

Ethnicity Reconstruction:

The Effect of Immigration on Ethnic-Linguistic Fractionalization in Latin America

Research in the field of ethnic politics is especially important in the context of Latin America, a region that continues to struggle with the incorporation and increasing participation of ethnic minorities, particularly Indigenous peoples, in the regional and national political arenas. Unfortunately, the correct approach to ethnicity research remains an unresolved contention between the Primordialist and Constructivist models. This study takes a crack at resolving this issue by empirically examining the effects of immigration on ethnic-linguistic fractionalization in 20 Latin American countries. The results indicate that immigrant ethnic identities are both fluid and endogenous, thereby validating the Constructivist model.

Holly Williams
Senior Honors Thesis
NYU Politics Dept., 2007

I. Introduction

There is little doubt that the study of ethnic demographics is an essential component in any assessment of political stability or social and economic growth. What remains open to debate however, is exactly how ethnic groups are formulated. The idea that "ethnicity is socially constructed" is now commonplace among researchers, but whether or not ethnic identity can be actually re-constructed is the key distinguishing aspect between the Primordialist and Constructivist models.

In the symposium, "Cumulative Findings in the Study of Ethnic Politics," Stephen Van Evera presented the argument that while the constructivist claim that ethnic identities are socially constructed is obviously, and unquestionably correct, ethnic identities are quite hard to reconstruct once they are formed, and therefore the Primordialist approach to ethnic politics research and analysis –under which identity is fixed in the present and future- remains valid. In his analysis, Van Evera does acknowledge that there are rare occasions in which ethnicity reconstruction can take place, particularly during migration -as immigrant assimilation behavior and immigrant majority states are especially conducive to change. However, he also maintains that the ethnic identities of non-immigrant groups, even in immigrant-majority states, such as the Indigenous peoples of Latin America, have remained fixed.

This study aims to answer the question: how are immigrant groups -under a presumed high level of assimilation and fluid ethnic identification- affecting the overall ethnic-linguistic fractionalization of the countries in Latin America? If it is true that ethnic identity is reconstructed among immigrant groups, I hypothesized that there would be a fluctuation in ethnic-linguistic fractionalization that corresponds with diverse foreign-born

population sizes. In other words, I believed that countries with more immigrants would show less ethnic-linguistic fractionalization because immigrants *do* in fact reconstruct their identities and assimilate. Through this examination of immigrant groups in Latin America I also sought to answer the more theoretical questions laid out in the symposium regarding the correct method for ethnic politics research: Constructivist or Primordialist?

Further, limiting this study to the specific region of Latin American has led to more comprehensive questions regarding possible conditional changes affecting the assimilation behavior of immigrants in Indigenous-populated countries.

II. Literature Review

A definition of "ethnicity"

Before examining the two different models of ethnicity study, it is important to first define exactly what an "ethnic group" is. James D. Fearon writes that "a 'prototypical' ethnic group has the following features:

(a) Membership is reckoned primarily by descent by both members and non-members; (b) members are conscious of group membership and view it as normatively and psychologically important to them; (c) members share distinguishing cultural features; (d) these cultural features are valued by a majority of members; (e) the group has or at least 'remembers' a homeland; and (f) the group has a shared history as a group that is 'not wholly manufactured but has some basis in fact (g) the group is not a caste or caste-like group. (Country 201)

Although this definition is ripe with ambiguity, especially in regards to the rules of descent, it is the definition that will be used for this paper, partly because this very ambiguity allows of the inclusion of a broader range of ethnic groups that might otherwise have been excluded because of varying social definitions of what constitutes an "ethnic group."

Two models of Ethnicity Study

As for the individual identifications of the members of these "ethnic groups," the Primordialist model's basic assertion is that 1) individuals have a single ethnic identity and 2) this identity is fixed in the present and future. However, many primordialists have now come to allow for the idea that ethnic identities may have initially been "constructed through human intervention" (Chandra, Cumulative 7) but still maintain that once acquired, an individual's identity is fixed in the long term. Further, this fixity implies that ethnic groups serve as exogenous variables because they were created prior to political, social and economic results.

The Constructivist model, on the other hand, asserts that 1) individuals have multiple ethnic identities and 2) the identity with which individuals identify fluctuates depending upon a particular causal variable. In more general terms, Joane Nagel writes that "one's ethnic identity is a composite of the view one has of oneself as well as the views held by others about one's ethnic identity. As the individual (or group) moves through daily life, ethnicity can change according to variations in the situations and audiences encountered" (Constructing 154). According to this model, ethnic groups are thereby endogenous to a set of political, economic and social processes; ethnic identification (and therefore ethnic demography) will change in accordance with changes in these causal variables.

Most social scientists have come to discredit the Primordialist model on the basis that "identities" are in fact "social categories" and these categories are hardly natural, inevitable, nor unchanging, as primordialists have argued. On the contrary, constructivists point out that content and membership rules of these categories have changed over time. (Fearon, Violence 848) Further, in her paper, "What is Ethnic Identity and Does it Matter?"

Kanchan Chandra argues that "it can be easy for individuals to switch between ethnic identities within the constraint of the underlying set of attributes" (Identity 418). Chandra identifies these constraints of membership in social categories as attributes associated – or believed to be associated- with descent. These descent-based attributes can be genetically inherited (such as skin color and other physical features) or culturally and historically inherited (such as name, language, and origin of one's parents and ancestors). Attributes that are "believed to be associated with descent" are the product of a "credible myth of association" (Chandra, Identity 400) that may even reflect a non-existent connection. The placement of these "descent-based attributes" in particular ethnic categories depends on how they are processed and coded by society: the groups are socially constructed. An individual's skin color, for example, may be interpreted as "black" in the United States, but "brown" in Brazil, "the difference is the result different historical and institutional context that have taught individuals to instinctively code the same data in different ways" (Chandra, Identity 416).

While many descent based attributes are displayed on an individual's body, and thus unalterable, "individuals can change between identity categories, often quite rapidly, by combining and recombining elements from their set of attributes differently" (Chandra, Identity 415). In other words, an individual can choose to emphasize certain visible attributes over others, depending on their desire to identify, and be identified, with a certain ethnic group by a particular audience. A prime example of this ability for ethnicity construction and reconstruction would be a black person in the United States originally from Jamaica. This individual could choose to emphasize his/her "Caribbean" ancestry, could voluntarily place themselves in the broad category of "African American," or could emphasize their birth in a former British-colony and English speaking country, as a "West

Indian." The "constrained-change" element here is apparent: this individual can only reconstruct his/her identity under a limited range of categories based on descent-based attributes; he/she cannot be identified as Swedish (if "Swedish" is based on birth), for example (Chandra, Identity 416).

Immigration and Ethnicity

According to the argument set forward by Daniel Byman, immigrant ethnic groups in particular show an exceptionally high level of fluidity and endogeneity because 1) immigrants accept the duty to compromise their culture when immigrating into another culture 2) individuals that leave their homeland voluntarily have weaker ethnic identities and 3) states that allow more immigration have more liberal and tolerant cultures and thus, the immigrants will be treated well and their identities unhardened by oppression (Van Evera 21). Under this assertion, immigrants show a lesser degree of "constrained change" than non-immigrants (i.e. face weaker/fewer constraints). Kanchan Chandra argues this point as well, stating that some descent-based attributes such as last name and birth place of one's ancestors can be changed "especially in the context of relocation or migration, when new histories can more easily be invented" (Identity 415).

In his paper, "Ethnic Structure and Cultural Diversity Around the World," James Fearon also discusses to the effect of immigration on ethnic demographics,

Latin America and the Caribbean are notable for the high proportion of the countries that are approximately partitioned between a majority group and a single minority group, usually 'mestizos' (or 'whites') and 'Indigenous peoples...Indigenous peoples...that were historically divided among many smaller tribes speaking diverse languages. A long history of assimilation and the numerical and political dominance of the settler populations has blurred these distinctions and made the common sense ethnic categories in many of these countries 'Indigenous' versus 'white/mestizo.' Exceptions are Guatemala and the Andean countries Bolivia, Peru, and Ecuador which are coded as having large Indigenous populations, along with noteworthy distinctions between whites and mestizos. (World 207)

Fearon reasons that the increase and dominance of the immigrant populations in most Latin American countries has resulted in the formation of two more general "catch-all" categories and a marked bi-polarization of ethnic demographics in the region.

This paper attempts to test the Constructivist model of ethnic identification by examining immigration -a social, political and economic process that should serve as a causal variable affecting the ethnic fractionalization of Latin American countries. Under Daniel Byman's assumption that immigrants are highly disposed to assimilate and reconstruct their identities, the countries examined in this study should demonstrate a negative correlation between the size of their immigrant populations and their level of ethnic fractionalization if the Constructivist approach is valid. This study should shed more light on the current methodological debate between ethnic politics researchers like James Fearon, Joane Nagel, Stephen Van Evera and Kanchan Chandra, as well as the other contributors to the symposium, "Cumulative Findings in the Study of Ethnic Politics."

III. Research Design

In Stephen Van Evera's paper he defends Primordialism by claiming that constructivists don't take into account the fact that the ethnic identities of the non-immigrant populations (even in immigrant-majority countries) are durable. He claims that even if it is assumed that immigrant population identities are fluid and endogenous, according to Daniel Byman's analysis of immigrant behavior, "immigrant states are different to start with, having aspects that are conducive to change" (Van Evera 3) and

therefore serve as one of the “rare conditions” needed for ethnicity reconstruction –an exception to the rule. I will attempt to address this question of immigrant assimilation in this study through empirical analysis, which is lacking in Van Evera's piece.

This research paper aims to empirically prove, through linear regression analysis, that countries with large immigrant populations are in fact an “exception to the rule,” and will show *less* ethnic-linguistic fractionalization because immigrants are highly disposed to assimilate and reconstruct their identities. Immigration will have a *negative* effect on ethnic-linguistic fractionalization, rather than *no* effect (which would still imply immigrant assimilation), because there is an assumption that immigrants assimilate into pre-existing majority ethnic groups. The basis of this assumption will be explained later in the paper.

The relationship between immigration and ethnic-linguistic fractionalization is expressed in the following testable hypothesis:

Null Hypothesis I: *Immigration will have a positive effect on ethnic-linguistic fractionalization.*

Alternative Hypothesis I: *Latin American countries with larger foreign-born populations will show lower degrees of ethnic-linguistic fractionalization, all other things held constant (ceteris paribus).*

Under the alternative hypothesis, the larger the size of the foreign-born population in a country, the less ethnic-linguistic fractionalization that country should show if indeed ethnic identities are endogenous to social processes and change over time. To fully comprehend this negative correlation between immigration and ethnic-linguistic

fractionalization, it is vital to understand just how ethnic-linguistic fractionalization is generated. The following ELF formula was first published in the Atlas Narodov Mira (1964) and later popularized by Charles Taylor and Michael Hudson (1972):

$$ELF_j = 1 - \sum_{i=1}^N s_{ij}^2$$

(where s_{ij} is the share of group i out of a total of n groups in country j)

In accordance with the ELF formula, if foreign-born populations in Latin American countries were *not* assimilating, then the s_{ij} , or the group shares of all non-immigrant ethnic groups would decrease in order to compensate for the rise in immigrant ethnic group shares, thereby *increasing* the ethnic-linguistic fractionalization of the countries. The chart below provides examples of how ethnic group shares function in the ELF formula:

Chart A: Fractionalization Examples
(Fearon, World 22)

Country	Structure	F
A	Perfectly homogeneous	0
B	2 groups, (.95, .05)	.10
C	2 groups, (.8, .2)	.32
D	2 groups, (.5, .5)	.50
E	3 groups, (.33, .33, .33)	.67
F	3 groups, (.55, .30, .15)	.59
G	3 groups, (.75, .20, .05)	.40
H	(.48, .01, .01, ...)	.76
I	(.25, .25, .25, .25)	.75
J	n groups, ($1/n, 1/n, \dots$)	$1 - (1/n)$

In following with the ELF formula, the Foreign-born population variable would need only to display zero effect on ethnic-linguistic fractionalization in order to prove that immigrants are assimilating (non-immigrant ethnic group shares would not be decreasing). If ELF actually *decreases* with increased immigration, one should reason that immigrants are assimilating into pre-existing majority ethnic groups (increasing the group share of the majority to the degree that the probability that any two people will be from a different ethnic group actually *decreases*). For example, the country of Cuba is reported by the CIA World Factbook to be 51% mulatto, 37% white and 11% black. If immigrants in Cuba assimilate into the majority mulatto group then they would increase that group's share (to maybe 75% of the population) and thereby *decrease* ELF by roughly 9% (as shown in countries F and G in **Chart A**). Theoretically, it is possible that ELF would *increase* if immigrants were assimilating into pre-existing minority ethnic groups, thereby raising the group share of blacks in Cuba, for example. If the Cuban blacks' group share increased from 11% to 33%, then the other two groups' shares would compensate and decrease, and then Cuba would be transformed into country E (as shown in **Chart A**) with a notable *increase* in ELF by around 8%.

By hypothesizing that ethnic-linguistic fractionalization *decreases* with an increase in immigration, there is an assumption that immigrants are assimilating into pre-existing *majority* ethnic groups. This assumption was based on existing literature by authors such as James Fearon, who claim there is an increasingly marked ethnic bi-polarization (between mestizo/white and Indigenous peoples) in many Latin American countries. This assumption is also consistent with the argument made by Daniel Byman that immigrants are prone to reconstruct their identities because they assimilate into countries that treat immigrants well, so their identities are unhardened by oppression. It is safe to assume

that immigrants would not choose to assimilate into minority ethnic groups that face more political and social obstacles than ethnic majorities; there would have been no motivation to assimilate out of their own immigrant minority ethnic group.

This study's findings are particularly relevant to Latin America, because many of the countries with large Indigenous populations and a high degree of ethnic fragmentation, such as Mexico, Peru and Guatemala, also show low levels of immigration, which in a broader view, could mean that the antiquated idea that Indigenous groups would eventually gain political and social access through assimilation with white populations rather than through continued political conflict, has virtually no chance. Not to mention the fact that while many immigrant populations do compromise and blend with other settlers, they often deal harshly with Indigenous peoples (Van Evera 3).

IV. Data Description

Table 1 contains the descriptive statistics for the variables in the dataset. It is important to note that the countries used in the dataset are the 20 main countries of Latin America, including Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, Haiti, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, The Dominican Republic, Uruguay, and Venezuela. Suriname, Guyana, French Guiana or Belize were not included in the dataset because they are not considered part of “Latin America” according to the Demographic Bulletin of Latin America, constructed by the Economic Commission of Latin America and the Caribbean (ECLAC).

The dependent variable, Ethnic-Linguistic Fractionalization, is the probability that any two citizens from a given country will be drawn from a different ethnic group. The ELF index was constructed in 1961 and 1985 by Phillip Roeder, using the Taylor and Hudson formula. Consequently, my time-series-cross-section dataset only includes just the two reported time periods. It is also important to note that Roeder uses “none of the groupings reported in the sources when data on sub-groups are available...” For example, he does not use catch-all “Indigenous peoples” groups, but identifies racial distinctions within ethno-linguistic groups such as Afro-Colombian vs. Euro-Colombians. In all of the models of this study the elf dependent variable was transformed into $\text{elf}2 = \log \text{elf}/(1-\text{elf})$, in order to allow variation in the dependent variable along the real line.

The causal or independent variable, Foreign-born percentage of total country population, was formulated by the Demographic Bulletin published by ECLAC (Economic Commission for Latin America and the Caribbean) under the United Nations and reflects population estimates for 1960 and 1985 (an average of the estimates for 1980 and 1990). The first control variable, Latitude, shown in **Model 2**, was included in the dataset in response to Daniel Posner's comments that "The further a country is from the equator (i.e. the higher the latitude) the lower its ethnic diversity" (14). The CIA World Factbook reported the mean latitude for the countries used in this study.

Illiteracy is another control variable included in the dataset, as shown in **Model 3**, and was taken from the published results of the UNESCO Institute for Statistics, and represents the estimated percentage of illiterate adults 15-24 years of age in 1970 and 1985. This variable was inspired by Stephen Van Evera, who identifies "mass literacy" as a factor affecting ethnic identification: "Ethnic identities harden when mass literacy is achieved... because mass literacy allows the identity to become stored in writing and

purveyed in common form to a mass audience" (1). Finally, it is also argued in a paper by Alberto Alesina that "Ethnic fragmentation is higher in poorer countries..."(157). Thus, Gross Domestic Product (GDP) per capita is also included in the dataset (**Model 4**). The G.D.P. variable is based on the Penn World Table's country reports for the years of 1961 and 1985.

An ethnic conflict variable is also included in the dataset (**Model 5**) in response to Stephen Van Evera's claim that violent conflict hardens ethnic identities "by enhancing the emotional impact of recorded national memories" (2). The ethnic conflict variable is a scaled combination of the protest levels and rebellion levels compiled by the Minority At Risk Project for the time periods of 1950-1969 and 1970-1989, in order to reflect any pre-conflict buildup and aftermath effects. The MAR protest index is on a scale from 0 to 5 (0 representing no reported incidents and 5 representing large demonstrations of greater than 100,000 people). The conflict index is on a scale of 0 to 7 (0 representing no reported incidents and 7 representing protracted civil war fought by rebel military units with base areas).

Including these control variables in the dataset takes into account the possibility that Illiteracy, Latitude, G.D.P., or Ethnic Conflict may be factors (besides immigration) affecting the ethnic-linguistic fractionalization in a given country in Latin America. Including these variables in my linear regression models allows me to discount other factors that may be causing variation in the dependent variable and pinpoint my independent variable (Foreign-born population) as the appropriate causal variable.

Causal Model I:

Estimation Equation: {where j is country and t is time period}

$$ELF_{jt} = \hat{\beta}_{jt} + \hat{\beta}_1 foreign_{jt} + \hat{\beta}_2 latitude_{jt} + \hat{\beta}_3 illiteracy_{jt} + \hat{\beta}_4 gdp_{jt} + \hat{\beta}_5 conflict_{jt}$$

$$H_0: \hat{\beta}_1 > 0 \quad H_1: \hat{\beta}_1 < 0$$

Table 1: Descriptive Statistics

Variable	Comments	# of Obs.	Mean	Std.Dev.	Min.	Max.
elf	Ethnic-Linguistic Fractionalization (i.e. the probability that any two people will be drawn from a different ethnic-linguistic group)	40	0.43	0.20	0.01	0.76
elf2	elf variable transformed $z = \log \text{elf}/(1-\text{elf})$	40	0.63	0.37	0.01	1.42
foreign	Foreign-born percentage of total population	35	2.47	2.64	0.30	12.70
latitude	Reported latitude of CIA World Factbook	40	16.49	8.85	2.00	34.00
illiteracy	Percentage illiterate of total population, 15-24 year olds	40	25.15	17.29	4.20	78.00
gdp	Gross Domestic Product per capita	38	2340.37	1774.33	360.03	6429.76
conflict	Scaled combination of protest and rebellion indices	32	2.94	3.10	0.00	11.00
protest	Highest observed level of protest	32	1.81	1.33	0.00	4.00
rebellion	Highest observed level of rebellion	32	0.69	1.47	0.00	6.00

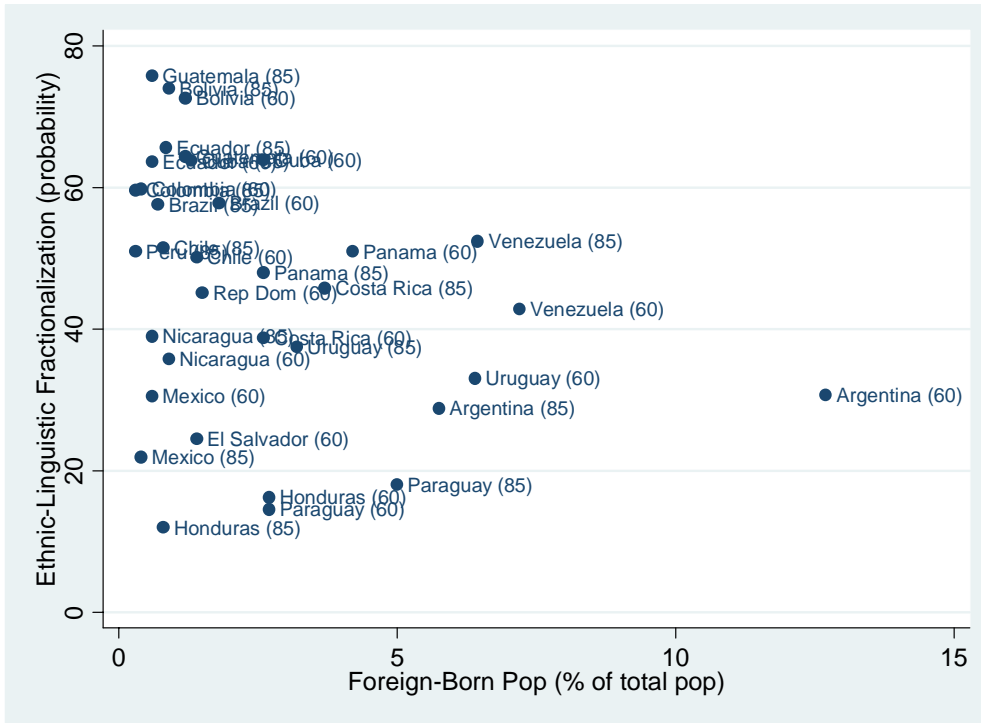
ind_pop	Indigenous population percentage of total population	32	15.23	21.77	0.20	71.20
ind_foreign	Interaction variable	30	15.19	20.18	0.14	75.60

**Table 2: Ethnic-Linguistic Fractionalization
(20 Countries of Latin America, 1960s and 1980s)**

ELF2	Model 1 Coefficient (Std. Err)	Model 2 Coefficient (Std. Err)	Model 3 Coefficient (Std. Err)	Model 4 Coefficient (Std. Err)	Model 5 Coefficient (Std. Err)
Foreign Born Pop.	-.045*** (.015)	-.031** (.017)	-.035** (.017)	-.026* (.016)	-.013 (.016)
Latitude		-.010*** (.003)	-.011*** (.004)	-.010*** (.004)	-.012*** (.005)
Illiteracy			-.002 (.005)	.003 (.006)	.006 (.007)
G.D.P.				.000 (.000)	5.02e-06 (.000)
Ethnic Conflict					.014 (.217)
Constant	.776 (.082)	.909 (.076)	.970 (.137)	.778 (.196)	.719 (.217)
Number of Obs.	35	35	35	33	29
Prob > F	0.007	0.000	0.000	0.003	0.008
R-squared	0.114	0.176	0.181	0.223	0.269

Note: In each cell, the top number is the coefficient and the robust standard error is in parentheses.
 *** Denotes $p < 0.01$, ** $p < 0.05$, * $p < 0.10$; one-tailed test

Graph 1: Ethnic-Linguistic Fractionalization vs. Foreign-Born Population



V. Results Analysis

Table 2:

As shown in **Model 1**, when considered without any control variables the Foreign-born population variable is highly significant, at the $p < .01$ level. The Foreign-born population variable also displays a negative coefficient sign, which confirms that null hypothesis I, $H_0: \hat{\beta}_1 > 0$, should be rejected. Therefore, the larger the size of the immigrant population in a Latin American country the more ethnically homogenous that country is likely to be (keeping in mind the assumption that immigrants assimilate into majority ethnic groups). The Foreign-born population variable maintains a negative coefficient in all of the models of **Table 2**, confirming that although the relationship becomes statistically insignificant in **Model 5**, immigration remains consistently negatively correlated with ethnic-linguistic fractionalization in Latin America.

It is important to note that **Model 1** does not include any control variables, which may account for the extremely high confidence level ($p < .01$). Not surprisingly, **Model 1** displays a very weak predictive power (R-squared), at about 11%. The best R-squared value in terms of the Foreign-born population – as the predicted causal variable- is in **Model 3**, displaying a 22% predictive power while still maintaining a 90% confidence level in the negative relationship between Foreign-born population and ethnic-linguistic fractionalization.

The small standard errors shown in all of the models of **Table 2** imply that there is little sampling fluctuation in the data, which is important for a confident rejection of null hypothesis I. Further, all of the models in **Table 2** also demonstrate a roughly 0 to 1% probability that the predicted values will fall outside the F-test. Thus, there is an extremely high probability that the variances of the data sets are statistically different, which also supports a confident rejection of null hypothesis I in all five models.

Model 2 reveals that adding a second independent variable, Latitude, decreases the significance of the relationship between Foreign-born population and Ethnic-Linguistic Fractionalization to the 5% level. However, even with Latitude included, null hypothesis I should still be confidently rejected. Latitude, as a control variable, remains individually very significant, at $p < .01$, in all 5 models of **Table 2**. This implies that Latitude does have a significant effect on Ethnic-Linguistic Fractionalization, even when several other factors are accounted for. The negative coefficient sign of the Latitude variable proves that countries farther away from the equator are in fact more ethnically homogenous.

As more control variables are added in each model of **Table 2**, the confidence level of the correlation between Foreign-born population and Ethnic-Linguistic Fractionalization decreases. With the inclusion of Illiteracy, Foreign-born population

remains significant at the 5% level, but is then reduced to the 10% level with the inclusion of Gross Domestic Product in **Model 4**. The Foreign-born population variable finally becomes a statistically insignificant factor ($p > .10$) with the addition of the Ethnic Conflict variable in **Model 5**. Therefore, null hypothesis I should be confidently rejected up until **Model 5**: immigration is significantly negatively correlated with ethnic-linguistic fractionalization even when taking into account Latitude, Illiteracy and Gross Domestic Product, but not when factoring in Ethnic Conflict.

It is important to note that Illiteracy, G.D.P. and Ethnic Conflict are not individually significant variables under the T-test: they each display a confidence level less than 90%. The revelation that Illiteracy, Gross Domestic Product and Ethnic Conflict are not proving to be significantly correlated with the Ethnic-Linguistic Fractionalization of a given country in Latin America is especially compelling due to the fact that they were noted in published literature as being significant factors in the ethnic demographics of any given country. Thus, this project has revealed that in the region of Latin America, mass literacy does not reduce ethnic fragmentation, poorer countries are not more likely to be ethnically divided, and countries displaying high levels of ethnic conflict are not comparatively more ethnically fractionalized.

The findings of this study may be the result of working with such a specific region of the world. As shown in **Table 2**, the number of observations is quite low in all five models, which may be impacting the significance of the variables. Further, by only examining the 20 main countries of Latin America, it is impossible to make any global/broad statements from these findings. These results may also reflect unreliable and/or flawed data, as is common in researching a region lacking the resources to devote to demographic and sociological research.

Graph 1:

Graph 1 displays the cross-section results of **Model 1**; just by eye-balling the graph one can see an apparent negative correlation between immigration and Ethnic-Linguistic Fractionalization. The graph also paints a picture of the overall ethnic demography and immigration statistics of the region of Latin America. It comes as no surprise, based on basic general knowledge of the region, that countries such as Guatemala, Bolivia, Ecuador and Peru are clustered together with a high degree of Ethnic-Linguistic Fractionalization and a small Foreign-born population, whereas Argentina and Venezuela display both lower Ethnic-Linguistic Fractionalization and larger immigrant populations. The two time periods used in the study (1960s and 1980s), are shown in parentheses to the right of the country names in the graph, and as shown, there is very little change in ELF over-time, but a noticeable decline in immigration over-time in many cases. **Graph 1** also exhibits several outlier countries, such as Argentina and Venezuela. It was in part due to these outliers that the dependent variable was transformed -by finding $\log \text{elf}/(1-\text{elf})$ - in order to expand the variable range to ∞ .

VI. Experimental Implications

A hypothetical experiment based on **Model 1** –the most statistically significant model for $\hat{\beta}_1$ - has been employed below in order to better illustrate the findings of this study. The fact that immigrants were found to cause a *decline* in Ethnic-Linguistic Fractionalization has huge implications for the ethnic demographics and political and social future of the region of Latin America.

$$E\hat{L}F_{jt} = \hat{\beta}_0 + \hat{\beta}_1(\overline{foreign})_{jt}$$

$$E\hat{L}F_{jt} = .78 + -.045(2.47) = .78 - .11 = 0.67$$

The results of this experiment indicate that a country with a Foreign-born population of roughly 2.47% (the mean Foreign-born population score) will also demonstrate an estimated 67% probability that any two persons in that country will be drawn from a different ethnic-linguistic group.

$$E\hat{L}F_{jt} = \hat{\beta}_0 + \hat{\beta}_1(\overline{foreign} + \sigma)_{jt}$$

$$E\hat{L}F_{jt} = .78 + -.045(5.11) = .78 - .23 = 0.55$$

The above experiment reveals that a country with a Foreign-born population of roughly 5.11% (one standard deviation away from the mean Foreign-born population score) will also exhibit an estimated 55% probability that any two persons in that country will be drawn from a different ethnic-linguistic group. Therefore, a country with an increase in the size of the Foreign-born population by about 2.64% (one standard deviation) will show a roughly 12% decrease in Ethnic-Linguistic Fractionalization.

This experiment implies that if a country such as Brazil, with a 57.8% Ethnic-Linguistic Fractionalization score in the 1960s, increased the size of the foreign-born population by about 2.47% -by encouraging and/or permitting more immigration- then its ethnic demographics would look a lot more like a country such as the Dominican Republic, which displayed a much lower Ethnic-linguistic Fractionalization score of about 45% in the 1960s.

$$E\hat{L}F_{jt} = \hat{\beta}_0 + \hat{\beta}_1(\overline{foreign} + 2\sigma)_{jt}$$

$$\hat{ELF}_{jt} = .78 + -.045(7.75) = .78 - .35 = 0.43$$

The above experiment demonstrates that a country with a Foreign-born population of roughly 7.75% (two standard deviations away from the mean Foreign-born population score) will also display an estimated 43% probability that any two persons in that country will be drawn from a different ethnic-linguistic group. Therefore, a country with an *increase* in the size of the Foreign-born population by about 5.28% (two standard deviations) will show a roughly 24% *decrease* in Ethnic-Linguistic Fractionalization.

This experiment implies that if a highly fractionalized country, such as Bolivia in the 1980s (with an ELF score of 74%), were to increase its Foreign-born population by about 5.28% -perhaps through immigration enticements- then its ethnic-linguistic fractionalization score would *decrease*. Bolivia's ethnic demographics would then look a lot more like a country such as Peru, which exhibited an Ethnic-Linguistic Fractionalization score of about 51% in the 1980s.

VII. Indigenous Peoples

An Indigenous population variable has also been included in this study as part of an interaction variable that exposes any possible changes in immigrant assimilation tendencies when entering into countries with distinct sizes of Indigenous populations. The foundation of this interaction variable derived from the theory that immigrant behavior will be quite different in Latin American countries with larger Indigenous populations for reasons alluded to previously, including the idea that 1) settler populations "deal harshly" with Indigenous peoples and 2) immigrants face more "change

constraints" when attempting to assimilate into societies made up of Indigenous populations that have unique dialects, accents, traditions and visible group identifiers: the range of possible identification categories is far more restricted.

Therefore, as there is an assumption that immigrants cannot/do not assimilate into Indigenous groups, ethnic-linguistic fractionalization will actually *increase* as the size of the foreign-born population increases in countries with large Indigenous populations: immigrants will increase non-indigenous minority group shares and/or not assimilate at all. The conditional effect of the size of the Indigenous population on the relationship between immigration and ethnic-linguistic fractionalization is expressed in the following testable hypothesis:

***Null Hypothesis II:** The size of the Indigenous population will have no conditional effect on the negative relationship between immigration and ethnic-linguistic fractionalization.*

***Alternative Hypothesis II:** As the size of the Indigenous population increases, immigration will have an increasingly positive effect on ethnic-linguistic fractionalization, all other things held constant.*

In order to test the above conditional relationship hypothesis, an Indigenous population variable and an interaction variable, `ind_foreign`, were included in **Model 7** of **Table 3**. The data for these variables derived from the Statistical Abstract of Latin America and the ECLAC (Economic Commission of Latin America and the Caribbean) Demographic Bulletin population estimates for the 1960s and 1980s. **Table 1** contains the descriptive statistics for the two variables. However, it is the interaction variable in

particular that will reflect any possible variation in the effect of the size of Foreign-born population on Ethnic-linguistic Fractionalization depending on the size of the Indigenous population in Latin American countries.

Causal Model II:

Estimation Equation 2: {where j is country and t is time period}

$$ELF_{jt} = \hat{\beta}_0 + \hat{\beta}_1 foreign_{jt} + \hat{\beta}_2 latitude_{jt} + \hat{\beta}_3 illiteracy_{jt} + \hat{\beta}_4 gdp_{jt} + \hat{\beta}_5 conflict_{jt} + \hat{\beta}_6 ind_{jt} + \hat{\beta}_7 ind_{jt} \cdot foreign_{jt}$$

$$\frac{\partial ELF_{jt}}{\partial foreign_{jt}} = \hat{\beta}_1 + \hat{\beta}_7 ind_{jt}$$

$$H_0: \hat{\beta}_1 = \hat{\beta}_7 = 0 \quad H_1: \hat{\beta}_1 < 0 \text{ and } \hat{\beta}_7 > 0$$

Table 3: Ethnic-Linguistic Fractionalization in Indigenous-Populated Countries (20 Countries of Latin America, 1960s and 1980s)

ELF2	Model 6 Coefficient (Std. Err)	Model 7 Coefficient (Std. Err)
Foreign Born Pop.	.003 (.016)	-.003 (.018)
Latitude	-.014*** (.005)	-.014*** (.005)
Illiteracy	-.003 (.004)	-.004 (.004)
G.D.P.	3.54e-06 (.000)	3.60e-06 (.000)
Ethnic Conflict	.012 (.015)	.014 (.016)
Indigenous Pop.	.012*** (.002)	.009** (.005)
Indigenous _Foreign (Interaction)		.004 (.005)
Constant	.726 (.139)	.742 (.147)
Number of Obs.	28	28
Prob > F	0.000	0.000
R-squared	0.681	0.688

Note: In each cell, the top number is the coefficient and the robust standard error is in parentheses.

*** Denotes $p < 0.01$, ** $p < 0.05$, * $p < 0.10$; one-tailed test

Graph 2: Interaction Variable (Ind_Foreign)

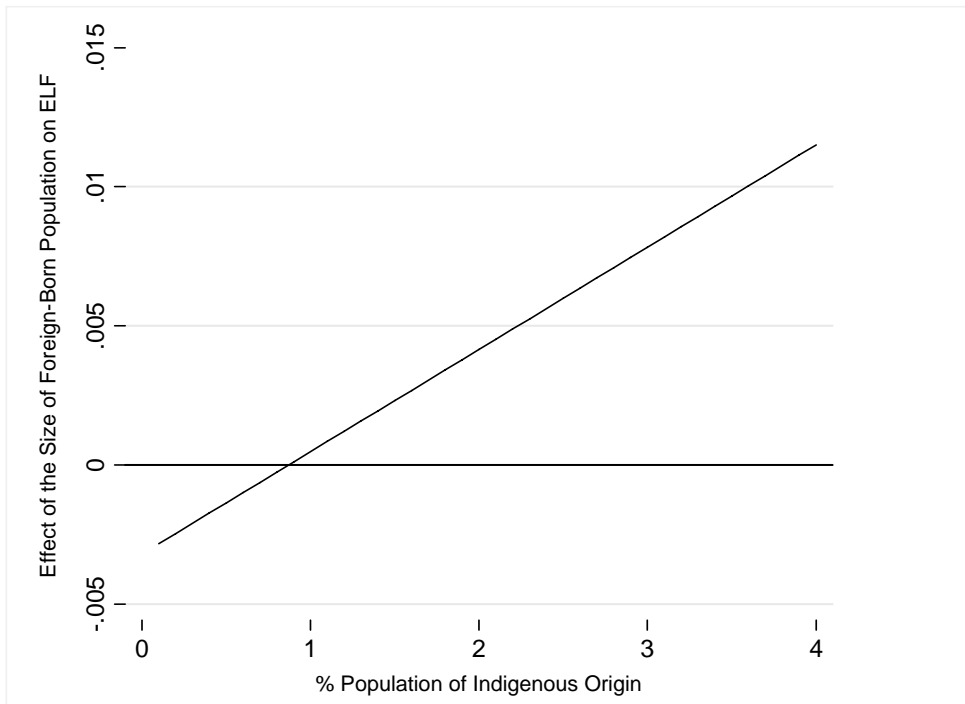


Table 3:

As shown in **Table 3**, the Indigenous population variable has a positive coefficient sign and is highly significant, at the 1% level in **Model 6** and the 5% level in **Model 7**. These results reveal, unsurprisingly, that Latin American countries with larger Indigenous populations are more ethnically fragmented. Unfortunately, when the Indigenous Population variable is included in the model, the Foreign-born population variable becomes insignificant (and actually displays a positive coefficient sign in **Model 6**), but this is most likely due to the skewed influence of the Indigenous Population variable and the small number of observations. However, the added independent variable does increase the predictive power (R-squared) of the last two models to a high 68%. Again, this bigger R-squared is a reflection of the overwhelming influence that the Indigenous population would obviously have on the ethnic fragmentation of any given country in Latin America.

What is really important to note in **Table 3**, is the positive coefficient sign of the interaction variable in **Model 7**. Although both the individual and joint tests of $\hat{\beta}_1$ and $\hat{\beta}_7$ were statistically insignificant, the results of **Model 7** indicate that null hypothesis II, $H_0: \hat{\beta}_1 = \hat{\beta}_7 = 0$, should be rejected: the size of the Indigenous population *does* have an effect on the relationship between immigration and ethnic-linguistic fractionalization in Latin American countries. The positive coefficient sign of the interaction variable, *ind_foreign*, suggests two possibilities, 1) that countries with a larger percentage of Indigenous peoples will mean a harder or even impossible process of assimilation for immigrants and/or 2) that immigrants actually assimilate into *minority* groups (most likely white/mestizo) rather than majority Indigenous groups, thereby increasing ethnic-linguistic fractionalization due to the increase in minority group shares with the addition of foreigners.

The outcome of **Model 7** comes as no great surprise, as the Indigenous peoples of Latin America -whether technically in the majority or not- have historically faced more social and political obstacles than mestizo/white groups, so immigrants would logically avoid assimilation into Indigenous groups. The results also substantiate prior claims of friction/fragmentation between settler/immigrant and Indigenous populations in Latin America. Unfortunately, the tests did not display real statistical significance so the results remain ambiguous; the results may be truly insignificant, or an accurate reflection of Latin American populations and merely suffer from flawed and/or incomplete data.

Graph 2:

As displayed in **Graph 2**, when the Indigenous population is less than approximately 1% of the total country population, the Foreign-born population variable

has a *negative* effect on ethnic-linguistic fractionalization. However, the Indigenous population need only be .9% of the total country population for the interaction variable to begin to show a positive coefficient, thereby reversing the negative correlation between immigration and ethnic-linguistic fractionalization. A continuation of the graph would reveal that increasing immigration (or a 1% increase in the size of the foreign-born population) in a country with an Indigenous population that is, for example, 10% of the total population, will result in a roughly 3% increase in ethnic-linguistic fractionalization. Although statistically insignificant, the positive coefficient sign of the interaction variable (expressed in the positive slope of **Graph 2**) implies that as the size of the Indigenous population *increases*, immigration has an increasingly *positive* effect on Ethnic-linguistic Fractionalization: null hypothesis II should be rejected.

VIII. Future Research

This study dealt with a very specific region, which unfortunately means that although the Constructivist model was proven to be the correct research method, this conclusion can only be applied to Latin America. A more global research paper, including many other regions and countries, would make a much stronger and broader statement about the Constructivist and Primordialist models.

Another failure of this study was the lack of data reflecting where immigrants were migrating from, and into which countries certain groups attempted to assimilate. This information would greatly enhance the depth of the study, as it is well known that countries such as Argentina and Brazil have very large European immigrant populations, in contrast to countries such as Venezuela and Chile that mostly deal with migrant

agricultural workers originating from other South American countries (Alvarado 102).

These factors, originating country and chosen destination, would undoubtedly affect the "change constraints" that influence assimilation tendencies of immigrants.

A time-series-cross-section panel was used in this study in order increase the number of observed cases. However, there were only two published ELF dates, 1961 and 1985, so the time-series analysis was not very valuable. Future research would greatly improve with the inclusion of more time periods so that any observed difference in Ethnic-Linguistic Fractionalization over-time could be more accurately examined. In general, extending the region of analysis, outside of Latin America, to increase the number of cases and thereby improve the statistical significance of the models would also result in a paper that could make more robust statements about the assimilation tendencies of immigrants, and their effect on Ethnic-Linguistic Fractionalization in various countries around the world.

VIV. Conclusion_____

The results of this study reveal that immigrants do in fact assimilate. Further, not only do they assimilate, but countries with larger immigrant populations actually display *lower* levels of ethnic-linguistic fractionalization because immigrants are assimilating into majority ethnic groups. With growing global trends in international migration, the malleability of immigrant ethnic identities is especially important. Therefore, the Constructivist model of ethnic politics research is the more appropriate -ethnicity has proven to be both fluid and endogenous to social and political processes. However, if the results of **Model 7** are taken into account, there is an apparent clash between immigrants

and Indigenous groups in Latin America. These findings leave behind many unanswered questions, particularly for a region that suffers from sustained political and social strife in respect to its Indigenous peoples.

References:

Alesina, Alberto. "Fractionalization." *Journal of Economic Growth*, Vol.8 (2003), pp.155-194.

Alvarado, Iván González and Hilda Sánchez. "Migration in Latin America and the Caribbean: A view from the ICFTU/ORIT." *Labour Education*, No. 129 (2002/4), pp.101-108.

Byman, Daniel. "Forever Enemies? The Manipulation of Ethnic Identities to End Ethnic Wars." *Security Studies*, Vol.9, No.3 (Spring 2000), pp.149-190.

Central Intelligence Agency, *The World Factbook*.

Chandra, Kanchan, ed. "Cumulative Findings in the Study of Ethnic Politics." *Symposium in APSA-CP* (Winter 2001).

Chandra, Kanchan. "What is Ethnic Identity and Does it Matter?" *Annu. Rev. Polit. Sci.* 2006, 9:397-424.

Fearon, James D. "Ethnic Structure and Cultural Diversity by Country." *Journal of Econ. Growth* 8 (2):2003

Fearon, James D. "Ethnic Structure and Cultural Diversity Around the World: A Cross-National Data Set on Ethnic Groups." Unpublished, Stanford University, 2002.

Fearon, James D. and David Laitin. "Violence and the Social Construction of Ethnic Identity." *International Organization*, Vol.54, No.4 (Autumn, 2000), pp.845-877.

Minorities At Risk Project, Center for International Development and Conflict Management, University of Maryland. 2004

Nagel, Joane. "Constructing Ethnicity: Creating and Recreating Ethnic Identity and Culture." *Social Problems*, Vol.41, No.1, Special Issue on Immigration, Race and Ethnicity in America. (Feb., 1994), pp.152-176.

Parenti, Michael. "Ethnic Politics and the Persistence of Ethnic Identification." *The American Political Science Review*, Vol.61, No.3 (Sep., 1967), pp.717-726.

Penn World Tables, Center for International Comparisons of Production, Income and Prices (CIC), University of Pennsylvania. 2006

Posner, Daniel. "The Implications of Constructivism for Studying the Relationship Between Ethnic Diversity and Economic Growth."

Roeder, Philip G. 2001. "Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985." February 16. <<http://weber.ucsd.edu/~proeder/elf.htm>>. (Oct. 2006)

United Nations, E.C.L.A.C. "Boletín Demográfico: Migración Internacional en América Latina." No.65 Centro Latinoamericano y Caribeño de Demografía (CELADE) January 2000.

U.N.E.S.C.O. Institute for Statistics, Literacy and Non Formal Education Sector. Estimates and Projections of Youth Illiteracy For Population Aged Between 15 and 24 Years Old. July 2002 Assessment.

Van Evera, Stephen. "Primordialism Lives!" APSA-CP: Newsletter of the Organized Section in Comparative Politics of the American Political Science Association, Vol. 12, No. 1 (Winter 2001): 20-22.