Econometric Theory, **33**, 2017, 739–754.
doi:10.1017/S0266466615000353

IDENTIFYING RESTRICTIONS FOR FINITE PARAMETER CONTINUOUS TIME MODELS WITH DISCRETE TIME DATA

Jason R. Blevins
Ohio State University

This paper revisits the question of parameter identification when a linear continuous time model is sampled only at equispaced points in time. Following the framework and assumptions of Phillips (1973), we consider models characterized by first-order, linear systems of stochastic differential equations and use a priori restrictions on the model parameters as identifying restrictions. A practical rank condition is derived to test whether any particular collection of at least $\lfloor n/2 \rfloor$ general linear restrictions on the parameter matrix is sufficient for identification. We then consider extensions to incorporate prior restrictions on the covariance matrix of the disturbances, to identify the covariance matrix itself, and to address identification in models with cointegration.

1. IDENTIFICATION AND ALIASING IN CONTINUOUS TIME MODELS

This paper develops identification results for first-order, linear systems of stochastic differential equations of the form

$$\mathrm{d}y(t) = Ay(t)\,\mathrm{d}t + \zeta(\mathrm{d}t) \tag{1}$$

in the case where observations of $y(t)$ are only available at discrete, equispaced points in time separated by intervals of length h. Here, $y(t)$ is a stationary $n \times 1$ random vector, A is a parameter matrix whose elements are real numbers, and $\zeta(\mathrm{d}t)$ is a vector of white noise innovations with covariance matrix $\Sigma\,\mathrm{d}t$. The exact discrete time process $y_t = y(th)$ corresponding to the continuous time system in (1) is the vector autoregressive (VAR) process

$$y_t = By_{t-1} + \varepsilon_t, \tag{2}$$

$$B = \exp(hA) \equiv \sum_{j=0}^{\infty} \frac{(hA)^j}{j!} = I + hA + \frac{(hA)^2}{2!} + \frac{(hA)^3}{3!} + \dots, \tag{3}$$

I am grateful to John Geweke, Robert de Jong, Matt Masten, Tucker McElroy, Peter Phillips, and three anonymous referees for useful comments and discussions. Address correspondence to Jason R. Blevins Ohio State University, Department of Economics, 1945 N High St., 410 Arps Hall, Columbus, OH 43210, USA; e-mail: blevins.141@osu.edu

where $\varepsilon_t = \int_0^h \exp(sA)\zeta(th - \mathrm{d}s)$ is a serially independent process with covariance matrix $\Omega = \int_0^h \exp(sA)\Sigma\exp(sA^\top)\,\mathrm{d}s$. See McCrorie (2000) for a formal derivation.

We approach the identification problem from the same viewpoint as Phillips (1973), that economic models can provide natural identifying restrictions in the form of restrictions on the parameter matrix A which can serve to rule out otherwise observationally equivalent parameter matrices. He noted that in the absence of such information there are a countably infinite number of other matrices in $\mathbb{R}^{n\times n}$ that are observationally equivalent to A in discrete time. Hansen and Sargent (1983) showed that there was identifying information arising from the positive definiteness of the covariance matrix Σ, but still A is unidentified in general without restrictions.

The fundamental difficulty is that the matrix equation (3) does not, in general, admit a unique solution (Coddington and Levinson, 1955, Ch. 3.1). In other words, the matrix exponential is not necessarily injective. However, for identification purposes we do not need uniqueness *per se*—or even a unique real solution—but only uniqueness within some set $\mathcal{A} \subseteq \mathbb{R}^{n\times n}$ of matrices that satisfy our assumptions and prior restrictions (see, e.g., Koopmans, 1949; Hurwicz, 1950). Failing to rule out multiple observationally equivalent structural matrices A results in an incompleteness which may yield uninterpretable estimates. Hamerle, Nagl, and Singer (1991) accordingly described this identification problem as “essential” for dynamic models in the social sciences but concluded that it had largely been ignored.

Our results build primarily on the work of Phillips (1973). He showed that in an $n \times n$ system, as few as $\lfloor n/2 \rfloor$ linear restrictions on A could be sufficient for identifying A and illustrated this in an example with $n = 3$, which was also discussed in Phillips (1972), where he showed that a single zero restriction on an element of A was sufficient to identify the entire matrix A.[1] Then, in two theorems he established formal rank conditions under which, for general $n \times n$ systems, homogeneous linear restrictions on a particular row or column of A could be used to identify that row or column of A.

This paper reinforces and extends the results of Phillips (1973) in several directions. First, in Section 2, we revisit the problem under the same assumptions and derive a practical rank condition involving only identified quantities that can be used to determine whether any particular $\lfloor n/2 \rfloor$ linear restrictions on A of a fully general form are sufficient for identification. We illustrate this result in an example with $n = 3$ and then consider an extension to models with cointegration. Then, in Section 3, we show that restrictions on the covariance matrix Σ can serve as identifying conditions for A and we consider joint identification of A and Σ.

As a practical matter, we note that in applications the coefficient and covariance matrices are typically functions of some underlying, lower-dimensional parameters, say $A(\theta)$ and $\Sigma(\theta, \mu)$. We focus on identification of A and Σ themselves, but one could also consider identification of (θ, μ) directly, which may either simplify or complicate the problem.

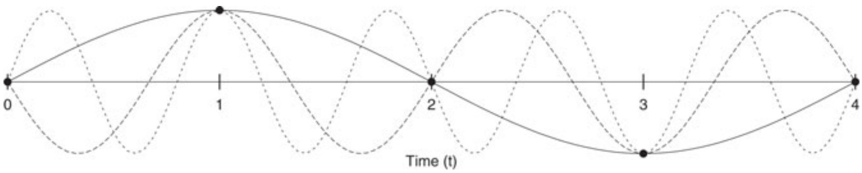

FIGURE 1. The Aliasing Problem in the Frequency Domain.
Plots of $f(t;m) = \sin(2\pi mt)$ for $t \in [0,4]$ and $m = 1/4$ (solid), $m = -3/4$ (long dash), and $m = 5/4$ (short dash).

The identification problem we address is also known as the aliasing problem and is perhaps most obvious in the frequency domain. Figure 1 depicts three sine waves which have different frequencies but are observationally equivalent when sampled at discrete intervals of length $h = 1$. A related phenomenon occurs in models characterized by (1), where multiple matrices A may be observationally equivalent when sampled at discrete intervals of length h. Whether this is the case depends critically on the spectrum of the parameter matrix A.

An early interest in continuous time regression models (Phillips, 1970, 1972; Sims, 1971; Geweke, 1978) provided the initial motivation for studying this identification problem (Phillips, 1973; Hansen and Sargent, 1983). The question remains relevant due to continued interest in continuous time models in economics, including recent theoretical and empirical work involving continuous time games (Doraszelski and Judd, 2012; Arcidiacono, Bayer, Blevins, and Ellickson, 2012; Schiraldi, Smith, and Takahashi, 2012). Our results are also relevant to applications in many other fields such as macroeconomics (Bergstrom, 1988; Bergstrom and Nowman, 2007), finance (Baxter and Rennie, 1996), and input-output analysis (Sinha and Lastman, 1982). See Yu (2014) for a survey of work by Phillips involving continuous time models in a variety of settings.

There are several other known sufficient conditions for identification in the general model and in some special cases that take the form of identifying restrictions on the matrix B (Culver, 1966; Cuthbert, 1972, 1973) and alternative or irregularly spaced sampling schemes (Cuthbert, 1973; Singer and Spilerman, 1976; Hansen and Sargent, 1983). However, even in light of these existing results there is significant scope for expanding the set of known identifying restrictions. We briefly outline some other known sufficient conditions for identification below. In the interest of brevity, we refer readers interested in estimation methods to recent surveys by Sørensen (2004), Fan (2005), Aït-Sahalia (2007), Bandi and Phillips (2009), McCrorie (2009), and Phillips and Yu (2009).

Mathematical treatments of the aliasing problem have taken a "top down" approach and focused on finding conditions on B that are sufficient for identification of A. For example, it follows from results of Culver (1966) that if all eigenvalues of $B = (b_{ij})$ are positive real numbers and no Jordan block belonging to any eigenvalue is repeated, then A is identified. In the special case where B is a discrete time Markov transition matrix corresponding to an infinitesimal generator matrix A, other sufficient conditions are $\min_i b_{ii} > 1/2$ (Cuthbert, 1972) and $\det(B) > e^{-\pi}$ (Cuthbert, 1973). However, in cases where an underlying economic model is defined from the "bottom up", it may be difficult to determine when such conditions on the matrix B are consistent with the assumptions and implications of the economic model.

There are also two important results that are useful when the researcher can control the sampling interval h. It is well known that for every parameter matrix A, there exists an interval $\overline{h}$ such that A is identified when $h \leq \overline{h}$ (Cuthbert, 1973; Singer and Spilerman, 1976; Hansen and Sargent, 1983). Furthermore, if the discrete time process can be sampled at two intervals h_1 and h_2, where $h_2 \neq h_1 k$ for some integer k, then A is identified (Singer and Spilerman, 1976, 5.1). Unfortunately, the applicability of these results is limited: the researcher may not have sufficient a priori knowledge about the matrix A to determine the value of $\overline{h}$ and in many studies the sampling frequency is pre-determined and fixed (e.g., quarterly or annual) and cannot be chosen by the researcher.

The cases that remain are important and frequently encountered and they are the focus of the remainder of this paper. Researchers routinely work with datasets for which they have no control over how the observations are sampled. When n is moderately large or when the known restrictions on A are inhomogeneous or involve restrictions across rows or columns of A, existing conditions for identification may not be practical or even applicable. Establishing new conditions for identification is therefore important for a large subset of empirical studies using continuous time models. In the following, we derive conditions for identification via restrictions on the entire matrix A and show how to apply them in the contexts of a simple continuous time regression model.

2. IDENTIFICATION VIA LINEAR RESTRICTIONS ON THE MATRIX A

Let $\mathcal{A} \subseteq \mathbb{R}^{n\times n}$ denote the parameter space of admissible matrices A that satisfy our maintained assumptions, stated below, as well as a collection of linear prior restrictions on A, to be chosen by the researcher. Let $A^0 \in \mathcal{A}$ denote the parameter matrix which generated the observable data. The starting point for our analysis is the population matrix $B^0 = \exp(hA^0)$ from the discrete time model in (3), which is identified from discrete time observations $\{y_t\}$.

DEFINITION. *A^0 is* identified *in $\mathcal{A}$ if A^0 is the unique solution to (3) in $\mathcal{A}$ for $B = B^0$.*

To show that A^0 is identified, we must establish that the assumptions and prior restrictions that define $\mathcal{A}$ are such that the matrix exponential equation $B^0 = \exp(hA)$ has a unique solution $A = A^0$ in $\mathcal{A}$. We proceed using the same framework and two main assumptions as Phillips (1973).

Assumption 1. The matrix A^0 has distinct eigenvalues, all of which have negative real parts.

Assumption 2. The eigenvalues of the matrix A^0 do not differ by an integer multiple of $2\pi i/h$.

Assumption 1 is sufficient for A^0 to be diagonalizable and we exploit this in our proofs. The second part of the assumption, that the eigenvalues have negative real parts, corresponds to a stationarity assumption. This assumption rules out models with cointegration, where A^0 has a repeated eigenvalue equal to zero, as well as models where A^0 is a scalar multiple of the identity matrix. Although stationary models are the main focus of this section, we also consider cointegration briefly in Section 2.2.

Assumption 2 only rules out models where two or more eigenvalues are congruent modulo $2\pi i/h$. Our proofs make use of the particular structure of the spectral decomposition of alternate solutions $A \neq A^0$ under this assumption. We note that the set of matrices A^0 in $\mathbb{R}^{n\times n}$ which do not satisfy Assumptions 1 and 2 is negligible.[2]

Even under these assumptions, A^0 is not identified in general without additional restrictions. Phillips (1973) demonstrated that $\lfloor n/2 \rfloor$ linear restrictions on A could be sufficient for identifying A. His Theorems 1 and 2 establish rank conditions under which $n-1$ linear, homogeneous restrictions on a single row or column of A^0 are sufficient to identify that row or column. Specifically, these restrictions were of the form $R_i a_i^0 = 0$ where $(a_i^0)^\top$ is the i-th row of A^0 and R_i is a $k \times n$ matrix that imposes k restrictions on a_i^0. Analogous restrictions can be imposed on the columns of A^0. He showed that if the matrix $R_i(A^0)^\top$ has rank $n-1$, then a_i^0 is identified and that if $n-1$ rows or columns of A^0 are identified, then the remaining row or column is also identified.

We derive a practical rank condition for establishing identification of the entire matrix $A^0 \in \mathcal{A}$ with $\lfloor n/2 \rfloor$ or more linear restrictions which can be inhomogeneous and can involve elements of A^0 in arbitrary rows and columns. Formally, we define $\mathcal{A}$ to be the set of $n \times n$ matrices A^0 that satisfy Assumptions 1 and 2 and a collection of k linear restrictions of the form

$$R\,\mathrm{vec}(A^0) = r, \qquad \textbf{(4)}$$

where R is a $k \times n^2$ matrix, r is a $k \times 1$ vector, and $\mathrm{vec}(A^0)$ denotes the vector obtained by stacking the columns of A^0. This includes all possible linear, (potentially) inhomogeneous restrictions on A^0.[3]

Our approach is to restate the identification problem in terms of a complete system of linear equations in n^2 unknowns—based on the vectorization of A^0—which allows us to precisely characterize a rank condition that is sufficient for identification. At least $\lfloor n/2 \rfloor$ equations are implications of the prior restrictions on A^0 in (4), provided by the researcher. The remaining equations are derived from known properties of the inverse mapping from B^0 to $\mathcal{A}$, which we turn to now.

Phillips (1973) showed that under Assumptions 1 and 2 any real solution A is related to A^0 by

$$A = V(\Lambda + D)V^{-1} = A^0 + VDV^{-1}, \qquad \textbf{(5)}$$

where Λ is a diagonal matrix containing the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A^0, V is a nonsingular matrix whose columns are the eigenvectors of A^0, and D is a diagonal matrix of the form

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & -M \end{bmatrix}, \tag{6}$$

with $M = \text{diag}(m) = \text{diag}(m_1, \ldots, m_\rho)$, $m_i = \frac{2\pi i}{h}\tilde{m}_i$, and $\tilde{m}_i \in \mathbb{Z}$ for $i = 1, \ldots, \rho$ where 2ρ is the number of complex eigenvalues[4] of A^0. Without additional restrictions on A^0, there may be multiple real solutions A with different complex eigenvalues, corresponding to matrices $M \neq 0$. Yet, it is apparent from the representation in (5) and (6) that all possible real solutions A have the same eigenvectors and real eigenvalues, meaning that those features are identified.

LEMMA 1. *If Assumptions 1 and 2 hold, then V, ρ, and all real eigenvalues $\lambda_1, \ldots, \lambda_{n-2\rho}$ of A^0 are identified, where V is the matrix of eigenvectors of A^0 and 2ρ is the number of complex eigenvalues of A^0. Furthermore, if $\rho = 0$, then A^0 is identified.*

Any matrix $A \in \mathcal{A}$ that satisfies $\exp(hA) = B^0$ must also satisfy the restrictions in (4), which implies $R\text{vec}(VDV^{-1}) = 0$. Therefore, the restrictions on A^0 imply

$$R(U \otimes V)\text{vec}(D) = 0, \tag{7}$$

where $U = V^{-\top}$ denotes the transpose of V^{-1}, $U \otimes V$ is a nonsingular $n^2 \times n^2$ matrix, $\text{vec}(D)$ is an $n^2 \times 1$ vector of unknowns, and hence 0 is the $k \times 1$ vector of zeros. This system is necessarily rank deficient unless there are $k \geq n^2$ restrictions. However, we can also use the structure of D in (6) to complete the system in (7). Many of the elements of D are zeros and the nonzero elements can be determined from the ρ complex eigenvalues and their complex conjugates. The following Lemma establishes that ρ restrictions can indeed be sufficient for identification of A^0. The proof, along with all subsequent proofs, appears in the Appendix.

The Lemma makes use of an $n^2 \times n^2$ permutation matrix P defined so that

$$P\,\text{vec}(D) = \left[-m_\rho \ldots -m_1 \; m_\rho \ldots m_1 \; 0 \ldots 0\right]^\top. \tag{8}$$

In other words, the first n rows of P are the vectors $e_{n^2}^\top, e_{(n-1)n-1}^\top, e_{(n-2)n-2}^\top, \ldots, e_1^\top$, where $e_i^\top$ denotes row i of the $n^2 \times n^2$ identity matrix. Since the locations of the complex eigenvalues in Λ and $\Lambda + D$ agree, this definition of P ensures that the first 2ρ elements of $P\text{vec}(\Lambda)$ are the 2ρ complex eigenvalues of A^0, which are followed by the $n - 2\rho$ real eigenvalues, which are in turn followed by the remaining $n(n-1)$ elements of Λ, which are all zeros, in any order.

LEMMA 2. *Suppose Assumptions 1 and 2 are satisfied and define the partitioned matrix*[5]

$$\Phi = \left[\Phi_{1:2\rho} \; \Phi_{2\rho+1:n^2}\right] \equiv \begin{bmatrix} R(U \otimes V)P^\top \\ I_\rho \; I_\rho \; 0_{n^2-2\rho} \end{bmatrix},$$

https://doi.org/10.1017/S0266466615000353 Published online by Cambridge University Press

where P is a permutation matrix defined as in (8). If rank$(\Phi_{1:2\rho}) = 2\rho$, where 2ρ is the number of complex eigenvalues of A^0, then A^0 is identified.

Remark. By Assumption 1, the matrix of eigenvectors V and the number of complex eigenvalues 2ρ are identified and therefore all quantities needed to check the rank condition in Lemma 2 are identified.

The following theorem gives a sufficient rank condition in terms of $\delta = \lfloor n/2 \rfloor$ rather than ρ. The condition may be stronger than necessary, but is still relatively mild; the main benefit is that the condition may be more useful in practice than Lemma 2 because it does not require determining ρ in advance.

THEOREM 1. *Suppose Assumptions 1 and 2 are satisfied and define the partitioned matrix*

$$\Psi = \left[\Psi_{1:2\delta}\ \Psi_{2\delta+1:n^2}\right] \equiv \begin{bmatrix} R(U \otimes V)P^\top \\ I_\delta\ I_\delta\ 0 \end{bmatrix},$$

where P is a permutation matrix defined as in (8). If rank$(\Psi_{1:2\delta}) = 2\delta$, then A^0 is identified.

Remark. As noted in an example of Phillips (1973, p. 357), the necessary conditions alone are enough to identify all structures except those in a set of zero Lebesgue measure. This intuition extends directly to the general case. In other words, role of the rank condition on the matrix Ψ in Theorem 1 (and Φ in Lemma 2) is to rule out a measure-zero subset of $\mathcal{A}$ for which A^0 is not identified.

Therefore, the model is generically identified (i.e., for almost every $A^0 \in \mathcal{A}$) even without the rank condition provided that rank$(R) \geq \delta$.

2.1. Example: A Simple Continuous Time Regression Model

As an example, consider the system (1) in the $n = 3$ case:

$$\begin{bmatrix} \mathrm{d}y_1(t) \\ \mathrm{d}y_2(t) \\ \mathrm{d}y_3(t) \end{bmatrix} = \begin{bmatrix} a_{11}^0\ a_{12}^0\ a_{13}^0 \\ a_{21}^0\ a_{22}^0\ a_{23}^0 \\ a_{31}^0\ a_{32}^0\ a_{33}^0 \end{bmatrix} \begin{bmatrix} y_1(t)\mathrm{d}t \\ y_2(t)\mathrm{d}t \\ y_3(t)\mathrm{d}t \end{bmatrix} + \begin{bmatrix} \zeta_1(\mathrm{d}t) \\ \zeta_2(\mathrm{d}t) \\ \zeta_3(\mathrm{d}t) \end{bmatrix}.$$

The matrix $A^0 = (a_{ij}^0)$ contains 9 parameters. In this three-equation case Phillips (1972, 1973) showed that a single zero restriction on one of the elements A^0 could be sufficient to identify the remaining parameters. We revisit this example to illustrate how to apply the conditions of Theorem 1 to establish identification, leading to a condition that is identical to the one derived by (Phillips, 1973, pp. 357–358).

First, we note that

$$\mathrm{vec}(A^0) = \left[a_{11}^0\ a_{21}^0\ a_{31}^0 | a_{12}^0\ a_{22}^0\ a_{32}^0 | a_{13}^0\ a_{23}^0\ a_{33}^0\right]^\top.$$

We consider the single homogeneous restriction characterized by

$$R = \left[0\,0\,0 | 0\,0\,0 | 1\,0\,0\right] \quad \text{and} \quad r = \left[0\right],$$

https://doi.org/10.1017/S0266466615000353 Published online by Cambridge University Press

which restricts a^0_{13} to be zero (i.e., it excludes $y_3(t)$ from the equation for $dy_1(t)$).

Recall for any observationally equivalent matrix $A \neq A^0$, $A = A^0 + VDV^{-1}$ with

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & m_1 & 0 \\ 0 & 0 & -m_1 \end{bmatrix}.$$

For the example restriction matrix R given above, we have

$$R(U \otimes V) = \left[u_{31}v_{11}\ u_{31}v_{12}\ u_{31}v_{13} \middle| u_{32}v_{11}\ u_{32}v_{12}\ u_{32}v_{13} \middle| u_{33}v_{11}\ u_{33}v_{12}\ u_{33}v_{13} \right],$$

where $U = (u_{ij})$ and $V = (v_{ij})$. The permutation matrix P places the $(3,3)$, $(2,2)$, and $(1,1)$ elements of D first. These correspond to elements 9, 5, and 1 of $\text{vec}(D)$, so the permutation matrix is of the form $P^\top = \left[e_9\ e_5\ e_1 \middle| \dots \right]$.

To verify the rank condition of Theorem 1, we can check the determinant of

$$\Psi_{1:2} = \begin{bmatrix} u_{33}v_{13} & u_{32}v_{12} \\ 1 & 1 \end{bmatrix},$$

which is $\det(\Psi_{1:2}) = u_{33}v_{13} - u_{32}v_{12}$. Therefore, the 3×3 matrix A^0 is identified with only a single zero restriction arising from the exclusion of the third variable from the first equation as long as $u_{33}v_{13} \neq u_{32}v_{12}$.

Since the second and third columns of V are complex conjugate pairs of eigenvectors, we have $\det(\Psi_{1:2}) = \bar{u}_{32}\bar{v}_{12} - u_{32}v_{12} = -2\text{Im}(u_{32}v_{12})$, where Im denotes the imaginary part. It follows that identification fails only for matrices A^0 for which $\text{Im}(u_{32}v_{12}) = 0$, which is equivalent to the condition obtained by Phillips (1973) on p. 385.[6] He explored the implications of this condition for the matrix A^0 and showed that identification fails only for a set of matrices in $\mathcal{A}$ with Lebesgue measure zero. Hence, in this 3×3 example a single zero restriction is sufficient to identify almost all structural matrices A^0.

2.2. A Remark on Cointegration

We now briefly turn to models with cointegration, where A^0 has a zero eigenvalue with multiplicity $n - p$ and so $A^0 = FG^\top$, where F and G are $n \times p$ matrices of full column rank and where $G^\top y(t)$ are the p linear cointegrating relations. In previous work, Phillips (1991) showed that there is no aliasing problem for the long run equilibrium submatrix H in triangular systems of the form

$$\begin{aligned} y_1(t) &= Hy_2(t) + v_1(t), \\ \mathrm{d}y_2(t) &= v_2(\mathrm{d}t). \end{aligned}$$

Such systems are special cases of (1) where $A^0 = FG^\top$ with $F = [\, I\ 0 \,]^\top$ and $G = [\, -I\ H \,]$. Kessler and Rahbek (2004) showed that if all eigenvalues of A^0 are real and no elementary divisors of A^0 are repeated, then the cointegration spaces spanned by the columns of F and $G^\top$ are identified (even if F and G are not themselves identified). Here, we consider the following alternative to Assumption 1 for models with p cointegrating relations. We note that p is identified since the eigenvalues of B^0 are $\mu_j = \exp(h\lambda_j)$, where λ_j are the eigenvalues of A^0 (Gantmacher, 1959, p. 240).

Assumption 3. The matrix A^0 has p distinct nonzero eigenvalues and an eigenvalue equal to zero with algebraic and geometric multiplicity $n - p$.

THEOREM 2. *Suppose Assumptions 2 and 3 are satisfied and define* $\delta_p \equiv \lfloor p/2 \rfloor$. *If* $rank(\Psi_{1:2\delta_p}) = 2\delta_p$, *then* A^0 *is identified.*

Although the aliasing problem is potentially worse under Assumption 3, it provides new information about A^0 in the form of the value and multiplicity of $n - p$ eigenvalues. The net result is that we can reduce the minimum number of required restrictions to $\delta_p < \delta = \lfloor n/2 \rfloor$.

3. IDENTIFICATION VIA LINEAR RESTRICTIONS ON THE MATRICES A AND Σ

In the previous section, we focused on identification of the coefficient matrix A^0 using only information about the discrete time coefficient matrix B^0. However, Hansen and Sargent (1983) considered the problem of joint identification of both the coefficient and covariance matrices (A^0, Σ^0) using information contained in both the discrete time coefficient and covariance matrices (B^0, Ω^0). They showed that Ω^0 contains identifying information about A^0 over and above that contained in B^0.

McCrorie (2003) also studied joint identification of (A^0, Σ^0) by defining an augmented matrix Ξ^0 and exploiting a result of Van Loan (1978, Theorem 1) on the matrix exponential:

$$\Xi^0 \equiv \begin{bmatrix} A^0 & \Sigma^0 \\ 0 & (-A^0)^\top \end{bmatrix} \quad \text{and} \quad \exp(h\,\Xi^0) = \begin{bmatrix} B^0 & B^0\Omega^0 \\ 0 & (B^0)^{-\top} \end{bmatrix}.$$

If the eigenvalues of Ξ^0 are real and no elementary divisors are repeated (a necessary and sufficient condition for a unique matrix logarithm), then (A^0, Σ^0) is identified (McCrorie, 2003, Theorem 1).

We now extend Theorem 1 by considering general linear restrictions on A^0 and Σ^0 of the form

$$R\,\mathrm{vec}\left(A^0\ \Sigma^0\right) = r, \qquad \textbf{(9)}$$

where R is a restriction matrix of dimension $k \times 2n^2$. We show that despite having a second $n \times n$ matrix to identify, it can be sufficient for R to contain only $\delta = \lfloor n/2 \rfloor$ restrictions as before. This is possible for two reasons. First, under our assumptions it follows from Lemma 3 of Kessler and Rahbek (2004) that identification of A^0 is sufficient for identification of Σ^0. Second, since Ω^0 contains

https://doi.org/10.1017/S0266466615000353 Published online by Cambridge University Press

identifying information about A^0 it seems likely that restrictions on Σ^0 could provide identifying restrictions on A^0. As in Theorem 1, provided that a similar rank condition holds we can use the same number of restrictions—now potentially involving A^0, Σ^0, or both—to identify both matrices.

THEOREM 3. *Suppose Assumptions 1 and 2 are satisfied and for $\delta = \lfloor n/2 \rfloor$ and $L = \left[I_n \; 0_{n\times n}\right]$ define the partitioned matrix*

$$\Psi = \left[\Psi_{1:2\delta} \; \Psi_{2\delta+1:4n^2}\right] \equiv \begin{bmatrix} R(U \otimes LV)P^\top \\ I_\delta \; I_\delta \; 0_{\delta\times(4n^2-2\delta)} \end{bmatrix}.$$

If $\text{rank}(\Psi_{1:2\delta}) = 2\delta$, then both A^0 and Σ^0 are identified.

3.1. Numerical Examples

To give some numerical examples, we revisit a model used by Phillips (1972) for a series of Monte Carlo experiments. The particular coefficient and covariance matrices he used were

$$A^0 = \begin{bmatrix} -0.6 & 0.45 & 0 \\ 4 & -0.8 & -1.6 \\ 0 & 0.8 & -0.4 \end{bmatrix} \quad \text{and} \quad \Sigma^0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Therefore, the Ξ^0 matrix in this case is

$$\Xi^0 = \begin{bmatrix} A^0 & \Sigma^0 \\ 0 & (-A^0)^\top \end{bmatrix} = \left[\begin{array}{ccc|ccc} -0.6 & 0.45 & 0 & 1 & 0 & 0 \\ 4 & -0.8 & -1.6 & 0 & 1 & 0 \\ 0 & 0.8 & -0.4 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0.6 & -4 & 0 \\ 0 & 0 & 0 & -0.45 & 0.8 & -0.8 \\ 0 & 0 & 0 & 0 & 1.6 & 0.4 \end{array}\right].$$

We will show how A^0 and Σ^0 can be identified using a single prior restriction on an element of Ξ^0, corresponding to a single element of either A^0 or Σ^0.

Given the form of Ξ^0, its eigenvalues are those of A^0 and $-A^0$. Hence, the six eigenvalues of Ξ^0 are $\lambda_1 = -1.57$, $\lambda_2 = -0.12 + 0.37i$, $\lambda_3 = -0.12 - 0.37i$, $\lambda_4 = 1.57$, $\lambda_5 = 0.12 + 0.37i$, and $\lambda_6 = 0.12 - 0.37i$. These are the diagonal entries of the 6×6 matrix Λ. The corresponding 6×6 eigenvector matrix is

$$V = \left[\begin{array}{ccc|ccc} -0.36 & -0.34+0.09i & -0.34-0.09i & -0.42 & 0.35+0.13i & 0.35-0.13i \\ 0.77 & -0.29+0.38i & -0.29-0.38i & -0.56 & 0.45+0.37i & 0.45-0.37i \\ -0.53 & -0.81 & -0.81 & -0.12 & 0.72 & 0.72 \\ \hline 0 & 0 & 0 & -0.65 & 0.06i & -0.06i \\ 0 & 0 & 0 & 0.16 & 0.01+0.01i & 0.01-0.01i \\ 0 & 0 & 0 & 0.22 & 0.01-0.03i & 0.01+0.03i \end{array}\right].$$

https://doi.org/10.1017/S0266466615000353 Published online by Cambridge University Press

Again, given the form of Ξ^0 the first three columns of V are the eigenvectors of A^0 stacked above n zeros. Since the eigenvectors of Ξ^0 and $\exp(h\Xi^0)$ are the same, the matrices V and $U = V^{-\top}$ here are identified. The Ψ matrix in this case, which is also identified, is

$$\Psi = \begin{bmatrix} R(U \otimes LV)P^\top \\ 1 \;\; 1 \;\; 0_{1\times 34} \end{bmatrix} \quad \text{with} \quad P^\top = \left[e_9 \; e_5 \; e_1 \middle| e_2 \ldots e_{36} \right].$$

First, recall the restriction $a_{13}^0 = 0$ considered previously. This homogeneous restriction is now characterized by the 1×18 vector $R = [e_7^\top]$ rather than a 1×9 vector (although the leading nine elements are the same). In this case we have

$$\Psi_{1:2} = \begin{bmatrix} 0.45 + 0.46i & 0.45 - 0.46i \\ 1 & 1 \end{bmatrix}$$

and $\det(\Psi_{1:2}) = 0.92i \neq 0$. Similarly, the restriction $a_{31}^0 = 0$ would be sufficient.

We also consider restrictions on $\Sigma^0 = (\sigma_{ij}^0)$. Σ^0 was assumed to be the 3×3 identity matrix, which yields nine possible prior restrictions on individual elements of Σ^0. First, consider the zero restriction $\sigma_{21}^0 = 0$ represented by $R = [e_{11}^\top]$ and $r = [0]$. In this case,

$$\Psi_{1:2} = \begin{bmatrix} -0.08 + 11.23i & -0.08 - 11.23i \\ 1 & 1 \end{bmatrix}$$

and $\det(\Psi_{1:2}) = 22.46i \neq 0$, so the rank condition is again satisfied. Second, consider the restriction $\sigma_{11}^0 = 1$ represented by $R = [e_{10}^\top]$ and $r = [1]$. This yields

$$\Psi_{1:2} = \begin{bmatrix} -5.11 + 6.51i & -5.11 - 6.51i \\ 1 & 1 \end{bmatrix}$$

with $\det(\Psi_{1:2}) = 13.02i \neq 0$. In fact, in this example the rank condition holds for all nine of the possible single-element restrictions on Σ^0.

4. CONCLUSION

We have established a new rank condition for identification of the parameter matrix A^0 in a linear, first-order system of continuous time stochastic differential equations when only equispaced observations are available. The condition can be used to determine whether any particular set of $\lfloor n/2 \rfloor$ general linear restrictions on the $n \times n$ parameter matrix A^0 are sufficient for identification of A^0. We also considered two extensions. The first shows that we can identify both A^0 and Σ^0 using prior restrictions on A^0 or Σ^0 (or both). The second shows that in models with p cointegrating relations, as few as $\lfloor p/2 \rfloor$ restrictions can be sufficient to identify A^0.

This paper only considered linear restrictions on the parameter matrices. We also focused on models where the nonzero eigenvalues are distinct. Considering nonlinear restrictions implied by a structural model and allowing for repeated nonzero eigenvalues are natural directions for future work in this area.

NOTES

1. Here $\lfloor x \rfloor$ is the largest integer smaller than or equal to x.

2. An alternative assumption used by McCrorie (2003) and Kessler and Rahbek (2004) is that the eigenvalues of A^0 are real and that no Jordan block occurs more than once. Due to the monotonic relationship between the eigenvalues of A^0 and B^0 (with B^0 being the matrix exponential of A^0), it is also sufficient for A^0 to have distinct, real eigenvalues. In contrast, Assumptions 1 and 2 allow for complex eigenvalues provided that they are neither repeated nor congruent modulo $2\pi i/h$.

3. Phillips (1973) also discussed restrictions of the form $f_r^\top A^0 g_r = h_r$ with $h_r \in \mathbb{R}$ for $r = 1,\dots,k$, which are more general than restrictions on individual rows or columns of A^0 but less general than those we consider. As an example, consider the single restriction on a 2×2 matrix $A^0 = (a_{ij})$, represented by $R = [1,0,0,1]$ and $r = [0]$, which requires the sum of the diagonal elements to be zero: $R\text{vec}(A^0) = a_{11} + a_{22} = 0$. On the other hand, it is not possible to choose vectors f_1 and g_1 such that $f_1^\top A^0 g_1 = a_{11} + a_{22}$. To see this, note that $f_1^\top A^0 g_1 = f_{11}g_{11}a_{11} + f_{12}g_{11}a_{21} + f_{11}g_{12}a_{12} + f_{12}g_{12}a_{22}$ and we cannot simultaneously satisfy both the requirements that $f_{11}g_{11} = f_{12}g_{12} = 1$ and $f_{12}g_{11} = f_{11}g_{12} = 0$.

4. Note that there must be an even number of complex eigenvalues because they appear in complex conjugate pairs: if λ is a complex eigenvalue of A with eigenvector v, then $Av = \lambda v$ implies that $A\bar{v} = \bar{\lambda}\bar{v}$.

5. Here $\Phi_{i:j}$ denotes the matrix formed by columns $i, i+1, i+2, \dots, j$ of Φ.

6. Mapping our notation to that of Phillips (1973), we have $u_{33} = \bar{s}_3$, $u_{32} = s_3$, $v_{13} = \bar{k}_1$, and $v_{12} = k_1$ and so the condition becomes $\text{Im}(k_1 s_3) = 0$. I thank an anonymous referee for pointing out this equivalence.

REFERENCES

Aït-Sahalia, Y. (2007) Estimating continuous-time models using discretely sampled data. In T.P.R. Blundell and W.K. Newey (Eds.), *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*. Cambridge University Press.

Arcidiacono, P., P. Bayer, J.R. Blevins, & P.B. Ellickson (2012) Estimation of dynamic discrete choice models in continuous time. Working Paper 18449, National Bureau of Economic Research.

Bandi, F.M. & P.C.B. Phillips (2009) Nonstationary continuous-time processes. In Y. Aït-Sahalia and L.P. Hansen (Eds.), *Handbook of Financial Econometrics*, Volume 1, Chapter 3. Amsterdam: North Holland.

Baxter, M. & A. Rennie (1996) *Financial Calculus: An Introduction to Derivative Pricing*. Cambridge University Press.

Bergstrom, A.R. (1988) The history of continuous-time econometric models. *Econometric Theory* 4, 365–383.

Bergstrom, A.R. & K.B. Nowman (2007) *A Continuous Time Econometric Model of the United Kingdom with Stochastic Trends.* Cambridge: Cambridge University Press.

Coddington, E.A. & N. Levinson (1955) *Theory of Ordinary Differential Equations.* McGraw-Hill.

Culver, W.J. (1966) On the existence and uniqueness of the real logarithm of a matrix. *Proceedings of the American Mathematical Society* 17, 1146–1146.

Cuthbert, J.R. (1972) On uniqueness of the logarithm for Markov semi-groups. *Journal of the London Mathematical Society* s2-4, 623–630.

Cuthbert, J.R. (1973) The logarithm function for finite-state Markov semi-groups. *Journal of the London Mathematical Society* s2-6, 524–532.

Doraszelski, U. & K.L. Judd (2012) Avoiding the curse of dimensionality in dynamic stochastic games. *Quantitative Economics* 3, 53–93.

Fan, J. (2005) A selective overview of nonparametric methods in financial econometrics. *Statistical Science* 20, 317–357.

Gantmacher, F.R. (1959) *The Theory of Matrices*, Volume 1. New York: Chelsea.

Geweke, J. (1978) Temporal aggregation in the multiple regression model. *Econometrica* 46, 643–661.

Hamerle, A., W. Nagl, & H. Singer (1991) Problems with the estimation of stochastic differential equations using structural equations models. *Journal of Mathematical Sociology* 16, 201–220.

Hansen, L.P. & T.J. Sargent (1983) The dimensionality of the aliasing problem in models with rational spectral densities. *Econometrica* 51, 377–387.

Hurwicz, L. (1950) Generalization of the concept of identification. In T.C. Koopmans (Ed.), *Statistical Inference in Dynamic Models*. New York: John Wiley & Sons.

Kessler, M. & A. Rahbek (2004) Identification and inference for multivariate cointegrated and ergodic Gaussian diffusions. *Statistical Inference for Stochastic Processes* 7, 137–151.

Koopmans, T.C. (1949) Identification problems in economic model construction. *Econometrica* 17, 125–144.

McCrorie, J.R. (2000) Deriving the exact discrete analog of a continuous time system. *Econometric Theory* 16, 998–1015.

McCrorie, J.R. (2003) The problem of aliasing in identifying finite parameter continuous time stochastic models. *Acta Applicandae Mathematicae* 79, 9–16.

McCrorie, J.R. (2009) Estimating continuous-time models on the basis of discrete data via an exact discrete analog. *Econometric Theory* 25, 1120–1137.

Phillips, P.C.B. (1970) The structural estimation of stochastic differential equation systems. Master's thesis, University of Auckland.

Phillips, P.C.B. (1972) The structural estimation of a stochastic differential equation system. *Econometrica* 40, 1021–1041.

Phillips, P.C.B. (1973) The problem of identification in finite parameter continuous time models. *Journal of Econometrics* 1, 351–362.

Phillips, P.C.B. (1991) Error correction and long-run equilibrium in continuous time. *Econometrica* 59, 967–980.

Phillips, P.C.B. & J. Yu (2009) Maximum likelihood and Gaussian estimation of continuous time models in finance. In T. Mikosch, J.-P. Kreiß, R.A. Davis, and T.G. Andersen (Eds.), *Handbook of Financial Time Series*, pp. 497–530. Springer.

Schiraldi, P., H. Smith, & Y. Takahashi (2012) Estimating a dynamic game of spatial competition: The case of the U.K. supermarket industry. Working paper, London School of Economics.

Sims, C.A. (1971) Discrete approximations to continuous time distributed lags in econometrics. *Econometrica* 39, 545–563.

Singer, B. & S. Spilerman (1976) The representation of social processes by Markov models. *The American Journal of Sociology* 82(1), 1–54.

Sinha, N.K. & G.J. Lastman (1982) Identification of continuous-time multivariable systems from sampled data. *International Journal of Control* 35, 117–126.

Sørensen, H. (2004) Parametric inference for diffusion processes observed at discrete points in time: A survey. *International Statistical Review* 72, 337–354.

Van Loan, C.F. (1978) Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control* 23, 395–404.

Yu, J. (2014) Econometric analysis of continuous time models: A survey of Peter Phillips' work and some new results. *Econometric Theory* 30, 737–774.

APPENDIX

A. Proofs

A.1. Proof of Lemma 2

P, as defined in (8), is a permutation matrix, so $P^\top P = I$ and the restrictions in (7) can be restated as

$$R(U \otimes V)P^\top d = 0 \tag{A.1}$$

where $d = P\operatorname{vec}(D) = \left[-m_\rho \ldots -m_1\; m_\rho \ldots m_1\; 0 \ldots 0\right]^\top$ denotes the corresponding $n^2 \times 1$ permuted vector of elements of D. The system of k linear equations in (A.1) corresponds to the restrictions on A^0 the elements of d being the n^2 unknowns. We can complete the system of equations for $d = [-m^\top, m^\top, 0]^\top$ as

$$\Pi d = \begin{bmatrix} R(U \otimes V)P^\top \\ I_\rho \; I_\rho \; 0_{n^2-2\rho} \\ 0 \;\; 0 \;\; I_{n^2-2\rho} \end{bmatrix} \begin{bmatrix} -m \\ m \\ 0 \end{bmatrix} = 0,$$

where $m = [m_\rho, \ldots, m_1]^\top$ is a vector of ρ unknowns—differences between the complex eigenvalues of an arbitrary solution A and those of A^0. The first block row imposes the k restrictions on A in (A.1). The second block, of dimension $\rho \times n$, imposes the restrictions on the diagonal elements of $-M$ and M. The third block, of dimension $(n^2 - 2\rho) \times n$, restricts the remaining elements of d to be zero.

If Π has full column rank, then $-m = m = 0$ is the unique solution. Note that we can write Π as

$$\Pi = \begin{bmatrix} \Phi_{1:2\rho} & \Phi_{2\rho+1:n^2} \\ 0 & I_{n^2-2\rho} \end{bmatrix}$$

and hence a sufficient condition for Π to have full column rank is $\operatorname{rank}(\Phi_{1:2\rho}) = 2\rho$.

A.2. Proof of Theorem 1

The proof follows from Lemma 2 after noting that $\rho \le \delta$ and redefining the permutation matrix P and the dimensions of the unknowns accordingly.

A.3. Proof of Theorem 2

We first show that (5) also holds when we replace Assumption 1 with Assumption 3. We appeal to the more general form of the theorem of Gantmacher (1959, VIII.8), as restated by Singer and Spilerman (1976, Proposition 2): All solutions A of the equation $\exp(hA) = B^0$ are given by $A = \frac{1}{h}\ln B^0 = \frac{1}{h} H Q \ln J Q^{-1} H^{-1}$, where H is any nonsingular matrix which reduces B^0 to Jordan form, $B^0 = HJH^{-1}$, Q is an arbitrary nonsingular matrix that commutes with J, with $QJ = JQ$, and J is the Jordan normal form, a block-diagonal matrix consisting of the Jordan blocks of B^0.

Under Assumption 3, the Jordan normal form of A^0 is a diagonal matrix J with n Jordan blocks of order 1 along the diagonal. The p blocks $J_1 = [\lambda_1], \ldots, J_p = [\lambda_p]$ correspond to

https://doi.org/10.1017/S0266466615000353 Published online by Cambridge University Press

the p distinct nonzero eigenvalues and the $n-p$ blocks $J_{p+1}=\cdots=J_n=[0]$ correspond to the repeated eigenvalue $\lambda_0=0$, which has geometric multiplicity $n-p$ by assumption. Given this structure, the matrix J and all commuting matrices Q have the form

$$J=\left[\begin{array}{ccc|ccc} \lambda_1 & 0 & & & \dots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ & \ddots & \lambda_p & 0 & & \\ \hline & & 0 & \lambda_0 & \ddots & \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & & & 0 & \lambda_0 \end{array}\right] \quad \text{and} \quad Q=\left[\begin{array}{ccc|ccc} q_{11} & 0 & & & \dots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ & \ddots & q_{pp} & 0 & & \\ \hline & & 0 & q_{p+1,p+1} & \cdots & q_{p+1,n} \\ \vdots & & & \vdots & \ddots & \vdots \\ 0 & \dots & & q_{n,p+1} & \cdots & q_{n,n} \end{array}\right].$$

Q has at most p nonzero diagonal elements in the upper left block and at most $(n-p)^2$ nonzero elements in the lower right block. However, since $\lambda_0=0$, the product $QJ=JQ$ is independent of the lower right block (i.e., the values $q_{j,k}$ for $j,k=p+1,\dots,n$). Since $Q\ln JQ^{-1}$ is diagonal for any admissible Q, $\ln B^0$ is independent of H. In other words, for any H, Q is determined so we may choose $Q=I$ and $H=V$ as a normalization. Hence, we can indeed still appeal to (5) and the result follows from the proofs of Lemma 2 and Theorem 1, noting that there are at least $n-p$ known real eigenvalues and so the number of complex eigenvalues is $\rho\le\lfloor p/2\rfloor$.

A.4. Proof of Theorem 3

The proof of identification of A^0 is similar to the proofs of Lemma 2 and Theorem 1. We replace V with LV and add additional restrictions to account for the additional known relationships among the $2n$ eigenvalues of Ξ^0. We discuss these new details first to establish identification of A^0 and then show that Σ^0 is identified whenever A^0 is identified.

Let $\lambda_1,\dots,\lambda_n$ denote the n eigenvalues of A^0 and let $\eta_1,\dots,\eta_{2n}$ denote the $2n$ eigenvalues of Ξ^0. Given the structure of Ξ^0, we have $\eta_1=\lambda_1,\dots,\eta_n=\lambda_n$ and $\eta_{n+1}=-\lambda_1,\dots,\eta_{2n}=-\lambda_n$. Therefore, the $2n$ eigenvalues of Ξ^0 are determined by the n unknowns $\lambda_1,\dots,\lambda_n$. As before, we can appeal to a theorem of Gantmacher (1959, VIII.8) to show that any alternative real solution Ξ is related to Ξ^0 as $\Xi=V(\Lambda+D)V^{-1}=\Xi^0+VDV^{-1}$ where D is now a $2n\times 2n$ diagonal matrix of the form $D=\operatorname{diag}(0_{1\times(n-2\delta)},m^\top,-m^\top,0_{1\times(n-2\delta)},m^\top,-m^\top)$ for $m=\left[m_1\dots m_\delta\right]^\top$, $m_i=\frac{2\pi i}{h}\tilde{m}_i$, and $\tilde{m}_i\in\mathbb{Z}$ for $i=1,\dots,\delta$. Hence, D has $4n^2$ elements with at least $4n^2-4\delta$ zeros.

Note that $\operatorname{vec}(L\Xi^0)=\operatorname{vec}\left(A^0\ \Sigma^0\right)$ so the prior restrictions in (9) can be restated as $R\operatorname{vec}(L\Xi^0)=r$. Since $L\Xi=LV(\Lambda+D)V^{-1}=L\Xi^0+LVDV^{-1}$ and since any other admissible Ξ matrix must also satisfy the restrictions, $R\operatorname{vec}(L\Xi)=r$, it follows that $R(U\otimes LV)\operatorname{vec}(D)=0$. This leads to the full system of equations for a suitably permuted vector of unknowns $d=P\operatorname{vec}(D)$:

$$\Pi d=\left[\begin{array}{ccccc} \multicolumn{5}{c}{R(U\otimes LV)P^\top} \\ \hline I_\delta & I_\delta & 0 & 0 & 0 \\ 0 & I_\delta & I_\delta & 0 & 0 \\ 0 & 0 & I_\delta & I_\delta & 0 \\ 0 & 0 & 0 & 0 & I_{4n^2-4\delta} \end{array}\right]\begin{bmatrix}-m\\ m\\ -m\\ m\\ 0\end{bmatrix}=0.$$

The stated rank condition guarantees that Π has full column rank, so $-m = m = 0$ and A^0 is identified.

Recalling the definition of Ω^0, we have $\text{vec}(\Omega^0) = \left[\int_0^h \exp(sA^0) \otimes \exp(s(A^0)^\top)\, ds \right]$ $\text{vec}(\Sigma^0)$. Under Assumptions 1 and 2, by Lemma 3 of Kessler and Rahbek (2004), the matrix $\int_0^h \exp(sA^0) \otimes \exp(s(A^0)^\top)\, ds$ is nonsingular. If A^0 is identified then $\Omega \neq \Omega^0$ implies $\Sigma \neq \Sigma^0$ and so Σ^0 is identified.