

**MODELO ECONOMÉTRICO PARA EL
ANÁLISIS DEL DESPLAZAMIENTO A PARTIR
DEL CONFLICTO ARMADO EN EL MARCO
DEL PROYECTO TEMPO**



**Concepción: Ph.D. Béatriz Nates-Cruz *et al.*
Diseño y Realización: Ing. Sergio Alejandro Camargo Vargas**

**SISTEMA INTEGRADO DE ANÁLISIS TERRITORIAL
DOCTORADO EN ESTUDIOS TERRITORIALES
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS SOCIALES Y HUMANAS
UNIVERSIDAD DE CALDAS**

2015

Contenido

1. INTRODUCCIÓN.....	3
2. DATOS EMPLEADOS.....	4
3. MARCO TEÓRICO.....	6
4. ESTADÍSTICA DESCRIPTIVA.....	12
4.1 Definiciones de Variables.....	12
5. ANÁLISIS ECONOMETRICO.....	18
5.1 Estimación Educación.....	18
5.2 Estimación Socioeconómica.....	19
5.3 Estimación Actores Conflicto Armado.....	21
5.4 Modelo Econométrico Territorial.....	22
5.4.1 Análisis del Modelo Econométrico Territorial.....	25
6. AJUSTE DEL MODELO.....	26
7. CONCLUSIONES.....	27
8. BIBLIOGRAFÍA.....	28
9. ANEXOS.....	29

1. INTRODUCCIÓN

El presente documento pretende establecer la relación entre las variables sociales, económicas y los actores armados en los hogares víctimas del desplazamiento en los municipios de Aguadas, Pensilvania, Samaná y La Dorada, pertenecientes al Departamento de Caldas y caso de estudio del proyecto TEMPO. Para esto, se realizaron estimaciones con las diferentes variables en relación al grado de desplazamiento.

La hipótesis principal que existe, es que el desplazamiento en Caldas y más precisamente en los municipios pertenecientes al proyecto TEMPO sólo se ve afectado por la presencia de grupos armados en acción con los elementos estatales, viendo este efecto a través del empleo de amenazas, homicidios y la siembra de terror para obligar a las personas a huir, con el propósito de poder sembrar cultivos ilícitos y asegurar corredores de transporte en el área.

Una segunda dinámica, es dada por variables o razones socioeconómicas que igualmente varían en gran medida en época de conflicto y que en este trabajo son de especial análisis. Identificar esas condiciones en la población al momento de un desplazamiento de sus tierras, partiendo de las condiciones regionales de la población que las hace vulnerables a situaciones de violencia, es un factor de importancia en el presente estudio.

La ley 387 de 1997, ha promovido la adopción de medidas de prevención para el desplazamiento forzado por parte del estado más no ha sido suficiente para contrarrestar el número de personas desplazadas, la pérdida de activos, la desintegración familiar, las altas tasas de desempleo, el deterioro de la calidad de vida, entre otros (González, 2002; Mooney, 2005; Ibáñez, 2006). Por este motivo, se ha planteado el uso de la econometría, puesto que se tiene conocimiento de los factores que afectan de forma principal el desplazamiento, pero no de cómo este se ve afectado y afecta otros de gran importancia.

La econometría busca en este caso evaluar la interacción del desplazamiento con variables socioeconómicas, de educación y de actores armados. La evaluación busca de igual forma definir la intensidad que tiene cada variable en el desplazamiento para con esto lograr entender el fenómeno desde diversos puntos de vista y mirar así a futuro el comportamiento que pueda tener el fenómeno si se tiene en cuenta las variables y aspectos mencionados.

2. DATOS EMPLEADOS

Los datos empleados para este trabajo tienen su origen en las bases de datos referentes a Acción Social (entidad estatal asociada a la presidencia de la República para la acción social y cooperación internacional), la RNI (Red Nacional de Información), CODHES (Organización de Monitoreo de desplazamiento), RUPD (Registro Único de Población Desplazada), Gobernación de Caldas y la base generada por el SIAT (Sistema Integrado de Análisis Territorial de la Universidad de Caldas) para el Proyecto TEMPO.

Debido a que la información no se tiene para un periodo único, se hizo una selección de un periodo ideal (2001- 2005) en el cual se encuentra la mayoría de la información y es más relevante la variable de estudio es decir, el Desplazamiento, (figuras 1a y 1b) por ende se proyectó la información a este periodo para poder tener una correcta evaluación.

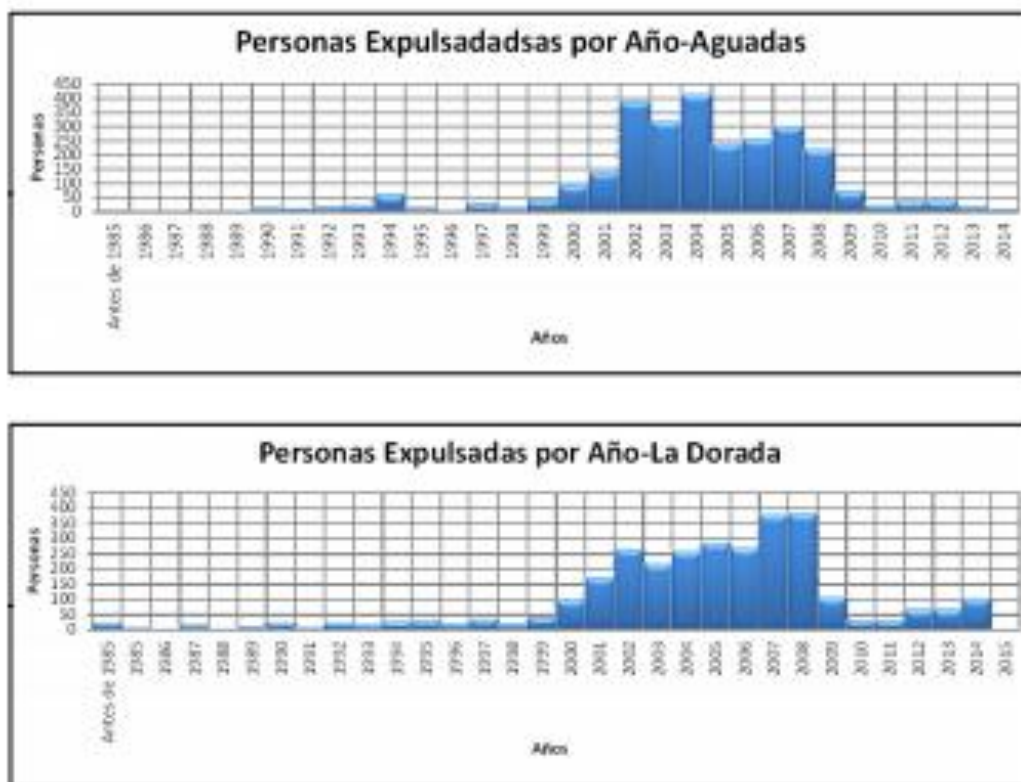


Figura 1a. Cantidad de expulsados por año para los municipios de Aguadas y la Dorada en el Departamento de Caldas. Información Proyecto TEMPO

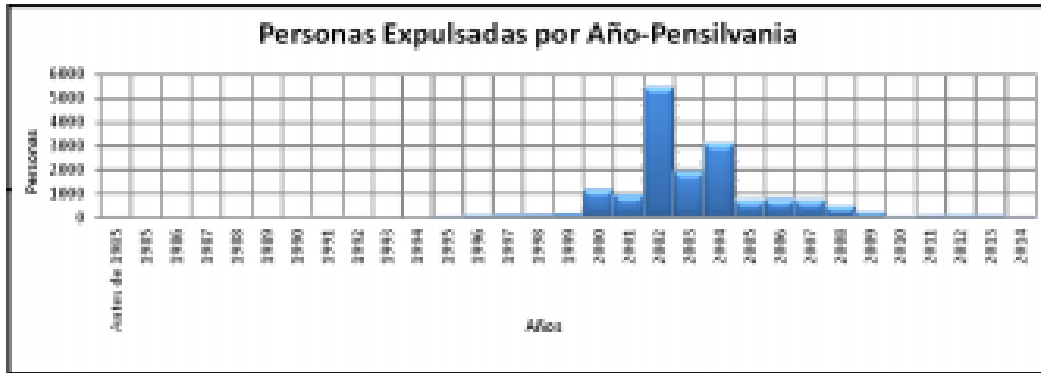


Figura 1b. Cantidad de expulsados por año para los municipios de Pensilvania y Samaná en el Departamento de Caldas. Información Proyecto TEMPO

Ya establecidas las fuentes y los datos a emplear, es posible avanzar en el análisis considerando las dinámicas y variables que interactúan en el fenómeno y su explicación.

3. MARCO TEÓRICO

3.1 RESUMEN ECONOMETRÍA (Basado en: Notas de econometría Sergio Camargo 2010-1, Gujarati Damodar, Econometría)

Un modelo econométrico se basa en un sistema de ecuaciones, pero tiene en cuenta la presencia de error en las observaciones y distorsiones.

A. MODELO DE REGRESIÓN LINEAL: fue introducido por Laplace y Gauss. Galton fue el primero en hablar sobre la regresión:

$$y_t = \beta_0 + \beta_1 x_t + e_t \quad t=1,2,\dots,n \text{ (Modelo de regresión lineal)}$$

Modelo de regresión lineal	
Partes	Supuestos
Y_t : Variable dependiente o explicada.	$E(e_t)=0$
X_t : Variables independientes o explicativas.	$E(e_t^2)=\sigma^2$
β_0, β_1 : Parámetros de la regresión	$E(e_t e_s)=0$.
e_t : Error del modelo	$e_t \sim$ Normal (los errores tienen distribución normal).

Tabla 1. Partes y supuestos del modelo de regresión

Es importante que se cumpla el último supuesto para poder realizar inferencia

El modelo de regresión se utiliza fundamentalmente para:

- ✓ Estimar valores de la variable explicada (Y) que no están dentro del rango muestral
- ✓ Para realizar predicciones

Lo que se desea encontrar es que:

- ✓ La diferencia (el error), entre el modelo de regresión lineal y las observaciones sea cero
- ✓ Tenga varianza constante
- ✓ Los errores sean independientes
- ✓ Errores giren alrededor de cero de acuerdo a una distribución normal

Modelo de regresión lineal en forma matricial

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

Los **supuestos para el modelo Lineal en forma matricial** son:

- ✓ $E(\mathbf{e})=0_{n \times 1}$
- ✓ $E(\mathbf{e}\mathbf{e}^T)=\sigma^2\mathbf{I}$. ($\sigma^2\mathbf{I} \rightarrow$ matriz de varianza de los errores)
- ✓ $\mathbf{e} \sim$ Multinormal (el vector de errores tienen distribución multinormal).

Donde \mathbf{I} es la matriz identidad de tamaño $n \times n$.

Los parámetros del modelo (β y σ^2) son valores desconocidos que debemos estimar:

$$\mathbf{Y}=\mathbf{X}\beta+\mathbf{e} \text{ (modelo en forma matricial)}$$

Donde, \mathbf{Y} y \mathbf{e} son vectores de tamaño $n \times 1$ y \mathbf{X} es una matriz $n \times 2$.

Métodos de estimación de parámetros

- ✓ **Método de Momentos:** contrasta la teoría con la práctica, igualando para despejar los parámetros a estimar
- ✓ **Mínimos Cuadrados Ordinarios:** busca minimizar los errores

$$\hat{\beta}=(X^T X)^{-1} X^T Y \text{ (Solución en forma matricial)}$$

NOTA:

- El estimador de β_0 se puede pensar como una corrección de la media de Y . Si la variable X tiene media cero, el estimador β_0 será la media de Y .
 - El estimador de β_1 es la covarianza estimada entre X e Y , dividido por la varianza estimada de X .
 - En todo el trabajo se supone que los valores x_i no son aleatorios sino fijos.
 - Se supone que $X^T X$ es no singular.
- ✓ **Máxima Verosimilitud:** busca hallar un estimador tal que se obtenga una mayor probabilidad o posibilidad de aparición de los datos en la muestra.

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{n} = \frac{\sum_{i=1}^n \hat{e}_i^2}{n}$$

Propiedades de los estimadores

- a. **Insesgamiento:** el parámetro estimado es igual al parámetro poblacional. Los estimadores de β_0 y β_1 son insesgados

$$E(\hat{\theta}) = \theta$$

- b. **Consistencia:** a medida que aumenta el tamaño de la muestra más cerca estará el parámetro estimado al parámetro población pues la varianza tiende a cero a mayor número de datos muestrales. Los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$ son estimadores consistentes.

- c. **Mínima varianza:** como su nombre lo indica busca minimizar la varianza. Los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$ son los estimadores de mínima varianza dentro de los estimadores lineales insesgados, lo cual es bueno porque mejora la calidad de las predicciones.

NOTA:

- **Inferencia:** para poder realizar inferencias sobre los Parámetros del Modelo de Regresión Lineal Simple se cumple el supuesto de normalidad sobre los errores, este supuesto además de servir para obtener estimadores puntuales, nos sirve para obtener estimadores por intervalo y realizar pruebas de hipótesis sobre ellos.
- **Predicción:** Si se necesita realizar predicción (o estimación) de Y_p usando el modelo de regresión lineal simple, solo se necesita conocer X_p y luego aplicar la ecuación:

$$\hat{y}_p = \hat{\beta}_0 + \hat{\beta}_1 x_p$$

Ajustes del Modelo

Una medida que nos indica el nivel de ajuste del modelo al fenómeno es el **coeficiente de determinación R^2** .

- $0 \leq R^2 \leq 1$.
- Si $R^2=0$ pésimo ajuste, indica que el modelo no explica, no se acerca en nada al fenómeno.
- Si $R^2=1$ buen ajuste, indica que el modelo explica perfectamente el fenómeno.

El R^2 lo que mide es: cuanta variabilidad de los datos recoge el modelo. La ecuación para el cálculo de R^2 está dada por:

$$R^2 = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2}$$

Importante:

- Si se observa la parte derecha de la ecuación a medida que los residuales son grandes, R^2 es pequeño.
- R^2 no tiene sentido si el modelo no tiene intercepto.

B. MODELO DE REGRESIÓN LINEAL MÚLTIPLE

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t \quad t=1,2,\dots,n. \text{ (regresión lineal múltiple)}$$

REGRESIÓN MÚLTIPLE	
Supuestos	Supuestos implícitos
$E(\varepsilon_t)=0$.	Los errores son no correlacionados con ninguna de las variables x_i .
$E(\varepsilon_t^2)=\sigma^2$. Para todo t .	
$E(\varepsilon_t \varepsilon_s)=0$. Para todo $t \neq s$, los errores son no correlacionados.	Las variables x_1, x_2, \dots, x_k son linealmente independientes, esto para que la matriz $X^t X$ sea no singular (o en otras palabras invertible).
Para realizar inferencia se agrega el supuesto de que ε_t tiene distribución normal (para todo t).	

Tabla 2. Supuesto del modelo de regresión lineal múltiple

Nota: Es importante notar que los dos primeros supuestos son necesarios para que los estimadores de MCO de $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ sean insesgados y de mínima varianza entre los lineales insesgados.

Modelo de Regresión múltiple en forma matricial

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

En forma compacta se reduce a:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon$$

Los **supuestos para la regresión múltiple en forma matricial** son:

- i) $E(\varepsilon) = 0$.
- ii) $E(\varepsilon\varepsilon') = \sigma^2 I$, I matriz identidad de tamaño $n \times n$.
- iii) Para realizar inferencia ε tiene distribución normal multivariada.

Los supuestos implícitos son:

- iv) $E(\varepsilon\mathbf{X}) = E(\varepsilon)E(\mathbf{X})$, no-correlación de ε con \mathbf{X} .
- v) \mathbf{X} tiene rango columna completo, $n \gg k+1$ (\gg léase mucho mayor).

Conceptos previos para entender las propiedades más importantes para el estimador de β en el modelo de regresión múltiple.

a. Esperanza de una Matriz Estocástica (aleatoria).

Dada $\mathbf{A} = \{a_{ij}\}_{i=1,2,\dots,n, j=1,2,\dots,m}$ matriz estocástica, se define la función esperanza de \mathbf{A} como:

$$E(\mathbf{A}) = \{E(a_{ij})\}$$

Propiedades: Sean \mathbf{A} y \mathbf{B} matrices estocásticas, α, θ y β matrices constantes y sea ϕ un escalar, todo definido de tal forma que las siguientes operaciones se puedan realizar.

- $E(\beta\mathbf{A}\theta + \mathbf{B}\alpha) = \beta E(\mathbf{A})\theta + E(\mathbf{B})\alpha$
- $E(\phi\mathbf{A}) = \phi E(\mathbf{A})$.

b. Matriz de Covarianza de un Vector Estocástico (aleatorio)

Dado un vector aleatorio $\mathbf{Y}^t = (y_1, y_2, \dots, y_n)$ con media $\mu^t = (\mu_1, \mu_2, \dots, \mu_n)$, se define la matriz de varianza – covarianza de la siguiente forma.

$$\text{Var}(\mathbf{Y}) = E\{(\mathbf{Y}-\mu)(\mathbf{Y}-\mu)^t\}$$

Propiedades:

- $V(\beta\mathbf{A}) = \beta V(\mathbf{A})\beta^t$.
- $V(\phi\mathbf{A}) = \phi^2 V(\mathbf{A})$.

Propiedades del Estimador $\hat{\beta}$ (de MCO y MV) en el Modelo de Regresión Lineal Múltiple.

- **Insesgamiento:** $\hat{\beta}$ es insesgado.
- **Mínima varianza:** $\hat{\beta}$ es de mínima varianza dentro de los lineales insesgados, lo cual quiere decir, que si realiza la estimación por intervalo para cualquiera de los β_i los intervalos obtenidos serían los de longitud más pequeña.
- **Consistencia:** el estimador es consistente

Teorema de Gauss – Markov.

El teorema de Gauss Markov resume las cualidades más importantes del estimador $\hat{\beta} = (X^t X)^{-1} X^t Y$.

- ✓ Es una función lineal de Y .
- ✓ Es un vector aleatorio con distribución muestral conocida (normal multivariada).

- ✓ Es insesgado.
- ✓ Su matriz de covarianza está dada por $\sigma^2(X'X)^{-1}$. Resulta ser de mínima varianza dentro de los lineales insesgados.

Estimación de σ^2

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n - (k + 1)}$$

Problemas del Coeficiente de Determinación

Presenta gran debilidad al buscar comparar dos o más modelos, en particular cuando uno de los modelos posee todas las variables independientes del otro y más, pues el R^2 crece a medida que se introducen nuevas variables predictoras en el modelo, no importando si la variable aporta nueva información en la explicación de la variable Y_t .

Por lo anterior, algunos autores proponen R^2 alternos que penalizan la entrada de nuevas variables en el modelo. Una de las propuestas es:

$$\tilde{R}^2 = 1 - \left(\frac{n-1}{n - (k+1)} \right) (1 - R^2)$$

Donde $k+1$ es el número de parámetros (o variables predictoras en el modelo).

4. ESTADÍSTICA DESCRIPTIVA

Se plantea el uso de la estadística para evaluar el comportamiento de los datos tanto en su conjunto como de forma individual, logrando así identificar variables que presentan errores y de igual forma definir a través del uso de la estadística las variables de mayor influencia en la variable de estudio (Desplazamiento). Para esto se emplea el software StatPlus, STATA y R, con lo que se realiza el análisis descriptivo inicial.

4.1 Definición de Variables

Se muestran el total de variables que de acuerdo a estudios previos como el planteado por Sagoyo, (2011) mediante un estudio de desplazamiento forzado para Colombia y el de Mosquera (2014) quien determina la influencia del desplazamiento en el trabajo, tienen una alta relevancia en el desplazamiento de una población hacia otros lugares de la geografía Colombiana y que de igual forma, en el Proyecto TEMPO se han tenido en cuenta para el análisis de los municipios de estudio.

No.	Nombre de variable	Notación	Tipo de variable
1	Municipio	MUN	Nominal
2	Topografía msnm	TOPO	Nominal
3	Área Municipio Km2	ARMUN	Nominal
4	Desplazamiento	DESP	Nominal
5	Retorno	RETO	Nominal
6	Zonas con Coca	ZC	Nominal
7	Población	POB	Nominal
8	Estudiantes Matriculados 2002	EM	Nominal
9	Indice Calidad Vida Urbano	ICV_U	Nominal
10	Indice Calidad Vida Rural	ICV_R	Nominal
11	Necesidades Básicas Insatisfechas Urbano	NBI_U	Nominal
12	Necesidades Básicas Insatisfechas Rural	NBI_R	Nominal
13	Tasa Mortalidad Infantil x 1000	TMI	Nominal
14	Salud Subsidiado	SS	Nominal
15	Salud Contributivo	SC	Nominal
16	Educ_Púb	EPUB	Nominal

17	Educ_Priv	EPRIV	Nominal
18	Cobertura Acue Urba	CAU	Nominal
19	Cobertura Acue Rural	CAR	Nominal
20	Homicidios	HOM	Nominal
21	Secuestros	SEC	Nominal
22	Hurtos	HUR	Nominal
23	% Analfabetismo funcional	AF	Nominal
24	% Desempleo	DESEM	Nominal
25	Grupos Armados Guerrilla	GAG	Nominal
26	Grupos Armados Paramilitar	GAP	Nominal

Tabla 3. Nombre, notación y tipo de las variables del modelo

En la tabla 3 se observa el nombre de las variables, la notación que se ha de usar durante el desarrollo del estudio y el tipo de variable. Siendo el Desplazamiento la variable explicada o dependiente (DESP) y el restante variables independientes o explicativas.

Se determinan a partir de las variables mencionadas en la tabla 3 unos estadísticos básicos (tabla 5) iniciales para cada variable en los que se puede observar que los datos en su mayoría presentan grandes valores de desviación estándar, esto debido a que no se cuenta con una cantidad significativa de datos o mediciones para cada una de las variables. Sin embargo, se pudo determinar que la mayoría de las variables presentan una distribución leptocurtica (figura 2), debido a que la curtosis es mayor a cero y unos valores de mediana cercanos a la media lo que indica el comportamiento semi gaussianano en la distribución de los datos, evitando con esto que sea necesario el uso de modelos de ajuste como el Logarítmico.

De igual forma, es posible observar que el sesgamiento es bajo lo que indica que las variables presentan un comportamiento adecuado. Sin embargo se observa que el coeficiente de variación es significativo, obedeciendo esto a la dispersión de los datos y que a su vez se ve reflejado en la desviación de los mismos, siendo esto generado por el tamaño de la muestra para cada variable (ver anexos estadísticas descriptivas).

Serie #4 (Desplazamiento)			
Recuento	4,00	Sesgamiento	0,33438
Media	529,5	Error Estándar del Sesgamiento	0,71714
Media LCL	-315,06283	Kurtosis	1,43269
Media UCL	1.374,06283	Error Estándar de la Kurtosis	0,58191
Varianza	281.709,66667	Sesgamiento Alterno (de Fisher)	0,57916
Desviación Típica	530,76329	Kurtosis Alternativa (de Fisher)	-2,75479
Error Estándar (de la Media)	265,38164	Coefficiente de Variación	1,00239
Mínimo	96,	Desviación Media	431,5
Máximo	1.185,	Segundo Momento	211.282,25
Rango	1.089,	Tercer Momento	32.473.827,
Suma	2.118,	Cuarto Momento	6,39558E+10
Error Estándar de la Suma	1.061,52657	Mediana	418,5
Suma de Cuadrados Total	1.966.610,	Error de la Mediana	166,30328
Suma de Cuadrados Ajustada	845.129,	Percentil 25% (Q1)	99,
Media Geométrica	302,59676	Percentil 75% (Q3)	849,
Media Armónica	176,85498	IQR	750,
Moda	#N/A	MAD (Median absolute deviation)	542,5
		Coefficiente de dispersión (COD)	1,03106

Tabla 5. Estadísticos Tipo. Estadísticos variable Desplazamiento.

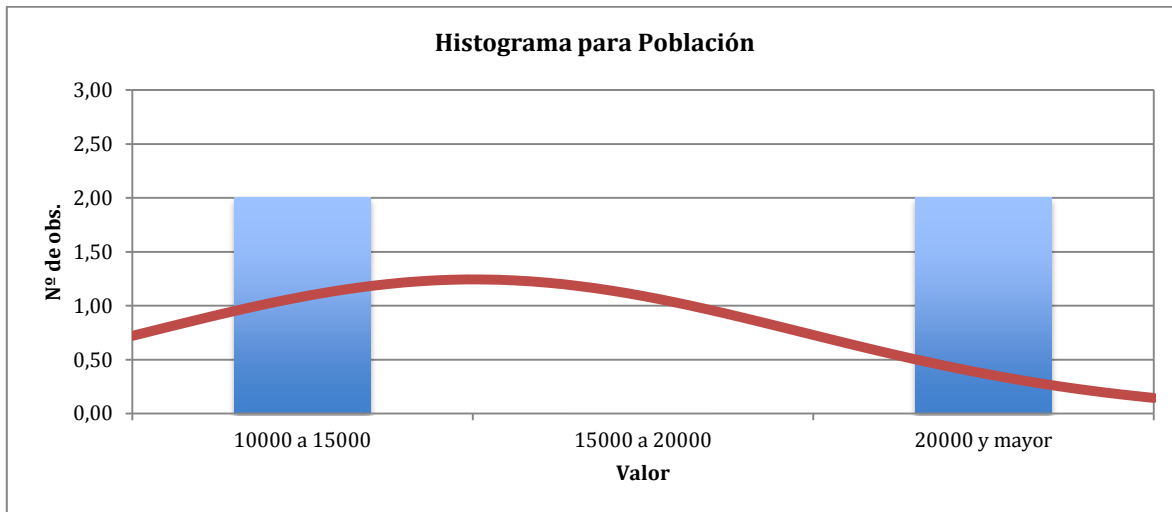


Figura 2. Histograma Poblacional Tipo. Histograma Poblacional para la variable Desplazamiento.

Mediante un algoritmo de correlación de Pearson se hizo el análisis de las variables, obteniendo como resultado una matriz (figuras 3a y 3b), con la que se pudo encontrar la existencia de diferentes intervalos de correlaciones Positivas y Negativas, para este estudio se han de tener en cuenta las variables positivas y las negativas de valores significativos puesto que estas responden y presentan un comportamiento asociado a la variable de estudio (DESP), a continuación se tienen en cuenta las variables de mayor peso, observando así la relación entre sí como se muestra a continuación:

Desplazamiento: Esta variable muestra una alta correlación con las variables RETO y ZC; se presenta igualmente una correlación moderada con las variables AF, TOPO, GAG y GAP; finalmente se presenta una correlación moderadamente baja con las variables NBI, SS y SC.

Topografía: Se evidencia una correlación alta con la variable AF; una correlación moderada con DESP, ZC, GAG, GAP y DESEM, y finalmente una correlación baja con EPUB.

Retorno: Alta correlación con las variables ZC, NBI, SC, HOM y GAP, una moderada con SS y una baja con SEC.

Zonas con Coca: Se tiene una alta correlación con DESP, AF y GAP; se presenta así mismo una moderada con las variables TOPO, ARMUN, NBI y HUR.

Analfabetismo: Se tiene una alta correlación con las variables TOPO, DESP y ZC; un interacción media con ARMUN, TMI y DESEM; finalmente la correlación baja se muestra para las variables HUR, GAG y GAP.

GAG y GAP: Muestra una alta incidencia en su relación con la mayoría de las variables, siendo de las más altas NBI, DESEM, DESP y RETO. En esta variable se implementó el uso de las Dummies para evaluar la interacción de esta variable en el territorio.

NBI: Teniendo en cuenta que se realizó una diferenciación en esta variable, indicando que se tiene un valor para las zonas Urbanas y otro para las Rurales, se evidencia que de forma conjunta se observa una alta correlación con la variable EPRIV; una relación media con RETO, ZC, ICV (General) y HOM; una baja interacción con DESP y finalmente de forma diferenciada se muestra una alta correlación de la variable SC para NBI_U y SS para NBI_R.

SS: Se muestra una alta interacción con las variables POB, EM, HOM, SEC y HUR; y una correlación media con RETO e ICV.

SC: se presenta una alta correlación con las variables POB, RETO, EM, GAG, GAP, HOM y HUR; una relación media con ICV y finalmente una baja con DESP.

De acuerdo al análisis realizado a partir de la matriz de correlación de Pearson se pudo observar que para las variables correlacionadas con la variable DESP, se presentan unas variables secundarias que las afectan, lo que influye en la variable a explicar, es por esto que se han de tener en cuenta las variables que tengan una alta relación con las que explican la variable de estudio o explicada y que ayudan en la mejora del modelo. A continuación se muestra la tabla 4 que muestra las variables resultantes luego del análisis:

No.	Nombre de variable	Notación	Tipo de variable
1	Municipio	MUN	Nominal
2	Topografía msnm	TOPO	Nominal
3	Área Municipio Km2	ARMUN	Nominal
4	Desplazamiento	DESP	Nominal
5	Retorno	RETO	Nominal
6	Zonas con Coca	ZC	Nominal
7	Población	POB	Nominal
8	Estudiantes Matriculados 2002	EM	Nominal
9	Indice Calidad Vida_Urbano	ICV_U	Nominal
10	Indice Calidad Vida_Rural	ICV_R	Nominal
11	Necesidades Básicas Insatisfechas_Urbano	NBI_U	Nominal
12	Necesidades Básicas Insatisfechas_Rural	NBI_R	Nominal
13	Tasa Mortalidad Infantil x 1000	TMI	Nominal
14	Salud Subsidiado	SS	Nominal
15	Salud Contributivo	SC	Nominal
16	Educ_Púb	EPUB	Nominal
17	Educ_Priv	EPRIV	Nominal
18	Homicidios	HOM	Nominal
19	Hurtos	HUR	Nominal
20	% Analfabetismo funcional	AF	Nominal
21	% Desempleo	DESEM	Nominal
22	Grupos Armados Guerrilla	GAG	Nominal
23	Grupos Armados Paramilitar	GAP	Nominal

Tabla 4. Nombre, notación y tipo de las variables del modelo resultante del análisis de Correlación de Pearson.

5. ANÁLISIS ECONÓMÉTRICO

Teniendo en cuenta lo aplicado por Sandoval et al, 2011 en donde se hace un estudio para los departamentos de Colombia, con el fin de explicar el desplazamiento a partir del uso de las variables asociadas a educación y condiciones sociales y teniendo en cuenta las variables resultantes del análisis de correlación, se hace una aproximación inicial de modelos.

5.1 Estimación Educación

Se plantea inicialmente de acuerdo a las variables que se tienen en la siguiente estimación.

$$DESP = \alpha + \beta_0 AF + \beta_1 EM + \beta_2 EPUB + \beta_3 EPRIV + \beta_4 DESEM + e$$

Luego de su aplicación se pudo definir que debido a que las variables AF y EPRIV son colineales

Dezplazamiento	Coef.
Analfabetismofuncional	0
educacionEstudiantesMatriculad	-.0685227
Educ_Pb	4.822158
Educ_Priv	0
Desempleo	-104.3417
_cons	2563.566

Figura 4. Coeficientes del Modelo de Estimación de Educación.

Se tiene entonces un modelo de estimación que explica la variable DESP a partir del siguiente modelo resultante:

$$DESP = \alpha + \beta_0 EM + \beta_1 EPUB + \beta_3 DESEM + e$$

donde se tiene que: $\beta_0 = -0.0685227$; $\beta_1 = 4.822158$; $\beta_3 = -104.3417$;
 $\alpha = 2563.566$; $e = 0,00000001$

Este resultado muestra que las variables EM, EPUB y DESEM presentan incidencia directa con el Desplazamiento, lo que concuerda con lo expresado por Sandoval et al,

2011, además se tiene un R^2 próximo a 1 lo que indica un buen ajuste del modelo. Sin embargo el comportamiento de EM y DESEM es inverso al del Desplazamiento como se muestra en la figura 5.

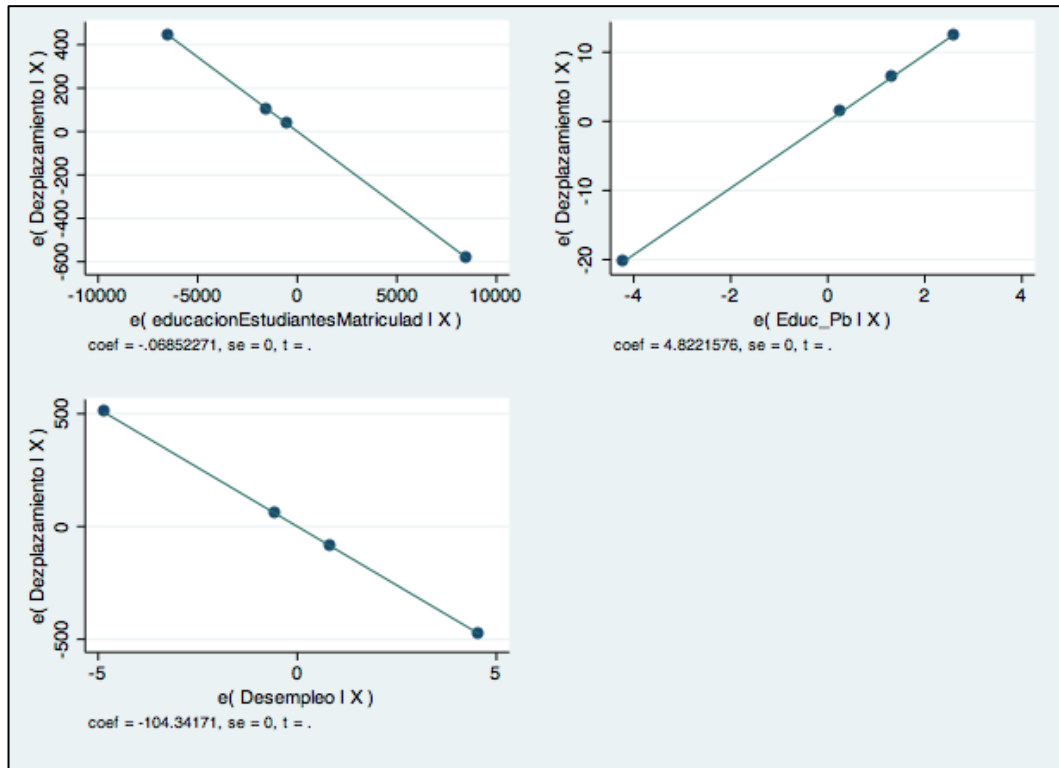


Figura 5. Comportamiento de las variables de Educación Significativas en el Modelo en relación con la variable Explicada.

5.2 Estimación Socioeconómica

Para esta estimación se plantea usar las variables finales obtenidas en la tabla 4 que no contemplan todas las propuestas por Sandoval et al (2011), pero si contiene las de mayor peso del análisis de correlación.

$$DESP = \alpha + \beta_0 NBI + \beta_1 TMI + \beta_3 DESEM + e$$

de acuerdo con lo planteado se tiene entonces que:

Desplazamiento	Coef.
NBI_U	0
NBI_R	-99.40877
TasaMortalidadInfantilx1000	81.66479
Desempleo	-119.4682
_cons	4258.037

Figura 6. Coeficientes del Modelo de Estimación Socioeconómica.

$$\beta_0 = -99.40877 ; \beta_1 = 81.66479 ; \beta_3 = -119.4682 ;$$

$$\alpha = 4258.037 ; e = 0,00000001$$

En este caso NBI y DESEM muestran una incidencia negativa o inversa en el Desplazamiento y TMI muestra una incidencia positiva (figura 7). El R^2 próximo a 1 lo que indica un buen ajuste del modelo.

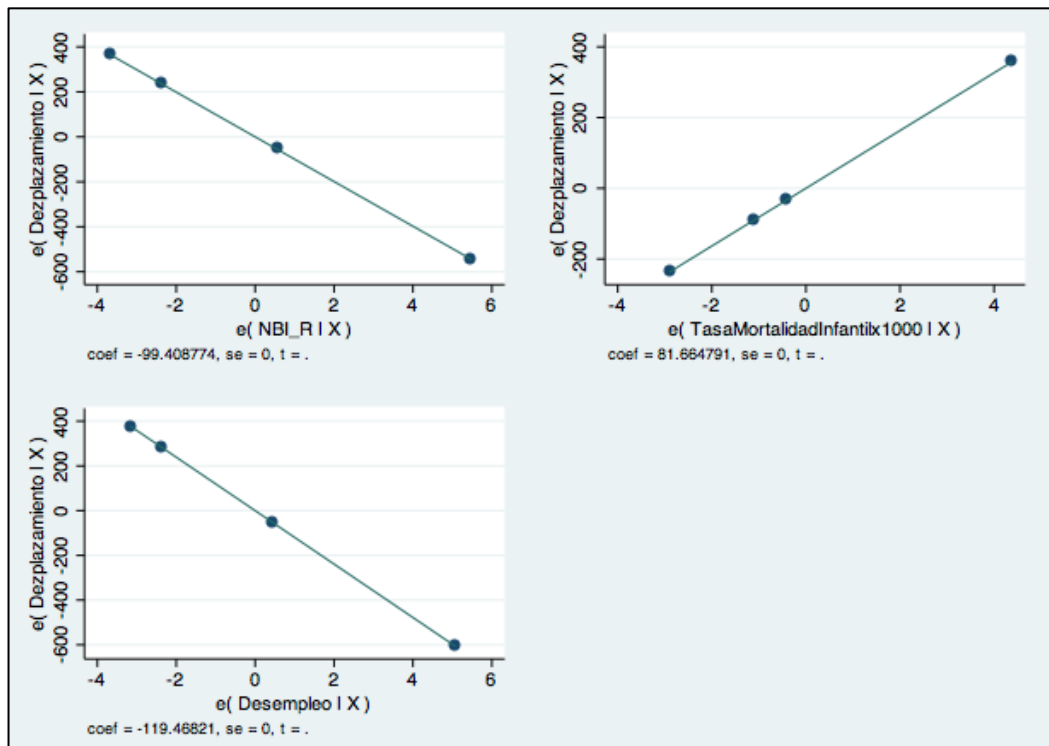


Figura 7. Comportamiento de las variables de Socioeconómicas Significativas en el Modelo en relación con la variable Explicada.

5.3 Estimación Actores Conflicto Armado

En esta se tienen en cuenta las variables que afectan la seguridad y que permiten evaluar la presencia del estado en las zonas de conflicto, teniendo en cuenta para ello las variables de mayor peso en este sentido, tomadas de la tabla 4 y donde se tienen las variables GAG y GAP definidas como Dummies por su información base.

$$DESP = \alpha + \beta_0 RETO + \beta_1 HOM + \beta_3 HUR + \beta_4 GAG + \beta_5 GAP + e$$

evaluando este modelo se obtiene el siguiente grafico:

Dezplazamiento	Coef.
Retorno	3.459222
Homicidios	0
Hurtos	0
GruposArmadosG	.0336346
GruposArmadosParami	.0363604
_cons	-2771.629

Figura 8. Coeficientes del Modelo de Estimación Factores del Conflicto Armado.

Es de resaltar que se evidencia al evaluar el modelo dos variables que presentan una alta colinealidad aun cuando se tiene un R^2 alto que predice un buen ajuste. Sin embargo, al presentar una prueba t se nota que para esas dos variables en relación a la variable explicada no son significativas. Por lo tanto, se puede aducir este fenómeno a la escasez de datos, por lo que se hace necesario excluir del modelo estas variables, quedando finalmente el modelo siguiente:

$$DESP = \alpha + \beta_0 RETO + \beta_1 GAG + \beta_3 GAP + e$$

donde los coeficientes son:

$$\beta_0 = 3.459222 ; \beta_1 = 0.0336346 ; \beta_3 = 0.0363604 ;$$

$$\alpha = -2771.629 ; e = 0,00000001$$

Siendo el Desplazamiento explicado entonces en su mayoría por la presencia de grupos armados y el retorno de la población a un área (figura 9), lo que concuerda con la mayoría de estudios planteados en este tema y que sugieren a los actores armados como las variables más influyentes tal como Sayago T (2011) mediante el uso de análisis del desplazamiento por gravedad.

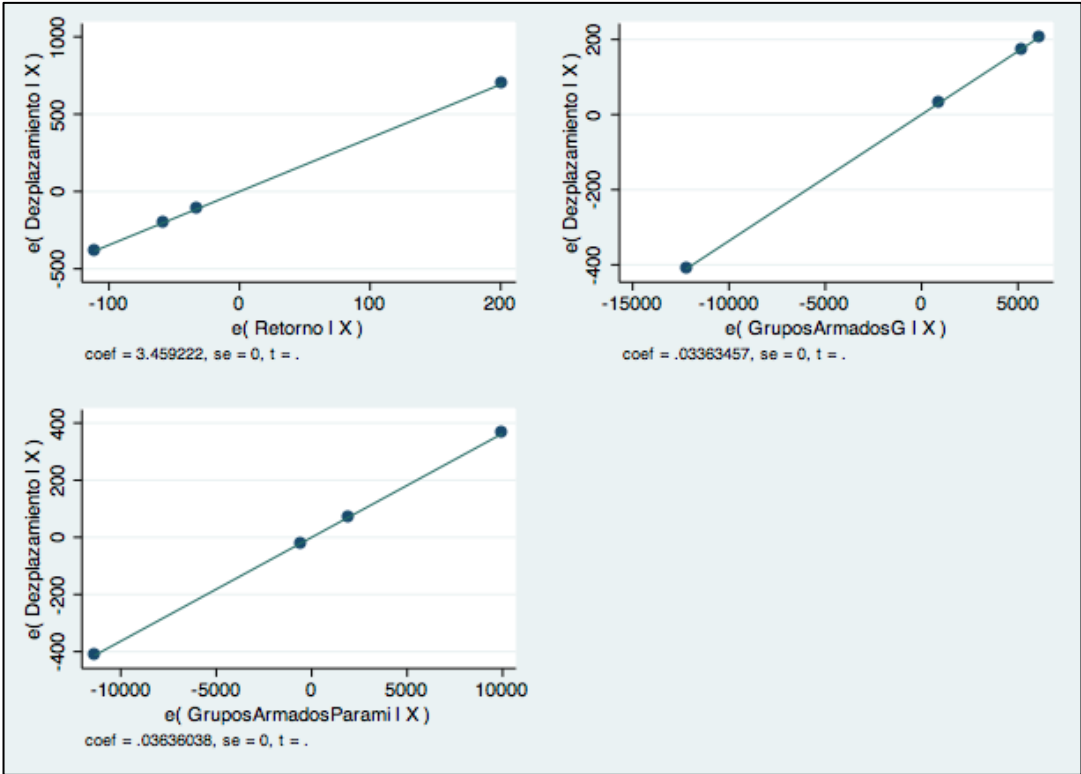


Figura 9. Comportamiento de las variables de Factores Conflicto Armado Significativas en el Modelo en relación con la variable Explicada.

5.4. Modelo Econométrico Territorial

Un elemento a tener en cuenta es la aglomeración de actores en las regiones de estudio lo que conlleva a una sumatoria que hace que el problema se agudice en una determinada área. Mediante este modelo se busca integrar el análisis econométrico junto con la perspectiva territorial, lo que nos permite inferir elementos de análisis en el proyecto TEMPO. Para este fin se hace uso inicial de las variables de población y desplazados, con el fin de medir la gravedad de expulsión que el desplazamiento causó e integrar a su vez el modelo econométrico que permita identificar las causas en una determinada región.

Teniendo presente las variables y su estadístico de correlación se plantea la siguiente formulación para determinar que tan intenso es el factor de desplazamiento de acuerdo a la actividad de los actores armados, tenido en cuenta su asociación con las variables significativas y que se asocian al modelo y las variables a trabajar.

$$IntDesp = \frac{Desplazados_i}{Población_i} \times 100$$

De acuerdo a lo obtenido se muestra la intensidad de desplazamiento que más se resalta es la referente al municipio de Samaná lo que concuerda con lo observado en este estudio y en el proyecto en general. Evaluando si la tasa de intensidad de desplazamiento es directamente proporcional a la tasa de intensidad de retorno se evalúa entonces esta situación mediante la siguiente formula:

Aplicando ahora el Modelo de Económico se logra obtener la siguiente estimación:

$$DESP = \alpha + \beta_0 ZC + \beta_1 GAG + \beta_3 GAP + e$$

donde los coeficientes son:

Dezplazamiento	Coef.
ZonasconCoca	.1107681
GruposArmadosG	-.0050747
GruposArmadosParami	-.0251333
_cons	1441.743

Figura 10. Coeficientes del Modelo de Estimación Económico Territorial.

$$\beta_0 = 0.1107681 ; \beta_1 = -0.0050747 ; \beta_3 = -0.251333 ;$$

$$\alpha = 1[441.743 ; e = 0,00000001$$

Empleando la formulación de intensidad de Desplazamiento en los resultados del modelo econométrico es posible obtener un resultado espacial, el cual permite observar los cambios en el territorio. Se puede observar (Figura 10), que la variable ZC es la que más significancia tiene al explicar la variable dependiente, si que esto de lado las variables GAP y GAG que son determinantes absolutos.

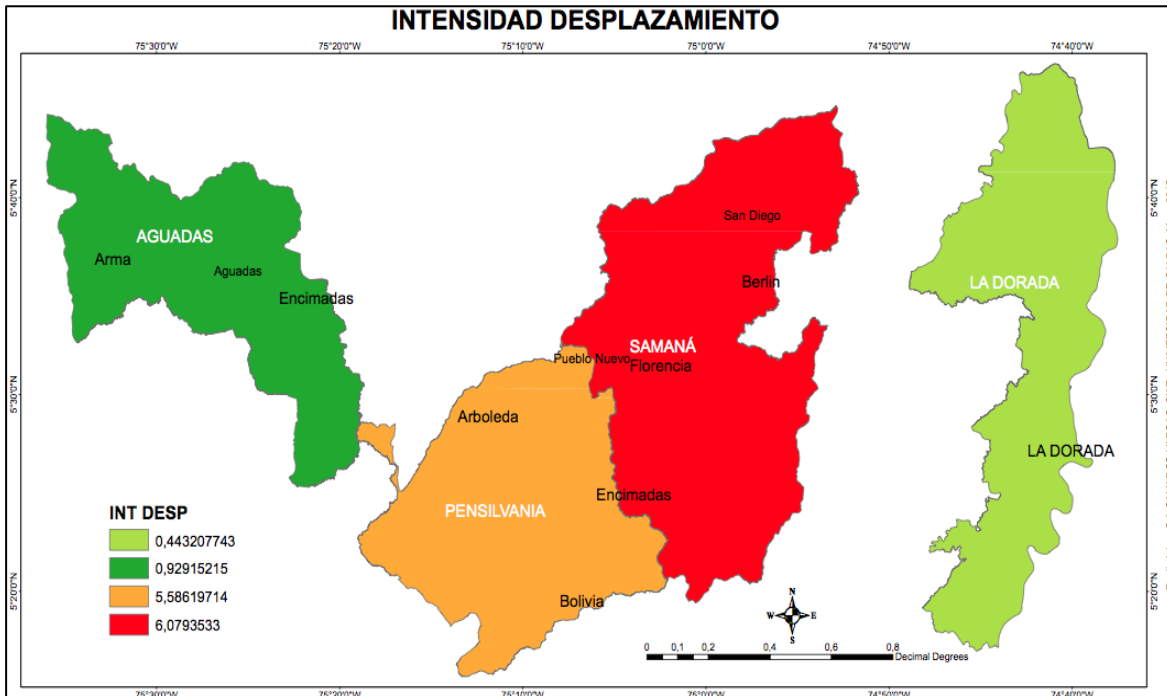


Figura 11. Intensidad del Desplazamiento dada por el Modelo Econométrico Territorial en los Municipios de Estudio del Proyecto TEMPO.

Para realizar un análisis completo se hizo pertinente emplear este mismo modelo tomando como variable explicativa el Retorno en su relación con el desplazamiento lo que llevo a plantear el modelo econométrico Territorial en relación con la siguiente ecuación:

$$IntReto = \frac{Retorno_i}{Desplazados_i} \times 1000$$

Retomando el modelo econométrico anteriormente expuesto y empleando este en la ecuación anterior para el desplazamiento, se puede entonces obtener la intensidad de las diferentes variables en la región de estudio y su actuar para la Intensidad de Retorno (Figura 12).



Figura 12. Intensidad de Retorno dada por el Modelo Econométrico Territorial en los Municipios de Estudio del Proyecto TEMPO.

5.4.1 Análisis del Modelo Econométrico Territorial.

Como un resultado interesante se puede observar que el municipio con mayor intensidad de desplazamiento es el mismo que presenta la menor intensidad de retorno siendo este Samaná, presentando un desplazamiento asociado a pérdida de tierras, de igual forma el municipio de Pensilvania presenta una alta tasa de retorno y así mismo una alta tasa de desplazamiento, lo que indica que en este municipio para el periodo de estudio el desplazamiento era momentáneo y no definitivo. El municipio de La Dorada por su parte evidencia ser el receptor de la mayor parte de los desplazados de municipios cercanos o adyacentes a este.

El modelo Econométrico Territorial permite definir que el desplazamiento va fuertemente ligado a las zonas con mayor sembrado de Coca y con presencia paramilitar, la presencia de grupos guerrilleros en este caso presenta una significancia media.

6. AJUSTE DE LOS MODELOS

De acuerdo a lo planteado inicialmente dentro de los supuestos de la econometría, es relevante nombrar que los modelos propuestos presentan el siguiente ajuste:

Prueba de hipótesis

$$\begin{aligned} H_0: & \text{El modelo ajusta los datos} \\ H_a: & \text{El modelo no ajusta los datos} \end{aligned}$$

Regla de decisión

$$\begin{aligned} \hat{\alpha} < \alpha & \rightarrow \text{Rechazo } H_0 \\ \hat{\alpha} > \alpha & \rightarrow \text{No Rechazo } H_0 \\ \hat{\alpha} < \alpha & \text{ rechaza la hipótesis nula} \end{aligned}$$

Conclusión: Los Modelos Ajustan los Datos.

Mirando el coeficiente de determinación ajustado R^2 -corregido, podemos decir que el 95% de la variabilidad de DESP después de penalizar la entrada de nuevas variables y eliminado las que presentan colinealidad, dividiendo por los grados de libertad es explicada por la variabilidad de las variables explicativas para cada modelo, es decir el ajuste es bueno.

7. CONCLUSIONES

De los modelos empleados en este trabajo se tiene un correcto ajuste de todos, lo que implica que las variables empleadas muestran el comportamiento del Desplazamiento en ellas. Es notorio en los estadísticos descriptivos la necesidad de ampliar la muestra para así obtener una mejor aproximación y refinamiento en los modelos, así como evitar la colinealidad entre las variables.

Es relevante en el modelo de Estimación de Educación la relación que tiene de forma directa con las variables EM y EPUB, así como en una segunda instancia con la variable DESEM, siendo esta última variable importante de igual forma en el modelo de Estimación Socioeconómico, por que lo que se puede entonces inferir que el Desplazamiento tiene en un segundo plano una inferencia en el Desempleo de una región.

Importante es que las variables TMI y NBI en el periodo estudiado tiene una alta variación dentro del modelo de Estimación Socioeconómico y responden a lo planteado por Sandoval et al (2011) en donde se define el desplazamiento en Colombia a partir de estas variables.

El empleo del Modelo Econométrico Territorial determina a Samaná como un municipio de moderada desocupación, puesto que su intensidad de desplazamiento fue alta, pero su intensidad de retorno fue muy baja en comparación con un municipio como Pensilvania el cual tuvo una alta intensidad de desplazamiento y así mismo una alta intensidad de Retorno, lo que lleva a pensar que Pensilvania tuvo un retorno completo y/o una sobre tasa de retorno, lo que infiere un aumento en la población original que se tenía a comienzos del 2001; este caso es similar al observado en el municipio de La Dorada donde se tuvo una intensidad de Retorno significativa y que es posible ver si se tiene en cuenta el aumento de desempleo en este municipio.

Los factores armados tienen una relación significativa con los cultivos de coca, pero este también se explica desde la variable de Topografía, ya que es notorio que en las zonas de mayor elevación topográfica el modelo Econométrico Territorial arroja una mayor intensidad de Desplazamiento, puesto que estas zonas favorecen por sus pisos térmicos y suelo el sembrado de Coca.

Finalmente, el estudio Arroja un análisis ideal de las variables y define unos modelos predictivos que pueden explicar el comportamiento del fenómeno en un tiempo diferente al evaluado, siempre y cuando se cuente con la misma o mejor información.

9. BIBLIOGRAFÍA

Dirección Territorial de Salud Caldas. (2011). *Análisis de Situación de Salud Caldas*, Ministerio de Salud, Colombia.

Hernández, L. (2010). *Procesos de Retorno y Reubicación de dos Comunidades Victimizadas por el Desplazamiento en los Montes de María. Actores Sociales y Proyectos Políticos*, Universidad Nacional de Colombia. Tesis Maestría.

Mosquera, M. (2014). *Influencia del Desplazamiento Forzado en el Trabajo Infantil en Colombia*, Universidad del Valle.

Mora, L. y Mejía, A. (2010). *Agenda Ciudadana "Desarrollo Rural de Caldas Frente a la Infraestructura de Transporte" Documento Base*, Contraloría de la Republica, Colombia.

Mochon, F. (1992). *Un Estudio Empírico de los Desplazamientos Interprovinciales de Capital y Trabajo. Una Reformulación*, Estudios Regionales, N 48, México.

Ordoñez, H. (2011). *Notas de Estadística*, Universidad Nacional de Colombia

Sajoyo, J. (2011). *Desplazamiento Forzoso en Colombia: Expulsión y Movilidad, Dos Dinámicas que Interactúan*, Universidad Nacional De Colombia, Tesis Maestría.

Sandoval, L., Boton, S. y Botero, M. (2011). *Educación, Desigualdad y Desplazamiento Forzado en Colombia*, Revista facultad de Ciencias Económicas: Investigación y Reflexión, Universidad Nueva Granada.

8. ANEXOS

Contains data

obs: 4
vars: 26
size: 512

variable	storage name type	display format	value label	variable label
Municipio	str11	%11s		MUN
Poblacion	int	%10.0g		POB
Topografiamsnm	int	%10.0g		TOPO
reaMunicipioKm2	double	%10.0g		ARMUN
Dezplazamiento	int	%10.0g		DESP
Retorno	int	%10.0g		RETO
ZonasconCoca	double	%10.0g		ZC
educacionEstu~d	double	%10.0g		EM
Educ_Pb	byte	%10.0g		EPUB
Educ_Priv	byte	%10.0g		EPRIV
Analfabetismo~l	double	%10.0g		AF
ICV_U	double	%10.0g		ICV_U
ICV_R	double	%10.0g		ICV_R
NBI_U	double	%10.0g		NBI_U
NBI_R	double	%10.0g		NBI_R
Desempleo	double	%10.0g		DESEM
TasaMortal~1000	double	%10.0g		TMI
SaludSubsidiado	int	%10.0g		SS
SaludContribu~o	int	%10.0g		SC
CoberturaAcue~a	double	%10.0g		CAU
CoberturaAcue~l	double	%10.0g		CAR
Homicidios	int	%10.0g		HOM
Secuestros	byte	%10.0g		SEC
Hurtos	int	%10.0g		HUR
GruposArmadosG	byte	%10.0g		GAG
GruposArmados~i	byte	%10.0g		GAP

Variable	Obs	Mean	Std. Dev.	Min	Max
Municipio	0				
Poblacion	4	16703	6413.336	10332	23144
Topografia~m	4	1475	924.5489	176	2214
reaMunicip~2	4	586.925	121.8913	482.7	761
Desplazami~o	4	529.5	530.7633	96	1185
Retorno	4	248.25	182.6698	32	476
ZonasconCoca	4	5374.661	6206.905	0	10869.93
educacionE~d	4	9008.316	6597.8	4763.061	18851.85
Educ_Pb	4	7.5	3.109126	3	10
Educ_Priv	4	2.5	5	0	10
Analfabeti~l	4	30.25	6.270832	21.3	35.9
ICV_U	4	80.6425	1.268053	79.02	82.01
ICV_R	4	57.9775	5.201778	54.05	65.62
NBI_U	4	20.11	3.531279	15.05	23.13
NBI_R	4	33.255	7.496117	25.08	41.68
Desempleo	4	13.925	4.248431	10.5	20
TasaMor~1000	4	15.195	5.69024	8.26	21.88
SaludSubsi~o	4	4812.25	4101.274	1145	10691
SaludContr~o	4	819.25	573.8695	180	1407
CoberturaA~a	4	98	0.4898979	97.4	98.4
CoberturaA~l	4	47.725	14.4742	34.1	65
Homicidios	4	129.5	93.09672	43	247
Secuestros	4	10	6.97615	4	20
Hurtos	4	36	50.9575	0	109
GruposArma~G	4	0.25	0.5	0	1
GruposArma~i	4	0.75	0.5	0	1