature)	→	=	Not significant
← ←	Civil Engagement (social capital)	→	=	Not significant

These findings are based on 680 observations of municipal level data for two years. Given the significance of key determinants, interventions can be targeted at specific local weakness that vary by municipality. National level programs have significant impact, but at the margin targeted by CEO interventions may be more beneficial. The results above also suggest that programs should be targeted toward youth, school graduation and urban areas. Improving the quantity and quality of housing may be beneficial as well. Current CEO interventions are designed to increase job creation and income growth via increased foreign and domestic investment, training, and educational programs to increase labor productivity and financial education to manage and accumulate wealth or "bankability".

These interventions should have positive results. Formal sector jobs provide participation in health and social security programs as well as greater attachment to community, all lowering the likelihood of immigrating. Because a large share of migration from Guatemala to the U.S. is families and unaccompanied minors, policies may be tailored to provide greater support for families, movement of female employment into the formal sector, and youth training and jobs.

To reiterate the implications for CEO Project activities:

- Activities that most quickly contribute to reduction in irregular migration to the US are those that create jobs in the formal sector, targeted at youth and improving school graduation, urban areas, and low-income municipalities.

- To the extent that individuals also open bank accounts and access other banking services and credit for housing and small business creation resulting in new jobs, accumulation of wealth and improved economic conditions, migration will be reduced.

- Investments in job creation in the formal sector, particularly in the private sector, not only generates higher levels of income and therefore lower migration, increases tax collections while providing access to social security and health services, but also ultimately reduces dependency on governmental assistance programs.

- Improved health and higher educational attainment also increase labor productivity and individual and family income, lessening migration.

- Other determinants, like violence, should be addressed by public efforts to address issues of insecurity, with success measured by a reduction in all measures of violence, but more specifically extortions in urban areas.

- Policies aimed at the lower quartile of the distribution of municipal GDP per capita should be crafted carefully to improve income and socioeconomic conditions while not providing incentives to migrate abroad. (Internal migration may be an alternative, thereby influencing the development of Intermediate Cities).

- Analysis at the municipal level makes it possible to design municipality specific interventions that could address local needs. The municipalities that may be targeted by CEO Project are not only those of the Guatemalan Western Highlands, but perhaps areas like Alta Verapaz and the border zone between Jutiapa and El Salvador, and between Zacapa and Honduras. Official data indicate these regions also show a significant number of migrants in relation to their population.

IX. NEXT STEPS FOR ANALYSIS

The analysis above identifies robust relationships between migration and municipal income per capita and measures of wealth, which are within the scope of the USAID | CEO project. There are also relationships between migration and violence, health status, and strength of family migration network, which may call for interventions beyond the scope of the CEO Project.

The first report, Kemme (2019), identified the basic determinants of migration and concluded that the migration decision is an individual economic decision with limited evidence of other factors influencing that decision. The analysis in this report, with an additional year of data that allows panel estimation techniques, finds that the decision to migrate is still an economic decision, but local conditions also play a significant role. Median age (youth), urban density, and low municipal GDP per capita are positively associated with migration. Improved health status and better housing conditions are negatively associated with migration.

Further on-going analysis may provide additional insights. Because there are lags in policy decisions, delivery of assistance and measurable impact, a longer consistent time series of determinants could provide powerful information to identify determinates and evaluate assistance impact at a more micro level. Additional data may allow the decomposition of the individual determinants to ascertain which individual component is most influential in the migration decision. In addition, the individual components of the migration determinant may allow policy makers to fine tune policies and target specific municipalities so that scarce assistance funds will have a greater impact.

It will be very important to revisit and update the official municipal level data in the data base. Additional years of data, both for the recent past and future years, for the variables in hand, will allow further analysis to examine dynamic relations between determinants already identified and changes in migration patterns over time.

Finally, to the extent that the study provides greater insight into the migration decision process, in the longer run, at the end of the CEO Project, this analysis may also provide a means of validating effectiveness of Project interventions in creating economic opportunity and thereby deterring migration.

X. REFERENCES

- Ambler, Kate (2019). "Migration and Remittances in Central America: New Evidence and Pathways for Future Research". IFPRI LAC Working paper 4.

- Bailey, John W. (2016). "Assessing Southern Border Security". Institute for Defense Analyses, IDA Paper NS P-5304, May 2016.

- Baker, Bryan (2017). "Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2014". Office of Immigration Statistics, Policy Directorate, U.S. Department of Homeland Security, https://www.dhs.gov/sites/default/files/publications/ Unauthorized%20Immigrant%20Population%20Estimates%20 in%20the%20US%20January%202014_1.pdf

- Baker, Bryan (2018). "Population Estimates: Illegal Alien Population Residing in the United States". January 2015, Homeland Security. Office of Immigration Statistics, Office of Strategy, Policy & Plans.

- Canales, Alejandro I. and Rojas, Martha Luz (2018). "Panorama de la Migracion internacional em Mexico y Centroamerica". UN, Economic Commission for Latin America and the Carribean.

- Carriquiry, Alicia and Malay, Majmundar (2013). "Options for Estimating Illegal Entries at the U.S.-Mexico Border". National Academies Press, 2013.

- Ceballos, F. and Hernández, M. (2020). "The Migration Propensity Index: An Application to Guatemala". International Food Policy Research Institute, Discussion Paper 01953.

- Clemens, Michael (2017). "Violence, Development, and Migration Waves: Evidence from Central American Child Migrant Apprehensions". Center for Global Development, Working Paper 459.

- Cohn, D., Passel, J., and Gonzalez-Barrera, A. (2017). "Rise in U.S. Immigrants From El Salvador, Guatemala and Honduras". Outpaces Growth from Elsewhere. Pew Research Center. https://www.pewresearch.org/hispanic/2017/12/07/rise-in-u-s-immigrants-

- Congressional Research Service (2019a). "Recent Migration to the United States from Central America: Frequently Asked Questions". CRS REPORT R45489. https://fas.org/sgp/crs/row/R45489.pdf

- Congressional Research Service (2019b). "Central American Migration: Root Causes and U.S. Policy". IN FOCUS IFIII5I, June 13. https://fas.org/sgp/crs/row/IF11151.pdf

- Del Carmen, Giselle and Sousa, Liliana D. (2018). "Human Capital Outflows: Selection into Migration from the Northern Triangle". Policy Research Working Paper 8334, World Bank Group.

- Demirgüç-Kunt, Asli; Klapper, Leora; Singer, Dorothe; Ansar, Saniya; and Hess, Jake (2018). "The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution". Washington, DC: World Bank. doi:10.1596/978-1-4648-1259-0. License: Creative Commons Attribution CC BY 3.0 IGO. (No data specific to Guatemala)

- Dinesen, Cecile; Ronsbo, Henrik; Juárez, Carla; González, Mariano; Estrada Méndez, Miguel Ángel; and Modvig, Jens. (2013). "Violence and social capital in post-conflict Guatemala". Rev Panam Salud Publica. 2013; 43(3):162–8.
- Domínguez, Ana Sofia; Olmedo, Eduardo; and Rayo, Mariano (2018). "Migración y Remesas". Asociación de Investigación y Estudios Sociales (ASIES), Guatemala, February 2018.
- Economic Commission for Latin America and the Caribbean (2018). "UN-ECLAC: Atlas of Migration in Northern Central America". (LC/PUB.2018/23), Santiago, 2018.
- Feed the Future Program (2014). "Zone of Influence Baseline Report". Guatemala, July 2014.
- FUNDESA (2021). "Guatemala Competitiveness Index: Local Competitiveness Index". https://fundesa.org.gt/country-indices-and-evaluations/local-competitiveness-index
- Hiskey, Johnathan T., Abby Cordova, Mary Fran Malone and Diana M. Orces (2018). "Leaving the Devil You Know: Crime Victimization, US Deterrence Policy, and the Emigration Decision in Central America". Latin American Research Review, vol. 53, no. 3.
- Hiskey, Jonathan T., Abby Cordova, Diana Orces, and Mary Fran Malone. (2016). "Understanding the Central American Refugee Crisis: Why They Are Fleeing and How U.S. Policies Are Failing to Deter Them". American Immigration Council, Special Report, February 1. https://www.americanimmigrationcouncil.org/research/understanding-central-american-refugee-crisis
- Inter-American Dialogue (2016). "Opportunities for My Community, Building Human and Economic Capital: A Strategy for Guatemala". The Dialogue – Leadership for the Americas (2016).
- International Food Policy Research Institute (2021a). "Monitoring of Migration Indicators, Baseline Results".
- International Food Policy Research Institute (2021b). "COVID-19 Study in Rural Guatemala, Impact of COVID-19 on food security and nutrition in the Western Highlands". www.ifpri.org
- IOM (2018). "Guatemala Country Profile". Available at: https://mic.iom.int/webntmi/guatemala/
- IOM (2016). "DTM Matriz de Seguimiento de Movilidad Humana". Guatemala 2016.
- IOM (2018). "Migration Flows Report in Central America, North America, and the Caribbean". San José, Costa Rica. Regional Office for Central America, North America and the Caribbean. Available at: https://rosanjose.iom.int/site/sites/default/files/Reportes/SITREP%207%20-%20ingl%C3%A9s.pdf
- IOM (2016). "Survey on International Migration of Guatemalan People and Remittances". Guatemala 2016. Available at: https://mic.iom.int/webntmi/descargas/descargasoim/ResumenEjecutivoEMRG2016.pdf
- IOM (2017). "Food Security and Emigration: Why people flee and the impact on family members left behind in El Salvador, Guatemala and Honduras". Guatemala, El Salvador and Honduras, 2017. Available at: https://docs.wfp.org/api/documents/WFP-0000019629/download/
- Kemme, David M. (2019). "Determinants of Immigration and Project Objectives". For the Creating Economic Opportunities Project for the United States Agency for International Development. Referenced in text as Kemme-CEO (2019).

- Kemme, David M. and Jorge Benavides (2021). "Synthetic measures of Illegal Immigration and Implications for Policy: The case of Guatemala". Paper presented at the International Political Science Association, World Congress of Political Science, July 2021.
- Kennedy, Peter (2006). "A Guide to Econometrics". Malden, MA: Blackwell Publishing (USA).
- Latin American Public Opinion Project (2018). "LAPOP Survey". Different countries. Results available at: https://www.vanderbilt.edu/lapop/data-access.php
- OECD (2019). "Revenue Statistics in Latin America and the Caribbean 2019". OECD Publishing, Paris, France. Available at: https://doi.org/10.1787/25666b8d-en-es
- Office of Immigration Statistics (2017). "Efforts by DHS to Estimate Southwest Border Security between Ports of Entry". Office of Immigration Statistics, Policy Directorate, U.S. Department of Homeland Security. Available at: https://www.dhs.gov/sites/default/files/publications/17_0914_estimates-of-border-security.pdf
- Orozco, Manuel (2017). "Patterns of Central American Migration". Inter-American Dialogue.
- Orozco, Manuel and Yansura, Julia (2014). "Understanding Central American Migration: The Crisis of Central American Child Migrants in Context". Inter-American Dialogue, August 2014.
- Passel, Jeffrey S. and D'Vera, Cohn (2016). "Unauthorized immigrant population stable for half a decade". Pew Research Center. Available at: http://www.pewresearch.org/fact-tank/2016/09/21/unauthorized-immigrant-population-stable-for-half-a-decade/
- Population Estimates (2018). "Illegal Alien Population Residing in the United States: January 2015". Homeland Security, Office of Immigration Statistics, Office of Strategy, Policy and Plans, December 2018.
- Rencher, Alvin (1998). "Multivariate Statistical Inference and Applications". New York: John Wiley & Sons.
- United Nations (2016). "International Migration and Development, Report of the General Secretary, to the General Assembly". 71st Session, Globalization and Interdependence.
- UNDP (2018). "Human Development Index 2018". United Nations Development Program. Available at: http://hdr.undp.org/en/countries/profiles/GTM
- Warren, Robert (2017). "Zero Undocumented Population Growth Is Here to Stay and Immigration Reform Would Preserve and Extend These Gains". Journal on Migration and Human Security, Volume 5, Number 2 (2017).
- World Bank (2018). "Moving for Prosperity: Global Migration and Labor Markets". Policy Research Report. The World Bank Group.
- World Bank (2016). "Guatemala: Country Partnership Framework for the period FY17-20: Marco de Alianza para La Republica de Guatemala para el periodo 2017-2020 (Spanish)". Washington, D.C.: World Bank Group. Available at: http://documents.worldbank.org/curated/en/538891487180272560/Marco-de-alianza-el-pais-para-La-Republica-de-Guatemala-para-el-period-2017-2020

- World Bank (2016). "MetaData of the Climate Change Knowledge Portal". Data available at: https://climateknowledgeportal.worldbank.org/themes/custom/wb_cckp/resources/data/Metadata.pdf

- World Bank (2016). "Climate Knowledge Portal: GUATEMALA". Data available for the country at: https://climateknowledgeportal.worldbank.org/country/guatemala/climate-data-historical

XI. APPENDIX I: DATA DESCRIPTION

Description of the variables used for this report:

CATEGORY	LABEL	VARIABLE	METRICS	DESCRIPTION	2018: DATA YEAR	2019: DATA YEAR	SOURCE
CONTROL	CEOAttended	CEO Attended	binary	Dummy: 1 = Attended, 0 = Not Attended	2018	2019	CEO Project
SyntheticMIGRANTS	Returnees	Returnees	people	Number of returnees documented by IOM	2018	2019	International Organization for Migration
SyntheticMIGRANTS	Remittances	Remittances	US dollars	Amount of Remittances registered by the Bank of Guatemala	2018	2019	Bank of Guatemala
EconOppIncome	GDPperCapita	GDP per capita	US dollars	GDP per capita, calculated as US dollars a year (PPP, 2012 numbers)	2017	2019	FUNDESA
EconOppIncome	Poverty	Poverty Rate	percentage	Share of population living under the Poverty line (national definition)	2014	2014	National Institute for Statistics
EconOppIncome	LaborForce	Labor Force	people	Number of working population in the age range from 15 to 64	2018	2019	National Institute for Statistics
EconOppIncome	Graduates	Graduates	percentage	Secondary Level Graduates as share of the Population in Secondary Level Age	2018	2019	Ministry of Education
EconOppIncome	CellPhones	Cell Phone Users	units	Number of active Cell Phone Users	2017	2019	Superintendence of Telecomunications
EconOppWealth	DepositsAcc	Deposits Accounts	units	Number of deposits accounts reported by the Superintendence of Banks	2018	2019	Superintendence of Banks
EconOppWealth	DepositsAm	Depostis Ammount	US dollars (thousands)	Amount of money reported as Deposits by the Superintendence of Banks	2018	2019	Superintendence of Banks
EconOppWealth	SavingsAcc	Savings Accounts	units	Number of savings accounts reported by the Superintendence of Banks	2018	2019	Superintendence of Banks
EconOppWealth	SavingsAm	Savings Ammount	US dollars (thousands)	Amount of money reported as Savings by the Superintendence of Banks	2018	2019	Superintendence of Banks
EconOppWealth	QuantiHousing	Housing Quantitative Deficit	percentage	Number of Households with Quantitative Deficit as share of Total Households	2018	2019	National Institute for Statistics

CATEGORY	LABEL	VARIABLE	METRICS	DESCRIPTION	2018: DATA YEAR	2019: DATA YEAR	SOURCE
EconOppWealth	QualiHousing	Housing Qualitative Deficit	percentage	Number of Households with Qualitative Deficit as share of Total Households	2018	2019	National Institute for Statistics
EconOppWealth	Electrific	Electrification	percentage	Number of Households with Electificaton Connection as share of Total Households	2016	2017	Ministry of Energy
HealthStatus	SSAffiliates	Social Security Affiliates	people	Number of workers that contribute to Social Security	2017	2019	Guatemalan Institute for Social Security
HealthStatus	ExpHealth	Public Expenditure in Health	GT Quetzales	Amount of Money registered in the Nation´s Budget as Public Expenditure in Health	2018	2019	Ministry of Finance
HealthStatus	ChronicMal	Chronic Malnutrition	percentage	Share of children under 5 years old that not meet potential Height for the actual Age	2015	2015	Ministry of Health
HealthStatus	Water	Water Coverage	percentage	Number of Households with access to Water as share of Total Households	2014	2019	National Institute for Statistics
HealthStatus	Sanitation	Sanitation Coverage	percentage	Number of Households with access to Sanitation as share of Total Households	2014	2019	National Institute for Statistics
Violence	Homicides	Homicides	units	Number of homicides reported by the National Police Force compared with population	2018	2019	National Police Force
Violence	Extortions	Extortions	units	Number of extortions reported by the National Police Force compared with the population	2018	2019	National Police Force
Violence	InFamVioMale	Intrafamiliar Violence Male	people	Number of intrafamiliar violence cases reported, Male victim	2017	2019	National Institute for Statistics
Violence	InFamVioFem	Intrafamiliar Violence Female	people	Number of intrafamiliar violence cases reported, Female victim	2017	2019	National Institute for Statistics
CivilInvolvement	AgroConflicts	Agro Conflicts	units	Number of conflicts registered that are related with Agricultural issues	2017	2019	Agricultre Affairs Secretary
CivilInvolvement	Voters	Voters	people	Number of people registered to vote in the last report	2018	2019	Electoral Supreme Tribunal
CivilInvolvement	Mangement	Local Management Evaluation	units	Score obtained in the Local Management Evaluation (0-100)	2016	2018	National Planning Secretariat
Climate	PreciptAvg	Average Precipitation	milimeters	Precipitation accumulated during	2018	2018	INSIVUMEH

CATEGORY	LABEL	VARIABLE	METRICS	DESCRIPTION	2018: DATA YEAR	2019: DATA YEAR	SOURCE
				the year, average 2010 to 2018			
Climate	Precipt	Precipitation	milimeters	Precipitation accumulated during the year, 2018	2018	2018	INSIVUMEH
Climate	TempAvg	Average Temperature	Celsius degrees	Average Temperature registered during the year, average 2010 to 2018	2018	2018	INSIVUMEH
Climate	Temp	Temperature	Celsius degrees	Average Temperature registered during the year, 2018	2018	2018	INSIVUMEH
Demographics	UrbanPop	Urban Population	percentage	Share of total population living in urban areas	2018	2019	National Institute for Statistics
Demographics	PopDensity	Population Density	habitants per km2	Number of Habitants per square kilometer of surface	2018	2019	National Institute for Statistics
Demographics	MedianAge	Median Age	years	Median Age of the population according to the Census	2016	2020	National Institute for Statistics

MUNICIPALITIES:

- Review the Database for consulting the complete list of Municipalities included in the analysis.
- Municipalities in the first Quartile of Municipal GDP per Capita

DEPARTMENT	MUNICIPALITY	GDP per capita
Escuintla	Nueva Concepción	$ 3,014.13
Sololá	Santa Catarina Palopó	$ 2,781.83
Sololá	Santa Cruz La Laguna	$ 2,703.77
Totonicapán	Santa María Chiquimula	$ 2,904.40
Totonicapán	Santa Lucía La Reforma	$ 2,337.79
Totonicapán	San Bartolo Aguas Calientes	$ 2,803.39
Quetzaltenango	Huitán	$ 3,036.09
San Marcos	San Miguel Ixtahuacán	$ 2,974.98
San Marcos	Concepción Tutuapa	$ 2,695.09
San Marcos	Tacaná	$ 2,844.19
San Marcos	Sibinal	$ 3,132.14
San Marcos	Tajumulco	$ 2,752.02
San Marcos	Ocós	$ 2,959.25
San Marcos	San José Ojetenam	$ 3,118.50
San Marcos	Sipacapa	$ 2,996.51
Huehuetenango	Chiantla	$ 2,926.84
Huehuetenango	Malacatancito	$ 2,973.72
Huehuetenango	Cuilco	$ 2,770.33
Huehuetenango	Nentón	$ 2,638.71
Huehuetenango	San Pedro Necta	$ 3,010.73
Huehuetenango	San Pedro Soloma	$ 3,033.68
Huehuetenango	San Ildefonso Ixtahuacán	$ 2,870.26
Huehuetenango	Santa Bárbara	$ 2,307.11
Huehuetenango	La Libertad	$ 2,852.95
Huehuetenango	San Miguel Acatán	$ 2,360.51
Huehuetenango	San Rafael La Independencia	$ 2,878.16
Huehuetenango	Todos Santos Cuchumatán	$ 2,566.90
Huehuetenango	San Juan Atitán	$ 2,589.42
Huehuetenango	Santa Eulalia	$ 2,393.92
Huehuetenango	San Mateo Ixtatán	$ 2,477.28
Huehuetenango	Colotenango	$ 2,511.11
Huehuetenango	San Sebastián Huehuetenango	$ 2,858.30
Huehuetenango	Tectitán	$ 2,663.20
Huehuetenango	San Juan Ixcoy	$ 2,309.30
Huehuetenango	San Sebastián Coatán	$ 2,428.50
Huehuetenango	Santa Cruz Barillas	$ 2,314.40
Huehuetenango	Aguacatán	$ 2,923.60
Huehuetenango	San Gaspar Ixchil	$ 1,405.80
Huehuetenango	San Gaspar Ixchil	$ 2,851.40
Huehuetenango	Santiago Chimaltenango	$ 2,745.60
Huehuetenango	Unión Cantinil	$ 2,951.30

DEPARTMENT	MUNICIPALITY	GDP per capita
Quiché	Chiché	$ 2,419.10
Quiché	Zacualpa	$ 2,009.10
Quiché	Chajul	$ 2,384.90
Quiché	Santo Tomás Chichicastenango	$ 2,912.70
Quiché	San Antonio Ilotenango	$ 2,682.20
Quiché	San Pedro Jocopilas	$ 2,172.00
Quiché	Cunén	$ 2,561.61
Quiché	San Juan Cotzal	$ 2,940.20
Quiché	Joyabaj	$ 2,151.70
Quiché	Nebaj	$ 2,893.20
Quiché	San Andrés Sajcabajá	$ 2,316.00
Quiché	San Miguel Uspantán	$ 2,159.00
Quiché	Sacapulas	$ 2,764.00
Quiché	San Bartolomé Jocotenango	$ 2,022.00
Quiché	Canillá	$ 2,175.30
Quiché	Chicamán	$ 2,192.00
Quiché	Playa Grande-Ixcán	$ 2,161.30
Baja Verapaz	Cubulco	$ 2,533.60
Baja Verapaz	Purulhá	$ 2,411.60
Alta Verapaz	Tamahú	$ 2,640.50
Alta Verapaz	San Miguel Tucurú	$ 2,355.90
Alta Verapaz	Panzós	$ 2,308.90
Alta Verapaz	Senahú	$ 2,279.90
Alta Verapaz	San Pedro Carchá	$ 2,378.70
Alta Verapaz	Lanquín	$ 2,451.50
Alta Verapaz	Santa María Cahabón	$ 2,404.60
Alta Verapaz	Chisec	$ 2,162.60
Alta Verapaz	Chahal	$ 2,334.40
Alta Verapaz	Fray Bartolomé De Las Casas	$ 2,518.50
Alta Verapaz	Santa Catarina La Tinta	$ 2,822.60
Alta Verapaz	Raxruhá	$ 2,237.90
Petén	San José	$ 3,042.90
Petén	San Andrés	$ 2,631.40
Petén	La Libertad	$ 2,467.20
Petén	Santa Ana	$ 2,634.20
Petén	Dolores	$ 2,796.10
Petén	San Luis	$ 2,599.10
Petén	Sayaxché	$ 2,326.80
Jalapa	San Pedro Pinula	$ 2,657.60
Jalapa	San Carlos Alzatate	$ 3,106.50
Jalapa	Comapa	$ 2,798.90
Jalapa	Conguaco	$ 3,124.30

XII. APPENDIX 2: ANALYSIS

1. Regression Analysis

A. All Departments and Municipalities

Table 1a

Dependent Variable: MIGPC1 spec 13
Method: Panel Least Squares
Date: 11/06/21 Time: 11:48
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.117753	0.082451	-1.428163	0.1542
GDPPERCAPITA	-2.73E-05	1.12E-05	-2.436090	0.0154
VIOLENCE_PC1	0.160043	0.068449	2.338143	0.0200
CELLPHONES	1.71E-06	6.92E-07	2.470330	0.0140
HEALTH_PC1	-1.414196	0.312802	-4.521063	0.0000
HOUSING_PC1	-3.704773	1.569242	-2.360868	0.0188
YOUTH_PC1	-0.468548	0.067515	-6.939961	0.0000
DEMOG4_PC1	0.137379	0.051438	2.670752	0.0079

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.975224	Mean dependent var	0.000878
Adjusted R-squared	0.949005	S.D. dependent var	1.247312
S.E. of regression	0.281669	Akaike info criterion	0.610236
Sum squared resid	25.86409	Schwarz criterion	2.932482
Log likelihood	140.9607	Hannan-Quinn criter.	1.509609
F-statistic	37.19452	Durbin-Watson stat	4.023952
Prob(F-statistic)	0.000000		

Table 1b

Dependent Variable: MIGPC1 spec 13a with
econop_income
(reported in text)

Method: Panel Least Squares
Date: 11/08/21 Time: 16:42
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.229099	0.070754	-3.237978	0.0013
ECONOPINC_PC1	-0.081153	0.040611	-1.998274	0.0465
VIOLENCE_PC1	0.150970	0.068346	2.208928	0.0279
CELLPHONES	1.75E-06	6.96E-07	2.507942	0.0126
HEALTH_PC1	-1.489997	0.318906	-4.672215	0.0000
HOUSING_PC1	-3.501391	1.572827	-2.226177	0.0267
YOUTH_PC1	-0.462875	0.067866	-6.820398	0.0000
DEMOG4_PC1	0.136873	0.051649	2.650076	0.0084

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.975079	Mean dependent var	0.000878
Adjusted R-squared	0.948705	S.D. dependent var	1.247312
S.E. of regression	0.282497	Akaike info criterion	0.616102
Sum squared resid	26.01625	Schwarz criterion	2.938348
Log likelihood	138.9897	Hannan-Quinn criter.	1.515475
F-statistic	36.97145	Durbin-Watson stat	4.023952
Prob(F-statistic)	0.000000		

Table 2a

Dependent Variable: MIGPC1 spec12
Method: Panel Least Squares
Date: 11/06/21 Time: 11:46
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.038964	0.103584	-0.376160	0.7070
GDPPERCAPITA	-2.80E-05	1.11E-05	-2.520547	0.0122
VIOLENCE_PC1	0.161798	0.067600	2.393470	0.0173
CELLPHONES	1.70E-06	6.87E-07	2.473592	0.0139
HEALTH_PC1	-1.414241	0.309402	-4.570882	0.0000
HOUSING_PC1	-3.389767	1.576668	-2.149957	0.0323
YOUTH_PC1	-0.469601	0.076413	-6.145586	0.0000
DEMOG4_PC1	0.139519	0.050805	2.746194	0.0064
INFRA_PC1	0.077497	0.025078	3.090308	0.0022
MANAGEMENT	-0.001912	0.002156	-0.886680	0.3759

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.975985	Mean dependent var	0.000878
Adjusted R-squared	0.950265	S.D. dependent var	1.247312
S.E. of regression	0.278168	Akaike info criterion	0.585016
Sum squared resid	25.07028	Schwarz criterion	2.920686
Log likelihood	151.4346	Hannan-Quinn criter.	1.489587
F-statistic	37.94656	Durbin-Watson stat	4.023952
Prob(F-statistic)	0.000000		

Table 2b

Dependent Variable: MIGPC1 spec 12a with econop_income

Method: Panel Least Squares
Date: 11/08/21 Time: 16:39
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.149287	0.092456	-1.614678	0.1074
ECONOPINC_PC1	-0.083017	0.040523	-2.048673	0.0413
VIOLENCE_PC1	0.152377	0.067521	2.256751	0.0247
CELLPHONES	1.74E-06	6.92E-07	2.516822	0.0123
HEALTH_PC1	-1.492962	0.315997	-4.724613	0.0000
HOUSING_PC1	-3.205614	1.579737	-2.029207	0.0433
YOUTH_PC1	-0.460796	0.077181	-5.970347	0.0000
DEMOG4_PC1	0.138902	0.051028	2.722067	0.0068
INFRA_PC1	0.076316	0.025169	3.032214	0.0026
MANAGEMENT	-0.002063	0.002176	-0.948034	0.3438

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.975827	Mean dependent var	0.000878
Adjusted R-squared	0.949938	S.D. dependent var	1.247312
S.E. of regression	0.279080	Akaike info criterion	0.591564
Sum squared resid	25.23498	Schwarz criterion	2.927234
Log likelihood	149.2344	Hannan-Quinn criter.	1.496135
F-statistic	37.69280	Durbin-Watson stat	4.023952
Prob(F-statistic)	0.000000		

Table 3a (no infrastructure)

Dependent Variable: MIGPC1 spec14
Method: Panel Least Squares
Date: 11/06/21 Time: 11:56
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.064291	0.104608	-0.614588	0.5393
GDPPERCAPITA	-2.81E-05	1.13E-05	-2.496903	0.0130
VIOLENCE_PC1	0.160109	0.068481	2.338003	0.0200
CELLPHONES	1.76E-06	6.95E-07	2.537697	0.0116
HEALTH_PC1	-1.428324	0.313412	-4.557343	0.0000
MANAGEMENT	-0.001815	0.002184	-0.831075	0.4065
HOUSING_PC1	-3.904777	1.588325	-2.458425	0.0145
YOUTH_PC1	-0.438300	0.076728	-5.712371	0.0000
DEMOG4_PC1	0.137233	0.051463	2.666641	0.0080

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.975277	Mean dependent var	0.000878
Adjusted R-squared	0.948956	S.D. dependent var	1.247312
S.E. of regression	0.281803	Akaike info criterion	0.611089
Sum squared resid	25.80924	Schwarz criterion	2.940047
Log likelihood	141.6740	Hannan-Quinn criter.	1.513061
F-statistic	37.05383	Durbin-Watson stat	4.023952

Table 4a (GDP per capita and GDPQ1_18 Dummy variable spec 13)

Dependent Variable: MIGPC1 spec 13 with dgdpq25_18
Method: Panel Least Squares
Date: 12/15/21 Time: 16:19
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.131569	0.079576	-1.653377	0.0992
GDPPERCAPITA	-1.94E-05	1.09E-05	-1.779821	0.0760
VIOLENCE_PC1	0.134996	0.066210	2.038921	0.0423
CELLPHONES	1.71E-06	6.67E-07	2.561278	0.0109
HEALTH_PC1	-1.385601	0.301770	-4.591585	0.0000
HOUSING_PC1	-2.170847	1.543931	-1.406051	0.1607
YOUTH_PC1	-0.334985	0.070309	-4.764498	0.0000
DEMOG4_PC1	0.110748	0.049896	2.219583	0.0271
DGDPQ25_18	0.231429	0.045924	5.039406	0.0000

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.977020	Mean dependent var	0.000878
Adjusted R-squared	0.952555	S.D. dependent var	1.247312
S.E. of regression	0.271687	Akaike info criterion	0.537975
Sum squared resid	23.98954	Schwarz criterion	2.866932
Log likelihood	166.2406	Hannan-Quinn criter.	1.439946
F-statistic	39.93576	Durbin-Watson stat	4.023952
Prob(F-statistic)	0.000000		

Table 4b (Spec 13 DGDPQ25_18 dummy interaction with income)

Dependent Variable: MIGPC1
Method: Panel Least Squares
Date: 12/16/21 Time: 14:10
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.138790	0.079622	-1.743121	0.0823
GDPPERCAPITA	-1.90E-05	1.09E-05	-1.741671	0.0825
VIOLENCE_PC1	0.136499	0.066174	2.062745	0.0399
CELLPHONES	1.74E-06	6.67E-07	2.608226	0.0095
HEALTH_PC1	-1.400888	0.301666	-4.643841	0.0000
HOUSING_PC1	-2.302259	1.538556	-1.496376	0.1355
YOUTH_PC1	-0.343001	0.069687	-4.922017	0.0000
DEMOG4_PC1	0.110259	0.049895	2.209834	0.0278
GDPPERCAPITA*DGDPQ25_18	8.64E-05	1.71E-05	5.053627	0.0000

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.977029	Mean dependent var	0.000878
Adjusted R-squared	0.952575	S.D. dependent var	1.247312
S.E. of regression	0.271632	Akaike info criterion	0.537565
Sum squared resid	23.97971	Schwarz criterion	2.866523
Log likelihood	166.3782	Hannan-Quinn criter.	1.439537
F-statistic	39.95250	Durbin-Watson stat	4.023952
Prob(F-statistic)	0.000000		

Table 4c (Spec 13 using econopinc_pc1 and DGDPQ25_18 dummy)

Dependent Variable: MIGPC1 spec 13 econ op
with dgdpq25
Method: Panel Least Squares
Date: 12/15/21 Time: 16:22
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.210435	0.068265	-3.082635	0.0022
ECONOPINC_PC1	-0.055857	0.039437	-1.416352	0.1576
VIOLENCE_PC1	0.127911	0.066001	1.938014	0.0535
CELLPHONES	1.73E-06	6.71E-07	2.582985	0.0102
HEALTH_PC1	-1.436482	0.307425	-4.672620	0.0000
HOUSING_PC1	-2.004520	1.543278	-1.298871	0.1949
YOUTH_PC1	-0.329120	0.070412	-4.674211	0.0000
DEMOG4_PC1	0.109793	0.050041	2.194075	0.0289
DGDPQ25_18	0.234961	0.045895	5.119511	0.0000

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.976938	Mean dependent var	0.000878
Adjusted R-squared	0.952387	S.D. dependent var	1.247312
S.E. of regression	0.272169	Akaike info criterion	0.541521
Sum squared resid	24.07476	Schwarz criterion	2.870479
Log likelihood	165.0490	Hannan-Quinn criter.	1.443493
F-statistic	39.79106	Durbin-Watson stat	4.023952
Prob(F-statistic)	0.000000		

Table 4d (Spec 13 using econopinc_pc1 and DGDPQ25_18 dummy interaction with income)

Dependent Variable: MIGPC1
Method: Panel Least Squares
Date: 12/16/21 Time: 14:18
Sample: 2018 2019
Periods included: 2
Cross-sections included: 339
Total panel (unbalanced) observations: 672

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.214905	0.069574	-3.088867	0.0022
ECONOPINC_PC1	-0.067842	0.040039	-1.694396	0.0911
VIOLENCE_PC1	0.140280	0.067164	2.088599	0.0375
CELLPHONES	1.76E-06	6.83E-07	2.575681	0.0104
HEALTH_PC1	-1.474621	0.313123	-4.709406	0.0000
HOUSING_PC1	-2.399240	1.573643	-1.524640	0.1283
YOUTH_PC1	-0.373874	0.070985	-5.266961	0.0000
DEMOG4_PC1	0.115683	0.051041	2.266455	0.0241
DGDPQ25_18*ECONOPINC_PC1	-0.136241	0.037477	-3.635334	0.0003

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.976052	Mean dependent var	0.000878
Adjusted R-squared	0.950557	S.D. dependent var	1.247312
S.E. of regression	0.277348	Akaike info criterion	0.579220
Sum squared resid	24.99968	Schwarz criterion	2.908177
Log likelihood	152.3822	Hannan-Quinn criter.	1.481192
F-statistic	38.28415	Durbin-Watson stat	4.023952
Prob(F-statistic)	0.000000		

B. CEO Project Departments only (6 Departments)

Table 1a

Dependent Variable: MIGPC1 spec13_ceo6_only
Method: Panel Least Squares
Date: 11/06/21 Time: 13:29
Sample: 2018 2019 IF CEOATTENDED=1
Periods included: 2
Cross-sections included: 132
Total panel (unbalanced) observations: 263

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.518407	0.385298	-1.345472	0.1809
GDPPERCAPITA	-6.73E-05	2.90E-05	-2.325642	0.0217
VIOLENCE_PC1	0.442348	0.195806	2.259115	0.0256
CELLPHONES	3.00E-06	1.30E-06	2.300533	0.0231
HEALTH_PC1	-2.382806	0.793466	-3.003033	0.0032
HOUSING_PC1	-8.598438	4.485067	-1.917126	0.0575
YOUTH_PC1	-0.894493	0.159447	-5.609981	0.0000
DEMOG4_PC1	0.230578	0.108819	2.118902	0.0361

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.973737	Mean dependent var	0.392233
Adjusted R-squared	0.944509	S.D. dependent var	1.606523
S.E. of regression	0.378442	Akaike info criterion	1.199653
Sum squared resid	17.75906	Schwarz criterion	3.087598
Log likelihood	-18.75439	Hannan-Quinn criter.	1.958373
F-statistic	33.31497	Durbin-Watson stat	3.984848
Prob(F-statistic)	0.000000		

Table 1b

Dependent Variable: MIGPC1 spec13_ceo6_only_ econop_income

Method: Panel Least Squares
Date: 11/14/21 Time: 12:38
Sample: 2018 2019 IF CEOATTENDED=1
Periods included: 2
Cross-sections included: 132
Total panel (unbalanced) observations: 263

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.825563	0.360324	-2.291168	0.0236
ECONOPINC_PC1	-0.181716	0.108746	-1.671010	0.0972
VIOLENCE_PC1	0.415994	0.197288	2.108560	0.0370
CELLPHONES	3.06E-06	1.35E-06	2.267296	0.0251
HEALTH_PC1	-2.555407	0.818123	-3.123498	0.0022
HOUSING_PC1	-8.756618	4.574982	-1.914022	0.0579
YOUTH_PC1	-0.875144	0.164028	-5.335324	0.0000
DEMOG4_PC1	0.229282	0.110266	2.079358	0.0396

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.973195	Mean dependent var	0.392233
Adjusted R-squared	0.943364	S.D. dependent var	1.606523
S.E. of regression	0.382326	Akaike info criterion	1.220078
Sum squared resid	18.12552	Schwarz criterion	3.108022
Log likelihood	-21.44025	Hannan-Quinn criter.	1.978798
F-statistic	32.62325	Durbin-Watson stat	3.984848
Prob(F-statistic)	0.000000		

11. Principal Components Analysis

Table 1: Immigration Latent Variable (Proxy)

Principal Components Analysis

Date: 10/23/21 Time: 13:30
Sample: 2018 2019
Included observations: 673
Balanced sample (listwise missing value deletion)
Computed using: Ordinary correlations
Extracting 2 of 2 possible components

Eigenvalues: (Sum = 2, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	1.551683	1.103366	0.7758	1.551683	0.7758
2	0.448317	---	0.2242	2.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2
REMITTANCES	0.707107	-0.707107
RETURNEES	0.707107	0.707107

Ordinary correlations:

	REMITTANCES	RETURNEES
REMITTANCES	1.000000	
RETURNEES	0.551683	1.000000

Table 2a: Economic Opportunity Income

Principal Components Analysis
Date: 10/23/21 Time: 14:42
Sample: 2018 2019
Included observations: 673
Balanced sample (listwise missing value deletion)
Computed using: Ordinary correlations
Extracting 4 of 4 possible components

Eigenvalues: (Sum = 4, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	2.304580	1.532176	0.5761	2.304580	0.5761
2	0.772404	0.213542	0.1931	3.076984	0.7692
3	0.558862	0.194709	0.1397	3.635846	0.9090
4	0.364154	---	0.0910	4.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3	PC 4
GDPPERCAPITA	0.545950	0.022540	-0.542348	0.638192
POVERTY	-0.528739	0.438378	0.276783	0.672050
LABORFORCE	0.414507	0.856685	0.120596	-0.282368
GRADUATES	0.500558	-0.270938	0.784032	0.247646

Ordinary correlations:

	GDPPERCAPITA	POVERTY	LABORFORCE	GRADUATES
GDPPERCAPITA	1.000000			
POVERTY	-0.585328	1.000000		
LABORFORCE	0.434267	-0.265456	1.000000	
GRADUATES	0.444992	-0.519797	0.326260	1.000000

Table 2b: Economic Opportunity Wealth

Principal Components Analysis
Date: 10/23/21 Time: 14:48
Sample: 2018 2019
Included observations: 673
Balanced sample (listwise missing value deletion)
Computed using: Ordinary correlations
Extracting 6 of 6 possible components

Eigenvalues: (Sum = 6, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	3.995071	2.975868	0.6658	3.995071	0.6658
2	1.019203	0.047877	0.1699	5.014274	0.8357
3	0.971326	0.960676	0.1619	5.985599	0.9976
4	0.010650	0.007830	0.0018	5.996250	0.9994
5	0.002820	0.001890	0.0005	5.999070	0.9998
6	0.000930	---	0.0002	6.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
SAVINGSACC	0.498580	-0.033804	-0.011300	0.696236	0.325245	0.399524
SAVINGSAM	0.499581	-0.026753	0.002170	0.098229	-0.854266	-0.101384
DEPOSITSACC	0.499876	-0.024964	-0.007329	-0.089540	0.374800	-0.775212
DEPOSITSAM	0.498725	-0.026913	-0.000732	-0.705353	0.154585	0.478674
QUANTIHOUSING	0.034789	0.723306	-0.689625	-0.000882	-0.004582	0.003547
ELECTRIFIC	0.044984	0.688202	0.724037	0.008112	0.007223	0.002555

Ordinary correlations:

	SAVINGSACC	SAVINGSAM	DEPOSITSACC	DEPOSITSAM	QUANTIHOUSING	ELECTRIFIC
SAVINGSACC	1.000000					
SAVINGSAM	0.995902	1.000000				
DEPOSITSACC	0.996016	0.997425	1.000000			
DEPOSITSAM	0.989417	0.994964	0.997155	1.000000		
QUANTIHOUSING	0.051934	0.048267	0.055974	0.049971	1.000000	
ELECTRIFIC	0.058011	0.072533	0.067168	0.070178	0.028594	1.000000

Table 3: Health Status

Principal Components Analysis
Date: 10/23/21 Time: 14:09
Sample: 2018 2019
Included observations: 672
Balanced sample (listwise missing value deletion)
Computed using: Ordinary correlations
Extracting 3 of 3 possible components

Eigenvalues: (Sum = 3, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	2.016657	1.045800	0.6722	2.016657	0.6722
2	0.970857	0.958371	0.3236	2.987514	0.9958
3	0.012486	---	0.0042	3.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3
CHRONICMAL	-0.168755	0.985319	0.025831
EXPHEALTH	0.695468	0.137601	-0.705259
SSAFFILIATES	0.698460	0.101052	0.708479

Ordinary correlations:

	CHRONICMAL	EXPHEALTH	SSAFFILIATES
CHRONICMAL	1.000000		
EXPHEALTH	-0.105280	1.000000	
SSAFFILIATES	-0.140806	0.986865	1.000000

Table 4: Violence

Principal Components Analysis
Date: 10/23/21 Time: 14:23
Sample: 2018 2019
Included observations: 678
Balanced sample (listwise missing value deletion)
Computed using: Ordinary correlations
Extracting 4 of 4 possible components

Eigenvalues: (Sum = 4, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	3.358873	2.866349	0.8397	3.358873	0.8397
2	0.492525	0.402874	0.1231	3.851398	0.9628
3	0.089651	0.030700	0.0224	3.941049	0.9853
4	0.058951	---	0.0147	4.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3	PC 4
INFAMVIOFEM	0.517038	0.335885	-0.569923	-0.543177
INFAMVIOMALE	0.478109	0.647988	0.390151	0.446437
EXTORTIONS	0.505926	-0.443231	0.604517	-0.426784
HOMICIDES	0.498119	-0.520422	-0.396901	0.568778

Ordinary correlations:

	INFAMVIOFEM	INFAMVIOMALE	EXTORTIONS	HOMICIDES
INFAMVIOFEM	1.000000			
INFAMVIOMALE	0.903283	1.000000		
EXTORTIONS	0.788079	0.680926	1.000000	
HOMICIDES	0.781037	0.634926	0.924263	1.000000

Table 5: Civic Engagement (proxy for strength of civil society, social capital)

Principal Components Analysis
Date: 10/23/21 Time: 13:40
Sample: 2018 2019
Included observations: 680
Computed using: Ordinary correlations
Extracting 2 of 2 possible components

Eigenvalues: (Sum = 2, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value
1	1.208017	0.416034	0.6040	1.208017
2	0.791983	---	0.3960	2.000000

Eigenvectors (loadings):

Variable	PC 1	PC 2
MANAGEMENT	0.707107	-0.707107
VOTERS	0.707107	0.707107

Ordinary correlations:

	MANAGEMENT	VOTERS
MANAGEMENT	1.000000	
VOTERS	0.208017	1.000000

Table 6: Drought / Climate

Principal Components Analysis
Date: 11/11/21 Time: 13:13
Sample: 2018 2019
Included observations: 680
Computed using: Ordinary correlations
Extracting 2 of 2 possible components

Eigenvalues: (Sum = 2, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	1.119927	0.239855	0.5600	1.119927	0.5600
2	0.880073	---	0.4400	2.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2
DROUGHT_PRECIP	0.707107	-0.707107
DROUGHT_TEMP	0.707107	0.707107

Ordinary correlations:

	DROUGHT_PRECIP	DROUGHT_TEMP
DROUGHT_PRECIP	1.000000	
DROUGHT_TEMP	0.119927	1.000000

Table 7a: Youth Population

Principal Components Analysis
Date: 11/06/21 Time: 10:58
Sample: 2018 2019
Included observations: 680
Computed using: Ordinary correlations
Extracting 2 of 2 possible components

Eigenvalues: (Sum = 2, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	1.500688	1.001375	0.7503	1.500688	0.7503
2	0.499312	---	0.2497	2.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2
MEDIANAGE	0.707107	-0.707107
GRADUATES	0.707107	0.707107

Ordinary correlations:

	MEDIANAGE	GRADUATES
MEDIANAGE	1.000000	
GRADUATES	0.500688	1.000000

Table 7b: Population Characteristics – Urban/Density

Principal Components Analysis
Date: 11/06/21 Time: 11:31
Sample: 2018 2019
Included observations: 673
Balanced sample (listwise missing value deletion)
Computed using: Ordinary correlations
Extracting 2 of 2 possible components

Eigenvalues: (Sum = 2, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	1.371197	0.742394	0.6856	1.371197	0.6856
2	0.628803	---	0.3144	2.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2
URBANPOP	0.707107	-0.707107
POPDENSITY	0.707107	0.707107

Ordinary correlations:

	URBANPOP	POPDENSITY
URBANPOP	1.000000	
POPDENSITY	0.371197	1.000000

Table 8: Local Housing Conditions

Principal Components Analysis
Date: 10/23/21 Time: 14:12
Sample: 2018 2019
Included observations: 679
Balanced sample (listwise missing value deletion)
Computed using: Ordinary correlations
Extracting 2 of 2 possible components

Eigenvalues: (Sum = 2, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	1.845123	1.690246	0.9226	1.845123	0.9226
2	0.154877	---	0.0774	2.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2
QUALIHOUSING	0.707107	-0.707107
QUANTIHOUSING	0.707107	0.707107

Ordinary correlations:

	QUALIHOUSING	QUANTIHOUSING
QUALIHOUSING	1.000000	
QUANTIHOUSING	0.845123	1.000000

Table 9: Local Infrastructure

Principal Components Analysis
Date: 10/23/21 Time: 14:21
Sample: 2018 2019
Included observations: 673
Balanced sample (listwise missing value deletion)
Computed using: Ordinary correlations
Extracting 3 of 3 possible components

Eigenvalues: (Sum = 3, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	1.834670	1.238238	0.6116	1.834670	0.6116
2	0.596433	0.027536	0.1988	2.431103	0.8104
3	0.568897	---	0.1896	3.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3
WATER	0.572459	0.769333	0.283579
SANITATION	0.583236	-0.138982	-0.800325
ELECTRIFIC	0.576304	-0.623546	0.528264

Ordinary correlations:

	WATER	SANITATION	ELECTRIFIC
WATER	1.000000		
SANITATION	0.419671	1.000000	
ELECTRIFIC	0.404382	0.427840	1.000000