An Econometric Approach for Modeling Population Change in Doña Ana County, New Mexico

Thomas M. Fullerton Jr.[1,*], Adam G. Walke[1], Diana Villavicencio[2]

[1]Department of Economics and Finance, University of Texas at El Paso, El Paso, USA
[2]Economic and Market Analysis Division, United States Department of Housing and Urban Development, Chicago, USA
*Corresponding author: tomf@utep.edu

Abstract An econometric model using time series analysis techniques is employed to model and forecast population changes in Doña Ana County, New Mexico. The model focuses on the interplay between economic and demographic variables. Individual, cointegrated equations are generated to account for the components of population change - births, deaths, net domestic and net international migration. Birth and death equations prove easier to model because of stable changes from period to period in relation to income levels and national demographic trends. Net migration equations were more difficult to model as economic conditions, specifically labor market conditions, influence changes over time. Predefined exogenous variables are used to generate out-of-sample simulations for the individual components of population change. Using those results, total population projections are estimated until the year 2018. Doña Ana County is projected to witness a slowdown in population growth, primarily as a consequence of increased domestic out-migration.

Keywords: population economics, regional economics, applied econometrics, migration, forecasting

1. Introduction

Located within the international border region of the southwestern United States, the Las Cruces metropolitan economy exhibits interesting demographic tendencies. Within the border region, Las Cruces has long been recognized as one of the most dynamic urban economies in the area. It is also the second largest metropolitan economy in New Mexico.

The Las Cruces metropolitan statistical area is comprised by Doña Ana County. Population growth in this urban economy is highly variable. That variability occurs because net migration frequently represents a large component of population change in Las Cruces. As recently as 2010, net migration exceeded residential natural increase as a contributor of population change. Migratory flows are functionally dependent upon relative regional economic conditions, both cyclical and structural, and can fluctuate tremendously (Cebula & Alexander, 2006; Cebula & Clark, 2011; Cebula & Clark, 2013; Lim, 2011).

A substantial amount of research has been conducted for the border region between México and the United States, yet analyses of border population growth have been relatively infrequent (Fullerton & Barraza de Anda, 2008). Of the research that does exist, Estrella (1992) and Tuiran (1992) point out that information about migration across borders is minimal and often confusing. In the case

of Doña Ana County, the interaction of structural demographic and cyclical economic factors on demographic expansion have yet to be studied econometrically. The study herein expands upon the Peach and Williams (1994) examination of the patterns and components of demographic change for the El Paso-Juarez-Las Cruces region by developing an econometric model of population change in Doña Ana County, New Mexico. The paper is organized as follows. Section two provides a review of relevant literature. Section three discusses the data. The methodology utilized is discussed in section four. Section five presents empirical findings. A summary and closing remarks are provided in the concluding section.

2. Literature Review

2.1. Population Modeling

Population modeling at a county level is challenging because counties are subject to greater volatility in their growth patterns than nations (Smith, 1984). In general, the ability to forecast population fluctuations is greater for large places than for small places (Tayman & Swanson, 1996). One reason for the difficulty of forecasting subnational demographic changes is that migration, the most volatile of demographic components (Davis, 1995), is frequently the largest contributor to population growth at the regional level (Cushing & Poot, 2004). For a border-

region urban economy such as Las Cruces, demographic forecasting is especially complicated because of often substantial fluctuations associated with international migration (Davis, 1995).

Some research indicates that modeling natural increase and net migration separately improves forecast accuracy (Plaut, 1981). However, regional and local data on the components of population change are sometimes incomplete. In particular, time-series data on migration at the county level are limited and data disaggregated into cohorts are not always available. Measuring births and deaths is a more straightforward process because yearly changes are less severe (Plaut, 1981). In the absence of complete data, net migration can be calculated as a residual by subtracting natural increase from population change (Shryock & Siegel, 1980). The latter approach is taken by Fullerton, et al. (2014) in modelling population change in Arkansas. The major determinants of natural increase and net migration are discussed below.

2.1.1. Natural Increase

Population change due to births and deaths within a region is related to a variety of regional, national, and international factors including economic conditions. Some studies indicate that children are "inferior goods" and, therefore, as real incomes increase individuals will have fewer children (Fukuchi & Yamaguchi, 1997; Isen & Stevenson, 2010). Women will have fewer children because the opportunity cost of having children increases as their value of time is increased by their wage rate or their education (Shultz, 2005). In addition, future fertility rates may be positively affected by the number of migrants seeking employment. That is, as more migrants enter a labor market the number of families having children increases.

Although natural increase is closely tied to the age-sex distribution, the aggregation of data can more clearly explain long-term historical trends and foreseeable structural changes (Smith, 1997). Plaut (1981) suggests using simple natural increase models because changes to births and deaths usually occur in relatively gradual manners. The simple birth and death specifications of Fullerton and Barraza de Anda (2008) that include local and national trends display good statistical diagnostics.

2.1.2. Net Migration

Economic factors tend to be the most dominant determinants of migration. Accordingly, studies of migration patterns can be analyzed using relative wage, income, and labor market data for different countries and regions (Bolton, 1985; Fullerton, 2001; Hernández-Murillo et al., 2011). Empirical evidence indicates that no single functional form or modeling strategy is superior to all others (Cushing & Poot, 2004).

Regional economic and geographic studies have helped clarify why migration does or does not occur by analyzing push factors from the sending area and pull factors from the receiving area. Business cycles affect migration at the origin and destination (Lee, 1966). A lack of economic opportunity in the form of lower real wages pushes individuals out of their place of origin. On the other hand, job availability and higher real wage rates act as pull factors that motivate immigration (Kazlauskienė & Rinkevičius, 2006; Djafar & Hassan, 2012). Stark and

Taylor (1991) argue that migrants from Mexico to the United States not only seek to improve absolute income levels, but also relative positions in reference to origination areas.

Harris and Todaro (1970) model the decision to relocate between rural and urban areas based upon expected real wages and the probability of obtaining employment in the destination market. Although migrants are risk averse, wage premiums induce mobility even if faced with the probability of being unemployed. Using a modified Harris-Todaro model, Plaut (1981) finds that net migration to Texas is positively affected by the probability of obtaining employment, which is proxied by the ratio of job vacancies to the number of unemployed persons. Similarly, Fullerton et al. (2014) provide evidence that the ratio of employment in Arkansas to employment nationwide strongly affects net migration to that state.

Cebula and Alexander (2006) demonstrate that economic factors are significant determinants of net interstate migration, while positive (amenities) and negative (disamenities) quality-of-life factors display mixed results. In general, warm-climate retirement destinations have increased in popularity (Cebula & Alexander, 2006; Anderson & Gerber, 2008). Las Cruces has a growing economy and is located in a warm weather region of the United States.

2.2. Doña Ana County

Peach and Williams (1994) use decennial data from the United States Census Bureau to extrapolate future population trends in Doña Ana County. The population projections are disaggregated by age, sex, and ethnicity. Although unemployment rates in border regions tend to be relatively high and variable, total employment in metropolitan border counties often grows at faster rates than in other regions (Peach & Williams, 2000). By considering local and national labor market conditions from both sides of the border, this study complements the earlier analyses by using econometric techniques to examine the interplay between economic and demographic variables across domestic and international boundaries.

3. Data

Basic components of population change include births, deaths, and net migration. Annual frequency estimates of those data are used to model population change in the county of Doña Ana, New Mexico. Several different national and regional socio-economic variables are used as regressors.

Unlike Peach and Williams (1994) which uses decennial population counts, annual data from the Bureau of Economic Analysis (BEA, 2014) and other sources are used in this study. Records for births and deaths are collected from 1971 to 2012 from the New Mexico Indicator–Based Information System maintained by the New Mexico Department of Health (NMDH, 2014). The data are restricted to residents of New Mexico. United States birth and mortality statistics are obtained from the Center for Disease Control and Prevention National Vital Statistics System (CDC, 2014). Data for births in the sample range from 1969 to 2012. Mortality records used for the sample also begin in 1969 and end in 2012.

Data for net international migration to Las Cruces come from Moody's Analytics for the period from 1991 through 2013 (White, 2014). In contrast, net domestic migration data are not readily available for most counties in the United States. Consequently, net domestic migration is determined residually from the other series in the sample (Shryock & Siegel, 1980).

Data on economic variables used as regressors come from several sources. Personal income data for Doña Ana County are for 1969 to 2012 (BEA, 2014a). Personal income for Doña Ana County is adjusted for inflation using the personal consumption expenditures deflator (BEA, 2014b). National and local employment data also cover the 1969 to 2012 (BEA, 2014b). A ratio of local employment to national employment is used to measure labor market conditions. In addition, several variables for Mexico are obtained for the same time period (WB, 2014). These data are population, GDP, and the GDP deflator. Table 1 lists definitions and the units of measure for the variables employed in the analysis.

Table 1. Variable Definitions

Mnemonic	Description	Units
DABIR	Doña Ana County resident births	Thousands
DADEA	Doña Ana County resident deaths	Thousands
DAEMP	Doña Ana County total full- and part-time employment	Thousands
DANDM	Doña Ana County net domestic migration	Thousands
DANI	Doña Ana County natural increase	Thousands
DANIM	Doña Ana County net international migration	Thousands
DANM	Doña Ana County net migration	Thousands
DAPOP	Doña Ana County total population	Thousands
DARPYPCE	Doña Ana County real personal income	Thousands of 2009 $
DIFF	United States – Mexico GDP per capita differential	Millions of 2009 $
DPUBIRUP	Scaled national birth to population ratio	NA (ratio)
DPUDEAUP	Scaled national death to population ratio	NA (ratio)
EMP	Doña Ana County / United States employment ratio	NA (ratio)
USBIR	United States births	Thousands
USDEA	United States deaths	Thousands
USEMP	United States total full- and part-time employment	Thousands
USPOP	United States population	Thousands

4. Methodology

This study represents an extension of work originally documented in Peach and Williams (1994) for the Las Cruces region. It utilizes different aspects of the econometric approaches to analyzing population growth detailed in Plaut (1981), Fullerton and Barraza de Anda (2008), and Fullerton et al. (2014). Least squares regression analysis is used to quantify the various components of population growth. Separate models for the components of population change often exhibit greater forecast accuracy (Plaut, 1981). Out-of-sample simulations are used to examine if that is also the case for Doña Ana County. Equations (1) through (4) are identities while equations (5) through (8) are stochastic. Current year population for Doña Ana County is the sum of its prior year value, natural increase, and net migration as shown in Equation (1).

$$dapop_t = dapop_{t-1} + dani_t + danm_t \qquad (1)$$

Equation (2) represents natural increase at time t. Natural increase is calculated as the difference between births and deaths. Fertility and mortality records are accurate at the county level, allowing for reliable natural increase measurement. Historically, in the short-to-medium run, fertility changes are small and tend not to hamper forecast capabilities (Plaut, 1981).

$$dani_t = dabir_t - dadea_t \qquad (2)$$

At the regional level, net migration can have a greater effect on population change than natural increase. In Equation (3), net migration in Doña Ana is the sum of net domestic and net international migration.

$$danm_t = dandm_t + danim_t \qquad (3)$$

In the absence of time series data for net migration flows, historical net migration is calculated indirectly using the residual method (Davis, 1995). That is, net migration is the difference between the change in population from one time period to the next and natural increase. A disadvantage of this method is that migration rates are extremely volatile over time (Davis, 1995) and errors can occur as a consequence of errors in population counts. In addition, international migrants can either settle or "float" in areas immediately along the border. Nonetheless, this technique is preferred in regional economics to modeling total population because it has historically provided accurate forecasts (Plaut, 1981). For this study, net international migration estimates are available, while net domestic migration estimates are generated using Equation (4).

$$dandm_t = \left[\left(dapop_t - dapop_{t-1} \right) - dani_t \right] - danim_t \quad (4)$$

The framework for analyzing the combined effect of economic variables on the variability of births, deaths, and net migration is completed using generalized least squares estimation procedures. Although demographic-economic interaction approaches do not always yield forecasts that outperform the outcomes of more simple specifications, such approaches provide clear theoretical foundations to population projections, provide consistent economic and demographic projections, and are useful analytical tools (Smith, 1984).

Moreover, the structural analysis of demographic-economic interactions using time-series methods allows for the extrapolation of these data by exploiting the relationship that exists over time for any single variable (Asteriou & Hall, 2011). However, demographic variables

tend to be non-stationary and variables like births or deaths may not have the same statistical moments in 1969 as in 2012. That is, the means of those series, at the very least, are not constant over time, making it more difficult to reliably model these time series using historical data. In addition, previous period disturbances may affect subsequent error term values, otherwise known as serial correlation. For those reasons, two sets of models are specified for Las Cruces using level and differenced data. In the differenced equations, each variable is differenced one or more times to obtain stationarity.

The specifications for the stochastic equations are described below. Several tests and evaluation criteria are used to choose the final specification for each equation. This includes the Autocorrelation Function (ACF), the Q-statistic (Hoff, 1983), and the Breusch-Godfrey Lagrange Multiplier test (Breusch & Godfrey, 1981). Finally, the Deviance Information Criterion (DIC) selection tool is utilized to choose the functional form. The DIC allows comparing models that have high explanatory powers (Spiegelhalter, 2002; Xiao et al., 2007).

As noted in Booth (2006), regional fertility time series are conditioned by national trends and are generally non-stationary. Accordingly, a national birth to national population ratio, scaled by the population of Doña Ana County, is included in the births specification shown in Equation (5). Additionally, Doña Ana County real personal income is included as an explanatory variable. Some studies indicate that the curve relating fertility and income flattens out or even rises at higher income levels, while other data indicate a strong negative relationship between income and fertility (Becker, 1960). In this study, births are predicted to be inversely correlated with real personal income.

$$\begin{aligned} \mathrm{dabir}_t &= \alpha_o + \alpha_1 \mathrm{dabir}_{t-1} + \beta_0 \mathrm{dapop}_t * \left(\mathrm{usbir}_t / \mathrm{uspop}_t \right) \\ &+ \beta_1 \mathrm{dapop}_{t-1} * \left(\mathrm{usbir}_{t-1} / \mathrm{uspop}_{t-1} \right) + \dots \\ &+ \beta_n \mathrm{dapop}_{t-n} * \left(\mathrm{usbir}_{t-n} / \mathrm{uspop}_{t-n} \right) + \theta_0 \mathrm{darpypce}_t \\ &+ \theta_1 \mathrm{darpypce}_{t-1} + \dots + \theta_n \mathrm{darpypce}_{t-n} + \varepsilon_t \end{aligned} \quad (5)$$

Deaths in Las Cruces are modeled as a function of prior year deaths and the ratio of United States deaths to United States population, scaled by the population of Doña Ana County. In addition, a real personal income variable is included in Equation (6) to account for Doña Ana County economic development. As an area's economic conditions improve and standards of living rise, death rates tend to fall (Todaro & Smith, 2011).

$$\begin{aligned} \mathrm{dadea}_t &= \alpha_o + \alpha_1 \mathrm{dadea}_{t-1} + \beta_0 \mathrm{dapop}_t * \left(\mathrm{usdea}_t / \mathrm{uspop}_t \right) \\ &+ \beta_1 \left(\mathrm{dapop}_{t-1} * \left(\mathrm{usdea}_{t-1} / \mathrm{uspop}_{t-1} \right) + \dots \right. \\ &+ \beta_n \left(\mathrm{dapop}_{t-n} * \left(\mathrm{usdea}_{t-n} / \mathrm{uspop}_{t-n} \right) + \theta_0 \mathrm{darpypce}_t \\ &+ \theta_1 \mathrm{darpypce}_{t-1} + \dots + \theta_n \mathrm{darpypce}_{t-n} + \mu_t \end{aligned} \quad (6)$$

To model net migration for Doña Ana County, two separate stochastic equations are utilized. First, the specification for net domestic migration shown in Equation (7) is taken directly from Fullerton and Barraza de Anda (2008), where the main economic variable is a local to national employment ratio. Net domestic migration in Doña Ana County is expected to increase if local employment increases relative to employment in the United States as a whole.

$$\begin{aligned} \mathrm{dandm}_t &= \alpha_o + \alpha_1 \mathrm{dandm}_{t-1} \\ &+ \beta_0 \left(\mathrm{daemp} / \mathrm{usemp} \right)_t + \beta_1 \left(\mathrm{daemp} / \mathrm{usemp} \right)_{t-1} \\ &+ \dots + \beta_n \left(\mathrm{daemp} / \mathrm{usemp} \right)_{t-n} + w_t \end{aligned} \quad (7)$$

Second, an equation for net international migration is also specified. Migration flows are affected by push and pull factors. One standard approach posits income differentials and employment opportunities as primary determinants of international migration (Massey et al., 1994). Places located immediately along an international border are affected not only by changes outside the region, but outside the country as well (Fullerton, 2001). Population growth due to migration is attributed to wage differentials and employment opportunities as immigrants search for improved economic conditions. Equation (8) approximates the wage differential between countries by the gap between United States real per capita income, multiplied by the probability of employment, and real per capita gross domestic product (GDP) in Mexico (Fullerton & Barraza de Anda, 2008).

$$\begin{aligned} \mathrm{danim}_t &= \alpha_o + \alpha_1 \mathrm{danim}_{t-1} + \beta_0 \mathrm{diff}_t \\ &+ \beta_1 \mathrm{diff}_{t-1} + \dots + \beta_n \mathrm{diff}_{t-n} + v_t \end{aligned} \quad (8)$$

Statistical estimation diagnostics and the DIC outcomes help determine the model forms selected below. Out-of-sample model simulations also help assess the validity of the system of equations that is developed. All empirical results are discussed in the next section.

5. Results

The estimates for the four individual components of population change – births, deaths, net domestic migration, and net international migration – are compared using the Deviance Information Criterion (DIC) selection tool for Doña Ana County. The DIC is used to compare different methods when using an identical dependent variable. Equations for the four components of demographic change are estimated using level data and then compared with four equations that employ differenced data.

The DIC tool is based on a goodness of fit measure and model size (Berg et al., 2004). Less complex models with a smaller DIC are preferred. The DIC does not provide a strict criterion for model selection (Berg et al., 2004). Therefore,

$$n \ln 2\pi + 2n \ln \sigma + \left(\frac{1}{\sigma^2} \right) \sum_{t=1}^{n} \varepsilon^2 \quad (9)$$

is similar to the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) (Spiegelhalter et al., 2002; Xiao et al., 2007). In Equation (9), n is the number of observations, ε is the residual term, and σ, the standard deviation of the residuals.

Deviance Information Criterion results are reported in Table 2. Figures in bold indicate the preferred model for each dependent variable. The level form of the birth equation is preferred over the differenced equation. In addition, DIC results indicate that the differenced death equation is preferred to the level form equation. Further, a differenced net domestic migration equation is favored over the level form. Lastly, a level net international

migration specification for Doña Ana County is also favored on the basis of the DIC calculations. The overall conclusion arising from comparison of DIC statistics for the level and differenced models is that neither of these two models is systematically superior to the other across all cases.

Table 2. Deviance Information Criterion Results

Dependent Variables	Level Models	Differenced Models
Births	**-96.0793**	-87.1729
Deaths	-151.1367	**-153.9718**
Net Domestic Migration	101.3507	**87.3371**
Net International Migration	**-33.9843**	-23.5249

One potential drawback of models estimated using differenced series is that they may omit long-run relationships (Engle & Granger, 1987; Wooldridge, 2013). As noted above, most demographic data tend to change slowly. If a combination of level non-stationary series forms a stationary (cointegrated) equation then a long-run equilibrium exists, which can be appropriately modeled using data in level form. To determine whether this is the case for the four stochastic equations described above, cointegration tests are performed. First, Augmented Dickey-Fuller (ADF) unit root tests of the individual first-differenced series reject the unit root null hypothesis at the 5 percent level; all variables are of the same order of integration, I(1), and are cointegration candidates. A two-step Engle-Granger test checks for cointegration. Dickey-Fuller results for the equation error terms are conducted and the test statistics are compared against Davidson-MacKinnon critical values. Also, Ljung-Box Q-statistics

and Breusch-Godfrey Lagrange Multiplier statistics check for the existence of autocorrelation.

Testing the stationarity of the residuals with the Dickey-Fuller test for births, deaths, net domestic, and net international migration confirms that all four of the equations represent cointegrating vectors (the individual equation results are discussed below). The computed ADF test statistic for the residuals of the birth equation is greater in magnitude than the Davidson-MacKinnon critical value associated with the five percent significance level, so the unit root null hypothesis is rejected. This finding confirms the cointegration suspicions given the DIC results. Further, Lagrange Multiplier and Q-statistics both indicate the absence of autocorrelation.

The estimated birth equation for Doña Ana County confirms a long-run relationship between the dependent and independent variables (Table 3). The computed F-statistic confirms the explanatory power of the combined independent variables. Births in Las Cruces are found to contain a strong inertial component. $DABIR_{t-1}$ is statistically significant at the five percent level. In addition, the top-down relationship between national and regional changes is confirmed for Doña Ana County. Births in the United States are representative of births in Doña Ana County. Real personal income has the predicted negative effect on births within Doña Ana County and is significant at the 5 percent level. The finding supports the results found in other studies conducted using regional data (Fukuchi & Yamaguchi, 1997; Isen & Stevenson, 2010). For every one billion dollar increase in real personal income, there tends to be a reduction of 74 births per year.

Table 3. Doña Ana Births Level Equation

Dependent Variable: DABIR
Method: Least Squares
Sample (adjusted): 1972 2012
Included observations: 41 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.1030	0.07023	-1.4671	0.1508
DABIR(-1)	0.5685	0.1021	5.5675	0.0000
DPUBIRUP	0.7061	0.1459	4.8407	0.0000
DARPYPCE	-7.36E-08	1.90E-08	-3.8647	0.0004
R-squared	0.9871	Mean dependent variable		2.6267
Adjusted R-squared	0.9861	Akaike info criterion		-2.1547
S.E. of regression	0.07867	Schwarz criterion		-1.9876
Sum of squared residuals	0.2290	Hannan-Quinn criterion		-2.0939
Log likelihood	48.1709	Durbin-Watson statistic		1.6532
F-statistic	945.9313	Probability (F-statistic)		0.0000
Breusch-Godfrey LM stat. [a]	1.1896	Probability (LM statistic)		0.7555
Ljung-Box Q statistic [a]	0.9394	Probability (Q statistic)		0.8160
ADF statistic [b]	-5.4207	Probability (ADF statistic)		0.0001

[a] H_0: no serial correlation; three lags of the residuals are used to calculate the test statistics.
[b] H_0: the residuals have a unit root.

Output results for deaths in Doña Ana County are found in Table 4. Earlier versions of this equation exhibited signs of serial correlation. Consequently, an autoregressive term is included in the equation to correct for residual autocorrelation. The Q-statistic and the Breusch-Godfrey LM test statistic reported in Table 4 confirm that autocorrelation is not present in the final specification. Further, the computed Dickey Fuller test-statistic for the death equation surpasses the critical value at the five percent significance level. Accordingly, the unit root null hypothesis is rejected. This means the residuals are

stationary at the five percent significance level and that the dependent and independent variables are cointegrated.

Although the constant term does not satisfy the five percent significance criterion, all of the other coefficients do. Individually, the explanatory variables are significant at the 5 percent level. Deaths in Doña Ana County also exhibit a fairly strong inertial tendency pattern with approximately 760 people passing away for every 1,000 persons that die in the previous year. In addition, the scaled national death-to-population ratio is positively correlated with deaths in Las Cruces, confirming the

prevalence of national demographic trends throughout most regions of the country. The autoregressive term is lagged three periods to account for the persistent effects of past demographic and epidemiological dynamics on present mortality trends and the slow evolution of those

patterns over time. The variable describing real personal income is omitted from the specification reported in Table 4 because it did not help explain the variation in the dependent variable.

Table 4. Doña Ana County Deaths Level Equation

Dependent Variable: DADEA
Method: Least Squares
Sample (adjusted): 1975 2012
Included observations: 38 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.1116	0.06578	-1.6965	0.0989
DADEA(-1)	0.7589	0.1098	6.9114	0.0000
DPUDEAUP(-1)	0.2789	0.1179	2.3647	0.0239
AR(3)	0.4828	0.1531	3.1534	0.0034
R-squared	0.9894	Mean dependent variable		0.8828
Adjusted R-squared	0.9884	Akaike info criterion		-3.7619
S.E. of regression	0.03510	Schwarz criterion		-3.5895
Sum of squared residuals	0.04189	Hannan-Quinn criterion		-3.7006
Log likelihood	75.4763	Durbin-Watson statistic		2.3714
F-statistic	1055.357	Probability (F-statistic)		0.0000
Breusch-Godfrey LM stat. [a]	3.3932	Probability (LM statistic)		0.3349
Ljung-Box Q statistic [a]	3.2779	Probability (Q statistic)		0.1940
ADF statistic [b]	-7.9061	Probability (ADF statistic)		0.0000

[a] H_0: no serial correlation; three lags of the residuals are used to calculate the test statistics.
[b] H_0: the residuals have a unit root.

Table 5 shows the estimated output for the net domestic migration equation. The ADF test-statistic is larger in magnitude than the critical value at the five percent level indicating that the equation is cointegrated. Further, residual autocorrelation does not appear to be problematic for this equation. Net domestic migration is a function of net domestic migration in the last period, plus contemporaneous and one period lags of the ratio of employment in Doña Ana County to total United States

employment. A strong inertial component in net domestic migration is confirmed by the influx of 688 migrants that occurs subsequent to every 1,000 prior period migrants to the county. Also, Table 5 shows that the effects regional labor market conditions on net migration wear off fairly quickly. Although this equation does not exhibit optimal statistical fit characteristics, it does have good out-of-sample simulation properties associated with it.

Table 5. Doña Ana County Net Domestic Migration Level Equation

Dependent Variable: DANDM
Method: Least Squares
Sample (adjusted): 1992 2012
Included observations: 21 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.4389	6.8537	0.2099	0.8362
DANDM(-1)	0.6877	0.1936	3.5526	0.0024
EMP	194223.3	106689.7	1.8205	0.0863
EMP(-1)	-198052.0	105588.4	-1.8757	0.0780
R-squared	0.4509	Mean dependent variable		3.3416
Adjusted R-squared	0.3540	Akaike info criterion		5.2060
S.E. of regression	3.0019	Schwarz criterion		5.4050
Sum of squared residuals	153.1931	Hannan-Quinn criterion		5.2492
Log likelihood	-50.6631	Durbin-Watson statistic		1.3663
F-statistic	4.6540	Probability (F-statistic)		0.01497
Breusch-Godfrey LM stat. [a]	4.2433	Probability (LM statistic)		0.2364
Ljung-Box Q statistic [a]	3.7146	Probability (Q statistic)		0.2940
ADF statistic [b]	-4.8743	Probability (ADF statistic)		0.0010

[a] H_0: no serial correlation; three lags of the residuals are used to calculate the test statistics.
[b] H_0: the residuals have a unit root.

Table 6 shows the output for the estimated net international migration equation. The Dickey Fuller t-statistic is greater in absolute value than the Davidson-MacKinnon critical value at the 5 percent significance level indicating that the residuals are stationary. The LM and Q-statistics indicate that autocorrelation is not present. Individually and together all variables are statistically

significant. International migration flows to Doña Ana County contain an inertial component and are positively correlated with the previous year's migration flows.

Although that result is not surprising, the coefficients for thetwo- three-period lags of the real per capita GDP differential, scaled by the probability of employment, run counter to what is hypothesized. The lagged effect of the

GDP differentials on migration may be due to communication lags in the flow of information via network effects to migrant-sending areas (Ruyssen et al., 2014).The positive coefficient for the two-year lag supports the finding that improved economic conditions are driving forces behind migration (Lim, 2011). It is unexpectedly offset by the parameter estimate for the three-year lag.

The GDP differential lagged three periods is inversely correlated with international migration for Las Cruces. Its absolute magnitude exceeds that of the two-year lag coefficient. The negative effect does not support the hypothesis advanced for this variable and runs counter to what has been documented for the nearby El Paso metropolitan economy (Fullerton and Barraza de Anda, 2008). Because migration often entails substantial costs, it may not be a realistic option for potential migrants with very low incomes (Vogler and Rotte, 2000). According to this line of reasoning, the stimulative effect of reduced incomes in migrant-sending countrieson immigration to the United States is dampened by poverty constraints in the former areas. Another potential factor undermining the hypothesized positive effect of the GDP gap on international migration is heightened immigration restrictions in the United States. Mayda (2010) finds that GDP in migrant-sending countries often has a null or even positive effect on out-migration and hypothesizes that this counterintuitive result may be due to the intensity of immigration controls in migrant-receiving countries. Such restrictions raise the costs associated with entering the destination country and have been observed in the United States in recent years.

Table 6. Doña Ana County Net International Migration Level Equation

Dependent Variable: DANIM
Method: Least Squares
Sample (adjusted): 1992 2012
Included observations: 21 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.5504	0.3801	4.0791	0.0008
DANIM(-1)	0.5311	0.1304	4.0720	0.0008
DIFF(-2)	80.6674	30.8182	2.6175	0.0180
DIFF(-3)	-118.8520	33.1988	-3.5800	0.0023
R-squared	0.9015	Mean dependent variable		0.7810
Adjusted R-squared	0.8841	Akaike info criterion		-1.2385
S.E. of regression	0.1197	Schwarz criterion		-1.0396
Sum of squared residuals	0.2435	Hannan-Quinn criterion		-1.1953
Log likelihood	17.0045	Durbin-Watson statistic		1.8696
F-statistic	51.8626	Probability (F-statistic)		0.0000
Breusch-Godfrey LM stat. [a]	2.3165	Probability (LM statistic)		0.5094
Ljung-Box Qstatistic[a]	2.3914	Probability (Q statistic)		0.4950
ADF statistic[b]	-3.9557	Probability (ADF statistic)		0.0073

[a]H_0: no serial correlation; three lags of the residuals are used to calculate the test statistics.
[b]H_0: the residuals have a unit root.

To further examine the reliability of the equations, out of sample simulations are conducted to forecast population change five years into the future. Time series extrapolations are widely used because they are easily automated and reliable (Armstrong, 2001). Because the simulation period is beyond the sample, some assumptions are required for the independent variables (Pindyck and Rubinfeld, 1976) in order to ultimately calculate total population changes. The percentage changes for the forecasted independent variables are obtained from the United States Census Bureau (USCB, 2012), Moody's Analytics (White, 2014), HIS Economics (Handler and Behravesh, 2014), Woodward (2013), and the PRS Group (2013). The variables are used to simulate the four components of population change to 2018 period and the values for those components are then incorporated into Equation (1) to calculate total projected population. The forecasts from 2013 forward are shown in Table 7.

Table 7. Endogenous Variable Forecasts (in thousands), 2013-2018

Endogenous Variables (in thousands), 2013 - 2018						
	2013	**2014**	**2015**	**2016**	**2017**	**2018**
DABIR	3.125	3.196	3.266	3.334	3.407	3.487
% Change	1.76%	2.27%	2.19%	2.09%	2.20%	2.34%
DADEA	1.531	1.539	1.581	1.621	1.646	1.689
% Change	2.92%	0.48%	2.76%	2.54%	1.52%	2.63%
DANDM	-0.662	-2.147	-1.540	-0.659	-0.439	-0.581
% Change	282.71%	224.30%	-28.28%	-57.23%	-33.28%	32.30%
DANIM	0.371	0.442	0.467	0.441	0.392	0.236
% Change	307.74%	19.02%	5.64%	-5.47%	-11.08%	-39.81%
DAPOP	214.492	214.444	215.056	216.551	218.265	219.718
% Change	0.61%	-0.02%	0.29%	0.70%	0.79%	0.67%

As can be seen from Table 7 above, total births and deaths are both predicted to increase steadily through 2018.

Net domestic migration is the largest determinant of population change affecting Doña Ana County in the near

future. Net domestic out-migration from the county peaks at 2,147 in 2014, but will subside noticeably thereafter. Net international migration to the region will average391 from 2013 to 2018. In all, with the irregular movement and large outflows of domestic migrants and the relatively small influx of international migrants, Doña Ana County is projected to observe slow population growth in 2015, with a steady upward shift from 2016 to 2018. These simulated outcomes are similar to what has been previously documented for this metropolitan economy (White, 2014).

Forecasts of demographic trends in Doña Ana County are affected by international, national, and state economic and demographic conditions. Historically, that has meant that population growth in Las Cruces is subject to surprising changes. The system of equations developed herein allows for simulating different potential outlook scenarios that emanate from the individual segments modeled. This approach appears to offer a good complement to the approach earlier developed in Peach & Williams (1994).

6. Conclusion

Measuring and predicting population change in response to economic conditionsis often important forregional economic planning efforts. Of the substantial research that exists for the United States-Mexico border, econometric analyses of border population growth are fairly limited. The purpose of this paper is to develop such a model for Doña Ana County, New Mexico. Estimation and simulation results for the system of equations developed exhibit good properties.

Four sets of level and four sets of differenced equations are generated to analyze the components of population growth. Comparisons between the two types of equations using the Deviance Information Criterion (DIC) do not reveal a preferred modeling method. However, Augmented Dickey-Fuller cointegration tests support the application of level form equations. This is a key finding because modeling differenced series may omit long-run relationships (Engle & Granger, 1987; Wooldridge, 2013).

An out-of-sample simulation is developed using economic forecasts of the exogenous variables. Results from that exercise indicate that natural increase will remain positive because births outnumber deaths. Net domestic migration is predicted to remain negative throughout the forecast period, although the volume of net out-migration shrinks gradually after 2014. Net international migration is expected to decrease in response to international economic factors. Using these results, total population is projected through the year 2018. Under the scenario simulated, Las Crucesexhibitsone year of negative population change, with total population is expected to trend upward at a relatively slow rate.

Although the level set of econometric models used here appear to be useful for projecting population growth in Doña Ana County, the validity of the approach should be tested for other border regions. Border counties have generally been overlooked in terms of econometric analysis. Relatively volatile rates of growth for many of these counties mean that accurate modeling efforts are likely to encounter numerous challenges.

References

[1] Anderson, J.B. and Gerber, J. (2008). Fifty years of change on the U.S.-Mexico border: Growth, development, and quality of life. Austin, TX: University of Texas Press.

[2] Armstrong, J.S. (2001). Extrapolation for time-series and cross-sectional data. Principles of forecasting: A handbook for researchers and practitioners. Norwell, MA: Kluwer Academic Publishers.

[3] Asteriou, D. and Hall, S.G. (2011). Applied econometrics. New York, NY: Palgrave Macmillan.

[4] BEA (2014a). Local area personal income and employment. Washington, DC: U.S. Bureau of Economic Analysis.

[5] BEA (2014b). National income and product accounts tables. Washington, DC: U.S. Bureau of Economic Analysis.

[6] Berg, A., Meyer, R., and Yu, J. (2004). Deviation information criterion for comparing stochastic volatility models. Journal of Business and Economic Statistics, 22(1), 107-120.

[7] Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. International Journal of Forecasting, 22(3), 547-581.

[8] Becker, G. S. (1960). An economic analysis of fertility. Demographic and economic change in developed countries, a conference of the Universities-National Bureau Committee for Economic Research. New York, NY: National Bureau of Economic Research, distributed by Columbia University Press.

[9] Bolton, R. (1985). Regional econometric models. Journal of Regional Science, 25(4), 495-520.

[10] Breusch, T.S. and Godfrey, L. (1981). A review of recent work on testing for autocorrelation in dynamic simultaneous models. In D. Currie, R. Nobay and D. Peel (Eds.), Macroeconomicanalysis: Essays in macroeconomics and econometrics(pp. 63-105). London, UK: Croom Helm.

[11] CDC (2014). Vital statistics data. Available at: http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm

[12] Cebula, R.J. and Alexander, G.M. (2006). Determinants of net interstate migration, 2000-2004. Journal of Regional Analysis and Policy, 36(2), 116-123.

[13] Cebula, R.J. and Clark, J.R. (2011). Migration, economic freedom, and personal freedom: An empirical analysis. Journal of Private Enterprise, 27(1), 43-62.

[14] Cebula, R.J. and Clark, J.R. (2013). An extension of the Tiebouthypothesis of voting with one's feet: The Medicaid magnet hypothesis. Applied Economics, 45(31), 4575-4583.

[15] Cushing, B. and Poot, J. (2004). Crossing boundaries and borders: Regional science advances in migration modeling. Papers in Regional Science, 83(1), 317-338.

[16] Davis, H.C. (1995). Demographic projection techniques for regions and smaller areas, a primer. Vancouver, Canada: UBC Press.

[17] Djafar, F. and Hassan, M.K. (2012). Dynamics of push and pull factors of migrant workers in developing countries: The case of Indonesian workers in Malaysia. Journal of Economics and Behavioral Studies, 4(12), 703-711.

[18] Engle, R.F. and Granger, C.W.J. (1987). Co-integration and error correction: Representation, estimation, and testing. Econometrica, 55(2), 251-276.

[19] Estrella Valenzuela, G. (1992). The floating population of the border. In J.R. Weeks & R. Ham-Chande (Eds.), Demographic dynamics of the U.S.-Mexico border (pp. 187-200).El Paso, TX: Texas Western Press.

[20] Fukuchi, T. and Yamaguchi, M. (1997).An econometric analysis of a suburban city. Studies in Regional Science, 27(2), 1-31.

[21] Fullerton, T.M., Jr. (2001). Specification of a Borderplex econometric forecasting model. International Regional Science Review, 24(2), 245-260.

[22] Fullerton, T.M., Jr. and Barraza de Anda, M.P. (2008). Borderplex population modeling. Migraciones Internacionales, 4(3), 91-104.

[23] Fullerton, T.M., Jr. Ramirez, D.A. and Walke, A.G. (2014). An econometric analysis of population change in Arkansas. Oxford Journal, 9(1), 28-40.

[24] Handler, D. and Behravesh, N. (2014).A note on economic data versus economic vitality.US Economic Outlook, (July), 5-18, Lexington, MA: IHS Global Insight.

[25] Harris, J.R. and Todaro, M.P. (1970). Migration, unemployment, and development: A two-sector analysis. American Economic Review, 60(1), 126-142.

[26] Hernández-Murillo, R., Ott, L.S., Owyang, M.T. and Whalen, D. (2011). Patterns of interstate migration in the United States from the survey of income and program participation. Federal Reserve Bank of St. Louis Review, 93(3), 169-185.

[27] Hoff, J.C. (1983). A practical guide to Box-Jenkins forecasting. Belmont, CA: Lifetime Learning Publications.

[28] Isen, A. and Stevenson, B. (2010).Women's education and family behavior: Trends in marriage, divorce, and fertility (No. w15725). Cambridge, MA: National Bureau of Economic Research.

[29] Kazlauskienė, A. andRinkevičius, L. (2006). Lithuanian "brain drain" causes: Push and pull factors. Engineering Economics, 46(1), 27-37.

[30] Lee, E.S. (1966). A theory of migration. Demography, 3(1), 47-57.

[31] Lim, J. (2011). Does wage differential driven migration continue to exist? Tests on the role of regional economic structure in wage differential driven migration. Annals of Regional Science, 47(1), 213-233.

[32] Massey, D.S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A. and Taylor, J.E. (1994). An evaluation of international migration theory: The North American case. Population and Development Review, 20(4), 699-751.

[33] Mayda, A.M. (2010). International migration: A panel data analysis of the determinants of bilateral flows. Journal of Population Economics, 23(4), 1249-1274.

[34] NMDH. (2014). New Mexico's indicator-based information system (NM-IBIS).Santa Fe, NM: New Mexico Department of Health.

[35] Peach, J. and Williams, J.D. (1994).Demographicchange inthe El Paso-Juárez-Las Cruces region. Estudios Fronterizos, 34 (Julio-Diciembre), 117-138.

[36] Peach, J. and Williams, J.D. (2000). Population and economic dynamics on the U.S.-Mexican border: Past, present, and future. Chapter 4 in P. Ganster (Ed.), The U.S. -Mexican border environment: A road map to a sustainable 2020. San Diego, CA: San Diego State University Press.

[37] Pindyck, R.S. andRubinfeld, D.L. (1976). Econometric models and economic forecasts. New York, NY: McGraw Hill.

[38] Plaut, T.R. (1981). An econometric model for forecasting regional population growth. International Regional Science Review, 6(1), 53-70.

[39] PRS Group. (2013). Mexico country report. East Syracuse, NY: The PRS Group, Inc.

[40] Ruyssen, I., Evaert, G. and Rayp, G. (2014).Determinants and dynamics of migration to OECD countries in a three-dimensional panel framework.Empirical Economics 46(1), 175-197.

[41] Shryock, H.S., Siegel, J.S. and Associates (1980).The methods and materials of demography.Washington, DC: United States Government Printing Office.

[42] Smith, S.K. (1984). Population projections: What do we really know? BEBRMonographs, 1.Gainesville, FL: Bureau of Economic and Business Research University of Florida.

[43] Smith, S.K. (1997). Further thoughts on simplicity and complexity in population projection models. International Journal of Forecasting, 13(4), 557-565.

[44] Shultz, T.P. (2005). Fertility and income (Discussion paper no. 925). Yale University: Economic Growth Center.

[45] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64(4), 583-639.

[46] Stark, O. andTaylor, J.E. (1991). Migration incentives, migration types: The role of relative deprivation. Economic Journal, 101(408), 1163-1178.

[47] Tayman, J. and Swanson, D.A. (1996). On the utility of population forecasts. Demography, 33(4), 523-528.

[48] Todaro, M.P. and Smith, S.C. (2011). Economic development (11thed.). Boston, MA: Addison-Wesley.

[49] Tuiran, R.A. (1992). Households and emigration in the northern border: The case of Reynosa. In J.R. Weeks & R. Ham-Chande (Eds.), Demographic dynamics of the U.S. -Mexico border (pp. 165-186).El Paso, TX: Texas Western Press.

[50] USCB. (2012). Projected population by single year of age, sex, race, and Hispanic origin for the United States: 2012 to 2060.Washington, DC: United States Census Bureau.

[51] Vogler, M. and Rotte, R. (2000) The effects of development on migration: Theoretical issues and new empirical evidence. Journal of Population Economics, 13(3), 485-508.

[52] WB (2014). World development indicators database: Mexico. Washington, DC: World Bank.

[53] White, D. (2014). Las Cruces. Moody's Analytics Précis U.S. Metro. West Chester, PA: Moody's Economic & Consumer Credit Analytics.

[54] Woodward, M.C. (2013). The U.S. economy to 2022: Settling into a new normal. Washington, D.C.: U.S. Bureau of Labor Statistics.

[55] Wooldridge, J.M. (2013). Introductory econometrics: A modern approach (5thed.). Mason, OH: South-Western, Cengage Learning.

[56] Xiao, N., Zarnikau, J. and Damien, P. (2007). Testing functional forms in energy modeling: An application of the Bayesian approach to U.S. electricity demand. Energy Economics 29(2), 158-166.