Working Paper

AUGUST 2009

An Exploratory Time Series Analysis of Apprehensions and Linewatch Hours on the Southwest Border

DEREKH CORNWELL, Ph.D.*

Executive Summary

Whether heightened border enforcement reduces future undocumented migration is a question at the heart of most policy debates on U.S. immigration reform. To address this question, most researchers utilize publicly available data on border apprehensions and linewatch hours, which serve as proxies for the flow of unauthorized migration and the underlying enforcement effort, respectively. A standard assumption of most researchers is that apprehensions are a function of linewatch hours. It is conceivable, however, that linewatch hours may also be a function of apprehensions. Although some studies recognize the possibility of a reverse relationship, the implied hypothesis—namely, that apprehensions predict linewatch hours—has not been formally tested. This paper fills the gap by providing a formal time series analysis of the historical relationship between monthly apprehensions and linewatch hours between 1963 and 2004. Overall, the findings indicate that while apprehensions and linewatch hours were strongly correlated in the same time period, past linewatch hours were not strong predictors of future apprehensions, or vice versa. This suggests that, while a strong contemporaneous relationship between linewatch hours and apprehensions exists, the two series may be less useful for forecasting purposes. Although this paper does not assess enforcement effectiveness or deterrence, the preliminary results may assist future researchers by providing empirical justification for econometric specification decisions made when studying border enforcement issues.

* Analysis Division, Office of Immigration Statistics, U.S. Department of Homeland Security. The views expressed are solely those of the author and do not necessarily reflect those of the Department of Homeland Security or any of its component agencies.

Homeland Security

Office of Immigration Statistics
POLICY DIRECTORATE

Introduction

Whether heightened border enforcement reduces future undocumented migration is a question at the heart of most policy debates on U.S. immigration reform. A core assumption of policymakers and many scholars is that the level of the unauthorized migrant flow is at least partly a function of the intensity of the underlying enforcement effort. Most models of border enforcement effectiveness posit an inverse relationship between enforcement intensity, which is usually operationalized as the number of linewatch hours during a specified period, and the flow of unauthorized migration, usually proxied by the number of border apprehensions. To date, most empirical work in this area specifies a structural relationship between the flow of unauthorized immigrants and border enforcement. Typically, this modeling approach uses standard econometric techniques to estimate an enforcement elasticity by regressing apprehensions on linewatch hours while controlling for economic factors and other policy or environmental conditions. A positive and statistically significant coefficient on the linewatch hour variable is usually interpreted as evidence of an enforcement effect, which is then used to generate a lower bound on the underlying apprehension probability (Hanson and Spilimbergo, 1999a). It is this bound on the apprehension probability that is used to infer the deterrence effect of enforcement efforts on unauthorized migration.

The probability of apprehension is not directly observed and, at best, can only be indirectly estimated. The reasonableness of this indirect estimation, however, ultimately depends on the validity of the identifying restrictions employed in the original structural model. As such, the core identifying restriction is that apprehensions are a function of linewatch hours. It is entirely possible, however, that linewatch hours may also be a function of apprehensions in which case an endogenous relationship between these two key variables may exist. While some studies attempt to control for this possibility via instrumental variable estimation (e.g., Hanson and Spilimbergo, 1999a; Orrenius and Zavodny, 2003), the core identifying assumption that linewatch hours are exogenous to apprehensions has not been tested. Moreover, the time series properties of these two key variables have not been adequately addressed by the extant border enforcement literature.

The central goal of this paper is to fill this void by providing a formal time series analysis of the relationship between apprehensions and linewatch hours using a flexible Vector autoregression (VAR) approach. Analyzing the historical relationship between these two variables enables one to take a step back from some of the more restrictive aspects of structural models and examine the dynamics of the unauthorized immigrant flow and border enforcement over time. Specifically, the VAR approach enables one to examine the validity of the endogeneity/exogeneity assumptions implicit in existing structural models as well as investigate how change in one variable affects movement in the other over time. Although the approach adopted here is modest in that it does not estimate an enforcement elasticity, assess enforcement effectiveness and deterrence, or attempt to incorporate the influences of economic factors, it does provide an account of the dynamics in the data, which are missing from existing studies. Vector autoregression models also enable one to trace out the dynamic changes in variables over time. By explicitly testing assumptions and examining the dynamic relationship between apprehensions and linewatch hours, the identifying restrictions imposed in structural models can be better justified and accounted for.

The remainder of this paper is organized as follows. The second section provides a brief overview of the border enforcement literature and the core assumptions of the most commonly used empirical models. Section three discusses the data and methodology employed in this paper. The fourth section discusses the results of the analysis. Finally, the paper concludes with a discussion of the caveats, shortcomings, and possible future directions of this research.

II. Theoretical Overview

II a) Measuring the Flow of Unauthorized Migration and Enforcement

To gauge the size of the unauthorized migrant flow and the effectiveness of border enforcement in reducing it, at least two pieces of information are needed: 1) a count of the number who attempt to cross the U.S. border and do so successfully; and 2) the probability that a migrant is apprehended while attempting an unauthorized border crossing. These data do not exist. As a result, estimates of both the flow of undocumented migration and the probability of apprehension must be derived from alternative data sources. The data most commonly used to generate these estimates are the number of border apprehensions, which are recorded by U.S. Customs and Border Patrol (CBP). Although these data only capture the number of foiled attempts to cross the U.S. border, which is an unknown fraction of the total number of attempted crossings in a given period, they are the only directly observed components of the total flow of unauthorized migration. In the absence of alternative data sources, apprehensions are commonly used as a proxy measure of the unauthorized inflow. How the number of apprehensions changes over time, therefore, should be of key interest when assessing enforcement effectiveness.

The probability of apprehension is another important piece of the enforcement effectiveness puzzle. The reason for this is embedded within the economic logic of migration, which views the migration decision as an investment decision. The economic view of migration posits that an individual migrates if the expected discounted difference in the stream of income between the new and old location exceeds moving costs (Sjaastad, 1962; Todaro and Maruszko, 1987; Hanson and Spilimbergo, 1999a). Included in an individual's cost function is the probability of being apprehended while attempting an unauthorized border crossing. An individual who is apprehended incurs a cost in the form of lost wages, time in detention, and return to their home country. For non-Mexican migrants crossing the southwest border, in particular, this cost may be non-trivial. The probability of apprehension is thus the primary means by which costs are imposed on unauthorized migrants and a possible deterrent effect transmitted to would-be migrants contemplating unauthorized entry.

As is the case with the total unauthorized flow, the probability of apprehension is similarly unknown. Again, the unknown nature of this quantity stems from the fact that a total count of attempted border crossings does not exist. As discussed below, some researchers (notably Espenshade, 1990, 1994) have developed models that generate estimates of the probability of apprehension; however, the most common way to measure border enforcement in the empirical literature is to use linewatch hours recorded by CBP. Linewatch hours refer to the total number of hours devoted to patrolling the border and are thought to reflect the intensity of the underlying enforcement effort. Consequently, an increase in linewatch hours may reflect a more intensive enforcement effort that should, *ceteris paribus,* increase the (still unobserved) probability of apprehension. In this sense, linewatch hours are a readily available and convenient proxy for enforcement effort and may positively correlate with the probability of apprehension. For these reasons, linewatch hours are widely used as measures of enforcement effort in empirical studies.[1]

II b) Empirical Uses of Apprehensions and Linewatch Data

Espenshade (1990; 1994) is widely credited with pioneering one of the original uses of apprehensions data for empirical analysis. In particular, Espenshade developed what is known as the Repeated Trials Model (RTM), which can be used to generate time series estimates of the flow and apprehension probabilities of unauthorized migrants. To produce these estimates, Espenshade (1990) invokes three critical assumptions: 1) undocumented migrants continuously attempt to cross the

---

[1] Linewatch hours are used almost exclusively in the literature as measures of enforcement intensity. The main reason for this stems for the ready availability of linewatch hours data. It should be noted, however, that linewatch hours are not the sole determinant of enforcement intensity. In practice, enforcement intensity also encompasses the use of technology, infrastructure, new tactics, and prosecutorial strategies. Unfortunately, monthly time series data do not exist for these variables. Therefore, the linewatch hours series is the only enforcement proxy that can be formally examined using time series analysis.

border until they are successful; 2) once migrants successfully cross the border into the U.S., they do not return to Mexico during the period of analysis; and 3) the underlying probability that a migrant is apprehended while attempting a border crossing remains constant over time. Imposing these assumptions on apprehensions and re-apprehensions data, Espenshade develops a simple probability model that generates estimates of the total flow of undocumented migration as well as the associated apprehension probability in each time period.

The core strength of the RTM is its simplicity and tractability. Holding the probability of apprehension constant reduces computational complexity. By assuming that all migrants who attempt to cross the border will eventually be successful, there is no need to estimate an unknown discouragement parameter. This also applies to the unknown parameters for return migration. While constant apprehension probabilities and no return migration are not, as Espenshade (1990) notes, tenable assumptions, they do permit analysts to generate undocumented flow estimates rather easily. In this regard, the assumptions of the RTM enable one to economize on available data in a straightforward way.

This strength diminishes when the model is used to examine enforcement effectiveness. By assumption, the probability of apprehension remains constant, which implies that it is insensitive to changes in the intensity of the underlying enforcement effort. In the original model, only apprehension and re-apprehension data are used to generate estimates of the undocumented flow.[2] While the probability of apprehension is needed to generate the flow estimates, it is held constant for mathematical convenience. Moreover, the core assumption that migrants will continue to attempt unauthorized entry until they are successful appears to discard the prospect of deterrence or any meaningful enforcement effectiveness. If analysts are solely interested in generating estimates of the unauthorized migrant flow, the RTM may be suitable for that purpose. Difficulty arises, however, when the model is employed for use in empirical evaluations of enforcement effectiveness (e.g., Espenshade, 1994). In these situations, drawing substantive conclusions about enforcement from the apprehensions probability is conceptually problematic.

Rather than rely on the RTM, most empirical studies of unauthorized migration specify a multivariate model in which apprehensions are expressed as a function of economic and policy variables as well as enforcement. As noted above, linewatch hours are a convenient and intuitively appealing measure of enforcement effort that should positively correspond with the unobserved probability of apprehension. The working hypothesis of most studies is that, all else equal, the flow of unauthorized migration should be inversely related to enforcement intensity. What is less clear, however, is how identification of the relationship between enforcement and the flow of unauthorized migrants should proceed. In the absence of well-developed theories of enforcement effectiveness, the issue of how best to model linewatch hours (or other enforcement proxies) and migration remains an unresolved question.

The implications of these issues for empirical analysis become clearer upon closer inspection of the literature. As noted earlier, apprehensions are commonly expressed as a function of enforcement, but not vice versa. Early empirical work on enforcement effectiveness, for example, examined the impact of the 1986 Immigration Reform and Control Act (IRCA) on unauthorized migration (Donato et al., 1992; Bean et al., 1990). Bean et al. (1990) estimate a multivariate regression model of apprehensions on linewatch hours and several temporal and economic control variables to assess how the undocumented flow was impacted by IRCA. Treating linewatch hours as exogenous, the authors find that apprehensions declined in the immediate post-IRCA era. Perhaps more importantly, the authors also find that increases in Border Patrol linewatch hours increased border apprehensions between 1977 and 1989. Aside from the temporal restriction of the sample, the problem with this conclusion is that it ignores potentially serious issues of model identification. By treating enforcement as an exogenous variable, the model assumes that enforcement decisions are not influenced by the unauthorized flow. To the extent that the

[2] Espenshade (1990) then estimates a multivariate model to measure the marginal impact of economic and seasonal variables on the unauthorized immigrant flow. The probability of apprehension, therefore, is only used to generate data on the unauthorized flow. This estimation strategy effectively eliminates an empirical assessment of enforcement effectiveness.

level of enforcement is a policy variable, it is plausible that the enforcement decision may be a response to the unauthorized flow. Under this scenario, CBP may set its level of enforcement in conjunction with either the contemporaneous or past volume of the unauthorized flow, which implies that linewatch hours should not be regarded as exogenous to apprehensions. If linewatch hours and apprehensions are endogenous to each other, then the empirical analysis should take this dynamic-feedback relationship into account. A more formal examination of the exogeniety of linewatch hours to apprehensions is the central goal of this paper.

More recent work relaxes the assumption that enforcement is exogenous to unauthorized migration. In their study of undocumented migration, border enforcement, and relative wages, Hanson and Spilimbergo (1999a) use both traditional Ordinary Least Squares (OLS) and instrumental variable (IV) estimation to control for potential endogneity and calculate an enforcement elasticity. Using a multivariate model that includes economic and enforcement variables, instrumented by real U.S. government defense expenditures as well as lagged values of enforcement,[3] Hanson and Spilimbergo (1999a) produce an enforcement elasticity that ranges from 0.5, for standard OLS with adjustments for serial correlation, to 1.3, when IV estimation is used. Similarly, Orrenius and Zavodny (2003), in their study of amnesty programs and undocumented migration, use standard regression techniques as well as instrumental variable methods to derive an enforcement elasticity. For the period between 1969 and 1996, Orrenius and Zavodny take first-differences of their data and calculate an enforcement elasticity of 0.44 using OLS; the enforcement coefficient rises to 0.58 when IV estimation is used. Studies examining the impact of enforcement on illegal border crossing behavior adopt a similar empirical strategy. Amuedo-Dorantes and Bansak (2007), for example, estimate a conditional probit model and find that linewatch hours, when instrumented by the Drug Enforcement Agency's (DEA) annual budget, decreases the likelihood that a migrant will attempt a repeat border crossing.

What these studies share in common is the recognition that enforcement is a policy variable that may be a response to undocumented migration. Within a structural equations framework, it is this potential endogeneity that necessitates the use of IV techniques to achieve identification and obtain consistent estimates of the impact of enforcement on apprehensions. The use of instrumental variable estimators, however, is not without potential challenges and shortcomings. Chief among these is what is known as the weak instrument problem (Angrist and Krueger, 2001). A valid instrument should be highly correlated with the endogenous variable, only indirectly related to the dependent variable via this correlation, and be uncorrelated with the error term. When instruments are either weakly correlated with the endogenous regressors or too many instruments are used, the weak instrument problem arises. Under these circumstances, coefficients will be biased.[4] Even with valid instruments, however, IV estimation tends to be less efficient than OLS, and the magnitude of this inefficiency is inversely related to the correlation between the proposed instruments and the endogenous regressors (Judge et al., 1988; Kennedy, 1998).

The dual problem of weak instruments and inefficiency may be even more pronounced in the time series context. Most of the studies cited above utilize time series data on apprehensions and linewatch hours at the monthly level of aggregation. The most commonly used instruments—namely, annual defense expenditures and the DEA budget—are yearly aggregates that are temporally mismatched with monthly linewatch hours. In addition to the potential for biased inference, inconsistent temporal aggregation between instruments and the endogenous regressors can exacerbate inefficiency. Moreover, if the historical dynamics of apprehensions and linewatch hours are of interest, then IV estimation is inappropriate because it discards information from the original data series. Rather than rely on inefficient and potentially biased IV estimation of time series data, the argument here is that endogeneity should be tested rather than merely asserted or assumed.

[3] How strong lagged values of linewatch hours are for IV purposes is an interesting question that this paper only indirectly addresses. Nevertheless, using lagged values of endogenous variables as instruments is common in Generalized Methods of Moments (GMM) estimation, but still relies on a certain understanding of the time series properties of the data. Again, it is these time series properties that are of key interest here.

[4] The weak instrument problem is of concern primarily because it re-introduces bias into the parameter estimates, which was precisely what instrumental variable estimation was designed to overcome. In such situations, the analyst may be introducing weak instrument bias as a replacement for the bias produced by endogeneity.

A related issue concerns the interpretation of enforcement elasticities. Without strong theoretical guidance, it is not entirely clear how the enforcement effects found in the literature should be interpreted. As noted above, the empirical literature typically finds positive and statistically significant impacts of linewatch hours on apprehensions, though the magnitude of the relationship varies depending on the choice of estimation technique. A positive coefficient on linewatch hours may be consistent with enforcement effectiveness if it signifies that the unobserved probability of apprehension is increasing. As the probability of apprehension increases, the cost to unauthorized migrants attempting or contemplating a border crossing also rises. There are two related problems with this inference. The first is that the probability of apprehension remains an unknown quantity, which introduces a significant amount of uncertainty into any deterrence calculation.[5] Secondly, and more importantly for the project at hand, the estimated enforcement effect only refers to contemporaneous values of whatever enforcement variables are employed in the model. It is conceivable, however, that enforcement may operate with a time lag, as information about the intensity of enforcement efforts may not spread instantaneously across the border. Such a scenario raises the possibility of lagged enforcement effects. How past values of enforcement impact current values of the unauthorized flow (and vice versa), is thus a worthy question for empirical analysis and should be part of an exploratory time series analysis.

To address this question, the historical dynamics of the apprehensions and linewatch hour data are also examined in this paper. The focus on historical dynamics can assist in the development of more structurally-oriented theories of enforcement and unauthorized migration. Given the common supposition of instantaneous enforcement effects in the literature, the prospect of dynamic adjustment between linewatch hours and apprehensions is difficult to test without an explicit time series approach. Although the bivariate nature of this study precludes a rigorous test of any dynamic adjustment hypothesis, the VAR approach discussed below can provide an exploratory first step towards developing and testing dynamic theories of enforcement effectiveness and unauthorized migration.

III. Data and Methods

The crux of the argument presented above is that the time series properties of apprehensions and linewatch hours data have not been examined or incorporated into empirical models of border enforcement. As a result, identifying restrictions are rarely tested and the potentially dynamic behavior of variables over time is often ignored. The simple bivariate VAR model presented below is intended as a first-step towards filling this gap by providing an exploratory assessment of the historical dynamics of apprehensions and linewatch hours. To this end, this section discusses the variables and data used in the empirical analysis, presents the empirical model, and elaborates on the VAR methodology.

III a) Variables

The two variables of interest in this analysis are Border Patrol apprehensions and linewatch hours. As noted above, apprehensions data constitute the only directly observed components of the flow of unauthorized migration to the U.S and are widely used in empirical studies.[6] In this analysis, the

[5] In addition to the fact that the probability of apprehension is unobserved, it is unclear what the actual cost of apprehension is to an unauthorized migrant. An apprehension certainly imposes costs on migrants, but the actual quantification of these costs and how they compare with the benefits of successful entry into the U.S. is an open question. For Mexicans, in particular, it is conceivable that the costs imposed on them from temporary detention in the U.S. and eventual removal to Mexico is still well below the marginal benefits of entry into the U.S. This is an issue that deserves further study.

[6] While apprehensions refer to events and not individuals, they do appear to be a reasonable proxy for the flow of unauthorized migration. One particularly compelling reason for this is the strong seasonality present in apprehensions data. Apprehensions tend to peak in March and remain relatively high through November before they begin to decline through the winter months. The peaks and valleys of the seasonal pattern corresponds closely with the onset of the agricultural and construction seasons in the United States. The persistence of this seasonal pattern, while not conclusive, is strongly suggestive and offers intuitive support for the notion that apprehensions track the flow of unauthorized migration over time. This point has been reiterated by a number of researchers (e.g., Bean et al., 1990; Espenshade, 1990; and Hanson and Spilimbergo, 1999a, among others).

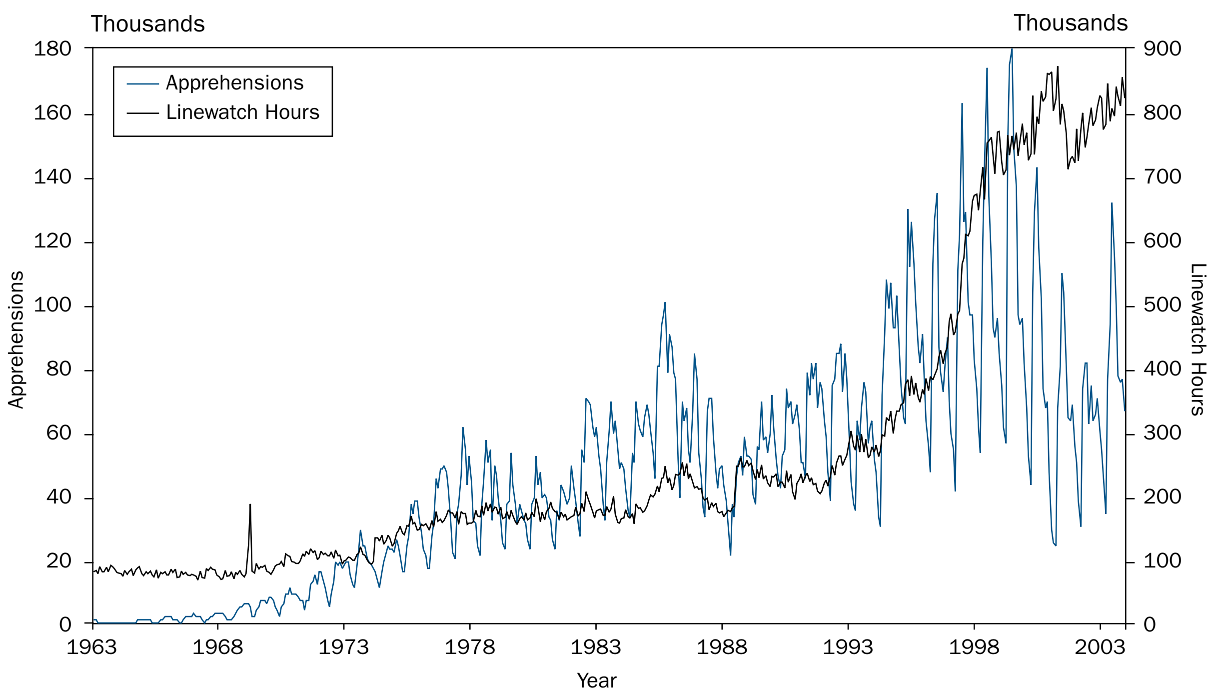
apprehensions variable constitutes the log number of monthly apprehensions along the Southwest border between July 1963 and September 2004. These data are currently collected by Customs and Border Protection (CBP), a component of the Department of Homeland Security (DHS).[7] The log number of monthly apprehensions is the measure of the unauthorized flow used in this paper. The log number of monthly linewatch hours between July 1963 and September 2004 is the enforcement variable used in this analysis. As noted above, linewatch hours constitute only one component of the total border enforcement effort, but are the only enforcement data available that is amenable to time series analysis. To control for the seasonality of the data, monthly indicator variables are also included in the analysis. Table 1 reports the descriptive statistics for the apprehensions and linewatch hours data. Figure 1 presents a time series graph of the two variables over the sample period.

Table 1.
Summary Statistics

Variable	Obs	Mean	Variance	Min	Max	J-B Statistic
Apprehensions	495	10.21	1.82	6.18	12.10	127.91*
Linewatch	495	12.22	0.53	11.13	13.68	32.59*

Note: The J-B Statistic refers to the Jarque-Bera test for normality in a data series. The null hypothesis of the test is that the series is distributed normally. Asterisks indicate the level of statistical significance of the resulting test statistic. *=p<.01. Both variables are reported in natural logs.

Figure 1.
Time Series Plot of Border Apprehensions and Linewatch Hours
July 1963 to September 2004



Before discussing methodology, it is important to clarify a few limitations of the data and the research design. The temporal frame for the apprehensions and linewatch hour data covers roughly forty years, which allows for robust examination of historical dynamics. While the length of the time

[7] Prior to 2004, the apprehensions and linewatch hours data were collected by the Immigration and Naturalization Service (INS). Linewatch hours data are not available after September 2004.

series sample is preferable for analytical purposes, the administrative nature of the data can create challenges. It is not uncommon for variables derived from administrative data to undergo definitional and standardization changes over time (Judson, 2005). Uncertainty regarding the reliability and validity of data definitions, reporting practices, and measurement consistency is an inherent limitation of historical administrative data.[8] Another potential drawback is the spatially-aggregated structure of the data. While monthly data are usually considered an acceptable level of time aggregation for most macro time series purposes, the lack of spatial disaggregation means that different sections of the Southwest border cannot be analyzed separately. Accordingly, what may be true of the entire Southwest border regarding the relationship between apprehensions and linewatch hours may not necessarily be true of specific sections of the border. The aggregate nature of the analysis means that operational and tactical changes made by CBP in certain sectors between 1963 and 2004 may not be accounted for in the analysis.[9]

A separate issue is the specification of the empirical model. Although the approach adopted here emphasizes flexibility in the model specification, the absence of covariates means that empirical findings may suffer from omitted variable bias. Despite this potential shortcoming, the exploratory nature of this study is intended only to provide an assessment of the historical dynamics of two variables that should be linked together theoretically. Even without exogenous variables, the residuals from the time series analysis of the bivariate model can be tested to ensure that they are white noise (random and not autocorrelated). A model that produces white noise residuals should help alleviate concern that the results from the bivariate model are spurious or seriously misspecified. This point notwithstanding, the results reported below in the findings section are only intended to help lay the groundwork for future work that incorporates time series dynamics and additional variables.

III b) Methodology

This section is divided into two subsections. The first subsection provides a general overview of the vector autoregression (VAR) approach to time series analysis. The second subsection discusses the preliminary diagnostic steps needed to formulate the VAR estimated in this paper.

III b1) Overview of Vector Autoregressions

The general VAR approach focuses on the underlying correlation and dynamic structure of time series data and does not assume that the correct structure of the variable relationships is known a priori (Brandt and Williams, 2007). In contrast to structural equation models in which most identifying assumptions are either applied a priori or are pre-tested, the VAR approach to modeling begins by examining the interrelated dynamics of the series. In particular, VARs emphasize the value of creating a complete dynamic specification of the series in a system of equations[10] (Brandt and Williams, 2007: 10). Accordingly, the choice of which variables are considered endogenous and exogenous, and their appropriate lag structures, are flexible in the initial model specification. Although the variables included in a model are based on theory, the correct structure of the underlying relationship is usually not assumed to be known in advance. The empirical estimation of VAR models typically involve multivariate autoregressive equations in which each variable is regressed on its past values and the past values of the other variables in the system. Model building then proceeds by testing for the appropriate lag structure in the system of equations. The resulting residuals are then regarded as exogenous shocks and are used

[8] With respect to historical linewatch hours data, the number of linewatch hours in a given time period may not reflect the impact of operational changes, such as pairing agents, that may have occurred over time in certain sectors.

[9] For example, new enforcement strategies and tactics have been implemented in certain sectors overtime, such as Operation Hold the Line in the El Paso sector in 1993 and Operation Gatekeeper in the San Diego sector in 1994. While the impact of these operations can be assessed with data at the sector level, the aggregate nature of the data used in this paper prevents an assessment of the impact of specific operations on overall apprehensions. With sector-level data, future work will examine the relationship between linewatch hours and apprehensions taking into account changes in CBP tactics.

[10] The basic idea behind VAR modeling is derived from the Wold decomposition whereby every dynamic time series can be partitioned into a set of deterministic and stochastic components. See Wold (1954) for the original proof and Hamilton (1994) for further discussion and elaboration on this topic.

to see how each equation responds to unanticipated changes in each variable. After accounting for these historical dynamics, inferences about causality (via Granger causality) and the endogenous structure of the data series can be tested (Brandt and Williams, 2007: 11).

To simplify the discussion, the reduced form bivariate VAR can be represented by the following two equations:

$$A_t = A_0 + \sum_{i=1}^{L_A} \beta_i A_{t-i} + \sum_{i=1}^{L_H} \alpha_i H_{t-i} + \sum_{i=1}^{11} \delta_i D_i + \varepsilon_i, \quad [1]$$

$$H_t = H_0 + \sum_{i=1}^{L_H} \alpha_i' H_{t-i} + \sum_{i=1}^{L'_A} \beta_i' A_{t-i} + \sum_{i=1}^{11} \delta_i' D_i + \nu_i. \quad [2]$$

In equation [1], A_t is the log number of apprehensions at month t; A_0 is an intercept; A_{t-i} is the log number of apprehensions i months ago; L_A is the maximum number of monthly lags for the apprehensions series; H_{t-i} is the log number of linewatch hours i months ago; L_H is the maximum number of lags for the linewatch hours series; D_i is an 11x1 vector of monthly dummy variables; ε_i is the vector of residuals assumed to be i.i.d. and white noise; finally, all Greek symbols are parameters. Equation [2] has the same structure as equation [1] but rearranges the key variables of interest. Specifically, H_t is the log number of linewatch hours at month t; H_0 is an intercept; H_{t-i} is the log number of linewatch hours i months ago; L'_H is the maximum number of lags for the linewatch hours series; A_{t-i} is the log number of apprehensions; L'_A is the maximum number of monthly lags for the apprehensions variable; D_i is an 11x1 vector of monthly dummy variables; and υ_i represent white noise residuals.

The two equations presented above are used to examine three aspects of the data: 1) the causal effects of endogenous variables on each other; 2) the amount of variance in one variable that can be attributed to changes in the other; and 3) the dynamic impact of changes in one variable on others in the model (Brandt and Williams, 2007: 22; Hamilton, 1994). The first issue is usually examined by an assessment of Granger causality while the second and third issues are addressed by examining the decomposition of forecast variance and impulse response analysis, respectively. These three components of the data analysis are briefly discussed in turn.

III b1a) Granger Causality

Tests of Granger causality are based on the concept of incremental forecasting value (Granger, 1969). Specifically, a variable X "Granger" causes a variable Y if Y can be better predicted from past values of X and Y together than from past values of Y alone, with other relevant information included in the prediction (Freeman, 1983; Pierce, 1977; Granger, 1969). In the present context, apprehensions will Granger cause linewatch hours if current values of linewatch hours are better predicted from past values of apprehensions and linewatch hours together than from past values of linewatch hours alone. Such tests have useful theoretical value. Theoretically, the intensity of border enforcement may be a response to the flow of unauthorized migrants. The flow of unauthorized migrants, in turn, may also respond to levels of border enforcement. Thus, there is the possibility that a dynamic feedback relationship exists between these variables over time. It is also possible that one variable Granger causes another but that this relationship is not reciprocal. Finally, it is also conceivable that neither variable predicts the other, in which case the variables may be temporally exogenous to each other. Granger causality can also assist with model construction. If one variable is helpful in statistically predicting another, then this information can be used to structure causal orderings and inform identification restrictions in systems of equations (Brandt and Williams, 2007: 22). Although interpretations of Granger causality must proceed with caution, the potential for statistically powerful and carefully constructed tests to contradict the exogeneity (or endogeneity) assumptions of established theories should be taken seriously (Freeman, 1983: 340; Geweke, 1978: 182).[11]

[11] Two alternative tests for Granger causality are those of Sims (1972) and Geweke et al., (1982). Sims' test pre-filters the data and uses lags and leads of the variables in the model while Geweks et al. also uses lags and leads of the variables but conducts the hypothesis tests only on the leads of the variable of interest. Monte Carlo studies suggest that both tests are inferior to the F-test as originally proposed by Granger (Freeman, 1983).

Although alternative testing procedures exist, the simplest, most common, and statistically powerful implementation of Granger causality involves an F-test on the block of lagged coefficients in equations [1] and [2].[12] To test whether linewatch hours Granger cause apprehensions, for example, one tests the hypothesis that the term $\sum_{i=1}^{L_H} \alpha_i H_{t-i}$ in equation [1] equals zero for a specified lag length. A rejection of the null hypothesis indicates that linewatch hours Granger cause apprehensions. In practice, this means that lagged apprehensions in conjunction with lagged enforcement is a better predictor of current apprehensions than lagged apprehensions alone. The same F-test is performed on the lagged values of apprehensions in equation [2] in order to assess whether apprehensions Granger cause linewatch hours. Both tests can be performed using OLS.

III b1b) Decomposition of Forecast Variance

How much impact the dynamic change in one variable has on the variance of a second variable is the question addressed by the decomposition of forecast variance. In VAR modeling, this technique is also known as innovation accounting because it assesses how orthogonal or independent shocks (innovations) in a variable lead to changes in another variable (Brandt and Williams, 2007: 23). Formally, the technique decomposes a variable's forecast error into a component that is attributable to the variable's own innovations and a portion that is due to changes in the second variable. The larger the percentage of forecast error attributed to the second variable, the more important that variable is in explaining or predicting the first variable. This process can also be interpreted as an examination of how much one variable accounts for the deviation from the forecasted value of a second variable. Of key interest here is what percentage of the deviation in apprehensions (linewatch hours) from their forecast is due to changes in linewatch hours (apprehensions). If, for example, linewatch hours account for a high percentage of the forecast variance in apprehensions, but not vice versa, this could be interpreted as evidence that enforcement helps predict the unauthorized flow. This result would also add credence to exogeneity and causal orderings in more structural models. Ultimately, such information is useful for testing theoretical specifications and obtaining historical dynamics.

To decompose the forecast of the VAR, the variance of the forecast errors from equations [1] and [2] must be computed. Technically, this is done by converting the VAR into a Vector Moving Average (VMA) representation in which the difference between the observed value of the vector of endogenous variables at time $t + s$ and the predicted values from the VAR are expressed as a function of the forecast errors over the current period back to period $s - 1$ (Brandt and Williams, 2007: 46). Transforming the original VAR into the VMA shows how the current forecast errors are dependent on past innovations. Once the forecast errors are estimated, however, they must be orthogonalized in order to separate a variable's own past influences from those of the other variables in the system. The most common orthgonalization procedure in this context, and the one utilized in this paper, is the Cholesky decomposition of the error covariance matrix, which is a generalization of a square root for a matrix. The Cholesky decomposition provides an accounting of the contemporaneous correlations among the innovations in the VAR and is a lower triangular matrix that can be used to analyze particular patterns of linear combinations of the residuals (Brandt and Williams, 2007: 47).[13]

---

[12] Freeman (1983) identifies four caveats that should be taken into account when interpreting the results of Granger causality tests. The first caveat involves the nature of the data used in the analysis. The release of economic data collected by Government agencies, for example, may proceed with a lag in which case analysts must decide whether decisions based on the data were made after or prior to public dissemination. A second caveat centers on the presence of a third variable that may explain both variables in the original bivariate VAR. While the existence of a single variable that would simultaneously account for both changes in apprehensions and linewatch hours at all lags and leads is unlikely, the possibility does exist. Third, the policymaking environment itself may create spurious orderings in Granger tests due to the non-experimental settings in which such tests are implemented. A final caveat relates to expectations on the part of actors in socioeconomic environments. If actors make policy decisions in anticipation of future events, then this can confound causal inference.

[13] For technical details on the Cholesky decomposition, see Hamilton (1994) or other standard time series textbooks. It should be noted, however, that the Cholesky decomposition has been criticized as a purely mechanical means of imposing identifying restrictions on the error covariance matrix of residuals from a VAR (e.g., Bernanke, 1986). The ordering of the variables used in the VAR is thus crucial when interpreting the results of residual analyses based on a Cholesky decomposition. Consequently, it is often advised that analysts investigate multiple variable orderings as a robustness check. In the present case, with only two variables in the VAR and a general absence of strong theory to assist with identification, the re-ordering of variables constitutes a normal robustness check. For the purposes of this paper, the Cholesky decomposition is a reasonable

In practice, this decomposition is computed and reported in percentage terms so that the total amount of variation in one variable due to another can be expressed in absolute terms. The shortcomings of this approach is that the interpretation of the results is not tied to the original scale of the variables and there is no measure of uncertainty with respect to the impact of the shocks to the VAR. The impulse response analysis described briefly below provides an alternative interpretation of the VAR innovations and corrects for these drawbacks.

III b1c) Impulse Response Analysis

As noted above, the residuals from an estimated VAR should be random noise (white noise), which constitutes a random shock or surprise innovation to the equations in the system. The analysis of the impulse response function (IRF) is an alternative way of using these residuals to assess how changes in one variable affect another over time. Impulse response functions are equivalent to dynamic multiplier analyses in ARIMA modeling and can be used to interpret the impacts of shocks to the vector moving average representation of the VAR. Since every stationary finite-order autoregressive time series model can be re-expressed as an infinitely lagged moving average model, the original VAR can be re-centered around its equilibrium values and the impact of shocks to the series assessed. The IRF standardizes the dynamic responses of the equations to the exogenous innovations so that each shock represents a one-standard deviation change, which also retains the scales of the variables in the fitted VAR (Brandt and Williams, 2007: 37).

The Cholesky decomposition is again used to impose identifying restrictions on the residual covariance matrix. The IRF can then be interpreted as the magnitude of a response in equation i for a shock to variable j at time $t > s$.[14] There is a corresponding impulse response for a shock to each variable in the system, which means that a system with m variables will have m^2 impulse responses (including the response of each variable to its own innovations) (Brandt and Williams, 2007: 39). In the present case, this means that the response of apprehensions to a shock in linewatch hours (and vice versa) can be traced over time. Specifically, one can investigate how apprehensions (linewatch hours) change after they have been affected by a surprise innovation in linewatch hours (apprehensions). The response of a stationary variable to a one-time shock should eventually dampen out to zero over time. Whether a response to an innovation is statistically different from zero is assessed by constructing confidence intervals around a variable's time path. In line with classical hypothesis testing, a variable's response is considered statistically insignificant if the confidence interval includes zero for a given time horizon. The confidence intervals presented in the subsequent analysis were calculated using a Monte Carlo integration technique implemented in the Regression Analysis of Time Series (RATS) statistical software and is discussed in Sims and Zha (1999).

III b2) Preliminary Testing: Unit Roots and Lag Length

Before the results of the empirical analysis are presented, it is first important to touch on two preliminary diagnostic tests that are commonly employed in VAR analyses—namely, unit root testing and lag-length specification. Both issues are discussed in turn.

Testing for non-stationarity in the form of unit roots has become a staple of time series econometrics. The reason for this stems from the challenges that non-stationary data pose for inference

---

estimation strategy. In general, however, theory should be used as much as possible to identify the error covariance matrix. With a Cholesky decomposition, this can be done by ordering the variables based as much as possible on theory (with the leading variables placed first the model) and then subsequently altering the variable ordering to check robustness. A more intuitively appealing approach is the use of structural VAR models that use more theoretically-motivated decompositions of the covariance matrix. Bernanke (1986), Sims (1986), and Blanchard and Quah (1989), for example, all propose alternative restrictions and decompositions based more on economic theory. The key problem with structural VARs is that convincing identifying restrictions can be difficult to find and defend; furthermore, the rank conditions necessary to identify structural VARs can be difficult to validate in practice (Hamilton, 1994: 333-335). Nevertheless, a more structurally minded approach to modeling apprehensions and linewatch hours is a future goal of this research.

[14] Formally, the IRF can be expressed as $\partial y_i(t + s) / \partial e_j(s) = cij(t)$, where y_i is the value of the endogenous variable at time t + s; e_j is the exogenous shock at time s; and $c_{ij}(t)$ represents the element of the ith row, jth column of the C_t matrix, which is defined as an infinitely-lagged moving average representation of the VAR depicted in equations [1] and [2] above.

using standard statistical techniques such as OLS. It is well established, for example, that OLS produces spurious results when applied to data with unit roots. What OLS is really estimating in such situations are common trends and not the underlying relationships between two or more variables. Inadequately accounting for unit roots can lead to estimates that appear to be significant and meaningful but in reality are meaningless and insignificant (Granger and Newbold, 1974; Hamilton, 1994). If data are non-stationary, transformations such as differencing are often employed to induce stationarity. While differencing a data series is common in ARIMA modeling, it is less common in VARs and is usually discouraged because such transformations discard long-run information. On the other hand, if testing reveals the likely presence of unit roots and the absence of error correction, then the standard statistical tests used for conducting Granger tests will still be valid when applied to differenced data (Brandt and Williams, 2007: 36).

To test for unit roots in the apprehensions and linewatch hours data, both series were first tested individually to determine the appropriate lag structure. This is important because tests for unit roots are often sensitive to lag-length specification (Hamilton, 1994). In accordance with standard practice, both data series were pre-tested using three information criteria to identify an appropriate lag length. The three selection criteria used were the Akaike Information Criterion (AIC), Schwartz Bayesian Information Criterion (BIC), and the Hannan-Quinn Information Criterion (HQIC). Table 2 reports

Table 2.
Lag Length Specification Tests For Unit Roots Analysis

Variable: Apprehensions				Variable: Linewatch Hours			
	Selection Criteria				Selection Criteria		
Lags	AIC	BIC	HQIC	Lags	AIC	BIC	HQIC
1.	–2.788	–2.763	–2.778	1.	–5.079	–5.053	–5.069
2.	–2.797	–2.763	–2.784	2.	–5.108	–5.074	–5.095
3.	–2.858	–2.815	–2.841	3.	–5.124	–5.082	–5.108
4.	–2.895	–2.844	–2.875	4.	–5.153	–5.102	–5.133
5.	–2.912	–2.852	–2.889	5.	–5.150	–5.090	–5.126
6.	–2.906	–2.838	–2.879	6.	–5.156	–5.087	–5.129
7.	–2.966	–2.889	–2.936	7.	–5.163	–5.086	–5.133
8.	–3.106	–3.020	–3.072	8.	–5.161	–5.075	–5.128
9.	–3.275	–3.180	–3.237	9.	–5.156	–5.061	–5.118
10.	–3.305	–3.201	–3.264	10.	–5.150	–5.046	–5.109
11.	–3.468	–3.356	–3.424	11.	–5.209	–5.097	–5.165
12.	–4.011	–3.890	–3.963	12.	–5.317	–5.196	–5.270
13.	–4.006	–3.876	–3.955	13.	–5.331	–5.201*	–5.280*
14.	–4.022	–3.883	–3.967	14.	–5.331	–5.192	–5.276
15.	–4.023	–3.875	–3.965	15.	–5.326	–5.178	–5.268
16.	–4.020	–3.863	–3.958	16.	–5.334	–5.177	–5.273
17.	–4.027	–3.861	–3.962	17.	–5.329	–5.163	–5.264
18.	–4.026	–3.851	–3.957	18.	–5.323	–5.148	–5.254
19.	–4.022	–3.838	–3.950	19.	–5.322	–5.138	–5.249
20.	–4.025	–3.831	–3.949	20.	–5.316	–5.123	–5.240
21.	–4.054	–3.852	–3.975	21.	–5.311	–5.109	–5.232
22.	–4.060	–3.849	–3.977	22.	–5.306	–5.095	–5.223
23.	–4.061	–3.840	–3.974	23.	–5.325	–5.105	–5.238
24.	–4.151*	–3.921*	–4.061*	24.	–5.337*	–5.107	–5.246

Note: AIC refers to the Akaike Information Criterion; BIC refers to the Schwarz Bayesian Information Criterion; and HQIC refers to the Hannan-Quinn Information Criterion. For all three information criteria, smaller values indicate better model fit.

* = optimal lag length for the indicated series and chosen statistic.

Table 3.

Unit Root Test Results for Variables in Levels

				ADF Critical Values			
Level Variable	**Lags**	**Trend**	**Test Statistic**	**1%**	**5%**	**10%**	**Unit Root**
Apprehensions	24	Yes	–2.0906	–3.981	–3.421	–3.133	Yes
Apprehensions	24	No	–3.5782	–3.446	–2.868	–2.570	No
Linewatch Hours . . .	13	Yes	–1.8562	–3.981	–3.421	–3.133	Yes
Linewatch Hours . . .	13	No	0.3975	–3.446	–2.868	–2.570	Yes
Linewatch Hours . . .	24	Yes	–1.7048	–3.981	–3.421	–3.133	Yes
Linewatch Hours . . .	24	No	0.4468	–3.446	–2.868	–2.570	Yes
				DF–GLS Critical Values			
				1%	**5%**	**10%**	
Apprehensions	24	Yes	0.558	–2.58	–1.95	–1.62	Yes
Apprehensions	24	No	1.598	–2.58	–1.95	–1.62	Yes
Linewatch Hours . . .	13	Yes	2.240	–2.58	–1.95	–1.62	No
Linewatch Hours . . .	13	No	2.808	–2.58	–1.95	–1.62	No
Linewatch Hours . . .	24	Yes	2.103	–2.58	–1.95	–1.62	No
Linewatch Hours . . .	24	No	2.723	–2.58	–1.95	–1.62	No

Note: The table reports results for both the Augmented Dickey-Fuller (ADF) test as well as the GLS-detrended unit root test proposed by Elliott et al. (1996). Critical values for the ADF test are based on Dickey and Fuller (1981) while those of the DF-GLS test are based on the Monte Carlo simulations presented in Elliott et al. (1996).

the results from the lag length tests. For the apprehensions series, all three information criteria indicate that 24 monthly lags is appropriate. The results for the linewatch hours series are not unanimous, as the AIC suggests a lag length of 24 while the BIC and HQIC center on 13 lags. Given these results, the subsequent unit root tests were performed for 24 and 13 month lags of linewatch hours.

Table 3 presents the results of two unit root (non-stationary) tests performed on the apprehensions and linewatch hour variables. The first is the familiar Augmented Dickey-Fuller (ADF) test estimated with and without a trend term in the augmented regression. The results for this test are displayed in the upper half of Table 3. With respect to apprehensions, the ADF test results for the presence of a unit root are mixed. When a trend term is included in the augmenting regression, the test statistic fails to reject the null of non-stationarity. By contrast, the test statistic clearly rejects the null hypothesis of a unit root at the 1% level when the trend term is excluded. This result would not be surprising if apprehensions are stationary, in which case including an erroneous trend term would bias the test statistic against the alternative hypothesis. In Figure 1 above, the apprehensions series appears to be growing in magnitude over time, which might suggest a time trend. On the other hand, the volatility of the series is high, which might suggest that mean-reversion is a reasonable approximation. Overall, the ADF test results are mixed regarding the existence of a unit root.

The upper half of Table 3 also presents the results for linewatch hours at both 13 and 24 lags. Regardless of the lag length specification and whether a trend term is included, the ADF test uniformly fails to reject the null hypothesis of a unit root at conventional levels of significance. Looking again at Figure 1 above, the linewatch series does appear to be growing steadily over time, which suggests that some form of de-trending may be necessary. The pattern depicted in Figure 1 is consistent with the fact that the border enforcement effort has intensified since the mid-1990s.

A shortcoming of the ADF test is its relatively low power against stationary alternatives to a near unit root process (Elliott et al., 1996). This weakness may explain the failure to reject the unit root null with respect to the linewatch hours series in particular. To address this issue, the potentially more powerful Dickey Fuller Generalized least Squares (DF-GLS) unit root test proposed by Elliott et al. (1996) was performed for both the apprehensions and linewatch hours series. The results from this battery of tests are reported in the bottom half of Table 3. While the test fails to reject the unit root null for the apprehensions data, the DF-GLS results uniformly reject the null hypothesis for linewatch hours at either the 1% or 5% levels depending on the lag length and type of de-trending. Considered together, the results from the ADF and DF-GLS tests paint a mixed picture with respect to the presence of unit roots in the data. If de-trending is inappropriate for the apprehensions data, then the original ADF test without a time trend may be appropriate; this result would indicate that apprehensions are stationary. Linewatch hours, by contrast, show a very clear upward trend without the large swings in volatility evident in apprehensions. This suggests that the GLS de-trending procedure implemented in the DF-GLS test is appropriate.[15]

Overall, the results regarding the existence of unit roots are mixed. This, unfortunately, is not an uncommon result when using different unit root tests and alternative lag lengths. To guard against the possibility that the empirical results presented below are spurious, all analyses are performed on the data in levels as well as first differences and any changes in results noted.

After testing for unit roots, the next step in the process involves specifying an appropriate lag length for the bivariate VAR estimated in this paper. As with the unit root tests, the results from VAR analyses can be sensitive to the choice of lag length. Lag specifications that are too short can result in spurious findings of Granger causality and can produce residuals that are not white noise

Table 4.
Lag Length Specification Tests For Bivariate VAR

	Selection Criteria		
Lags	AIC	BIC	HQIC
1.	–1061.6479	–1045.0712	–1055.1527
2.	–1121.4909	–1088.4064	–1108.5694
3.	–1171.8794	–1122.3569	–1152.6014
4.	–1223.4653	–1157.5757	–1197.9017
5.	–1245.3583	–1163.1732	–1213.5807
6.	–1272.3599	–1173.9519	–1234.4409
7.	–1275.6119	–1161.0547	–1231.6252
8.	–1326.7859	–1196.1539	–1276.8058
9.	–1394.5734	–1247.9422	–1338.6757
10.	–1452.7968	–1290.2428	–1391.0578
11.	–1492.1664	–1313.7673	–1424.6638
12.	–1546.1547	–1351.9891	–1472.9671
13.	–1779.1401	–1569.2900*	–1700.3500*
14.	–1782.0478	–1556.5893	–1697.7303
15.	–1783.5900*	–1542.6050	–1693.8275
16.	–1775.2080	–1518.7840	–1680.0880
17.	–1772.9481	–1501.1669	–1672.5524
18.	–1765.5206	–1478.4675	–1659.9345
19.	–1759.0805	–1456.8419	–1648.3904
20.	–1755.4187	–1438.0823	–1639.7123
21.	–1751.8995	–1419.5542	–1631.2656
22.	–1756.7219	–1409.4578	–1631.2508
23.	–1752.6959	–1390.6044	–1622.4789
24.	–1757.5197	–1380.6936	–1622.6495

Note: AIC refers to the Akaike Information Criterion; BIC refers to the Schwarz Bayesian Information Criterion; and HQIC refers to the Hannan–Quinn Information Criterion. For all three information criteria, smaller values indicate better model fit.

* = optimal lag length for the indicated series and chosen statistic.

[15] More sophisticated models that deal with irregular temporal patterns, increasing seasonality, and growing volatility also exist. Examples of such models, for example, are the ARCH or GARCH techniques typically used in studies of volatility. If the unauthorized flow is becoming more volatile over time, then techniques that model this behavior may be of interest. Work in this area is reserved for future projects, as this paper is largely exploratory.

(Hamilton, 1994). Conversely, lag specifications that are too long reduce the efficiency of VAR estimates. A common practice is to select the optimal lag length based on the same information criterion methodology used when testing for unit roots. Table 4 presents the results of the lag length specification tests for the bivariate VAR. According to the BIC and HQIC, a model with 13 lags is appropriate; the AIC indicates that 15 lags is optimal. As a further robustness check, both a 13 and 15 lag model were estimated with no differences in substantive results.

IV Findings

This section presents the results from the estimated VAR model represented in equations [1] and [2]. In line with the discussion above, the presentation discusses the Granger causality results, the decomposition of forecast variance, and the impulse response analysis.

IVa) Granger Causality

Table 5 presents the results from several Granger causality tests. The upper half of the table reports the results for the variables in levels while the bottom half of the table shows the results for the data in first-differences. The first column of Table 5 lists the dependent variable while the second column lists the proposed predictor variable for a given VAR. A statistically significant F test indicates that the predictor variable Granger causes the dependent variable. For each bivariate test, which corresponds to the individual rows of the table, the last column indicates whether Granger causality exists. All reported test include controls for seasonality.[16]

The most striking result that emerges from Table 5 is the absence of Granger causality for any of the bivariate specifications examined. Starting with the results for the data in levels, reported in the upper half of the table, none of the tests indicate that one variable can be considered endogenous to lagged values of another. The test results reported in the first row of the table comes closest to rejecting the null hypothesis of no Granger causality with an F statistic of 1.47 (p-value=.13). Nonetheless, the reported test statistic is still below the level considered statistically significant in standard practice. Substantively, this means that lagged values of enforcement, as measured by linewatch hours, do not aid in forecasting current apprehensions. Nor do apprehensions appear to predict current values of enforcement. For the 13-lag model with apprehensions predicting linewatch hours, the F statistic is 1.11 (p-value=.35), which is well below conventional levels of statistical significance. The pair of Granger tests for the 13-lag model in levels suggests that lagged linewatch hours are exogenous to current apprehensions and vice versa. The history of enforcement does not appear to predict future apprehensions for the time period spanning 1963 to 2004. The results suggest that feedback between apprehensions and linewatch hours does not occur and that the dynamics of the two variables are only related contemporaneously, a point discussed further below. If the historical dynamics between enforcement and the flow of unauthorized migration are only contemporaneously correlated, this suggests that the predictive power of previous enforcement efforts at the border may fade over time. Similarly, previous levels of the unauthorized flow do not appear to forecast current levels of enforcement.

---

[16] A random walk with drift is often the best representation of a trend in economic time series (Nelson and Plosser, 1982). For this reason, time trends are often left out of VAR models because the presence of a unit root in the autoregressive component of the VAR means that a separate time trend becomes an estimate of a quadratic trend; the inclusion of an intercept picks up the linear trend. With additional lags (or other variables) included in the model, the VAR tends to overfit the sample, which leads to poor forecasts (RATS User's Guide, 2007: 344). This point notwithstanding, the inclusion of a time trend is occasionally justified by the need to control for unobserved components that may impact the variables of interest. Technological innovations, for example, are often proxied by time trends in empirical analysis of economic growth. It is conceivable that increased use of technology along the Southwest border by CBP could also be modeled in this fashion. To test for this possibility, as well as the overall sensitivity of the Granger causality analysis, the VAR models were estimated with and without the trend term. The results are mildly sensitive to the inclusion of the trend term. For example, including a time trend changes the statistical significance of the Granger tests for the model using lags of linewatch hours to forecast apprehensions (p<.10). On the other hand, the coefficient on the trend term is approximately zero and is statistically insignificant (p<.30) at conventional levels. While a substantive interpretation of the time trend coefficient is difficult to make, the results do strongly suggests that the trend term is an erroneous regressor in the VAR. Accordingly, the subsequent analysis refers to the VAR models without the time trend.

Table 5.
Granger Causality Test Results

Dependent Variable	Predictor Variable	Lag Length	F Statistic	P-Value	Granger Causality
Test in Levels					
Apprehensions	Linewatch Hours	13	1.47	.13	No
Linewatch Hours . . .	Apprehensions	13	1.11	.35	No
Apprehensions	Linewatch Hours	14*	1.45	.14	No
Linewatch Hours . . .	Apprehensions	14*	1.02	.43	No
Apprehensions	Linewatch Hours	15*	1.35	.18	No
Linewatch Hours . . .	Apprehensions	15*	0.84	.62	No
Test in Differences					
Apprehensions	Linewatch Hours	13	1.34	.19	No
Linewatch Hours . . .	Apprehensions	13	0.97	.48	No
Apprehensions	Linewatch Hours	14	1.15	.32	No
Linewatch Hours . . .	Apprehensions	14	1.42	.14	No
Apprehensions	Linewatch Hours	15	1.25	.23	No
Linewatch Hours . . .	Apprehensions	15	1.32	.19	No

Note: All tests include seasonal dummy variables as exogenous regressors. The * indicates that tests in levels were performed by including the reported number of lags in the regression but retricting the F-test to the first 13 lags. This procedure is suggested when series in levels may have unit roots.

Without further investigation, the results for the levels variables reported in Table 5 are not easily reconciled with most theoretical expectations in the border enforcement literature. In particular, the finding that apprehensions are exogenous to past linewatch hours is surprising. It is certainly conceivable that the results are a product of lag misspecification or inadequately modeled unit roots. To examine this possibility, the robustness of the non-causality findings are also reported in Table 5. Tests using a lag length of 14 and 15 months were estimated and are reported in the upper half of the table (labeled with an asterisk). These tests are based on modifications to the standard F tests used to test for Granger causality and provide correct test statistics in the presence of unit roots (Dolado and Lutkepohl, 1996).[17] The results for this battery of tests confirm those of the unmodified tests in levels. Specifically, the results indicate the absence of Granger causality for the apprehensions and linewatch hours series. Again, the linewatch hours series comes closest to predicting apprehensions (p-values=.14 and .18 for the 14 and 15-lag models, respectively), but are still below conventional significance levels. As an additional check, the Granger tests were also performed on first differences of the data. The results from these tests are presented in the bottom half of Table 5. Since the data are in log differences, these tests examine whether the growth rate of apprehensions predicts the current growth rate of linewatch hours and vice versa. The results mirror those reported in the upper half of the table—namely, that neither series appears to significantly predict the other.

Given the results reported in Table 5, it is important here to emphasize that tests for Granger causality are useful only in the prediction and forecasting of variables and should not be used to draw inferences about structural parameters. After all, Granger tests can only establish whether one variable precedes another temporally and if the correlation is statistically significant. Granger tests cannot determine if a relationship is non-spurious; to determine whether a relationship is spurious requires theory.

This caveat notwithstanding, the results discussed above may be useful for theory building and testing, as the historical dynamics of apprehensions and linewatch hours could inform more structurally-minded examinations of border enforcement effectiveness. The results, for example, suggest that apprehensions and linewatch hours are contemporaneously correlated but do not have a

---

[17] The modified tests are conducted by including additional lags in the VAR (one for each unit root) and then restricting the F test to the lags minus the lag for the unit roots.

strong historical relationship. Evidence for contemporaneous correlation can be found in the variance decomposition analysis presented below, but it can also be demonstrated more directly by supplemental F tests that include current values of the variables in addition to the lags. Although these results are not reported in Table 5, the standard Granger tests that also include contemporaneous values are statistically significant for both apprehensions and linewatch hours. Formally, the results indicate that linewatch hours are contemporaneously related to apprehensions with an F statistic of 2.6 (p-value=.001) and that apprehensions correlate with linewatch hours with an F statistic of 2.3 (p-value=.005). These findings are consistent with previous research that finds a strong relationship between enforcement and apprehensions in the same time period. Substantively, this supports the idea that linewatch hours have an immediate impact on apprehensions in the same time period. On the other hand, when only lags are included in the F tests, the estimates are never statistically significant at conventional levels. Given this paper's focus on dynamics, this result means that the history of border enforcement may be of limited use when trying to predict future apprehensions and vice versa.

It is not entirely clear what implications these results have for the model specifications used in the existing literature. At first glance, the finding of non-causality indicates a much weaker relationship between the two variables than is usually supposed. On the other hand, if apprehensions and linewatch hours are most strongly related to each other contemporaneously, then this suggests that both variables cannot be considered exogenous to each other in non-dynamic models, which reinforces the appropriateness of IV estimation. Nevertheless, the failure to find evidence of either strict endogeneity (i.e., one variable predicting another but not vice versa) or of feedback (both variables predicting each other) is puzzling and can provide a basis for further theorizing about the historical dynamics of enforcement and the unauthorized flow.

IVb) Decomposition of Forecast Variance

As noted above, the decomposition of forecast variance is used to examine how much the fitted VAR deviates from the actual values of the vector of endogenous variables. What percentage of a variable's deviation from its forecasted value is attributable to another variable provides additional insight into historical relationships. Specifically, the decomposition procedure is often used in conjunction with Granger causality tests to assess endogeneity relationships and predictive power. If apprehensions and linewatch hours are exogenous to each other, as indicated by the Granger results, then the empirical expectation is that innovations in one variable do not explain the forecast variance in the other. Evidence for contemporaneous correlation exists when one variable begins to explain the forecast variance in the other with a time lag. This occurs because the correlation takes time to work through the lags in the system.

Table 6 reports the results of the forecast error decomposition for the VAR model of apprehensions and linewatch hours. When interpreting the results, the ordering of the variables is important because the decomposition assumes that all the variance in the initial period is due entirely to the first variable in the ordering. As the forecast horizon expands, the other variables in the system begin to exert their influence. In Table 6, for example, the apprehension variable is first in the ordering, which means that it explains all of its forecast variance in the initial period. It is possible that results may be sensitive to potentially arbitrary variable orderings. Without strong theoretical guidance, a common recommendation is to switch the variable orderings in order to check for robustness. Results from the re-ordered model are also discussed below.

The first column in Table 6 lists the steps in the forecast with each step corresponding to one month. Thus, the first step represents the first month of the forecast while the thirty-sixth step represents the thirty-sixth month. The total forecast horizon covers three years. The next two columns in the table report the percentage of forecast variance in the apprehensions series explained by apprehensions and linewatch hours, respectively. Because the VAR accounts for all forecast variance, each row sums to 100%. As noted, the apprehensions series accounts for all of its own forecast variance in the first month while the linewatch hours series accounts for none. The last two columns in the table

Table 6.
Decomposition of Forecast Variance for VAR(13) Model

	Forecast Error % for Apprehensions Innovations in		*Forecast Error % for Linewatch Hour Innovations In*	
Step	**Apprehensions**	**Linewatch Hours**	**Apprehensions**	**Linewatch Hours**
1.	100.000	0.000	3.674	96.326
2.	99.999	0.001	4.313	95.687
3.	99.991	0.009	3.800	96.200
4.	99.950	0.050	3.995	96.005
5.	99.867	0.133	4.744	95.256
6.	99.417	0.583	5.078	94.922
7.	98.730	1.270	5.794	94.206
8.	97.966	2.034	6.013	93.987
9.	96.961	3.039	6.646	93.354
10.	95.859	4.141	6.636	93.364
11.	95.489	4.511	6.554	93.446
12.	95.272	4.728	6.602	93.398
13.	95.728	4.272	6.887	93.113
14.	96.137	3.863	7.082	92.918
15.	96.417	3.583	7.044	92.956
16.	96.595	3.405	7.119	92.881
17.	96.669	3.331	7.294	92.706
18.	96.625	3.375	7.472	92.528
19.	96.500	3.500	7.699	92.301
20.	96.314	3.686	7.834	92.166
21.	96.045	3.955	8.014	91.986
22.	95.813	4.187	8.080	91.920
23.	95.729	4.271	8.113	91.887
24.	95.734	4.266	8.181	91.819
25.	95.888	4.112	8.300	91.700
26.	96.050	3.950	8.408	91.592
27.	96.178	3.822	8.458	91.542
28.	96.270	3.730	8.533	91.467
29.	96.318	3.682	8.636	91.364
30.	96.328	3.672	8.755	91.245
31.	96.304	3.696	8.886	91.114
32.	96.254	3.746	8.990	91.114
33.	96.182	3.818	9.093	90.907
34.	96.126	3.874	9.169	90.831
35.	96.117	3.883	9.235	90.765
36.	96.143	3.857	9.309	90.691

Note: The table reports forecast variance decompositions for a 36-month period.

report the forecast error in linewatch hours due to innovations in apprehensions and linewatch hours themselves. Thus, in the first period, linewatch hours account for 96% of their own forecast variance while apprehensions account for roughly 4%. Focusing on the first set of columns, it is apparent that linewatch hours do not explain much of the forecast variance in apprehensions. After one year (step 12), for example, linewatch hours only account for roughly 5% of the variance in apprehensions. After two years (step 24), this percentage drops to approximately 4%. Despite slight deviations, the overall pattern appears fairly stable for most of the forecast horizon after one year. By the thirty-sixth month, only about 4% of the apprehensions series is explained by innovations in linewatch hours.

At first glance, apprehensions do appear to explain more variance in the linewatch hours series, as shown in the last two columns of Table 6. After one year, for example, 6.6% of the forecast variance in linewatch hours is accounted for by changes in the apprehensions series. At the end of two years (step 24), this number increases to about 8% before rising to a little over 9% by the end of the forecast horizon (step 36). While the explained forecast variance changes fairly steadily, it appears that apprehensions may have slightly more influence on linewatch hours than linewatch hours have on apprehensions. The problem is that this conclusion may be sensitive to the order in which the variables entered the equation. The results reported in Table 6 assume that changes in apprehensions occur first. This assumption is similar to the idea that linewatch hours may be a reaction to the flow of unauthorized migration. Again, given the bivariate structure of the VAR, and without strong theoretical guidance, it can be problematic to base inference on this particular variable ordering. After all, it also seems reasonable to posit that apprehensions may react to unanticipated changes in enforcement intensity, which would place linewatch hours at the front of the variable ordering.

Table 7 reports the decomposition results when linewatch hours are placed first in the system of equations. Overall, the results suggest that the order of the variables makes a difference. In Table 7, the variance in the linewatch hours forecast is decomposed first and is reported in the first pair of columns. What is immediately noticeable is how little variance is accounted for by apprehensions. After one year, for example, less than 1% of the forecast variance in linewatch hours is due to changes in apprehensions. After two years, only 1.2% of the forecast variance can be attributed to apprehensions. By the end of the third year, only 1.6% of the forecast variance in linewatch hours can be attributed to apprehensions. The results reported in the last pair of columns indicate that linewatch hours account for about 11% of the variance in apprehensions after one year. This percentage remains fairly steady at step 24 before declining slightly to about 10.8% by the end of the sample period.

Comparing the results from Tables 6 and 7 reveals that the variance decomposition analysis is sensitive to the order of the variables. What also emerges from the analysis, however, is that unexpected changes in one variable do not appear to have a particularly large impact on the innovations in the other variable. If apprehensions are entered first in the equation, no more than 4.7% of its variance is explained by linewatch hours, and this occurs after one year. In addition, less than 10% of the forecast error in linewatch hours is explained by apprehensions. Similarly, when linewatch hours appear first in the ordering, no more than 1.6% of its forecast variance is attributable to apprehensions after three years. Table 7 also shows that linewatch hours never account for more than about 11% of the variance in apprehensions. These results indicate that the magnitudes of the variance decompositions are relatively small and fairly stable over time. Altogether, the results from this section seem consistent with the idea that apprehensions and linewatch hours are contemporaneously correlated, as the impacts of one variable on the other are small and operate with a lag. These findings also gel with the gist of the Granger results discussed earlier.

IVc) Impulse Response Analysis

As mentioned above, the impulse response function (IRF) enables one to analyze the response of one variable to a random shock in another variable while maintaining the original units of the data as well as providing an estimate of uncertainty. The results presented here are based on a Cholesky decomposition of the estimated residual covariance matrix of the estimated VAR. Substantively, the IRF is useful because it can provide a more statistically principled means of measuring a variable's response to changes in another. In the present context, the IRF can help determine how quickly apprehensions (linewatch hours) adjust after being shocked by an unanticipated change in linewatch hours (apprehensions). Such tests can provide support for substantive hypothesis tests with respect to variable dynamics over time. If, for example, theory suggests that a change in enforcement intensity should reduce apprehensions over time, this expectation can be tested using the IRF. It is important to remember, however, that the results presented here are purely exploratory and are intended to assist with theoretical development by giving an account of the dynamic behavior of apprehensions and linewatch hours.

Table 7.
Decomposition of Forecast Variance for VAR(13) Model

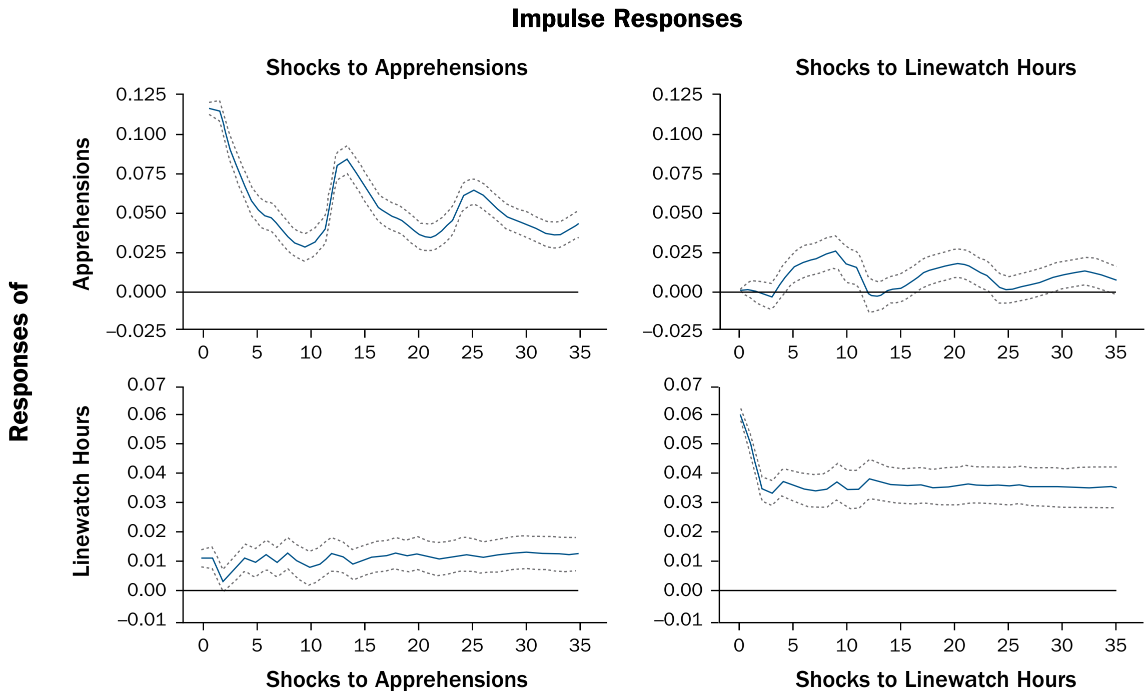
Step	*Forecast Error % for Linewatch Hours Innovations in*		*Forecast Error % for Apprehensions Innovations In*	
	Linewatch Hours	**Apprehensions**	**Linewatch Hours**	**Apprehensions**
1.	100.000	0.000	3.674	96.326
2.	99.941	0.059	3.774	96.226
3.	99.830	0.170	3.593	96.407
4.	99.833	0.167	3.363	96.637
5.	99.669	0.331	3.765	96.235
6.	99.621	0.379	4.760	95.240
7.	99.396	0.604	6.034	93.966
8.	99.364	0.636	7.263	92.737
9.	99.138	0.862	8.658	91.342
10.	99.175	0.825	10.107	89.893
11.	99.219	0.781	10.727	89.273
12.	99.228	0.772	11.208	88.792
13.	99.149	0.851	10.312	89.688
14.	99.100	0.900	9.494	90.506
15.	99.129	0.871	9.088	90.912
16.	99.120	0.880	8.861	91.139
17.	99.068	0.932	8.859	91.141
18.	99.014	0.986	9.071	90.929
19.	98.932	1.068	9.390	90.610
20.	98.892	1.108	9.767	90.233
21.	98.829	1.171	10.218	89.782
22.	98.815	1.185	10.613	89.387
23.	98.815	1.185	10.820	89.180
24.	98.798	1.202	10.906	89.328
25.	98.758	1.242	10.672	89.328
26.	98.723	1.277	10.396	89.604
27.	98.711	1.289	10.216	89.784
28.	98.688	1.312	10.109	89.891
29.	98.652	1.348	10.091	89.909
30.	98.607	1.393	10.145	89.855
31.	98.555	1.445	10.251	89.749
32.	98.515	1.485	10.388	89.612
33.	98.476	1.524	10.549	89.451
34.	98.449	1.551	10.683	89.317
35.	98.426	1.574	10.748	89.252
36.	98.399	1.601	10.753	89.247

Note: The table reports forecast variance decompositions for a 36-month period.

Figure 2 presents the impulse responses when the apprehensions series is included as the first variable in the ordering of the contemporaneous covariance matrix of residuals. When interpreting the graph, this means that the initial shock to linewatch hours (presented in the upper right corner) is zero. The upper left corner of Figure 2 shows how apprehensions respond to its own shock, which is defined as one standard deviation of the apprehensions equation residuals. The time horizon for the impulse response analysis is thirty-six months and is recorded on the x-axis of each individual graph. The dotted lines shown in Figure 2 represent 68% confidence intervals, which, according to Sims and Zha (1999), give a more accurate summary of the central tendency of the response. The upper left corner of Figure 2 shows how the apprehensions series adjusts to its own innovations. Formally,

Figure 2.

Impulse Response Analysis of VAR(13) Model in Levels



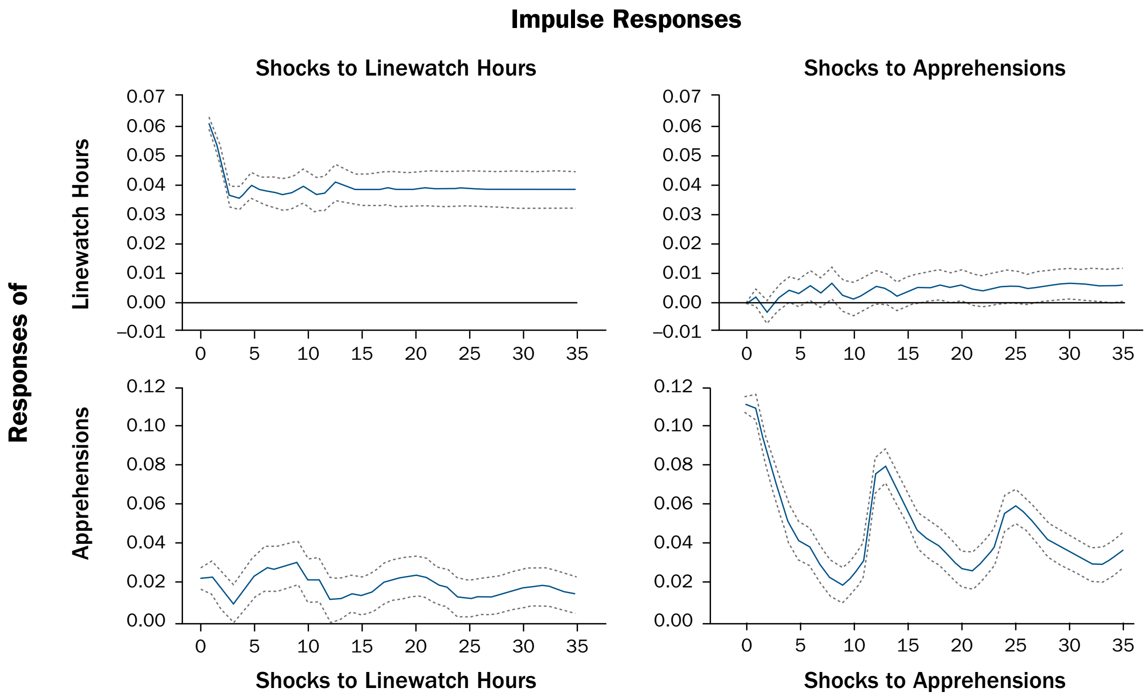
a 0.125 unit shock (one standard deviation) leads to a fairly rapid decay up until about ten months before the series begins to rise again and then dampen. For a longer forecast horizon, the series should eventually return to zero. The undulating pattern for apprehensions depicted in the upper left corner of Figure 2 may indicate un-modeled seasonality as well as the possibility of a unit root. Despite the inclusion of seasonal indicator variables in the VAR, the volatility of the seasonal pattern may have increased over the sample period. This may suggest that more sophisticated seasonal adjustments are needed. The possibility of a unit root is addressed by examining the IRF in first differences and is briefly discussed later in the section.

The upper right corner of Figure 2 shows how apprehensions respond to a shock in linewatch hours. The response appears to show some of the same seasonal variation noted previously but is not statistically different from zero after roughly twenty-four months. Furthermore, shocks to linewatch hours may have a lagged effect on apprehensions, as the confidence interval for the response trajectory is not different from zero until after about four months. After four months, however, apprehensions begin to rise before returning to zero after about a year. The pattern of the impulse response may be consistent with a delayed enforcement effect after an innovation occurs in the apprehensions series. The response series also exhibits some the same sine wave-like patterns evident in the previous graph. Again, this could point to more complex seasonality, unit roots, or both. The bottom left corner of Figure 2 shows the response of linewatch hours to an innovation in apprehensions. Specifically, a 0.125 unit shock to apprehensions appears to lead to a sustained 0.01 increase in linewatch hours. This response might indicate a small but persistent increase in enforcement after an unanticipated increase in apprehensions. Once again, the fact that the response is persistent and does not dampen out to zero over the forecast horizon may indicate the presence of a unit root.

To check the sensitivity of the VAR results, the IRF was also estimated for the alternative model placing linewatch hours first in the variable ordering. Figure 3 reports the impulse responses for the alternative model. The graph in the top right corner of the figure depicts the response of linewatch

Figure 3.

Impulse Response Analysis of VAR(13) Model in Levels



hours to a shock in apprehensions. Because the lower confidence interval always encompasses zero, it appears that enforcement does not significantly respond to apprehensions. This finding contrasts slightly with the one suggested by Figure 2 in which enforcement had a small but statistically significant response to apprehensions. Although the general relationship is similar in the two figures—namely, a weakly positive response of enforcement to apprehension shocks—the statistical significance is not robust. The graph in the bottom left corner of Figure 3 maps the response of apprehensions to innovations in linewatch hours. Here the general findings are similar to those reported in the upper right corner of Figure 2. Specifically, the results indicate that the response of apprehensions follows an undulating pattern with slight declines followed by slight increases. The confidence interval never encompasses zero, which indicates that the response, while small in magnitude (about 0.02 units for a 0.06 unit shock in linewatch hours) is statistically significant. Overall, the basic contours of the impulse responses summarized in Figures 2 and 3 are similar, though the statistical significance appears to be sensitive to variable orderings. The evidence presented in the IRFs is broadly consistent with the notion that apprehensions and linewatch hours are contemporaneously correlated. In this sense, the general findings seem broadly in harmony with those presented in the previous two subsections.

As an additional robustness check, the impulse responses for the VAR model in first differences are also presented. The differenced data are presented in order to guard against possible unit roots and because some empirical models [notably, Orrenius and Zavodny (2003)] estimate structural models in first differences. Figure 4 reports the impulse responses when apprehensions are affected by a shock to the growth rate in linewatch hours. The response pattern suggests that the changes in enforcement may have a delayed impact on changes in apprehensions because the response does not rise above zero until about four months. The response quickly dies off, however, and is not significantly different from zero for most of the forecast horizon. Some of the same movements in the undulating pattern seen in the previous graphs are still evident, though the magnitude of the fluctuations are smaller. This suggests that while unit roots may not be present, seasonality may still be an issue.

Figure 4.

Impulse Response Analysis of VAR(13) Model in First Differences

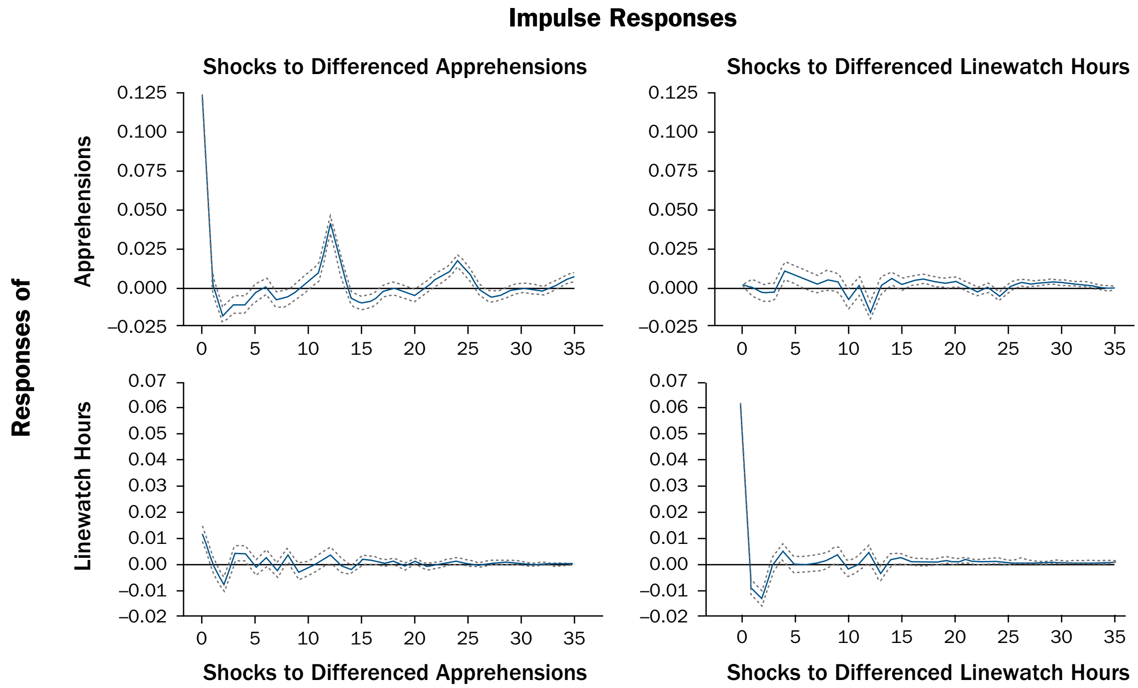
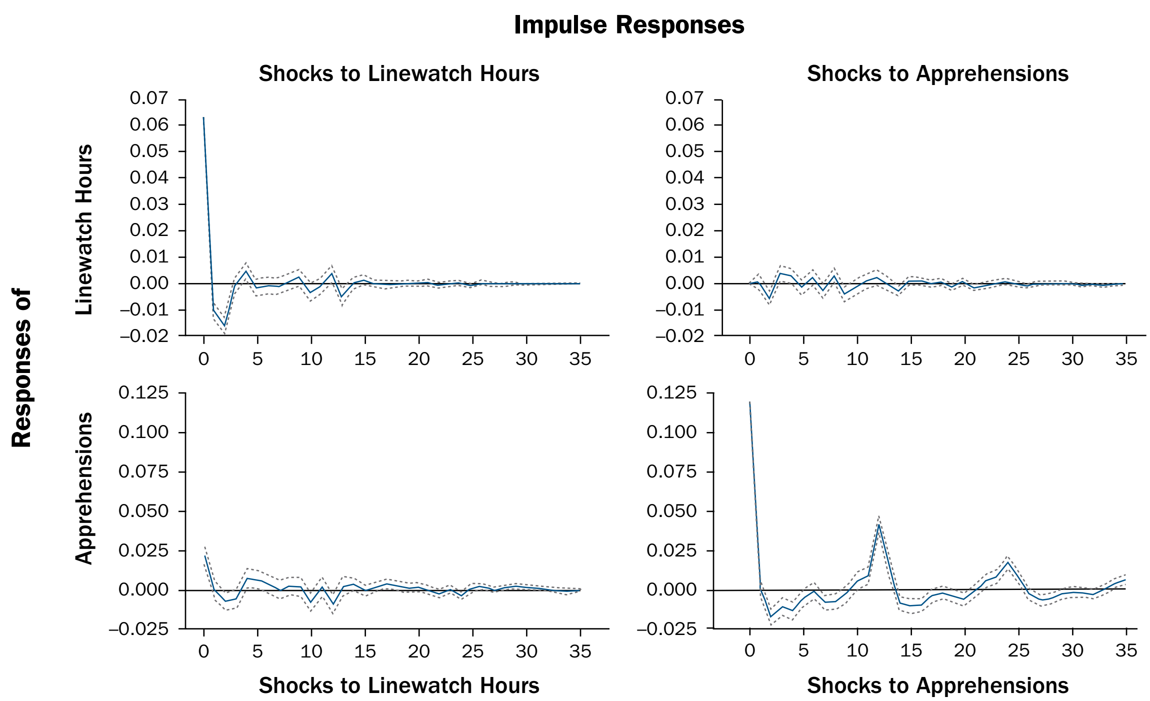


Figure 5.

Impulse Response Analysis of VAR(13) Model in First Differences



The bottom left corner of Figure 4 depicts the reaction of linewatch hours to an apprehensions innovation. The graph shows that a 0.125 unit shock in the growth rate of apprehensions leads to a rapid decay in the growth rate of linewatch hours. Figure 5 illustrates the impulse responses when the variable ordering is reversed. Focusing on the upper right and lower left graphs, the response patterns look broadly similar to the graphs in Figure 4. In particular, a shock to the growth rate in apprehensions leads to a slightly delayed negative response in linewatch hours in the first few months. On the other hand, the response is not significantly different from zero for most of the forecast horizon. Similarly, a shock to the growth rate of linewatch hours leads to a rapid decline in the growth rate of apprehensions before the response dampens to zero. Ultimately, the similarity of the response functions in Figures 4 and 5 suggests that the historical relationship between apprehensions and linewatch hours in relatively insensitive to variable orderings.

IVd) Residual Analysis

The final tests presented here involve testing the residuals for model misspecification. A critical requirement of VAR analysis is that the estimated residuals can be characterized as white noise. Failure to obtain white noise residuals suggests that relevant information, in the form of un-modeled lags, is missing from the VAR results. To ensure that the VARs examined in this paper are adequately specified, the autocorrelation function (ACF) of the residuals from equations [1] and [2] were graphed and tested for white noise via the standard Q-statistic. Figure 6 presents the ACF for the residuals from equation [1] with apprehensions as the dependent variable. The Q-statistic of 33.55 (p-value=.30) fails to reject the null hypothesis of white noise at conventional levels of statistical significance. Similarly, Figure 7, which graphs the residuals and reports the Q-statistic for the model with linewatch hours as the dependent variable, also suggests that misspecification is not a problem. Specifically, the figure reports a Q-statistic of 25.48 (p-value=.71) that is well below statistical significance at conventional levels. The lag specification used in this paper, therefore, appears to be adequate.

Figure 6.
Residuals and Q-Statistic for Apprehensions as a Dependent Variable

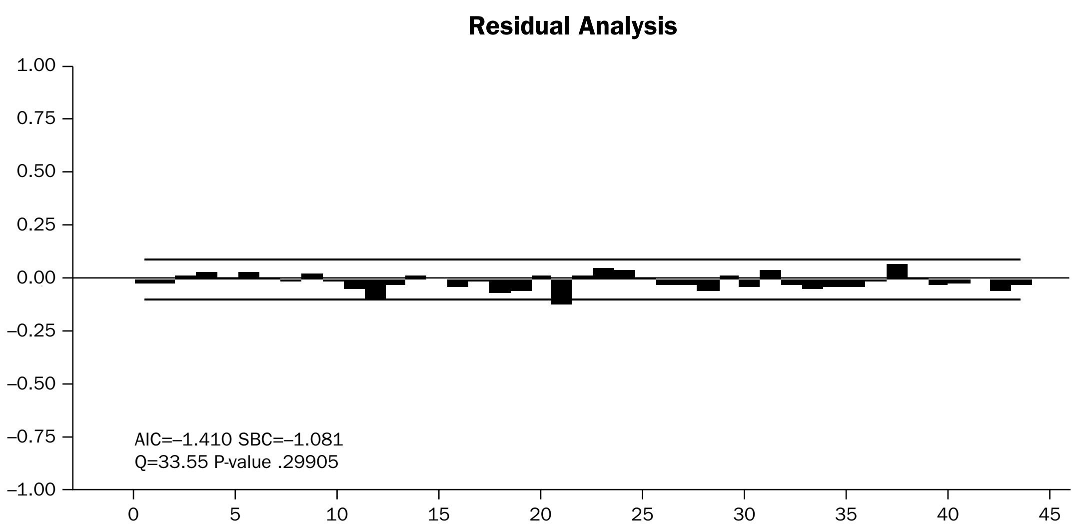
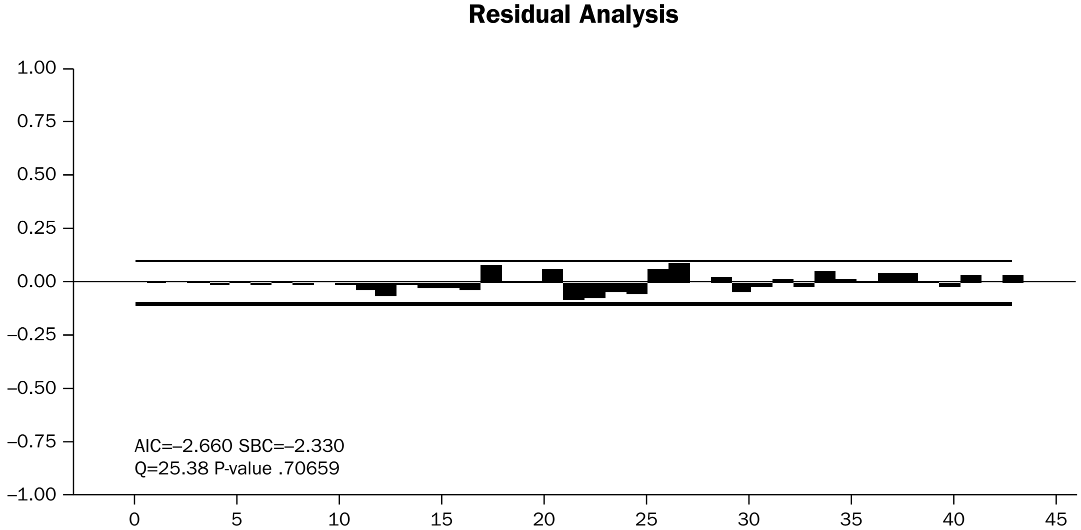


Figure 7.
Residuals and Q-Statistic for Linewatch Hours as Dependent Variable



V Conclusion

Most models of border enforcement and unauthorized migration estimate an enforcement elasticity by regressing border apprehensions on CBP linewatch hours data. In early versions of these models, identifying restrictions pertaining to the exogeneity of apprehensions and enforcement were imposed based on an implicit structural understanding the underlying relationship (Bean et al., 1990). Generally, this meant that enforcement was regarded as exogenous to apprehensions. In more recent studies, the exogeneity restriction was relaxed and the impact of linewatch hours on apprehensions assessed via instrumental variables estimation (Hanson and Spilimbergo, 1999a; Orrenius and Zavodny, 2003). What both sets of models lacked, however, was a formal test of the exogeneity assumptions as well as an accounting of the historical dynamics in the data.

The results presented in this paper provide a first step towards analyzing the historical relationship between apprehensions and linewatch hours. Specifically, this analysis helps answer questions that should be of interest when assessing models of border enforcement effectiveness as they relate to time. First and foremost, forecasting apprehensions based on past linewatch hours, and vice versa, appears to be more problematic than one might expect. The most striking result in this context is the absence of Granger causality for any of the empirical specifications examined. Substantively, this means that the history of linewatch hours does not help forecast current numbers of apprehensions, or vice versa. In fact, the weight of the evidence from the Granger tests, variance decompositions, and, to a lesser extent, the impulse response analysis, suggests that the two series are only contemporaneously correlated. From a practical standpoint, this suggests that apprehensions and linewatch hours are strongly related in the same time period, but that the relationship weakens when past values of these variables are used for forecasting purposes. This finding may provide useful information for structural modeling efforts. In particular, it strongly suggests that IV estimation is appropriate when apprehensions are expressed as a function of linewatch hours in the same time period. Past values of either variable, however, are considerably weaker when used for prediction purposes. This underscores the need to think about the dynamics in the data seriously. It is important to ask, for example, why two series that are so often linked together in discussions of enforcement effectiveness are rather weak predictors of each other historically.

Despite this key finding, several caveats should be noted. First, the results here cannot be used to assess enforcement effectiveness. As noted above, one should not interpret the lack of a strong forecasting relationship as evidence against an enforcement effect. After all, in structural models, linewatch hours are usually specified to have a contemporaneous impact on apprehensions. The time series results in this paper suggest that the correlation between apprehensions and linewatch hours is statistically strong in the same time period, which is consistent with existing studies that find a contemporaneous enforcement effect. Additionally, the possibility of unit roots (non-stationarity) in both data series means that the results from the impulse response analyses should be viewed skeptically. Thus, even though the impulse analysis revealed a low magnitude of response between apprehensions and linewatch hours over time, this does not rule out an enforcement effect. Instead, the results highlight the importance of testing and accounting for unit roots in the apprehensions and linewatch hours data in order to guard against possibly spurious relationships and inferences. While this analysis has tried to take the unit root issue seriously by considering alternative estimation strategies—i.e., by using differenced data and modified F tests for Granger causality—much more sophisticated and rigorous approaches, such as cointegration analysis, are available and should be examined.

A second major issue involves the absence of covariates. There is no assurance that the results of bivariate tests will hold in the multivariate context. It is possible that the results of this analysis will not be robust when additional variables are added to the system of equations. Theoretically, a third variable may predict both apprehensions and linewatch hours in which case the bivariate Granger test results may be overturned by a supplemental test that incorporates additional variables (i.e., a block exogeneity test). The incorporation of additional variables, however, requires more intense theorizing and clear specification of dynamic relationships. A potentially useful candidate theory is that proposed by Hanson and Spilimbergo (1999b), which sees the supply of and demand for border enforcement as a function of economic shocks to certain sectors. If both apprehensions and linewatch hours can be expressed as responses to variations in economic conditions, then this may also help explain the absence of Granger causality reported in this paper. These points notwithstanding, at the very least, the results presented in this paper raise several issues that should serve as motivation for future work on the dynamics of border enforcement and the flow of unauthorized migration.

Works Cited

Amuedo-Dorantes, C. and Bansak, C. 2007. The Effectiveness of Border Enforcement in Deterring and Postponing Repetitive Illegal Crossings Attempts. Draft Manuscript.

Angrist, J.D. and Krueger, A.B. 2001. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *The Journal of Economic Perspectives*. 15(4):69-85.

Bean F.D., Espenshade, T.J., White, M.J., Dymowski, R.F. 1990. "Post-IRCA Changes in the Volume and Composition of Undocumented Migration to the United States: An Assessment Based on Apprehensions Data." Pp. 111-158 in *Undocumented Migration to the United States*, edited by F.D. Bean, B. Edmonston, and J.S. Passel. Santa Monica, CA: RAND.

Bernanke, B. 1986. Alternative explanations of the money-income correlation. In *Carnegie-Rochester conference series on public policy*. Amsterdam: North-Holland.

Blanchard, O. and Quah, D. 1989. The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*. 79:655-673.

Brandt, P.T. and Williams, J.T. 2007. *Multiple Time Series Models*. Sage Publications: Thousand Oaks, CA.

Dolado, J.J. and Lutkepohl, H. 1996. Making Wald tests work for cointegrated VAR Systems. Econometric Reviews. 15:369-386.

Donato, K.M., Durand, J., and Massey, D.S. 1992. Mexican Migration to the United States: A Critical Review. *Latin American Research Review*. 27(2):3-42.

Elliott, G., Rothenberg, T.J., and Stock, J.H. 1996. Efficient Tests for an Autoregressive Unit Root. *Econometrica*. 64(4):813-836.

Espenshade, T.J. 1994. Does the Threat of Border Apprehension Deter Undocumented U.S. Migration? *Population and Development Review*. 20:871-892.

Espenshade, T.J. 1990. Undocumented Migration to the United States: Evidence from a Repeated Trials Model. In Bean, F.D., Edmonston, B., and Passel, J.S. [eds] *Undocumented Migration to the United States IRCA and the Experience of the 1980s*. Urban Institute Press: Washington, D.C.

Freeman, J.R. 1983. Granger causality and the time series analysis of political relationships. American Journal of Political Science. 27:327-358.

Geweke, J. 1978. Testing the exogeneity specification in the complete dynamic simultaneous equation model. *Journal of Econometrics*. 7:163-185.

Geweke, J., Meese, R. and Dent, W. (1982). Comparing alternatives tests of causality in temporal system," *Journal of Econometrics*. 21:161-194.

Granger, C.W.J. and Newbold, P. 1974. Spurious regressions in econometrics. *Journal of Econometrics*. 2:111-120.

Granger, C.W.J. 1969. Investigating causal relations by econometric models and crosspectral methods, *Econometrica*. 37:424-438.

Hamilton, J.D. 1994. *Time Series Analysis*. Princeton University Press: Princeton, NJ.

Hanson, G. and Spilimbergo, A. 1999a. Illegal Immigration, Border Enforcement, and Relative Wages: Evidence from Apprehensions at the U.S.-Mexico Border. *The American Economic Review*. 89:1337-1357.

Hanson, G. and Spilimbergo, A. 1999b. Political Economy, Sectoral Shocks, and Border Enforcement. NBER Working Paper 7315.

Judge, G.C., Hill, R.C., Griffiths, W.E., Lutkepohl, H., Lee, T. 1988. *Introduction to the Theory and Practice of Econometrics*. John Wiley and Sons: New York, NY.

Judson, D.H. 2005. Computerized Record Linkage and Statistical Matching. In Kempf-Leonard, K. [ed] *Encyclopedia of Social Measurement*. ElSevier Academic Press. Oxford Press: Oxford, UK.

Kennedy, P. 1998. *A Guide to Econometrics*. MIT Press: Massachusetts, MA.

Nelson, C.R. and Plosser, C.I. 1982. Trends and Random Walks in Macroeconomic Time Series. *Journal of Monetary Economics*, 10:139-162.

Orrenius, P. and Zavodny, M. 2003. Do Amnesty Programs Reduce Undocumented Migration? Evidence from IRCA. *Demography*. 40:437-450.

Pierce, D.A. 1977. Relationships—and lack thereof between economic time series, with special reference to money and interest rates. *Journal of the American Statistical Association*. 72:11-22.

Sims, C.A. and Zha, T. 1999. Error bands for impulse responses. *Econometrica*. 67: 1113-1156.

Sims, C.A. 1986. Specification, estimation, and analysis of macroeconomic models. *Journal of Money, Credit and Banking*. 18(1):121-126.

Sims, C.A. 1972. Money, income, and causality. *American Economic Review*. 62:540-552.

Sjaastad, L.A. 1962. The Costs and Returns of Human Migration. *The Journal of Political Economy*. 70:80-93.

Todaro, M.P. and Maruszko, L. 1987. Illegal Migration and U.S. Immigration Reform: A Conceptual Framework. *Population and Development Review*. 13(1):101-114.

Wold, H. 1954. A Study in the Analysis of Stationary Time Series (2nd Edition). Uppsala, Sweden.