Identification through Heteroskedasticity.[1]

Roberto Rigobon

Sloan School of Management, MIT and NBER

First version: December, 1999.

This version: April, 2000.

[1]I thank Olivier Blanchard, Frank Fisher, Robert Hall, Vladimir Kliouev, Guido Kuersteiner, Andrew Lo, Whitney Newey, Brian Sack, Ken Rogoff, Enrique Sentana, Min Shi, Tom Stoker, and Mark Watson for helpful comments and suggestions. I thank the seminar participants at the MIT Macro seminar, the MIT Finance seminar, the Harvard International seminar, the BYU seminar, the International Seminar at Rochester, and the Econometrics lunch at MIT. An earlier version of the results of the paper were also presented at the International Seminar at Princeton and at the Finance Seminar at Ohio University. I thank them for all their suggestions. All remain errors are mine. *Comments are welcomed* to 50 Memorial Drive, Room E52-447, Cambridge, MA 02142. email: *rigobon@mit.edu.* home page: *http://web.mit.edu/rigobon/www/.*

Abstract

In this paper, I develop a new identification method to solve the problem of simultaneous equations, based on heteroskedasticity of the structural shocks. I show that if the hetereoskedasticity can be described as a two-regime process, then the system is just identified under relatively general conditions. Identification is also discussed under more than two regimes, when the residuals exhibit ARCH behavior, and when there are aggregate shocks.

Additionally, an important result in the paper is that if the data exhibits heteroskedasticity, but it has been misspecified, the estimates are still consistent. The main condition for this to be the case is that the variance covariance matrices still have to satisfy the rank condition.

This methodology is applied to measure contagion across sovereign bonds between Argentina and Mexico. The estimates of the simultaneous parameters are stable to different definitions of the regimes.

JEL Classification Numbers: C30, F32, C10

Keywords: Identification Restrictions, Simultaneous Equations, Contagion, Bond Market.

Roberto Rigobon
Sloan School of Management, MIT
E52-447
Cambridge, MA 02142-1347
and NBER

1 Introduction

The question of identification when the model includes endogenous variables has been studied for several decades now.[1] The problem can be oversimplified to one in which the parameters of interest (unknowns) belong to a system of simultaneous equations (structural form). The parameters must be recovered from a linear transformation of the system known as the "reduced form". The number of equations obtained from the reduced form, however, is smaller than the number of unknown parameters; and therefore, the original parameters cannot be recovered without additional information.

The literature has solved the problem by offering supplementary constraints. Some of these assumptions are relatively uncontroversial especially in macroeconomic applications, such as normalization and independence of the structural shocks. Normalization refers to a change in the units in which the shocks are measured, while independent shocks are widely utilized in macroeconomic applications when either a recursive system is used, or there is a justifyable economic interpretation of the shocks.[2] These two assumptions, however, are usually not enough to fully solve the problem. Therefore, the literature has developed further restrictions, usually parameter constraints: exclusion, sign, long run, and relative variance restrictions. These assumptions had proven to be very useful in numerous applied problems. On the other hand, there are several circumstances in which they are difficult to justify.

In this paper, I discuss a new procedure to solve the identification problem of endogenous variables. This restriction is based on heteroskedasticity of the structural shocks, and the conditions in which it can be used are relatively general. Moreover, as I discuss later, this identifying restriction, in contrast with the traditional assumptions, is indeed testable.

I apply the methodology to measure the propagation, or "contagion", between Argentinean and Mexican Brady Bonds. This is a case in which none of the standard assumptions can be defended, thus leaving the problem of identification unsolved using the traditional techniques. Nonetheless, using the fact that Brady Bonds exhibit important conditional heteroskedasticity, it is possible to

[1] See Fisher [1976] for the most comprehensive treatment of the subject. See Haavelmo [1947] and Koopmans, et.al. [1950] for the seminal contributions.

[2] In general, this constraint is imposed because the definition, or the economic interpretation, of the shocks in the structural equations implies their independence: for example, nominal versus real shocks, permanent versus transitory shocks, or supply versus demand shocks.

estimate the contemporaneous relationship.

The paper is organized as follows: In section 2, the typical problem of identification is specified and the new identification assumption is discussed. The conditions under which the assumption can be used are developed. An application of the methodology is implemented to measure the propagation of shocks between sovereign bonds from Argentina and Mexico.

In section 3, the case where there are more than two regimes is analyzed. In general, the system is over identified. A GMM interpretation of the identification problem is addressed, and the extra equations are used to test some of the maintained assumptions. In this section, the timing of the regimes is imposed exogenously.

In section 4, necessary conditions for identification are derived for multivariate processes with unobservable common shocks. The presence of unobservable aggregate shocks is equivalent to relaxing the assumption of independence of structural shocks. The minimum number of covariance restrictions that achieves identification is discussed. In this section, only the order condition is analyzed.

In section 5, the question of consistency under misspecification of the heteroskedasticity is studied. Three cases are developed: first, when the number of regimes are correctly specified but not the timing of the regimes, or windows. Second, when the number of regimes is smaller than the actual number of regimes exhibited by the data. In these two cases, the estimates solve the same system of equations as with the correctly specified model. Therefore, consistency is assured if these types of misspecifications occur. An application of two alternative procedures to define the regimes is highlighted. The estimates are remarkably stable to alternative methodologies defining the regimes, which should be expected. On the other hand, when the data is homoskedastic and two or more regimes are assumed, the estimates are inconsistent. Simulations are presented to illustrate the magnitude of the bias.

Finally, in section 6, conclusions and extensions are presented.

2 New Identification Procedure

To illustrate the problem of identification assume we are interested in estimating the following system of equations:

$$p_t = \beta q_t + \varepsilon_t, \quad (1)$$

$$q_t = \alpha p_t + \eta_t, \quad (2)$$

where (1) is the demand equation, (2) is the supply equation, p_t is the observed price, q_t is the observed quantity, and ε_t and η_t are the structural shocks. The parameters of interest are α, β, and the variances of the structural shocks: σ^2_ε, σ^2_η. In this system of equations two assumptions, have been applied; normalization[3] and orthogonalization of structural shocks, $\sigma_{\varepsilon\eta} = 0$.

Equations (1) and (2) cannot be consistently estimated by OLS because of simultaneous equation bias: if α and β are different from zero, then the right hand side variables are correlated with the errors in both equations. Moreover, without further assumptions the parameters of interest cannot be consistently estimated by any method. Indeed in this example, the only statistics that can be obtained are parameters of the covariance matrix of the reduced form. Solving for p_t and q_t in equations (1) and (2):

$$p_t = \frac{1}{1-\alpha\beta}\left[\beta\eta_t + \varepsilon_t\right], \quad (3)$$

$$q_t = \frac{1}{1-\alpha\beta}\left[\eta_t + \alpha\varepsilon_t\right], \quad (4)$$

which implies a covariance matrix,

$$\Omega = \frac{1}{(1-\alpha\beta)^2}\begin{bmatrix} \beta^2\sigma^2_\eta + \sigma^2_\varepsilon & \beta\sigma^2_\eta + \alpha\sigma^2_\varepsilon \\ . & \sigma^2_\eta + \alpha^2\sigma^2_\varepsilon \end{bmatrix}.$$

The problem of identification is that the covariance matrix only provides three equations (the variance of p_t, the variance of q_t, and the covariance between p_t and q_t) and there are four unknowns: α, β, σ^2_η, σ^2_ε.

[3] The coefficients of p_t in the first equations, and of q_t in the second one are equal to one. This assumption is innocuous in the sense that it just sets the units in which the disturbances are measured.

The literature has solved this by imposing additional parameter constraints: (i) Exclusion restriction. This amounts to imposing either $\alpha = 0$, or $\beta = 0$.[4] (ii) Sign restrictions. The imposition of the sign on the slopes of the structural equations can achieve partial identification because the two inequalities imply a region of admissible parameters.[5] (iii) Long run constraints. When the structural form includes lagged dependent variables, it is possible to constrain the long run behavior of a particular shock. This assumption is equivalent to imposing that the sum of some of the lag coefficients is equal to zero.[6] (iv) Finally, constraints on the variances, for example that $\sigma_\eta^2/\sigma_\varepsilon^2$ is equal to some constant.

These restrictions have proven to be very useful in several applications. However, there are important economic problems in which non of them can be justified.

The purpose of this section is to offer a new identification method that is based on heteroskedasticity, and the two standard assumptions of normalization and independence of the structural shocks. Here, I discuss the case in which the heteroskedasticity can be described as two regimes, and in the following sections, I study further generalizations.

2.1 Identification under two regimes.

The maintained assumptions are the following: (i) the data exhibits two regimes in the second moments; (ii) it is known when the shift between the regimes occurs; (iii) the parameters (α and β) are stable through out the two regimes; (iv) and all variances are finite. Additionally, without loss of generality, assume that the first moments are equal to zero. Assumptions (iii) and (iv) are standard in macro-applications where traditional identification restrictions are used. Assumptions (i) and (ii) state that the data is heteroskedastic and that its form is known.[7]

Under these assumptions it is possible to estimate two covariance matrices, one for each regime.

[4] Technically this is implemented as a Cholesky decomposition of the variance covariance matrix. This restriction has been used in several applications, specially in the measurement of the transmission of monetary policy in US. The restriction, however, is assuming the problem of simultaneous equations away; if any of the contemporaneous parameters is zero then there is no problem of endogenous bias, and the estimation could have been done by OLS in the structural equations. Circumstances in which this assumption can be used are, therefore, limited.

[5] Even though a unique estimate cannot be obtained, at least an admissible interval is derived. See Fisher [1976].

[6] If it is known that one shock does not have permanent effects, then under some conditions, it is possible to obtain identification. For example, assume that nominal shocks are short lived, while real shocks are permanent. Imposing this constraint Blanchard and Quah [1989] and Shapiro and Watson [1988] were able to estimate the effects of aggregate shocks on aggregate activity and unemployment.

[7] These assumptions are relaxed in the following sections.

Denote the regimes as $s \in \{1, 2\}$, and define the variances of the structural shocks in regime s as $\sigma_{\varepsilon,s}$ and $\sigma_{\eta,s}$, then the covariance matrices of the reduced form are,

$$\Omega_s \equiv \begin{bmatrix} \omega_{11,s} & \omega_{12,s} \\ & \omega_{22,s} \end{bmatrix} = \frac{1}{(1-\alpha\beta)^2} \begin{bmatrix} \beta^2\sigma^2_{\eta,s} + \sigma^2_{\varepsilon,s} & \beta\sigma^2_{\eta,s} + \alpha\sigma^2_{\varepsilon,s} \\ . & \sigma^2_{\eta,s} + \alpha^2\sigma^2_{\varepsilon,s} \end{bmatrix}, \quad s \in \{1, 2\} \tag{5}$$

where Ω_s indicates the covariance matrix of the regime s.

In this new system of equations there are six unknowns: α, β, $\sigma^2_{\eta,1}$, $\sigma^2_{\varepsilon,1}$, $\sigma^2_{\eta,2}$, and $\sigma^2_{\varepsilon,2}$. But the two covariance matrices provide six equations. Therefore, this is a non linear system with six equations and six unknowns, and if not singular then the system is identified.[8] Note that no additional assumption is required to identify the structural form; there are no parameter restrictions, other than the stability of the coefficients.

This identification assumption is in some ways more general than the standard restrictions used in the literature. The traditional restrictions require assumptions that have to be economically justified. In some circumstances, those explanations are not entirely convincing and cannot be tested. On the other hand, the presence of heteroskedasticity in the structural shocks, can be tested on the reduced form using standard techniques. An alternative interpretation, is that this procedure assures identification on the alternative hypothesis that heteroskedasticity exists, but not under the null.[9] The assumption on the stability of the parameters cannot be tested in this case, but as discussed in section 3, when there are more than two regimes, even this assumption can be tested.

2.1.1 Conditions for identification

After solving for the variances in equations (5), α and β satisfy the following non-linear system of equations:

$$\beta = \frac{\omega_{12,s} - \alpha \cdot \omega_{11,s}}{\omega_{22,s} - \alpha \cdot \omega_{12,s}}, \quad s \in \{1, 2\}. \tag{6}$$

[8] Non linear systems, in general, can have more than one solution. However, because the matrices that are involved are positive semi definite it is the case that always there are two solutions to the system. One is the model as in equations (1) and(2). The second solution is solving for q_t in equation (1), and solving for p_t is equation (2).

[9] A Bayesian approach to identification can be also used, where the researcher can have a prior about the existence of heteroskedasticity in the data.

After further algebra α solves the quadratic equation:

$$[\omega_{11,1}\omega_{12,2} - \omega_{12,1}\omega_{11,2}]\,\alpha^2 - [\omega_{11,1}\omega_{22,2} - \omega_{22,1}\omega_{11,2}]\,\alpha + [\omega_{12,1}\omega_{22,2} - \omega_{22,1}\omega_{12,2}] = 0 \qquad (7)$$

There are two solutions to the quadratic equation. It is easy to show that if α, β is a solution to the system of equations, then $\alpha^* = 1/\beta$, $\beta^* = 1/\alpha$, is the other solution. Indeed, the solutions are the two possible ways in which the structural form can be written.[10] Thus, their structural interpretations are the same.

For the purpose of inference in this paper, the small sample distributions of α and β are derived by bootstrapping taken as given the asymptotic distributions of the covariance matrices. The following proposition summarizes the conditions under which α and β are identified.

Proposition 1 *Let p_t and q_t be described by equations (1) and (2), where the parameters (α and β) determining the law of motion are stable through out all realizations, and where the disturbances exhibit heteroskedasticity that can be described with two regimes.*

Then, if the covariance matrices satisfy

$$\det\left|\Omega_2 - \frac{w_{11,2}}{w_{11,1}}\Omega_1\right| \neq 0 \qquad (8)$$

the structural form is just identified.

Equation (8) is equivalent to

$$w_{11,1}w_{12,2} - w_{11,2}w_{12,1} \neq 0 \qquad (9)$$

Conditions (8) or (9) are similar to testing the rank condition when the order condition has been satisfied.[11] The existence of heteroskedasticity provides the order condition, and because the system is non-linear, the rank condition takes the form of equation (8). Even though condition

[10] One is given by equations (1) and (2), and the second one is given by solving for q_t in (1) and by solving for p_t in (2).

[11] In the standard linear problem, the order condition indicates that the number of equations has to be at least the number of unknowns. The rank condition requires that the number of linearly independent equations has to be equal to the number of unknowns. In practice, this is done by computing the rank of the matrix. In this case, because the system is non-linear the condition is somewhat more complicated.

(8) is not formally a rank condition, I use this name because its close parallel with the traditional literature on identification. From now on I refer to this constraint as either the "rank condition" or as the "covariance of the weighted difference".

Proof. Identification is not achieved only when equation (7) has no real solutions, or it has infinite solutions.

A real solution requires

$$[\omega_{11,1}\omega_{22,2}-\omega_{22,1}\omega_{11,2}]^2-4\left[\omega_{11,1}\omega_{12,2}-\omega_{12,1}\omega_{11,2}\right]\left[\omega_{12,1}\omega_{22,2}-\omega_{22,1}\omega_{12,2}\right]>0$$

After some algra this is equal to

$$\left[\omega_{11,2}^2\omega_{22,2}^2\right]\left[\theta_{11}-\theta_{22}\right]^2-\left[2\omega_{11,2}\omega_{22,2}\omega_{12,2}^2\right]\left[2\left(\theta_{11}-\theta_{12}\right)\left(\theta_{12}-\theta_{22}\right)\right]>0$$

where $\theta_{11}=\frac{\omega_{11,1}}{\omega_{11,2}}$, $\theta_{12}=\frac{\omega_{12,1}}{\omega_{12,2}}$, and $\theta_{22}=\frac{\omega_{22,1}}{\omega_{22,2}}$. This inequality is positive if both of the following inequalities are satisfied

$$\begin{aligned}\left[\omega_{11,2}^2\omega_{22,2}^2\right]-\left[2\omega_{11,2}\omega_{22,2}\omega_{12,2}^2\right] &> 0\\ \left[\theta_{11}-\theta_{22}\right]^2-\left[2\left(\theta_{11}-\theta_{12}\right)\left(\theta_{12}-\theta_{22}\right)\right] &> 0\end{aligned}$$

The first one is satisfied because the positive definite properties of the covariance matrix: $\omega_{11,2}\omega_{22,2}\left[\omega_{11,2}\omega_{22,2}-\right.$ 0. The second inequality is, after some algebra, equal to

$$\left[\theta_{11}-\theta_{12}\right]^2+\left[\theta_{22}-\theta_{12}\right]^2>0$$

which is always positive. Therefore, if the coefficients in the quadratic equation are different from zero, then the two roots are real.

The last requirement is to show when the quadratic equation does not have infinite solutions. This needs that either $\omega_{11,1}\omega_{22,2}-\omega_{22,1}\omega_{11,2}\neq 0$, or $\omega_{11,1}\omega_{12,2}-\omega_{12,1}\omega_{11,2}\neq 0$. Given the model generating the data, these two assumptions are not satisfied if the heteroskedasticity implies a proportional change of both structural shocks variances. In other words, when $\Omega_2=a\Omega_1$, for some scalar a. This is the only case in which the solution to the quadratic equation (7) has infinite solutions.

Note that if $\Omega_2 = a\Omega_1$ then $\det[\Omega_2 - a\,\Omega_1] = 0$, which can be tested by computing whether or not $\det\left[\Omega_2 - \frac{\omega_{11,2}}{\omega_{11,1}}\,\Omega_1\right] \stackrel{?}{=} 0$. By construction this is equivalent to asking if the covariance of the normalized difference is equal to zero: $\omega_{11,1}\omega_{12,2} - \omega_{11,2}\omega_{12,1} \stackrel{?}{=} 0$. The asymptotic properties of this statistic are better behaved than the one from the determinant, and in the empirical section this is what is implemented. ■

In conclusion, the identification is valid if the relative importance of the shocks changes across the regimes. This is the case, for example, of almost all macro-applications where ARCH or GARCH models are used.

The intuition of why identification is achieved is the following: the heteroskedasticity changes the region in which the errors are distributed, and when the parameters are stable it enlarges along the structural equations the ellipse in which the realizations are distributed.

The simplest intuition can be developed by first analyzing the case in which the variance changes for one shock. Assume that it is known that at some point in time there is an increase in the variance of the supply shocks. During that period, the "cloud" of realizations is going to widen along the demand curve as is depicted in figure 1. Comparing how the ellipse of the realizations has changed across the two samples allows one to determine the slope of the demand curve. In this particular case, because it has been assumed that the structural shocks are independent, then this is enough to estimate the slope of the supply curve.

Moreover, this explanation has an instrumental variable interpretation. The heteroskedasticity during periods of high volatility, increases the likelihood that the supply schedule shifts frequently. Thus, it is equivalent to a *probabilistic* instrument that is "moving" the supply equation.[12]

Finally, when both variances shift, then there is an expansion along both schedules. If the relative importance of shocks remains the same, then the two ellipses are proportional and no additional information is obtained from the heteroskedasticity, other than the magnitude of the shift. On the contrary, if the relative importance of the shocks changes, then the ellipse widens more along one of the schedules than the other and the problem is solved as before.

[Figure 1 here]

[12] In an earlier paper, I used a similar procedure (only one variance changes) to test for the stability of the parameters. I implement the test using an Instrumental Variable methodology. See Rigobon [1999].

The following three remarks are in order:

Remark 1 *It is not necessary to know which shocks becomes more important across the regimes. The fact that there is change is enough.*

The proof of proposition 1 only requires the variances to shift, not the knowledge of the direction of the shifts.

Remark 2 *The assumption of knowing when the regime shift occurs can be relaxed. The two covariance matrices can be estimated by implementing a regime shift regression as in Hamilton [1990].*

First, to obtain identification no assumption about the relationship between the two covariance matrices was imposed. If the covariances are correlated across the regimes, but satisfy condition (8) then they still provide enough equations for identification. In section 5 the issue of consistency when the heteroskedasticity is misspecified is addressed, and in section 5.4 a markov switching regime is used to estimate the high and low variance regimes.

Second, note that identification is achieved only under the alternative hypothesis; if there is no heteroskedasticity (the null hypothesis) there is only one regime, and there is no identification.

Remark 3 *The assumption that the structural shocks have to be independent across the equations can be relaxed, and the relative importance and the total net effect of the structural shocks can still be estimated.*

If the structural shocks are not independent, then they can always be writen as follows:

$$
\begin{aligned}
p_t &= \beta q_t + \varepsilon_t + \lambda \eta_t, \\
q_t &= \alpha p_t + \eta_t + \delta \varepsilon_t,
\end{aligned}
$$

where ε_t and η_t are independent. The reduced form is

$$\underbrace{\begin{bmatrix} 1 & -\beta \\ -\alpha & 1 \end{bmatrix}}_{A} \begin{bmatrix} p_t \\ q_t \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \lambda \\ \delta & 1 \end{bmatrix}}_{B} \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}$$

$$B^{-1}A \begin{bmatrix} p_t \\ q_t \end{bmatrix} = \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}.$$

In this case, if there are only two regimes the procedure identifies a normalization of $C \equiv B^{-1}A$. The net effect of the structural shocks and the relative importance of the shocks in each of the regimes still can be obtained $\left(\sigma_\varepsilon/\sigma_\eta\right)_s$. In this example, the parameters estimated are

$$\theta_1 = \frac{1}{1+\delta\lambda}\frac{\beta+\lambda}{1+\alpha\lambda}, \text{ and } \theta_2 = \frac{1}{1+\delta\lambda}\frac{\alpha+\delta}{1+\beta\delta}$$

Note that additional regimes (in the variance) do not achieve identification. Further regimes one give additional equations to estimate parameters θ_1 and θ_2.

The structural parameters cannot be recovered without more information on how the structural shocks are interrelated. In section 4, I return to these issues and discuss the necessary conditions for identification when there are common aggregate shocks, which is equivalent to this formulation.

2.1.2 Related literature

Before using the estimator in an application, at this time it is useful to discuss the relationship between the present paper and the current literature on identification using heteroskedasticity.

A closely related paper is Klein and Vella [2000b]. They discuss the problem of identification and estimation in a binary endogenous model when exclusion restrictions (or any other parameter restrictions) are not available. They show how the presence of heteroskedasticity achieves identification. Their procedure does not rely on the non-linearities of the binary model.[13] In fact, in Klein and Vella [2000a] , they study the case of the triangular model. They estimate semi-parametrically the heteroskedasticity and use the residual from the second equation as an additional regressor in the first equation as the intrument. As is argued in section 5, in the model presented in this

[13] See also Fiorentini and Sentana [1999].

paper it is not necessary to estimate the exact form of the heteroskedasticity to obtain consistent estimators.[14]

Rigobon [1999] discusses partial identification of simultaneous equation models with unobservable aggregate shocks. He is more concerned with developing a test for stability of parameters rather than identifying the system, and his identification depends on the presence of a particular form of the heteroskedasticity. Rigobon assumes that in the short run only one of the variances shifts. The methodology developed here does not require this assumption.

Finally, Sentana and Fiorentini [1999] study the problem of estimation in factor regressions when there is heteroskedasticity. They discuss the case in which the system is not triangular and the conditions in which identification is achieved. This is equivalent to solve the problem of identification in endogenous variables.

The estimation procedures are very different among all these papers. They share, however, the same intuition to solve the problem of endogenous variables; the heteroskedasticity adds equations to the system.

2.2 Brady Bonds: An example

In this section, I estimate the propagation of shocks between Mexican and Argentinean sovereign bonds. The data consists of the daily Brady Bond returns for Mexico and Argentina between January 1994 to December 1998 obtained from the Emerging Markets Bond Index Plus (EMBI+) index constructed by JPMorgan. The EMBI+ country indexes track total returns for traded external debt instruments in emerging markets. The indexes are computed by simulating holding a portfolio with the weights determined by risk, market capitalization, and liquidity considerations.

It should be clear that if Mexican shocks affect Argentina (through trade, for example), then Argentinean shocks should influence Mexico. This means that the price of both sovereign bonds are determined simultaneously. The two standard assumptions of normalization and independence of the structural shocks, are relatively easy to justify: the idiosyncratic shocks to each country (elections, social demonstrations, natural disasters, etc.) can be considered independent. The other traditional identification assumptions, however, are difficult to defend: (i) it is not reasonable

[14] See also Chen and Khan [1999] for a general solution of the problem of identification in sample selection model when the data exhibits heteroskedasticity.

to assume exclusion restrictions in one direction and not in the other one, as has already been argued; (ii) nor does it make sense to assume that one transmission is positive while the other one is negative (thus no sign restrictions can be imposed); (iii) there are no good reasons to assume that the shocks to one country are more persistent than the shocks to the other one (therefore, long run restrictions can not be enforced); (iv) and finally, it is difficult to substantiate an assumption about the relative importance of idiosyncratic shocks across the countries. This leaves the problem of identification unsolved with the standard techniques.

In figure 2, the two indexes are shown. As can be seen, the indexes are highly correlated; more than 99 percent in levels and more than 80 percent in daily returns.

[Figure 2 here]

In figure 3, the rolling variances and covariance of daily returns are shown, where a 20 days rolling window is used. Note that the Mexican crisis implied an increase in the conditional variance of both countries by more than 15 times. Similarly for the Russian crisis. The Asian crises, though, did not have a large impact on the conditional variances. The important conditional heteroskedasticity present in the data grants the use of the procedure here described to solve the problem of identification.

[Figure 3 here]

An important question in the measurement of the propagation of shocks (or "contagion") is the choice of the frequency.[15] In the international finance literature, it has been repeatedly highlighted that contagion is indeed a high frequency event. The US stock market collapse in October of 1997 during the Hong Kong crisis is a perfect example; the stock market felt and recovered in less than a week of the initial shock. In fact, in most of the cases, the impact occurs within two weeks and usually they disappear after one or two quarters.[16] Moreover, shifts in the second moments, which here is the instrument to identify the system, are also short lived (see figures 2 and 3). This cast important problems in the empirical implementation given that several theories of contagion

[15] See Pristker [1999] for a survey on the theoretical literature and Forbes and Rigobon [1999] for a survey on the empirical literature.

[16] The longest instance of contagion was the Mexican crisis in 1994, which lasted almost four months. The main reason for the "turbulence" in this case, however, was the uncertainty around the rescue package that was announced in mid January, but finally approved by the congress in March.

are related to macro fundamentals, such as trade, fiscal policy, etc. which are, at best, measured monthly, and in most of the cases quarterly. Hence, in practice, a contagion event might occur and no change in macro fundamentals were account for the propagation, not because they did not occur, but becuase they were not measured. Therefore, if we are up to identify the propagation of shocks, only a reduce form interpretation can be implemented. The coefficients from what I call the structural from (α and β) are the result of the complex transmission mechanisms that occur between fundamentals and market participants expectations that unfortunately cannot be meassured at the observed frequencies. This data problem does not preclude the estimation of α and β. It changes our interpretation of such coefficients. Even though in some of the specifications I control for all the macro fundamentals available at high frequencies (daily); US interest rates, US stock market returns, exchange rates of Mexico and Argentina, and their stock markets, it should be understood that they are not enough to determine the channels through which contagion is operating. In this paper, I am concerned with highlighting the use of the methodology for identification, and not with the structural interpretation of the coefficients. The choice of this data is threefold: (i) because the question of how to measure contagion continues to be open and this might be an alternative procedure for it; (ii) because the problem of identification of the strenght of contagion cannot be solved with the standard identification methods; (iii) and because the data exhibits important conditional heteroskedasticity that grants the use of the procedure developed here.

2.2.1 Estimation Procedure

The estimation procedure is the following: First, a VAR is run to control for lags and common observable shocks, such as US interest rates. Second, the residuals from the first step are used to compute the contemporaneous relationship between the countries. Is in this step when the definition of the high and low variance regimes (the heteroskedasticity) is required to identify the system.

The following reduced form is estimated first:

$$\left\{ \begin{array}{c} Arg_t \\ Mex_t \end{array} \right\} = c + \phi\left(L\right) \left\{ \begin{array}{c} Arg_t \\ Mex_t \end{array} \right\} + \phi US_t + \Phi\left(L\right) \left\{ \begin{array}{c} US_t \\ z_t \end{array} \right\} + \nu_t \tag{10}$$

where Arg_t and Mex_t are the Argentinean and Mexican daily returns, z_t are the country controls (exchange rates and stock market returns of Mexico and Argentina)[17], US_t is the daily return on holding the US 10 years bond, c is a constant, $\phi(L)$ and $\Phi(L)$ are lag operators (usually 5 or 10 lags), and ν_t represent the residuals. The 10 years bond is used given that this is close to the average maturity of the two outstanding Brady bonds. Several sensitivity analysis were performed, and choosing a different US bond did not change the results.

After regression (10) is estimated, the second step is to take the residuals and compute their contemporaneous relationship. Under the assumption of endogeneity, the residuals are a linear combination of the structural shocks which can be written as follows:

$$\begin{aligned}\nu_{1,t} &= \frac{1}{1-\alpha\beta}\left[\beta\eta_t + \varepsilon_t\right], \\ \nu_{2,t} &= \frac{1}{1-\alpha\beta}\left[\eta_t + \alpha\varepsilon_t\right].\end{aligned}$$

To implement the identification procedure, a high and low variance period has to be defined. Because periods of crisis have been associated with increases in volatility, this is the criterion used in this section. Five cases are studied: the Mexican crisis alone, the Asian crises, the Mexican and Asian crises together, the Russian crisis alone, and All three recent crises together; Mexican, Asian, and Russian.

The crises periods are as follows: The Mexican crisis started in December 19, 1994 when the abandonment of the fixed exchange rate occurred. The end of the crisis is set at March 31, 1995 when the markets calmed down after the US bailout. The South East Asian crises started with the speculative attack against Thailand (June 01, 1997) and ended after the Korean crisis calmed down at the end of January 1998 (January 31, 1998). Finally, the Russian crisis started with the collapse in the bond market at the beginning of August (August 01, 1998) and lasted until the end of the month (August 31, 1998).

The sample for each exercise is as follows: for the Mexican crisis, the sample runs from January 01, 1994 until March 31, 1995; for the Asian crises the sample starts in April 01, 1995 and ends in January 31, 1998; while for the Russian crisis the sample is from February 01, 1998 until December 31, 1998. I define the tranquil period as the part of the sample that does not belong to the crisis window.

[17] Sensitivity analysis was analyzed by including or not these fundamentals and the results remained the same.

Table 1, presents the results of estimating the covariance matrix for both regimes. The columns highlight the crisis under consideration. The first set of three rows indicates the covariance matrix in the low variance regime; the next three rows are the covariance matrix in the high volatility regime; and the bottom three rows show the increase across the two regimes. Note that the increase in the second moments is sizeable.

	Mexico	Asia	Mexico Asia	Russia	Mexico Asia Russia
Low Variance Regime					
Variance Argentina $(\cdot 10^{-4})$	0.6324	0.1248	0.5432	1.7854	0.7923
Covariance $(\cdot 10^{-4})$	0.3676	0.0962	0.3197	1.0827	0.4713
Variance Mexico $(\cdot 10^{-4})$	0.4113	0.1297	0.3616	0.8420	0.4584
Hig Variance Regime					
Variance Argentina $(\cdot 10^{-4})$	8.9847	0.5979	3.1016	7.4546	3.5428
Covariance $(\cdot 10^{-4})$	7.2500	0.4201	2.4627	4.9882	2.7200
Variance Mexico $(\cdot 10^{-4})$	8.9327	0.4165	2.9720	3.6873	3.0612
Change					
Variance Argentina	14.21	4.79	5.71	4.18	4.47
Covariance	19.72	4.37	7.70	4.61	5.77
Variance Mexico	21.72	3.21	8.22	4.38	6.68

Table 1: Change in the variance covariance matrix during several crises. Exogenous crisis window.

The next step is to determine if the estimates of the covariance matrices satisfy the rank condition. Taking the change in the second moments as given, the distribution of the rank condition is obtained using a bootstrap; 1000 series of residuals are computed using the asymptotic distribution of the covariance matrices estimated is table 1.[18] For each run, the "covariance of the weighted difference" is calculated. Table 2 summarizes some of the statistics of the bootstrapped distribution.

Covariance of the Weighted Difference	Mexico	Asia	Mexico Asia	Russia	Mexico Asia Russia
Average $(\cdot 10^{-4})$	0.2057	0.0047	0.0637	0.0522	0.0610
Standard Deviation $(\cdot 10^{-4})$	0.0756	0.0032	0.0144	0.0344	0.0137
Quasi t_stat	2.72	1.44	4.42	1.52	4.45
Mass below zero	0.00%	0.00%	0.00%	0.00%	0.00%

Table 2: Test of the Rank Condition. Exogenous crisis window.

As can be seen, if the quasi t_stats are used (remember that the small sample distributions are not normal) the "covariance of the weighted difference" cannot be rejected to be equal to zero during Asia and Russia, even though both have no mass below zero.

[18] The distributions of the bootstrapps are not shown for brevety. The results are provided upon request or can be found in my web page.

Using the estimated covariance matrix in each regime, estimates of $\hat{\alpha}$ and $\hat{\beta}$ solving equations (6) are obtained. Gieven the economic interpretation of the coefficients I take the roots that imply $|\alpha\beta| < 1$. Their distribution is obtained by the same bootstrapping performed before. In table 3 the results are shown.

	Mexico	Asia	Mexico Asia	Russia	Mexico Asia Russia
$\hat{\alpha}$					
Point Estimate	0.3247	0.1911	0.3139	0.7109	0.3088
Standard Deviation	0.1470	0.0747	0.0791	0.5512	0.0871
Quasi t_stat	2.21	2.56	3.97	1.29	3.55
Mass below zero	2.60%	0.50%	0.20%	14.60%	0.20%
$\hat{\beta}$					
Point Estimate	0.6474	0.6586	0.6495	0.4637	0.6493
Standard Deviation	0.1133	0.0468	0.0609	0.2439	0.0669
Quasi t_stat	5.72	14.08	10.67	1.90	9.70
Mass below zero	0.00%	0.00%	0.00%	5.60%	0.00%

Table 3: Estimation of the structural parameters. Exogenous Windows.

Given our assumptions, $\hat{\alpha}$ is the propagation from Argentina to Mexico and $\hat{\beta}$ is the transmission from Mexico to Argentina. The estimates are, in most cases, statistically different from zero, and relatively large which means that they are economically important.[19]

First, note that confirming the results from table 2, the parameters estimated during the Russian crisis are not significantly different from zero. The implication of failing the rank condition is that the confidence intervals of the parameters should be infinitely wide. This is because there exists a continuum of solutions to equation (7).[20] Moreover, as can be seen, it is not the case that the coefficient estimated is closer to zero, rather that the standard deviation is large (in comparison to the estimates in the other subsamples).

Second, the estimates are remarkably stable across sub-samples. I return to this point in the next section where a Hausman specification test is implemented.

[19] For example, in order to give some intuition of the order of magnitude of the coefficients, assume the two idiosyncratic shocks have the same variance. Then, if $\alpha = 0.10$ and $\beta = 0.10$ the correlation between the two prices would have been 19.8 percent. For the estimated values of $\alpha = 0.33$ and $\beta = 0.65$ the correlation is 78.1 percent.

[20] I owe this interpretation to Mark Watson.

3 Identification under more than two regimes.

Generalizations of the identification problem to more than two regimes should be straight forward. First, I study the case in which there are S regime shifts. Second, I discuss the case in which there is a continuum of states, or ARCH. Thirdly, I estimate the Brady Bond example using four regimes.

3.1 S regimes.

When the data exhibits multiple regimes, there is a covariance matrix for each of them, giving rise to the following set of equations:

$$\Omega_s \equiv \begin{bmatrix} \omega_{11,s} & \omega_{12,s} \\ & \omega_{22,s} \end{bmatrix} = \frac{1}{(1-\alpha\beta)^2} \begin{bmatrix} \beta^2\sigma^2_{\eta,s} + \sigma^2_{\varepsilon,s} & \beta\sigma^2_{\eta,s} + \alpha\sigma^2_{\varepsilon,s} \\ . & \sigma^2_{\eta,s} + \alpha^2\sigma^2_{\varepsilon,s} \end{bmatrix}, \quad s \in \{1,..,S\} \tag{11}$$

where this system has $3S$ equations (one covariance matrix per regime) and $2S+2$ unknowns: S times two variances for each regime, plus two parameters (α and β).

For S larger than two the system has more equations than unknowns. This can have three interpretations: Firstly, those equations can be thought as overidentifying restrictions, and the model can be tested. For example, the underlying assumption that α and β are stable through time can be tested if at least three regimes are present in the data.

Secondly, it can be a factor regression model. The left hand side variables of equations (11) are the estimates (or observables), the variances ($\sigma^2_{\eta,s}$ and $\sigma^2_{\varepsilon,s}$) are the unobservable factors, and the coefficients are the weights or factor loadings.[21,22]

Thirdly, it has a minimum distance interpretation, either as a NLS or as a GMM. Solving in equations (11) for α and β, the same relationship as equation (6) is obtained in each state.

$$\beta = \frac{\omega_{12,s} - \alpha \cdot \omega_{11,s}}{\omega_{22,s} - \alpha \cdot \omega_{12,s}}, \ s \in \{1,...,S\} \tag{12}$$

The parameters can be estimated as of to reduce the distance between the left hand side and the

[21] Factor analysis usually assumes that the $\omega_{ij,s}$'s are independent. It is unlikely, however, that this is the case in this setup. Therefore proper corrections have to be taken into consideration in the estimation procedure.

[22] See Sentana and Fiorentini [1999] for a discussion of factor regression model estimation under heteroskedasticity. They also briefly study the problem of identification of the factor weights when the variables are endogenous.

right hand side of equation (12) as a NLS,

$$\sum_s \left(\beta - \frac{\omega_{12,s} - \alpha \cdot \omega_{11,s}}{\omega_{22,s} - \alpha \cdot \omega_{12,s}}\right)^2 = 0, \tag{13}$$

or as GMM,

$$E_s \left\{\beta - \frac{\omega_{12,s} - \alpha \cdot \omega_{11,s}}{\omega_{22,s} - \alpha \cdot \omega_{12,s}}\right\} = 0. \tag{14}$$

Both the NLS and the GMM require important assumptions on the innovations of the system of equations (11) in order to assure that equations (13) or (14) make econometric sense.

Finally, it is possible to estimate a bivariate ARCH or GARCH model and construct a system similar to (11), where there is a set of equations for each period.

Again, this has the same interpretations as before, even though the overidentification interpretation assumes a very large set of constraints that are not independent. Note that by construction the observations are serially correlated.

3.2 Brady Bonds: $S = 4$ regimes.

Going back to the Brady Bonds example, assume there are four regimes: three crises and one tranquil period. The sample period runs from January 94 to December 98, the crises windows are defined as before, and the tranquil time is the rest of the sample. It is assumed that each crisis coincides with a different regime. After running the first step, four covariance matrices are estimated.

Regimes	Tranquil	Mexico	Asia	Russia
Variance Argentina $(\cdot 10^{-4})$	0.7922	8.9847	1.1502	7.4546
Covariance $(\cdot 10^{-4})$	0.4713	7.2499	0.7844	4.9882
Variance Mexico $(\cdot 10^{-4})$	0.4584	8.9327	0.6756	3.6873

Table 4: Variance Covariance Matrix for all the regimes.

Similarly as before, the Mexican and Russian crises imply important increases in the second moments, while the Asian crises mildly affected the covariance matrix. Moreover, note that the estimates of the variances are similar to those obtained in the previous section.

The four covariance matrices provide more equations than unknowns and the parameters are estimated using the GMM approach. The distribution of $\hat{\alpha}$ and $\hat{\beta}$ are obtained by bootstrapping as before. In table 5 the results are summarized.

	$\hat{\alpha}$	$\hat{\beta}$
Point Estimate	0.317345	0.639383
Standard Deviation	0.1919	0.0858
Quasi t_stat	1.65	7.52
Mass below zero	5.60%	0.20%

Table 5: Summary statistics of the bootstrapped distribution.

Note that the estimates are close to those obtained in the previous section. Moreover, it is not possible to reject the hypothesis that the parameters are the same between these ones and all previous five exercises. This is a Hausman Specification Test and its results are presented in table 6.

	α		β	
	Difference	Std. Deviation	Difference	Std. Deviation
Mexico	0.009018	0.123046	0.005694	0.085682
Asia	0.124586	0.176575	0.016905	0.057466
Mexico-Asia	0.001765	0.174657	0.007833	0.042220
Russia	0.395288	0.516776	0.178026	0.232380
All	0.006813	0.170784	0.007633	0.031756

Table 6: Specification tests.

There is no single case in which the test is rejected.

Implicitly, this is a test on the stability of α and β. As is discussed in section 5 the estimates are consistent to misspecification of the heteroskedasticity. Therefore, it should be expected that the estimates using two or four regimes are consistent if the only difference in the samples is the specification of the heteroskedasticity. However, if α or β are not stable across crises, then the shorter sample regressions (those from the previous section) are consistent, but the ones using the whole sample are not. Under this setup, the estimates using the whole sample and 4 regimes are efficient and consistent under the null, but inconsistent if α or β shift. On the other hand, the shorter regressions using only two regimes are consistent in both hypothesis. Note that the test on the stability of α and β is between crises and not between tranquil and crisis periods. This continues to be a maintained assumption.

4 Aggregate shocks

In the previous sections, the stochastic process is bivariate and there are no common shocks. In this section, I relax this assumption and discuss the necessary conditions to achieve identification.

I continue to assume normalization of all structural shocks and aggregate shocks, as well as their independence. Modify the structural form as follows:

$$p_t = \beta q_t + \gamma z_t + \varepsilon_t, \tag{15}$$

$$q_t = \alpha p_t + z_t + \eta_t, \tag{16}$$

where z_t is an unobservable common shock whose variance is σ_z^2.

This model not only encompasses the case in which there are aggregate omitted variables, but also the case in which the original structural shocks are correlated. In this case, z_t would be describing that correlation. The reduced form is then

$$p_t = \frac{1}{1-\alpha\beta}\left[\beta\eta_t + \varepsilon_t + (\beta+\gamma) z_t\right], \tag{17}$$

$$q_t = \frac{1}{1-\alpha\beta}\left[\eta_t + \alpha\varepsilon_t + (1+\alpha\gamma) z_t\right], \tag{18}$$

where the covariance matrix is

$$\Omega = \frac{1}{(1-\alpha\beta)^2}\begin{bmatrix} \beta^2\sigma_\eta^2 + \sigma_\varepsilon^2 + (\beta+\gamma)^2\sigma_z^2 & \beta\sigma_\eta^2 + \alpha\sigma_\varepsilon^2 + (1+\alpha\gamma)(\beta+\gamma)\sigma_z^2 \\ . & \sigma_\eta^2 + \alpha^2\sigma_\varepsilon^2 + (1+\alpha\gamma)^2\sigma_z^2 \end{bmatrix}.$$

In this case, the problem of identification cannot be solved using the heteroskedasticity alone. Each additional covariance matrix contributes with three equations, but also with three new unknowns; $\sigma^2_{\eta,s}$, $\sigma^2_{\varepsilon,s}$, and $\sigma^2_{z,s}$. Thus the system remains equally under-identified.

How to solve the problem when there are common shocks then? Allowing for more endogenous variables.

Assume that there are $s \in \{1, ..., S\}$ possible regimes and that the structural form is written as follows:

$$A_{NxN}\begin{bmatrix} x_{1,t} \\ \vdots \\ x_{N,t} \end{bmatrix} = \Gamma_{NxK}\begin{bmatrix} z_{1,t} \\ \vdots \\ z_{K,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \vdots \\ \varepsilon_{N,t} \end{bmatrix}, \tag{19}$$

where the assumption of orthogonalization implies,

$$\begin{aligned} E\left[z_{i,t}, z_{j,t}\right] &= 0 \quad \forall i \neq j, \quad i,j \in \{1, K\} \\ E\left[\varepsilon_{i,t}, \varepsilon_{j,t}\right] &= 0 \quad \forall i \neq j, \quad i,j \in \{1, N\} \\ E\left[z_{i,t}, \varepsilon_{j,t}\right] &= 0 \quad \forall i \neq j, \quad i \in \{1, K\}, \quad j \in \{1, N\}, \end{aligned} \tag{20}$$

and where $x_{n,t}$, $n \in \{1, ..., N\}$ are the N endogenous variables; $z_{k,t}$, $k \in \{1, ..., K\}$ are the K unobservable common (or aggregate) shocks, assumed to be independent but not identically distributed, and with variance $\sigma_{z,k,s}$ in state s; $\varepsilon_{n,t}$ are the structural shocks, assumed to be independent, but not identically distributed, and with variance $\sigma_{\varepsilon,n,s}$ in state s.

A_{NxN} is the matrix that describes the contemporaneous parameters, where the assumption of normalization has been already imposed:

$$A_{NxN} = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix}, \tag{21}$$

and Γ_{NxK} are the parameters from the common shocks. Normalization of the common shocks is also assumed; in this case, it implies a unitary impact on the first equation,

$$\Gamma_{NxK} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nk} \end{bmatrix}. \tag{22}$$

Proposition 2 *A multivariate system of N equations, with K unobservable aggregate shocks, described by equations (19), (20), (21) and (22) can be identified if the number of states (S) satisfies,*

$$S \geq 2\frac{(N+K)(N-1)}{N^2 - N - 2K}, \tag{23}$$

and if there is a minimum number of endogenous variables that satisfies

$$N^2 - N - 2K > 0. \tag{24}$$

Proof. Note that the proposition states a necessary condition, but not a sufficient one. Thus it is stating an order condition.

From equation (19), the number of equations is given by the covariance matrix in each regime. This provides $\frac{N(N+1)}{2}$ equations in each state. The total number of unknowns is as follows: The matrix A_{NxN} has $N(N-1)$ parameters; The matrix Γ_{NxK} has $K(N-1)$ parameters; the variances of the aggregate shocks in each state (K variances times S regimes) and the variances of the structural shocks in each regime (N variances times S regimes). Identification, then, requires

$$\begin{aligned} S \cdot \frac{N(N+1)}{2} &\geq N(N-1) + K(N-1) + S \cdot K + S \cdot N \\ S &\geq 2\frac{(N+K)(N-1)}{N^2 - N - 2K}. \end{aligned}$$

Inequality (23) indicates the minimum number of states required to obtain identification. Finally, in order for (23) to make sense, there is a minimum number of endogenous variables, which is given by,

$$N^2 - N - 2K > 0.$$

For example, the minimum number of endogenous variables necessary to solve the problem that motivated this section is $N > 2$. ■

Equation (24) is the "catch up" constraint. In fact, it indicates when one additional regime in the variance-covariance adds more equations than unknowns. In the example that motivated this section, $N = 2$ and $K = 1$, which implies that the inequality in not satistied and no further information is obtained from the heteroskedasticity. Moreover, interpreting the aggregate shocks as the sources of correlation between the structural shocks, this constraint requires that at least one of the covariance has to be zero. For example, when there are N structural shocks, it will require $K = \frac{N(N-1)}{2}$ aggregate shocks to describe all the cross-correlations. Note that in this case, however, the inequality (24) is not satisfied. In other words, in order for the heteroskedasticity to provide additional information then at least one covariance has to be restricted. In summary, it is the combination of the standard assumption (covariance restriction) and the heteroskedasticity what achieves identification.

There are several remarks that can be extracted from proposition 2: First, in the absence of aggregate shocks only two states are required to achieve identification. Independently of the number of endogenous variables N, when $K = 0$, the right hand side of equation (23) is equal to 2, and equation (24) is always satisfied.

Second, in the limit, when the number of endogenous variables is infinite, the system can be just identified with only two states, independently of any finite number of aggregate shocks.

Finally, if $K > 0$ and N is finite, the number of states required to achieve identification is always larger than two. With two states the constraint (23) is always violated.

5 Consistency under misspecification of the heteroskedasticity.

An important question that arises from the previous exercises is the issue of consistency when the heteroskedasticity is misspecified. In this section three cases are evaluated: (i) when the windows of the heteroskedasticity are wrongly specified but the number of regimes is correct, (ii) when the data has more regimes than the ones assumed in the specification, (iii) and when the data does not exhibits heteroskedasticity, but it is assumed erroneously that it has.

Without loss of generality, only the case in which there are no aggregate shocks is discussed. The generalization to common unobservables should be straight forward.

5.1 Misspecification of the windows.

Assume that the two returns are described by equations (1) and (2), and that the data exhibits heteroskedasticity with only two regimes. If the windows are misspecified, the computed covariance matrices are linear combinations of the true underlying covariance matrices. Denote

$$\begin{aligned}
\Omega_{r1} &= \lambda_{r1}\Omega_1 + (1-\lambda_{r1})\,\Omega_2, \\
\Omega_{r2} &= (1-\lambda_{r2})\,\Omega_1 + \lambda_{r2}\Omega_2,
\end{aligned}$$

where Ω_1 and Ω_2 are the true covariance matrices describing the heteroskedasticity, Ω_{r1} and Ω_{r2} are the estimated covariance matrices, and λ_{r1} and λ_{r2} are weights indicating how correct the windows are; when they are equal to one, the windows coincide with the true regimes.

Proposition 3 *Assume the original system satisfies the rank condition (8). If the misspecified heteroskedasticity satisfies the rank condition (8) then the model is identified and its estimators are consistent.*

Proof. After some algebra the two covariance matrices can be written in terms of the underlying variances:

$$\begin{aligned}\Omega_{r1} &= \frac{1}{(1-\alpha\beta)^2}\begin{bmatrix}\beta^2\sigma^2_{\eta,r1}+\sigma^2_{\varepsilon,r1} & \beta\sigma^2_{\eta,r1}+\alpha\sigma^2_{\varepsilon,r1}\\ . & \sigma^2_{\eta,r1}+\alpha^2\sigma^2_{\varepsilon,r1}\end{bmatrix},\\ \Omega_{r2} &= \frac{1}{(1-\alpha\beta)^2}\begin{bmatrix}\beta^2\sigma^2_{\eta,r2}+\sigma^2_{\varepsilon,r2} & \beta\sigma^2_{\eta,r2}+\alpha\sigma^2_{\varepsilon,r2}\\ . & \sigma^2_{\eta,r2}+\alpha^2\sigma^2_{\varepsilon,r2}\end{bmatrix},\end{aligned}$$

where

$$\sigma^2_{\eta,r1} = \lambda_{r1}\sigma^2_{\eta,1}+(1-\lambda_{r1})\,\sigma^2_{\eta,2} \quad \text{and} \quad \sigma^2_{\varepsilon,r1}=\lambda_{r1}\sigma^2_{\varepsilon,1}+(1-\lambda_{r1})\,\sigma^2_{\varepsilon,2} \tag{25}$$

$$\sigma^2_{\eta,r2} = (1-\lambda_{r2})\,\sigma^2_{\eta,1}+\lambda_{r2}\sigma^2_{\eta,2} \quad \text{and} \quad \sigma^2_{\varepsilon,r2}=(1-\lambda_{r2})\,\sigma^2_{\varepsilon,1}+\lambda_{r2}\sigma^2_{\varepsilon,2}. \tag{26}$$

Given that the original heteroskedasticity satisfied the rank condition ($\sigma^2_{\eta,1}\sigma^2_{\varepsilon,2}-\sigma^2_{\eta,2}\sigma^2_{\varepsilon,1}\neq 0$), there are two questions that have to be answered: when the misspecified model satisfies the rank condition, and when the estimates are consistent. After some algebra, Ω_{r1} and Ω_{r2} satisfy equation (8) if and only if

$$\sigma^2_{\eta,r1}\sigma^2_{\varepsilon,r2}\neq\sigma^2_{\eta,r2}\sigma^2_{\varepsilon,r1}.$$

Substituting by the definitions of the variances (equations 25 and 26) the rank condition is not satisfied if and only if

$$\lambda_{r1}=1-\lambda_{r2}.$$

In other words, the rank condition is not satisfied if the windows are so badly determined that they imply the same weights of the true regimes. Thus, the two computed matrices are identical. This is equivalent to test the rank condition on the misspecified model.

Assume the rank condition is satisfied, then the question is if the solution of the new system of equations is consistent. Substituting in equation (7) the estimated $\hat{\alpha}$ solves.

$$\frac{\Phi\beta}{(1-\alpha\beta)^3}\left(\hat{\alpha}^2-\left(\frac{1}{\beta}+\alpha\right)\hat{\alpha}+\frac{\alpha}{\beta}\right)=0 \tag{27}$$

where

$$\Phi=\left(\sigma_{\eta,1}^2\sigma_{\varepsilon,2}^2-\sigma_{\eta,2}^2\sigma_{\varepsilon,1}^2\right)\left(1-\lambda_{r1}-\lambda_{r2}\right)$$

Note that under the assumptions that the original heteroskedasticity satisfies the rank condition, and that $\lambda_{r1}\neq 1-\lambda_{r2}$, then Φ is different from zero. Equation (27) solves the exact same quadratic equation as the well specified model. Thus the consistency is assured if the covariance matrix is consistently estimated. The two solutions are α and $1/\beta$. Therefore, if the misspecified system satisfies the rank condition, then the estimates are consistent. ■

5.2 Under-specified number of regimes.

Assume the two returns are described by equations (1) and (2), and that the data exhibits heteroskedasticity with S^* regimes, where there are no restrictions to the form of the heteroskedasticity, nor on S^* being finite (thus ARCH models are included). For simplicity denote the variances of the structural shocks in each regime as follows:

$$\begin{aligned}\sigma_{\eta,s}^2&=\left(1+\lambda_{\eta,s}\right)\sigma_{\eta,0}^2\\ \sigma_{\varepsilon,s}^2&=\left(1+\lambda_{\varepsilon,s}\right)\sigma_{\varepsilon,0}^2\end{aligned}\quad\forall s\neq 0$$

where $\sigma_{\eta,s}^2$ and $\sigma_{\varepsilon,s}^2$ represent the variances of the idiosyncratic shocks in regime s, and $\lambda_{\eta,s}$ and $\lambda_{\varepsilon,s}$ are the changes of those variances relative to the variances in regime $s=0$.

The problem of consistency is analyzed in the worst of the circumstances; it is assumed that only two regimes are computed. The first regime corresponds to the first $\hat{s}<S^*$ regimes, and the

second regime corresponds to the second $S^* - \hat{s}$ regimes.

The covariance matrices of each of the regimes are given by:

$$\Omega_{r1} = \frac{1}{(1-\alpha\beta)^2} \begin{bmatrix} \beta^2 \frac{1}{\hat{s}} \sum\limits_{s<\hat{s}} \sigma^2_{\eta,s} + \frac{1}{\hat{s}} \sum\limits_{s<\hat{s}} \sigma^2_{\varepsilon,s} & \beta \frac{1}{\hat{s}} \sum\limits_{s<\hat{s}} \sigma^2_{\eta,s} + \alpha \frac{1}{\hat{s}} \sum\limits_{s<\hat{s}} \sigma^2_{\varepsilon,s} \\ . & \frac{1}{\hat{s}} \sum\limits_{s<\hat{s}} \sigma^2_{\eta,s} + \alpha^2 \frac{1}{\hat{s}} \sum\limits_{s<\hat{s}} \sigma^2_{\varepsilon,s} \end{bmatrix}$$

and

$$\Omega_{r2} = \frac{1}{(1-\alpha\beta)^2} \begin{bmatrix} \beta^2 \frac{1}{S^*-\hat{s}} \sum\limits_{s>\hat{s}} \sigma^2_{\eta,s} + \frac{1}{S^*-\hat{s}} \sum\limits_{s>\hat{s}} \sigma^2_{\varepsilon,s} & \beta \frac{1}{S^*-\hat{s}} \sum\limits_{s>\hat{s}} \sigma^2_{\eta,s} + \alpha \frac{1}{S^*-\hat{s}} \sum\limits_{s>\hat{s}} \sigma^2_{\varepsilon,s} \\ . & \frac{1}{S^*-\hat{s}} \sum\limits_{s>\hat{s}} \sigma^2_{\eta,s} + \alpha^2 \frac{1}{S^*-\hat{s}} \sum\limits_{s>\hat{s}} \sigma^2_{\varepsilon,s} \end{bmatrix}$$

which can be written as

$$\begin{aligned} \Omega_{r1} &= \frac{1}{(1-\alpha\beta)^2} \begin{bmatrix} (1+\lambda_{\eta,r1})\,\beta^2\sigma^2_{\eta,0} + (1+\lambda_{\varepsilon,r1})\,\sigma^2_{\varepsilon,0} & (1+\lambda_{\eta,r1})\,\beta\sigma^2_{\eta,0} + (1+\lambda_{\varepsilon,r1})\,\alpha\sigma^2_{\varepsilon,0} \\ & (1+\lambda_{\eta,r1})\,\sigma^2_{\eta,0} + (1+\lambda_{\varepsilon,r1})\,\alpha^2\sigma^2_{\varepsilon,0} \end{bmatrix} \\ \Omega_{r2} &= \frac{1}{(1-\alpha\beta)^2} \begin{bmatrix} (1+\lambda_{\eta,r2})\,\beta^2\sigma^2_{\eta,0} + (1+\lambda_{\varepsilon,r2})\,\sigma^2_{\varepsilon,0} & (1+\lambda_{\eta,r2})\,\beta\sigma^2_{\eta,0} + (1+\lambda_{\varepsilon,r2})\,\alpha\sigma^2_{\varepsilon,0} \\ & (1+\lambda_{\eta,r2})\,\sigma^2_{\eta,0} + (1+\lambda_{\varepsilon,r2})\,\alpha^2\sigma^2_{\varepsilon,0} \end{bmatrix} \end{aligned}$$

where

$$\lambda_{\eta,r1} = \frac{1}{\hat{s}} \sum_{s<\hat{s}} \lambda_{\eta,s} \quad \text{and} \quad \lambda_{\eta,r2} = \frac{1}{S^*-\hat{s}} \sum_{s>\hat{s}} \lambda_{\eta,s} \tag{28}$$

$$\lambda_{\varepsilon,r1} = \frac{1}{\hat{s}} \sum_{s<\hat{s}} \lambda_{\varepsilon,s} \quad \text{and} \quad \lambda_{\varepsilon,r2} = \frac{1}{S^*-\hat{s}} \sum_{s>\hat{s}} \lambda_{\varepsilon,s} \tag{29}$$

Proposition 4 *Assume the true heteroskedasticity is described by S^* regimes, and that at least two of those covariance matrices satisfy the rank condition (8).*

Assume that only two regimes have been used in the estimation, then if the following conditions are satisfied the system is identified and its estimates are consistent.

1. *The misspecified variace covariance matrices have to exhibit heteroskedasticity.*

2. *The misspecified variace covariance matrices have to satisfy the rank condition (8).*

Proposition 5 *and where the λ's are given by equations (28) and (29).*

Proof. The first assumption in the proposition is to guarantee that the original system can be identified if the heteroskedasticity is well speficied. In the ill specified model, identification is achieved if the relative volatilities change. This is equivalent to as if

$$\lambda_{\eta,r1} \neq \lambda_{\eta,r2} \quad \text{or} \quad \lambda_{\varepsilon,r1} \neq \lambda_{\varepsilon,r2} \tag{30}$$

Equation (30) indeed guarantees that the two estimated covariance matrices are different. In other words, that the order condition will be satisfied; there is heteropskedasticity in the estimated model.

The next question is, as before, what are the conditions for consistency. Substituting in equation (7) for the computed covariance matrices (Ω_{r1} and Ω_{r2}) the estimated $\hat{\alpha}$ satisfies,

$$\frac{\sigma_{\eta,0}^2 \sigma_{\varepsilon,0}^2 \Phi \beta}{(1-\alpha\beta)^3} \left(\hat{\alpha}^2 - \left(\frac{1}{\beta} + \alpha \right) \hat{\alpha} + \frac{\alpha}{\beta} \right) = 0 \tag{31}$$

where

$$\Phi = (1 + \lambda_{\varepsilon,r1})(1 + \lambda_{\eta,r2}) - (1 + \lambda_{\varepsilon,r2})(1 + \lambda_{\eta,r1}).$$

Note that if Φ is different from zero, then $\hat{\alpha}$ solves the same quadratic equation as the original model. Φ is different from zero if condition (30) is satisfied, and

$$\frac{\lambda_{\eta,r1}}{\lambda_{\eta,r2}} \neq \frac{\lambda_{\varepsilon,r1}}{\lambda_{\varepsilon,r2}}, \tag{32}$$

is also satisfied. Condition (32) indicates that the change in the variances across the misspecified regimes, cannot be proportional. In other words, this is equivalent to the rank condition discussed before. Again, the two roots solving equation (31) are α and $1/\beta$.

In summary, even though the assumed form of the heteroskedasticity implies a smaller number of regimes than those exhibited in the data, the system is identified and its estimates are consistent if and only if the order and rank conditions are satisfied. ■

5.3 No heteroskedasticity.

In this section it is assumed that the data does not exhibits heteroskedasticity, but it has been erroneously assumed that it has. I show that, generally, the rank condition is satisfied, but that the estimates are inconsistent.

A clear example in which this misspecification occurs is if the sample is split accordingly to the variance of one of the variables. In order to highlight the size of the inconsistency the following simulation was run: (i) 500 random bivariate realizations were generated with constant variances; (ii) assuming an α and β the series of p_t and q_t were computed; (iii) the sample was split accordingly to high and low volatility of p_t. Two splits are made, using two standard deviation and using one and a half; (iv) compute the covariance matrix in each of the sub-samples and estimate $\hat{\alpha}$ and $\hat{\beta}$; (v) compare the estimates with the true parameters.

Moreover, in order to indicate the biased obtained, another simulation is run in which the variance indeed increases (only one shock increases its variance by 2). The low and high volatility windows are set using the same procedure as before; split the sample accordingly to the variance of p_t. I use this simulation to compute the magnitude of the small sample bias and contrast it with the previous exercise.

I compare the average deviation between the true parameter and the estimated one in all three exercises. Several parameters were used in the simulations: α and β changed from 0.1 to 0.9. In figure 5, I show the absolute value in the deviation of the estimated α and the true one.[23]

[Figure 5 here]

As can be seen, the biases produced by the split are an order of magnitude larger than those generated by the sample bias. When the data exhibits heteroskedasticity (top-left panel), the maximum distance between the estimates and the true parameter is at most 0.004. Which, in fact, occurs for the maximum absolute values of α and β. On the other hand, in both exercises where the sample was split but the data was homoskedastic, it is frequent to find absolute deviations larger than 1!

It is important to highlight that it is not possible to show analytically that spliting the sample achieves identification. Moreover, it is not possible to show the opposite either: there are some

[23] The results for β are similar and therefore not shown for brevity.

circumstances in which the split in fact produces consistent estimators. This is the reason why in this section only a simulation is presented.

In conclusion, inconsistent estimators are likely to be obtained if the number of regimes assumed is larger than the ones the data truly exhibits. This should be taken, therefore, as a word of caution in empirical implementations.

5.4 Additional methods to define the windows.

As is argued in the previous section, a misspecification in the windows does not affect consistency. In this section, I study two alternative procedures to define the regime windows, and show that the estimates, generally, are not statistically different from those estimated in previous sections. First, I determine the regimes endogenously using a Markov switching model in the spirit of Hamilton [1990]. Second, I estimate the crisis window as those times in which the variance of the residuals is larger than 1.5 times the unconditional variance in the sample. I give most of the attention to the first one, and mainly present the results for the last one.

5.4.1 Switching regime.

In this section, the two regimes are determined endogenously by estimating a Markov switching regime model that allows for changes in variances. The sample is split in the same five subsamples studied before. The first step is to estimate regression (10) and then determine the high and low variance regime by implementing the EM algorithm explained is Hamilton [1990] and in Hamilton [1994], chapter 22.[24]

In table 7, the estimates of the covariance matrices are shown. The top row indicates the subsample considered. The next three rows are the variance covariance estimates in the low variance regime. The second set of three rows are the estimates for the high variance state. The last set of rows compute the increase across regimes.[25]

Note that, in general, the crisis regime is almost always 10 times more volatile than the tranquil state. Using the asymptotic distribution it is straight forward to show that the variance covariance across the states are statistically different.

[24] This step was computed either by forcing the means across regimes to be the same, or by allowing them to change, and by allowing the lags to increase from 5 days to 10 days. The results are qualitatively the same.

[25] The estimates of the means are not shown given that they are not interesting for the purpose of this paper.

Switching Regime	Mexico	Asia	Mexico Asia	Russia	Mexico Asia Russia
Regime of low variance					
Variance Argentina $(\cdot 10^{-4})$	0.594	0.179	0.265	0.316	0.273
Covariance $(\cdot 10^{-4})$	0.341	0.116	0.165	0.247	0.179
Variance Mexico $(\cdot 10^{-4})$	0.343	0.131	0.183	0.255	0.196
Regime of high variance					
Variance Argentina $(\cdot 10^{-4})$	7.871	1.903	4.766	12.420	6.382
Covariance $(\cdot 10^{-4})$	6.038	1.274	3.504	7.510	4.352
Variance Mexico $(\cdot 10^{-4})$	7.449	1.432	4.211	5.245	4.466
Change					
Variance Argentina	12.26	9.62	16.96	38.29	22.39
Covariance	16.71	10.03	20.26	29.46	23.30
Variance Mexico	20.73	9.94	21.99	19.60	21.76

Table 7: Variance Covariance Matrix for several crises and combination of crises. Regime switches are estimated endogenously.

In table 8 the results from testing the rank condition are summarized. Using the quasi t_stats the rank condition is not satisfied during the Asian and All the crises.

Switching Regime	Mexico	Asia	Mexico Asia	Russia	Mexico Asia Russia
Covariance of the Weighted Difference					
Point Estimate $(\cdot 10^{-4})$	1.530	0.118	0.550	2.153	0.228
Standard Deviation $(\cdot 10^{-4})$	0.578	0.088	0.219	0.734	0.168
quasi t_stat	2.65	**1.34**	2.51	2.93	**1.36**

Table 8: Rank Condition for several crises. Regimes are estimated with Markov Switching.

The distributions of $\hat{\alpha}$ and $\hat{\beta}$ are also determined by the bootstrapping. In table 9 these results are summarized.

Switching Regime	Mexico	Asia	Mexico Asia	Russia	Mexico Asia Russia
$\hat{\alpha}$					
Point Estimate	0.321	0.249	0.236	0.548	0.329
Standard Deviation	0.110	0.451	0.187	0.052	0.511
Quasi T_stat	2.92	0.55	1.27	10.53	0.64
$\hat{\beta}$					
Point Estimate	0.619	0.459	0.667	0.607	0.240
Standard Deviation	0.149	0.538	0.231	0.104	0.704
Quasi T_stat	4.17	0.85	2.89	5.83	0.34

Table 9: Estimation of the structural parameters for several crises.

As can be seen, during the Mexican crisis, both $\hat{\alpha}$ and $\hat{\beta}$ are statistically different from zero, while during the Asian crises the estimates are not. This confirms the results from table 8.[26] For

[26] The coefficients are "well" estimated during the Mexican, Mexico-Asia, and the Russian crises. The rejection

the Russian crisis the estimates arehighly significant. Note that the point estimates are close across the different windows. Moreover, comparisons between the estimates using the exogenous windows and the switching regime are preformed in table 10. Because it is not clear how these estimators are related, in order to perform the test, I assume the correlation between the estimates that maximizes the possibilities of rejection.[27] In other words, I choose the correlation that minimizes the standard deviation of the difference in the estimators. As can be seen, the hypothesis that the estimates are the same across the two methodologies determining the windows cannot be rejected.

	$\hat{\alpha}$		$\hat{\beta}$	
	Difference	Smallest Standard Deviation	Difference	Smallest Standard Deviation
Mexico	0.004099	0.037186	0.028286	0.035334
Asia	0.058056	0.376488	0.199158	0.490981
Mexico-Asia	0.077489	0.107759	0.017557	0.170172
Russia	0.163203	0.499199	0.143263	0.139875
All	0.020506	0.423731	0.409088	0.637265

Table 10: Exogenous vs. Regime switching windows.

5.4.2 Regimes as large changes in second moments.

The second exercise is to define the crisis windows as large changes in the conditional second moments. The procedure is as follows: First, the unconditional covariance matrix is computed from the residuals obtained from the first step. Second, the conditional variances and covariances are determined along rolling windows of 20 days. Third, the crisis periods are defined as those instances in which the conditional variances, or covariance, (any one of the three) are larger than the mean values (the unconditional) plus 1.5 times its standard deviation.

In figure 4, the rolling window covariance together with the crises windows are shown. Notice that the Asian crises are not included given that they represented a relatively small shift in the variances, as it has being discussed before.

[Figure 4 here]

from zero is not necessarily because the means are close to zero, but because the standard deviation of the estimates are almost 4 times larger than those obtained during the Mexican crisis.

[27] Computing the standard deviation in this way assures that if there is no rejection no test can be constructed such that the parameters are statistically different. Note that this variance is smaller than the one implied by a Hausman Specification Test.

In this section, I only explore the example where all the sample is used. Again, a VAR regression is run first, and then the windows are defined by splitting the sample between high and low variance of the residuals.[28] In table 11, the values of $\hat{\alpha}$ and $\hat{\beta}$, as well as their standard deviations are shown.

Endogenous Windows	$\hat{\alpha}$	$\hat{\beta}$
Point Estimate	0.315000	0.724991
Standard Deviation	0.146255	0.431501
T _stat	2.153771	1.680160

Table 11: Estimation of the structural parameters for the three crises: Mexican, South East Asian, and Russian crises. Crises windows are defined as 1.5 standard deviation increases in the conditional second moments.

The tests to determine if the estimates are statistically different from the previous two methodologies are shown in table 12. For the $\hat{\alpha}$ estimates, there are two cases in which the test is rejected; with the Mexican crisis using exogenous windows, and with the Russian crisis estimated using the switching regime. For the $\hat{\beta}$ estimates, there is only one case when the test is rejected; the Asian crises estimated with the switching regime. Remember that these rejections are obtained using the correlation that minimizes the variance, thus they should be taken cautiously.

	$\hat{\alpha}$		$\hat{\beta}$	
	Difference	Smallest Standard Deviation	Difference	Smallest Standard Deviation
Exogenous Windows				
Mexico	**0.009665**	**0.000767**	0.077117	0.318229
Asia	0.123938	0.071575	0.065906	0.384737
Mexico-Asia	0.001118	0.067193	0.074978	0.370618
Russia	0.395935	0.404938	0.260837	0.187596
All	0.006166	0.059143	0.075178	0.364562
Switching Regime				
Mexico	0.005566	0.03642	0.105403	0.282894
Asia	0.065882	0.304913	**0.265064**	**0.106244**
Mexico-Asia	0.078607	0.040566	0.057421	0.200447
Russia	**0.232732**	**0.094261**	0.117574	0.327471
All	0.01434	0.364588	0.484266	0.272703

Table 12: Tests to determine if the estimates are different among methodologies defining the windows. Endogenous vs. Exogenous windows and Regime switching.

In conclusion, it is fair to say that the estimates are remarkably stable across samples and methodologies defining the windows. This should be surprising taken into consideration that the estimates in this section are not even considering the Asian crises as part of the high volatility regime.

[28] For reasons of brevety, I do not report the estimates of the variance covariance matrix, nor the rank condition. It was satisfied.

6 Conclusions and extensions

This paper develops a new identification assumption that solves the problem of simultaneous equations, based on the heteroskedasticity of the structural shocks.

The paper presents three important results: First, proposition 1, states the circumstances where the order and rank conditions for identification are obtained. In summary, the order condition requires four assumptions: firstly, normalization; secondly, independence of the structural shocks; thirdly, the stability of the structural parameters (α and β in my model); and finally, the existence of at least two regimes of different variances. As is discussed in the paper, several macroeconomic and finance applications satisfy these conditions. These assumptions might look restrictive, but for example, in all macro applications in which VAR and recursive identifications are used, these assumptions have been already imposed. Moreover, GARCH specifications usually make the same set of assumptions. The identification method based on heteroskedasticity, therefore, can be used in a broad set of circumstances. General conditions for identification are discussed in section 4.

Second, in section 5 it is shown that consistency is achieved even if the heteroskedasticity is incorrectly specified. If the original data exhibits heteroskedasticity, the procedure achieves identification and consistent estimators if the rank condition is satisfied by the misspecified model. It is crucial, though, that the data exhibits some form of heteroskedasticity, otherwise, inconsistent estimates are likely to be obtained.

Third, the paper discusses the case in which there are more regimes than the ones required for just identification. These equations are used to construct a specification test. In the application developed in the paper, the overidentifying restrictions were not rejected.

Two future avenues of research are worth mentioning: (i) applying similar techniques to non linear models might improve the traditional methodologies that depend on the non-linearities to solve the problem of identification[29]; (ii) developing a Bayesian approach to identification.[30]

Finally, it is important to notice that the assumption of heteroskedasticity differs from the standard identification assumptions used in the literature in the sense that it can be tested in

[29] See Klein and Vella [2000a, b] as well as Sentana and Fiorentini [1999] for similar identification procedures in those models.

[30] The present methodology achieves identification only under the alternative hypothesis that heteroskedasticity exists. A Bayesian approach to identification can therefore be implemented if there is uncertainty about the existence of heteroskedasticity

a reduce form. The standard identification assumptions (usually parameter restrictions) have no implications on the reduce form. However, the existence of heteroskedasticity in the structural equations implies that the reduce form is also heteroskedastic. Therefore, it can be tested.

References

Blanchard, O. and Quah, D. (1989). The dynamic effects of aggregate demand and aggregate supply disturbances. *American Economic Review*, 79:655–73.

Chen, S. and Khan, S. (1999). $\sqrt{n}$-consistent estimation of heteroskedastic sample selection models. *University of Rochester, Mimeo.*

Fisher, F. M. (1976). *The Identification Problem in Econometrics.* Robert E. Krieger Publishing Co., New York, second edition.

Forbes, K. and Rigobon, R. (1999). Measuring contagion: Conceptual and empirical issues. *MIT mimeo.*

Haavelmo, T. (1947). Methods of measuring the marginal propensity to consume. *Journal of the American Statistical Association*, 42:105–122.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45:39–70.

Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press, Princeton, New Jersey.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46:1251–72.

Klein, R. and Vella, F. (2000a). Employing heteroskedasticity to identify and estimate triangular semiparametric models. *Rutgers mimeo.*

Klein, R. and Vella, F. (2000b). Identification and estimation of the binary treatment model under heteroskedasticity. *Rutgers mimeo.*

Koopmans, T., Rubin, H., and Leipnik, R. (1950). *Measuring the Equation Systems of Dynamic Economics*, volume Statistical Inference in Dynamic Economic Models of *Cowles Commission for Research in Economics*, chapter II, pages 53–237. John Wiley and Sons, New York.

Rigobon, R. (1999). On the measurement of the international propagation of shocks. *MIT Mimeo.*

Sentana, E. and Fiorentini, G. (1999). Identification, estimation and testing of conditional heteroskedastic factor models. *CEMFI mimeo.*

Shapiro, M. D. and Watson, M. W. (1988). *Sources of Business Cycle Fluctuations.* MIT Press, Cambridge, Mass.

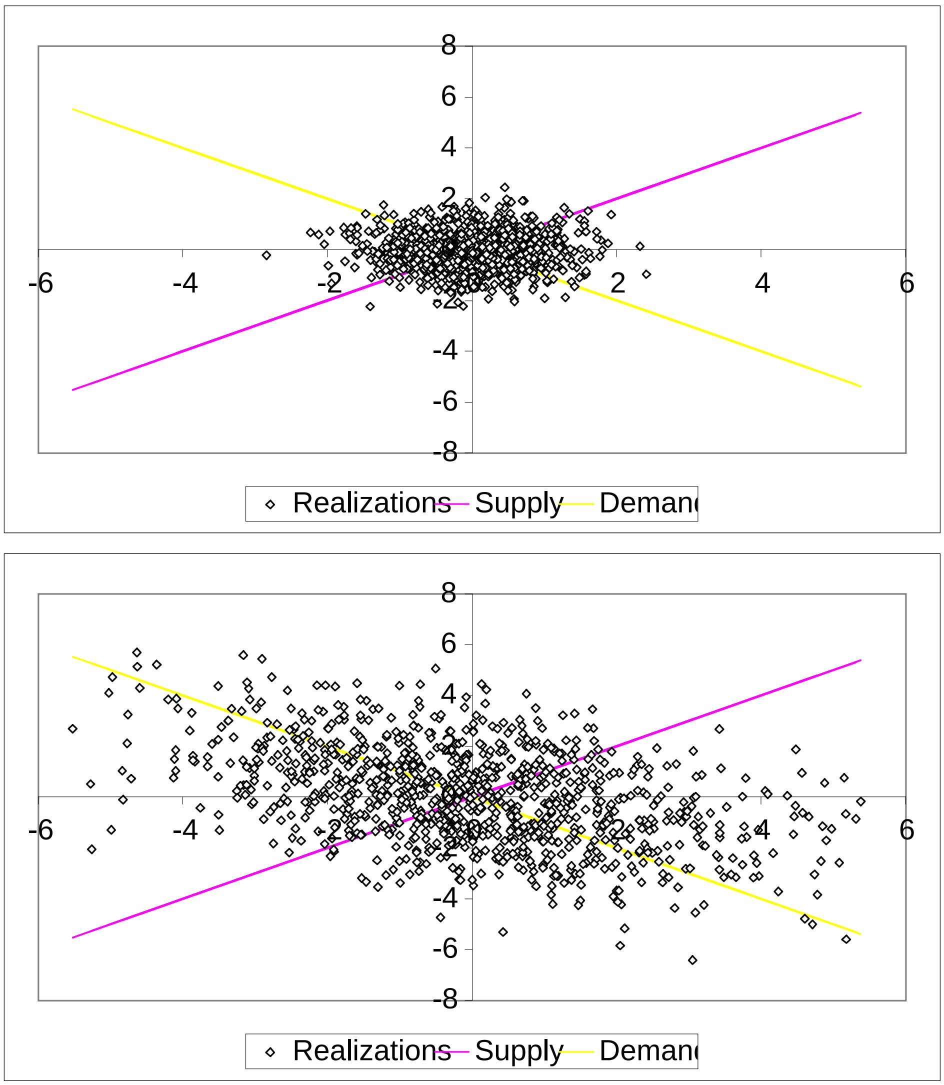

Figure 1: Identification Problem.

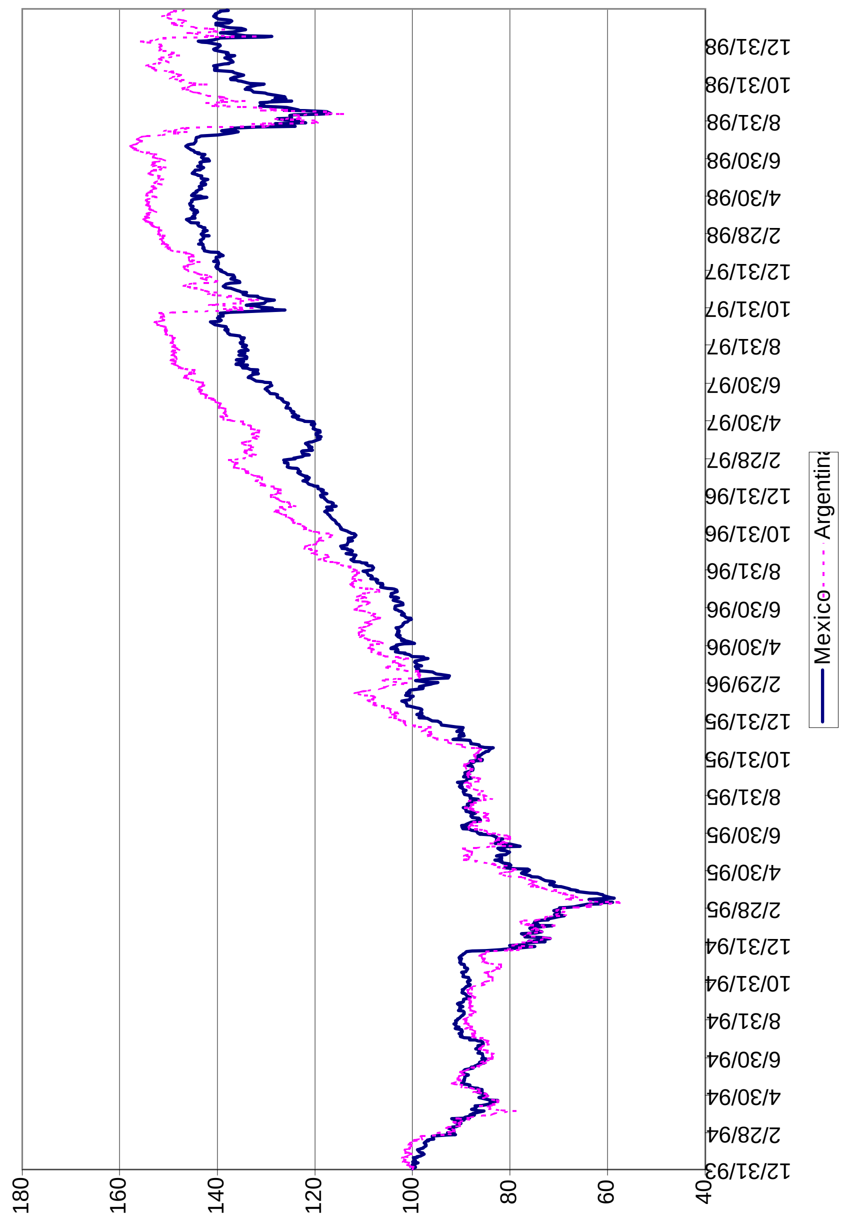

Figure 2: Argentina and Mexico Brady Index. Source JPMorgan.

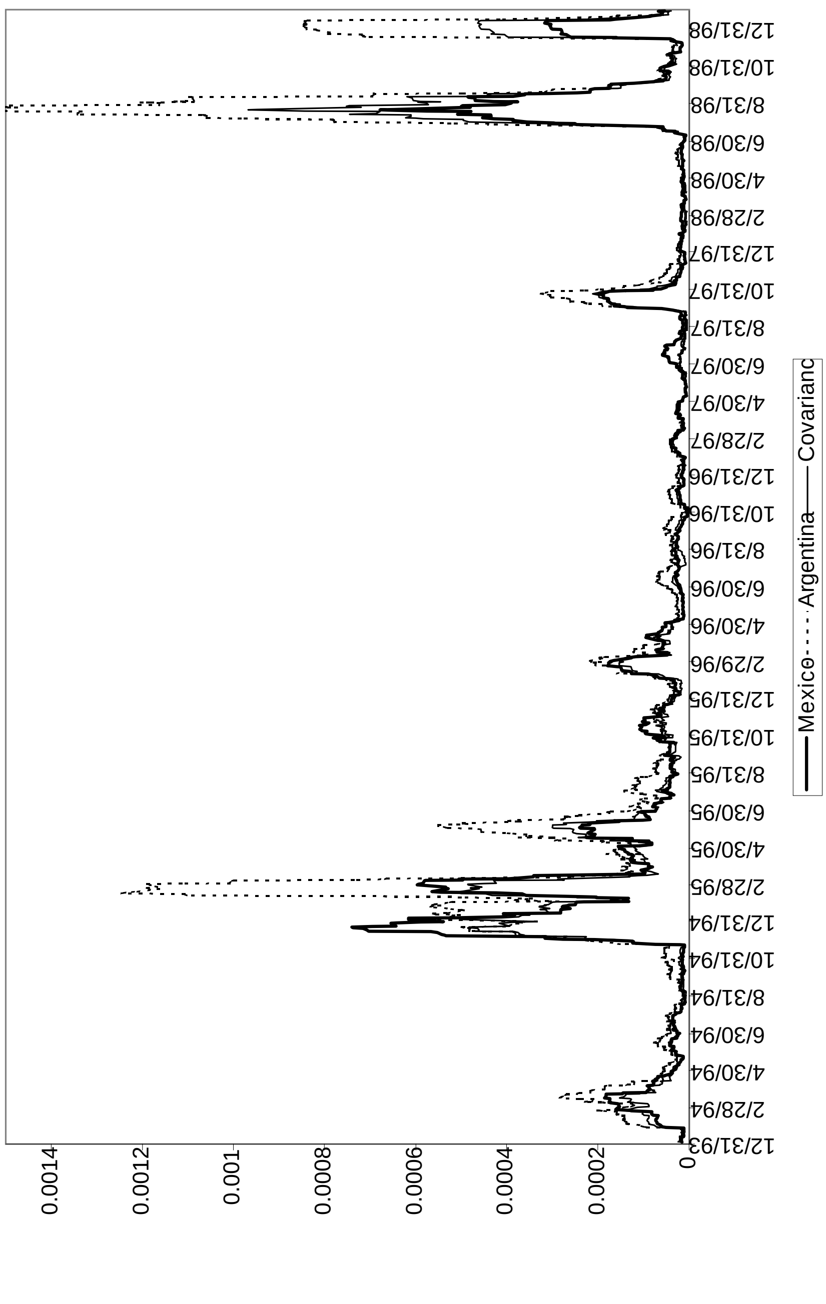

Figure 3: Rolling variances and covariance of daily returns on Argentinean and Mexican Brady Bonds. Window = 20 days.

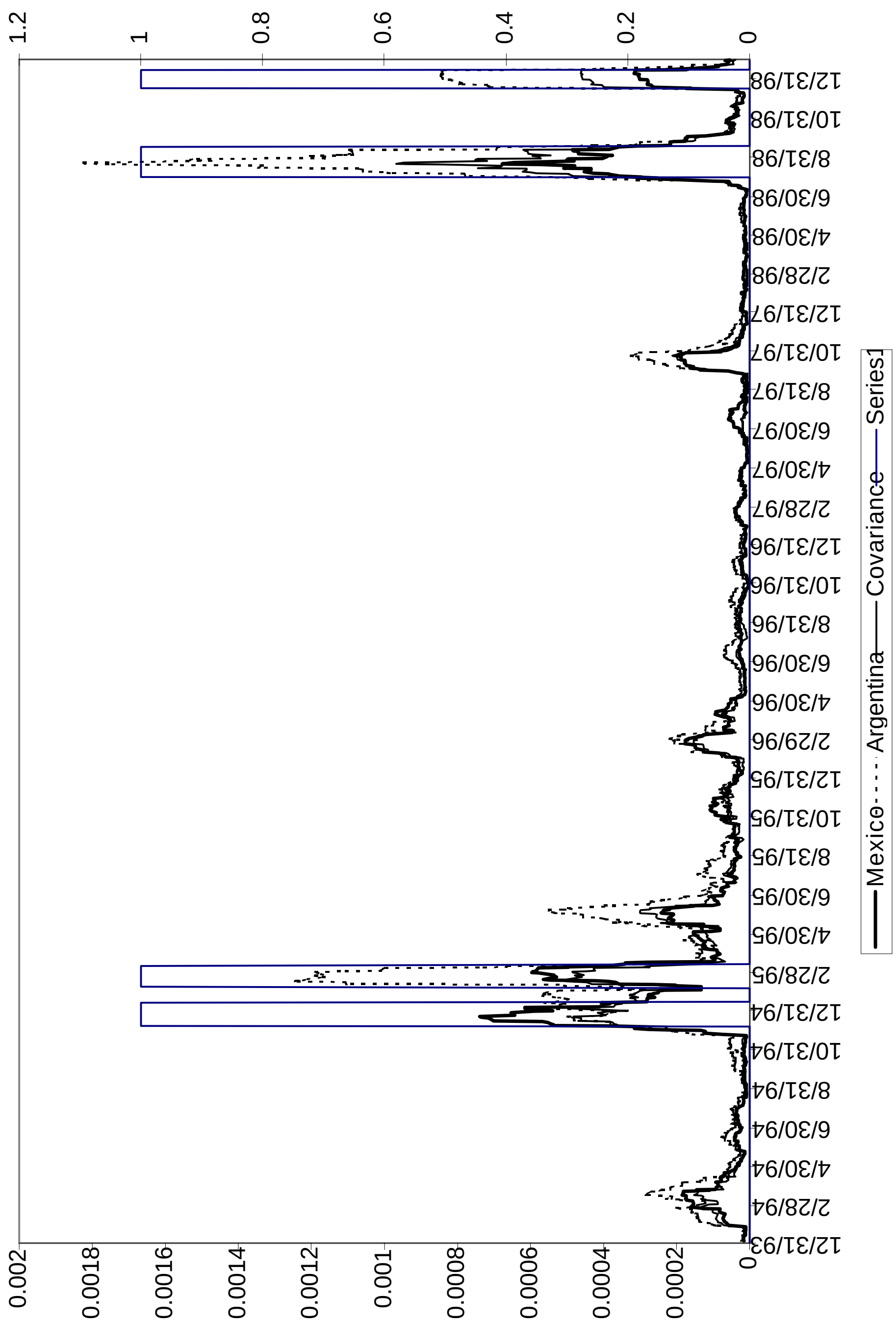

Figure 4: Crisis window defined as large shifts in second moments.

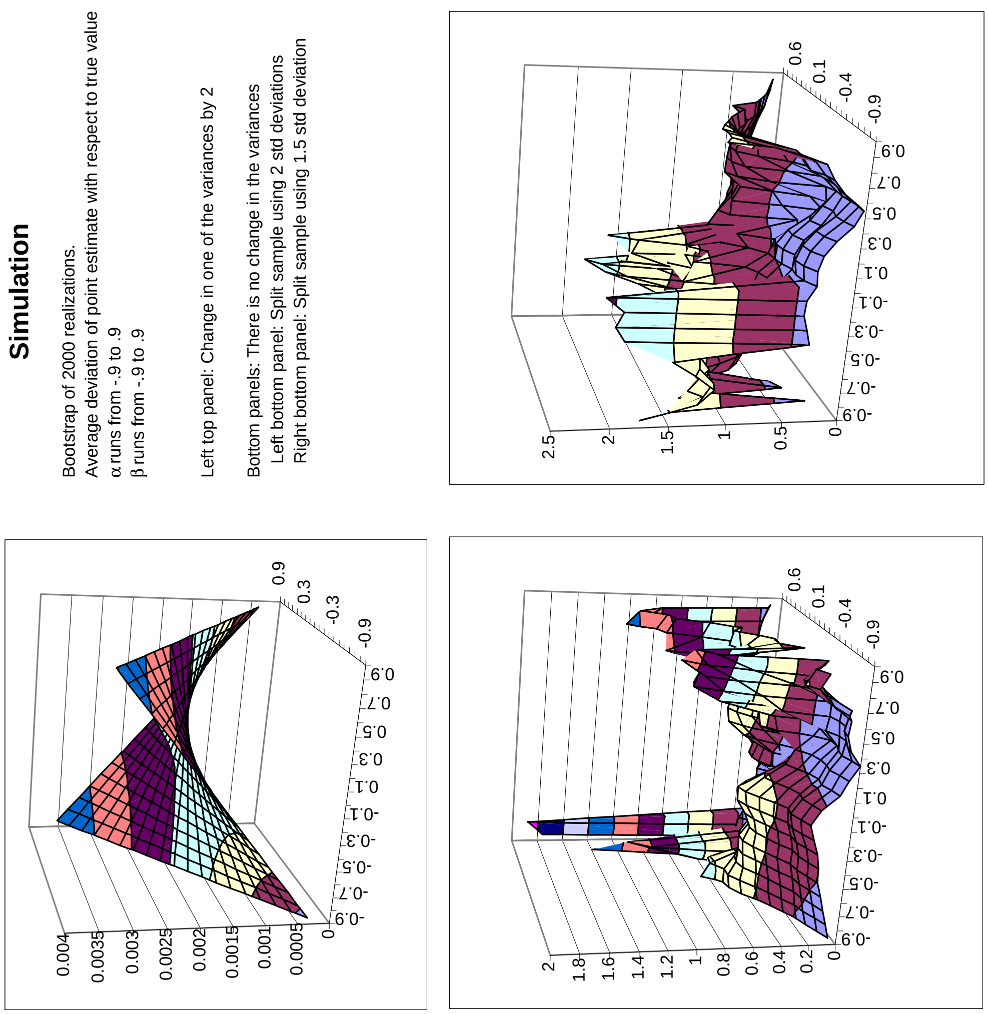

Figure 5: Evaluating the consistency of α when heteroskedasticity does not exists.