Available online at www.sciencedirect.com
ScienceDirect
Journal of Econometrics 136 (2007) 303–318
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Time reversibility of stationary regular finite-state Markov chains

William J. McCausland*

Centre Interuniversitaire de Recherche en Économie Quantitative and Centre Interuniversitaire de Recherche en Analyse des Organisations, Département de sciences économiques, Université de Montréal, C.P. 6128, succursale centre-ville, Montréal, Que., Canada H3C 3J7

Received 25 January 2005; accepted 28 September 2005
Available online 2 May 2006

Abstract

We propose an alternate parameterization of stationary regular finite-state Markov chains, and a decomposition of the parameter into time reversible and time irreversible parts. We demonstrate some useful properties of the decomposition, and propose an index for a certain type of time irreversibility, applicable to chains whose states are naturally ordered. Two empirical examples illustrate the use of the proposed parameter, decomposition and index. One, on gasoline price mark-ups, involves observed states. The other, on U.S. investment growth, features latent states.

JEL classification: C11; C13; C22; E32

Keywords: Finite-state Markov chains; Time reversibility; Bayesian inference; Hidden Markov models

1. Introduction

Intuitively, a random process is time reversible if the statistical properties of the process are the same as those of the same series running backwards through time. For discrete-time processes, a formal definition is the following.

Definition 1.1. A discrete-time random process $\{x_t\}$ is *time reversible* if for every positive integer K and all integers t and τ, the distributions of $(x_t, x_{t+1}, \ldots, x_{t+K})$ and $(x_{\tau-t}, x_{\tau-t-1}, \ldots, x_{\tau-t-K})$ are identical.

*Tel.: + 1 5143437281; fax: + 1 5143437221.
E-mail address: william.j.mccausland@umontreal.ca.

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2005.09.001

Several papers in economics find evidence of time irreversibility in observed time series. In the large literature on business cycle fluctuations, many authors document a tendency for downswings to be faster than upswings, an example of what is known in the literature as "steepness". Ramsey and Rothman (1996) survey much of this literature, and find evidence of time irreversibility in many macroeconomic variables. Chen and Kuan (2001) find that stock index returns are time irreversible, and Fong (2003) finds stock trading volume to be time irreversible. Noel (2003) analyzes retail gasoline markets in 19 Canadian cities and finds strong evidence, in some but not all markets, of cycles where prices tend to rise sharply and decline gradually.

Although time reversibility is quite restrictive, there are many examples of commonly used statistical models for discrete time processes that impose it. Time reversible processes include exchangeable processes, univariate stationary Gaussian processes and time discretized stationary univariate diffusions. The last two are often used as models for macroeconomic variables and asset returns, respectively.

Allowing time irreversibility is not only important for constructing more realistic statistical models. Time reversibility may be of economic interest in itself. Maskin and Tirole (1988) discuss a dynamic game in which two firms competing in an output market choose prices in each period. One set of equilibria features what the authors call "Edgeworth cycles", in which large price jumps are followed by more gradual falls in price as the two firms repeatedly undercut one another. Noel (2003) explicitly draws the connection between cycles in gasoline prices and Edgeworth cycles, and shows that the degree of market penetration of independent gasoline retailers helps predict the presence or absence of cycles in a market, in a way that is consistent with an extension by Eckert (2003) of the theory of Maskin and Tirole.

In markets where participants simultaneously observe noisy value-relevant signals and learn about the signal through time, price volatility and trading volume may feature rapid growth and slow decay. Sims (2003) discusses information processing capacity constraints and shows how these can generate non-instantaneous reactions to information. Peng and Xiong (2003) propose a model with these constraints that features not only the clustering and long memory of volatility (their objective) but also cycles of volatility with sharp increases and gradual decreases.

Economic agents often face adjustment costs, which can lead to time irreversible decision processes. When inflation is positive, the real prices of goods produced by firms with menu costs rise abruptly and fall gently. Capital investment decisions of firms may be more easily made than reversed.

Where asset price bubbles occur, prices in excess of risk-adjusted expected discounted dividends expand slowly and pop suddenly.

Tests for time reversibility have been proposed by various authors. Ramsey and Rothman (1996) introduce a time domain test and Hinich and Rothman (1998) propose a frequency domain test. Robinson (1991) describes an entropy based test that can be used to test for time reversibility. Chen et al. (2000) introduce a class of tests based on characteristic functions that do not require the existence of any moments.

In this paper, we are more concerned with characterizing the time irreversibility of processes rather than testing for their time reversibility. We focus on stationary regular finite-state Markov chains. These chains may be directly observed, but often they are incorporated as latent processes in more elaborate models, known as hidden Markov models. There has been much recent interest in these models, and computational advances for their analysis. See, for example, Scott (2002).

Stationary regular finite-state Markov chains can be reversible or not. When they are parameterized as described in Section 2, time reversibility is easy to check and the degree and nature of time irreversibility is transparent. To facilitate the characterization of time irreversibility, we introduce, in Section 3, a decomposition of the parameter into what we will call its time reversible and time irreversible parts. We offer graphical representations of the parameter and its decomposition, which promote intuition and help interpret the nature of a chain's time irreversibility. Many chains have naturally ordered states, and for such chains we propose an index for a certain type of time irreversibility.

When the number of states is greater than two, imposition of time reversibility amounts to a reduction in the dimension of the parameter space. This parsimony may be useful in some applications. Incorporating latent Markov chains is an easy way of generating models that can feature time reversibility or not, depending on values of the parameters.

Section 4 presents results from two empirical applications illustrating the use of the proposed parameter, decomposition and index. The first, on gasoline price mark-ups, involves directly observed states. The second, on U.S. investment data, features latent states. The results suggest that the proposed time irreversibility index for chains with naturally ordered states is an empirically interesting measure of time irreversibility.

Section 5 concludes and identifies directions for further research.

2. A new parameterization

A stationary regular finite-state Markov chain $\{s_t\}$ is usually parameterized by its Markov transition matrix P, which gives the conditional probabilities $\Pr[s_t = j | s_{t-1} = i]$. In this section, we first review some results on stationary regular finite-state Markov chains. We then introduce an alternate parameterization of these chains by the matrix Π of *joint* probabilities $\Pr[s_{t-1} = i, s_t = j]$. Finally, we point out some of the advantages of this parameterization.

2.1. A conventional parameterization

Let $\{s_t\}$ be a Markov chain with finite state space $\{1, \ldots, m\}$. Let P be its $m \times m$ transition matrix. That is, for all $i, j \in \{1, \ldots, m\}$,

$$P_{ij} = \Pr[s_t = j | s_{t-1} = i]. \tag{1}$$

We review the following important and well known results. See Iosifescu (1980), and especially Theorem 1.9, Proposition 4.1 and Theorems 4.2 and 4.4.

1. $\{s_t\}$ is regular[1] if and only if P is regular.[2]
2. If $\{s_t\}$ is regular, then

(a) there exists a unique $1 \times m$ row-stochastic[3] vector π, which we will call the *stationary distribution* of $\{s_t\}$, such that $\pi P = \pi$,

[1] A Markov chain is *regular* if it is irreducible and aperiodic.

[2] A Markov transition matrix P is *regular* if there exists an integer $n > 0$ such that for all states i and j, $(P^n)_{ij} > 0$.

[3] A real $m \times n$ matrix (or vector) P is *row-stochastic* if for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, $\sum_{k=1}^{n} P_{ik} = 1$ and $P_{ij} \geqslant 0$. The simpler term "stochastic" is often used, but "row-stochastic" is also used and helps to differentiate the term from a new one ("matrix-stochastic") that I introduce later.

(b) $\pi > 0$, and
(c) if $\{s_t\}$ is stationary, then for all $i \in \{1, \ldots, m\}$ and all t, $\pi_i = \Pr[s_t = i]$.

The results imply that a stationary regular finite-state Markov chain is fully described by its Markov transition matrix, and that the following parameter set indexes the stationary regular m-state Markov chains.

$$\mathbf{P} \equiv \{P \in \mathbb{R}^{m \times m} \colon P \text{ is row-stochastic and regular}\}. \tag{2}$$

2.2. An alternate parameterization

For the purposes of studying time reversibility and time irreversibility of stationary regular finite-state Markov chains, it is convenient to consider an alternate parameterization, one which gives joint state probabilities rather than conditional state probabilities.

Definition 2.1. Let $\{s_t\}$ be a stationary regular finite-state Markov chain. Define the *joint probability matrix* Π of $\{s_t\}$ as the $m \times m$ matrix such that for all $i,j \in \{1, \ldots, m\}$, $\Pi_{ij} \equiv \Pr[s_{t-1} = i, s_t = j]$.

For any stationary regular finite-state Markov chain, the Markov transition matrix P, the stationary distribution vector π and the joint probability matrix Π are related by the simple expression $\Pi = DP$, where D is the diagonal matrix with $D_{ii} = \pi_i$, $i = 1, \ldots, m$. We can also invert D to write $P = D^{-1}\Pi$. These claims follow immediately from the definition of conditional probability and the fact that $\pi > 0$, a consequence of regularity. Stationarity implies that row sums equal column sums and that both give the stationary distribution: $\pi = \iota'\Pi = (\Pi\iota)'$, where ι is an $m \times 1$ vector of ones.

We introduce three definitions that will be useful for describing the properties of Π. Let A be any real $m \times m$ matrix. The first definition is for matrices that specify valid joint probability mass functions on $\{1, \ldots, m\}^2$. We call A *matrix-stochastic* if for all $i,j \in \{1, \ldots, m\}$, $A_{ij} \geqslant 0$ and

$$\sum_{i=1}^{m} \sum_{j=1}^{m} A_{ij} = 1. \tag{3}$$

The concept of balance refers to the equality of row sums to column sums, which for matrix-stochastic A means that the two marginal distributions are identical. We say that A is *balanced* if for all $i \in \{1, \ldots, m\}$,

$$\sum_{j=1}^{m} A_{ij} = \sum_{j=1}^{m} A_{ji}. \tag{4}$$

The third definition is closely related to the regularity of Markov chains and their Markov transition matrices. We say that A is *regular*[4] if there exists an integer $n > 0$ such that for all $i,j \in \{1, \ldots, m\}$, $(A^n)_{ij} > 0$.

The following result states that Π must satisfy these three properties.

Result 2.1. *Let $\{s_t\}$ be a stationary regular finite-state Markov chain, and Π be its joint probability matrix. Then Π is matrix-stochastic, balanced and regular.*

[4]This is a generalization of the definition of regularity of Markov transition matrices to square matrices.

Proof. That Π is matrix stochastic follows from the definition of probability. Balance follows from stationarity, as $\sum_{j=1}^{n} \Pi_{ij}$ and $\sum_{j=1}^{n} \Pi_{ji}$ must both equal π_i. Since the elements of P are non-negative, regularity of P depends only on which of these elements are zero. Since each element of Π is zero if and only if the corresponding element of P is zero, regularity of P implies regularity of Π. $\square$

The next result states that the three properties are exhaustive.

Result 2.2. *Let $m \times m$ matrix Π be matrix-stochastic, balanced and regular. Then Π is the joint probability matrix for a stationary regular m-state Markov chain.*

Proof. Let $\pi \equiv (\Pi\iota)'$ and D be the diagonal matrix with $D_{ii} = \pi_i$, $i = 1, \ldots, m$. Regularity of Π rules out a row or column of zeros, so $\pi > 0$ and D is non-singular. Now let $P \equiv D^{-1}\Pi$.

Matrix stochasticity of Π implies row stochasticity of P and π. Regularity of Π implies regularity of P. Balance implies $\iota'\Pi = \pi$ and so $\pi P = \pi$. The stationary Markov chain with Markov transition matrix P is therefore regular and has π as its stationary distribution and $\Pi = DP$ as its joint probability matrix. $\square$

With the previous two theorems, we have established the following parameter set as an alternative to $\mathbf{P}$ for stationary regular finite-state Markov chains:

$$\mathbf{\Pi} = \{\Pi \in \mathbb{R}^{m \times m} \colon \Pi \text{ is matrix-stochastic, balanced and regular}\}. \tag{5}$$

An important advantage of the Π parameterization lies in the following result.

Result 2.3. *Let $\{s_t\}$ be a stationary regular finite-state Markov chain, and let Π be its joint probability matrix. Then $\{s_t\}$ is time reversible if and only if Π is symmetric.*

A generalization of this result to richer state spaces is well known in the Markov chain Monte Carlo literature. For a simple proof of the special case, see McCausland (2004a), a working paper version of this article.

Symmetry of Π is easy to check. Checking reversibility using only the Markov transition matrix P is more difficult.

An added advantage of the Π parameter over the P parameter is the transparency of many other characteristics of a Markov chain. While calculating π from P involves solving a system of equations, computing π from Π is a matter of finding row (or column) sums, a back-of-the-envelope calculation that one can do to a crude but useful approximation in one's head. Likewise, it is easier to compute P from Π than Π from P. We can compute the Markov transition matrix for the time reversed chain as easily as P. The magnitude, direction and nature of temporal asymmetries become easily discernible, as they mirror asymmetries of the Π matrix.

While garnering relevant information from a given value or distribution of Π is easier, providing a value or distribution for Π is more difficult, since the condition that row sums equal column sums is more difficult to observe that the condition that row sums equal one. In a paper in preparation, we show how to elicit hierarchical prior distributions for Π directly.

2.3. A graphical representation of Π

An important objective of this research is to facilitate the characterization of the time irreversibility of a process. To this end we propose a graphical representation of Π to promote intuition and to serve as an expositional tool. We construct a directed graph with the same information as Π. The graph has m vertices, labelled 1 through m, each representing one of the states of the underlying Markov chain. For all vertices $i,j \in \{1,\ldots,m\}$, there is directed edge of weight Π_{ij} from i to j. We will call Π_{ij} the *flow* from state i to state j, $\Pi_{ij} + \Pi_{ji}$ the *total flow* between states i and j, $\sum_{k=1}^{m} \Pi_{ik}$ the *outflow* from state i and $\sum_{k=1}^{m} \Pi_{ki}$ the *inflow* to state i.

Fig. 1 shows a 3×3 joint probability matrix and the corresponding graph. Line thicknesses are proportional to the corresponding elements of Π. The balance condition on Π is equivalent to the property that the inflow to a state equals the outflow from it. Balance is intimately related to stationarity, since the outflow from a state i gives the marginal probability $\Pr[s_{t-1} = i]$ and the inflow to i gives the marginal probability $\Pr[s_t = i]$.

We have seen that reversibility of $\{s_t\}$ is equivalent to the symmetry of Π. This is equivalent in turn to the property that for all states i and j, the flow from i to j is equal to the flow from j to i. This condition is known as *detailed balance*.

For $m = 2$, balance implies detailed balance or equivalently, symmetry of Π. Thus, all stationary regular two-state Markov chains are time reversible.

3. Time reversibility and a decomposition

In this section, we describe a decomposition of the joint probability matrix Π. We demonstrate several properties of the decomposition that are useful for testing for time reversibility and for characterizing time irreversibility. Finally, we introduce an index for a certain type of time irreversibility in applications where states have a natural order.

3.1. A decomposition

Any $\Pi \in \mathbf{\Pi}$ can be decomposed as $\Pi = X + L$, where

$$X \equiv (\Pi + \Pi')/2 \quad \text{and} \quad L \equiv (\Pi - \Pi')/2. \tag{6}$$

$$\Pi = \frac{1}{20}\begin{bmatrix} 0 & 2 & 3 \\ 1 & 1 & 3 \\ 4 & 2 & 4 \end{bmatrix}$$

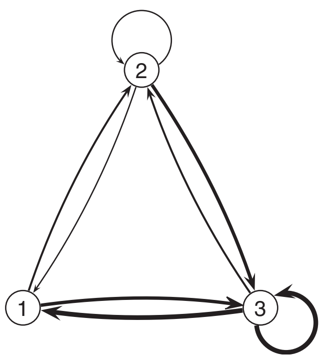

Fig. 1. A joint probability matrix Π and a graphical representation.

Anticipating future results, we will call X the time reversible part of Π and L the time irreversible part of Π. The symbols X and L are chosen because the symmetry and asymmetry of these letters resemble the symmetry and asymmetry of the matrices they represent.

The first important result is that the time reversible part of a joint probability matrix is the joint probability matrix for a time reversible stationary regular finite-state Markov chain.

Result 3.1. *Suppose $\Pi \in \mathbf{\Pi}$ and let $X = (\Pi + \Pi')/2$. Then $X \in \mathbf{\Pi}$ and $X' = X$.*

Proof. That X is matrix-stochastic and balanced is obvious. For all $i,j \in \{1,\dots,m\}$, $X_{ij} \geqslant 0$ and $X_{ij} > 0$ whenever $\Pi_{ij} > 0$, so X must be regular. $X' = X$ is obvious. $\square$

The next result summarizes properties that a joint probability matrix Π and its reversible part X have in common: the chains they govern have the same stationary distribution, the same state persistence $\Pr[s_t = i | s_{t-1} = i] = \Pi_{ii}/\pi_i$ in every state i, and the same total flow $\Pi_{ij} + \Pi_{ji}$ between any states i and j.

Result 3.2. *Suppose $\Pi \in \mathbf{\Pi}$ and let $X = (\Pi + \Pi')/2$ and $L = (\Pi - \Pi')/2$. Let π be the stationary distribution of the chain whose joint probability matrix is Π. Let P and P^X be the Markov transition matrices associated with Π and X. Then*

1. $\pi P^X = \pi P = \pi$,
2. *for all $i \in \{1,\dots,m\}$, $P^X_{ii} = P_{ii}$, and*
3. $X + X' = \Pi + \Pi'$.

Proof. For all $i \in \{1,\dots,m\}$, $\sum_{j=1}^m \Pi_{ij} = \sum_{j=1}^m X_{ij} = \pi_i$, so π gives the stationary distribution for both the Π and the X chains, and therefore $\pi P^X = \pi P = \pi$. In addition, for all $i \in \{1,\dots,m\}$, $\Pi_{ii} = X_{ii}$ and therefore $P_{ii} = \Pi_{ii}/\pi_i = X_{ii}/\pi_i = P^X_{ii}$. $X + X' = \Pi + \Pi'$ is obvious. $\square$

We now define the subset $\mathbf{X}$ of $\mathbf{\Pi}$ of joint probability matrices for time reversible stationary regular finite-state Markov chains.

$$\mathbf{X} \equiv \{\Pi \in \mathbf{\Pi} \colon \Pi \text{ is symmetric}\}. \tag{7}$$

The decomposition $\Pi = X + L$ leads to two obvious but useful corollaries of Result 2.3. Reversibility is equivalent to its joint probability matrix Π being equal to its reversible part ($\Pi = X$) and equivalent to its irreversible part being zero ($L = 0$).

The L matrix accounts for deviations from reversibility. The value L_{ij} is the deviation of Π_{ij} from X_{ij}, signed such that a positive value represents flow from i to j in excess of X_{ij}. Unlike X and Π, L is not a joint probability matrix. It satisfies three conditions, all easy to verify. The first is the antisymmetry condition $L' = -L$. Thus, the flow in excess of X_{ij} from i to j has the same magnitude and opposite sign as the flow in excess of $X_{ji} = X_{ij}$ from j to i, and there can be no excess flow from i to i. The second is that the row and column sums of L are zero, a kind of balance. It means that the sum of all excess flows out of i must be matched by excess flows into i. The third is that for all i and j, $-X_{ij} \leqslant L_{ij} \leqslant X_{ij}$, which ensures that the elements of Π are non-negative. We can think of $X_{ij} = X_{ji}$ as the capacity of excess flow of the "pipe" between i and j.

3.2. A graphical representation of X and L

To further facilitate the characterization of the time irreversibility of a chain, we offer graphical representations for the components of our decomposition. Let X and L be the time reversible and time irreversible parts of a joint probability matrix Π. We can construct an undirected graph with the same information as X. For all vertices $i, j \in \{1, \ldots, m\}$, there is an undirected edge of weight $X_{ij} = X_{ji}$ from i to j. The weight is the average of the flow Π from i to j and the flow Π_{ji} back from j to i. It is also the size of the bidirectional flow for the reversible chain with joint probability matrix X. Fig. 2 shows the time reversible part X of the joint probability matrix Π of Fig. 1, and the corresponding graph. Line thicknesses are proportional to the corresponding elements of X.

We can construct a directed graph with the same information as L. For all vertices $i, j \in \{1, \ldots, m\}$, there is an directed edge of weight $|L_{ij}| = |L_{ji}|$ between i and j. If L_{ij} is positive, the direction is from i to j; if negative, from j to i. Fig. 3 shows the time irreversible part L of the joint probability matrix Π of Fig. 1, and the corresponding graph. Line thicknesses are proportional to the corresponding elements of L.

Recall that edges from i to j represent flows from i to j in excess of X_{ij}. The condition that row and column sums must equal zero, or more intuitively, that excess inflows equal excess outflows, suggest an analogy with systems of fluids in pipes with no sources or sinks. The graphical representations of L in Figs. 3 and 4 show at a glance how excess flows circulate through the various states.

Suppose we have a stationary regular finite-state Markov chain with joint probability matrix Π. It is well known that the time reversed chain is also stationary, regular and Markov. Let Π_r be its joint probability matrix and let $X_r = (\Pi_r + \Pi_r')/2$ and $L_r = (\Pi_r - \Pi_r')/2$. It is easy to see that $\Pi_r = \Pi'$, $X_r = X$ and $L_r = L' = -L$. The graphical representation of the time reversed chain will have the same edges and weights as that of the original chain, with all directions reversed.

3.3. An index of time reversibility

In many applications, including the two examples in this paper, states have a natural order. We assume states are labelled in a manner consistent with this order. So, for example, if states are the indices of histogram bins of gasoline price mark-ups, state m represents the interval with the highest mark-ups and state 1 stands for the interval with

$$X = \frac{1}{40}\begin{bmatrix} 0 & 3 & 7 \\ 3 & 2 & 5 \\ 7 & 5 & 8 \end{bmatrix}$$

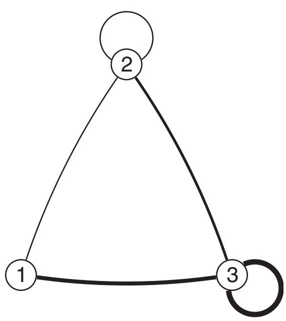

Fig. 2. Matrix X and a graphical representation.

$$L = \frac{1}{40}\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

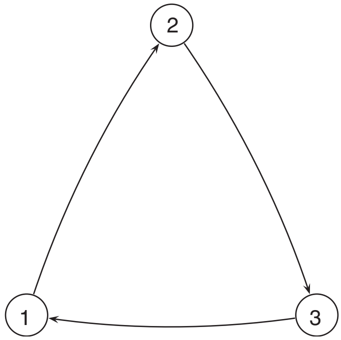

Fig. 3. Matrix L and a graphical representation.

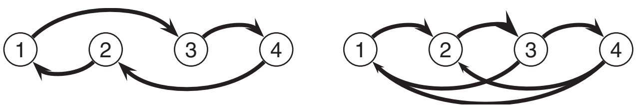

Fig. 4. Time irreversibility with and without zero total circulation.

the lowest. In this section, we describe an index for a certain type of time irreversibility, applicable to chains with ordered states. It is designed to capture the degree to which state transitions to higher states tend to be smaller and more frequent than those to lower states. The index can be negative, in which case jumps are larger and less frequent than falls.

Recall the examples from the introduction. Positive values of the index describe processes whose irreversibility is like that of economic growth series, where increases in growth are more frequent and smaller than decreases. Negative values describe processes like gasoline price mark-ups with Edgeworth cycles, where jumps in prices are less frequent and larger than falls.

Consider, by way of example, the Π matrix in Fig. 1. The sum of the elements above the diagonal ($\frac{8}{20}$) gives the probability of a jump to a higher state from one observation to the next. The sum of the elements below the diagonal ($\frac{7}{20}$), giving the probability of a fall, is smaller. However, the conditional expectation of the size of a jump, given a jump, is

$$(\Pi_{12} + 2\Pi_{13} + \Pi_{23})/(\Pi_{12} + \Pi_{13} + \Pi_{23}) = \tfrac{11}{8}, \tag{8}$$

while that of the size of a fall, given a fall, is

$$(\Pi_{21} + 2\Pi_{31} + \Pi_{32})/(\Pi_{21} + \Pi_{31} + \Pi_{32}) = \tfrac{11}{7}, \tag{9}$$

which is larger.

Take an interior state $i \in \{2, \ldots, m-1\}$ and consider the net outflow $C_i \equiv \sum_{j>i} L_{ij}$ from state i to higher states. Balance implies that C_i is also the net inflow $\sum_{j<i} -L_{ij}$ to i from lower states. Balance also implies that this flow must be offset by a net flow C_i from states higher than i directly to states lower than i. We will call C_i the *circulation through* i and $C \equiv \sum_{i=2}^{m-1} C_i = \sum_{j>i} L_{ij}$ the *total circulation*.

Note that C is also related to the relative probabilities of moving to a higher or lower state:

$$C = \sum_{j>i} L_{ij} = \sum_{j>i} (\Pi_{ij} - \Pi_{ji})/2 = (\Pr[s_t > s_{t-1}] - \Pr[s_t < s_{t-1}])/2. \tag{10}$$

For the L matrix in Fig. 3, corresponding to the Π matrix in Fig. 1, we calculate the circulation C_2 through state 2 and the total circulation C to be $L_{23} = \frac{1}{40}$.

Clearly, time reversibility implies $C = 0$, but the converse is not true for $m > 3$. Fig. 4 represents on the left a non-zero L matrix for which $C = 0$. Negative circulation through state 2 is offset by positive circulation through state 3. On the right is a representation of an L matrix whose total circulation is positive: both interior states have positive circulation.

We see that C is a useful index for a particular type of time irreversibility where circulation through a state tends to be in the same direction across states. The results of the next section suggest that C might be empirically relevant.

In many applications, states do not have an unambiguous natural order. In some cases, several orders may apply. Take for example a Markov mixture of normals process in which both mean and variance are state dependent. We can order latent states by mean, variance or marginal probability. In such cases, it may be interesting to consider circulation properties in terms of one or more of these orders. Even in applications with no natural order, the order induced by the marginal probabilities of states may yield an interesting index of circulation.

Other measures of departure from reversibility not relying on states having a order may be useful. For example, we can take $|L_{ij}|$ or $|L_{ij}|/X_{ij}$ as an index of irreversibility attributable to the flows between states i and j. We might measure the degree of irreversibility of the chain by aggregating either of these indices by, for example, summing or taking the maximum.

4. Empirical examples

The first empirical example investigates the time irreversibility of gasoline price mark-ups. We have 267 weekly observations of retail price r_t and wholesale price w_t for gasoline from November 27, 1989 to September 25, 1994. r_t is an average for a sample of gasoline stations in Windsor, Ontario, Canada. w_t is the price charged for large scale purchases of unbranded gasoline at the terminal in Toronto, Ontario. The data, collected by the government of Ontario, are the same as those used in Eckert (2002).

We divide the mark-up r_t/w_t into six bins according to Table 1 and model the evolution of the mark-up bin s_t as a stationary regular 6-state Markov chain.

We choose a prior on P with independent rows, each having a Dirichlet distribution. The Dirichlet parameter associated with P_{ij} is 2 if $i = j$ and 1 otherwise. We note that P is regular with probability one. The chain is irreversible with probability one, but since the density of P is invariant to state relabelling, the prior is neutral with respect to the direction of cycles in the L matrix.

We use the BACC software, described in Geweke (1999) and McCausland (2004b), to generate a posterior sample of 100,000 draws of P and then construct posterior samples for

Table 1
Mark-up bins

State	Range
1	$r_t/w_t < 1.0$
2	$1.0 \leqslant r_t/w_t < 1.1$
3	$1.1 \leqslant r_t/w_t < 1.2$
4	$1.2 \leqslant r_t/w_t < 1.3$
5	$1.3 \leqslant r_t/w_t < 1.4$
6	$1.4 \leqslant r_t/w_t$

Table 2
Posterior mean and standard deviation of the elements of P, π and Π in the gasoline mark-up example

P	0.4647	0.2048	0.1329	0.0660	0.0660	0.0656
	0.0828	0.5092	0.0411	0.2647	0.0613	0.0409
	0.0117	0.1811	0.5275	0.2329	0.0234	0.0234
	0.0085	0.0173	0.2704	0.5948	0.0923	0.0168
	0.0373	0.0768	0.1112	0.4794	0.2211	0.0741
	0.0719	0.0735	0.0723	0.1416	0.2851	0.3556
	0.1239	0.1010	0.0847	0.0619	0.0621	0.0616
	0.0392	0.0719	0.0283	0.0630	0.0339	0.0281
	0.0116	0.0413	0.0535	0.0453	0.0164	0.0162
	0.0083	0.0121	0.0404	0.0446	0.0263	0.0118
	0.0360	0.0511	0.0596	0.0947	0.0781	0.0496
	0.0669	0.0680	0.0675	0.0893	0.1165	0.1232
π	0.0508	0.1567	0.2747	0.3848	0.0868	0.0462
	0.0222	0.0322	0.0356	0.0408	0.0192	0.0181
Π	0.0251	0.0099	0.0063	0.0031	0.0031	0.0032
	0.0128	0.0812	0.0063	0.0406	0.0094	0.0063
	0.0032	0.0491	0.1461	0.0636	0.0064	0.0063
	0.0032	0.0066	0.1032	0.2300	0.0353	0.0064
	0.0032	0.0066	0.0095	0.0412	0.0198	0.0064
	0.0032	0.0033	0.0032	0.0063	0.0127	0.0176
	0.0167	0.0058	0.0043	0.0030	0.0031	0.0032
	0.0064	0.0257	0.0045	0.0096	0.0052	0.0043
	0.0031	0.0102	0.0307	0.0132	0.0044	0.0044
	0.0031	0.0046	0.0135	0.0379	0.0102	0.0044
	0.0032	0.0046	0.0053	0.0107	0.0101	0.0047
	0.0032	0.0033	0.0032	0.0043	0.0063	0.0121

π, Π, X and L. Tables 2 and 3 show the posterior mean and standard deviation for each element of the matrices P, π, Π, X and L. Only the upper triangle of X and L are shown, the elements below the diagonal being redundant ($X_{ij} = X_{ji}$ and $L_{ij} = -L_{ji}$).

Table 4 gives, for the circulation in states 2–5 and the total circulation, the posterior mean and standard deviation, and the posterior probability that the circulation is negative. Fig. 5 displays a histogram for the posterior sample of total circulation. There is strong evidence of negative total circulation and negative circulation through state 3, and

Table 3
Posterior mean and standard deviation of the elements of X and L in the gasoline mark-up example

X	0.0251	0.0114	0.0047	0.0032	0.0032	0.0032
		0.0812	0.0277	0.0236	0.0080	0.0048
			0.1461	0.0834	0.0080	0.0048
				0.2300	0.0383	0.0063
					0.0198	0.0096
						0.0176
	0.0167	0.0053	0.0029	0.0023	0.0024	0.0025
		0.0257	0.0061	0.0058	0.0038	0.0030
			0.0307	0.0121	0.0037	0.0029
				0.0379	0.0094	0.0034
					0.0101	0.0046
						0.0121
L	0	−0.0015	0.0016	−0.0000	−0.0000	−0.0000
		0	−0.0214	0.0170	0.0014	0.0015
			0	−0.0198	−0.0016	0.0016
				0	−0.0029	0.0001
					0	−0.0031
						0
	0	0.0030	0.0024	0.0020	0.0020	0.0020
		0	0.0050	0.0049	0.0031	0.0024
			0	0.0056	0.0032	0.0025
				0	0.0046	0.0027
					0	0.0030
						0

Table 4
Posterior moments for the circulation through individual states and for total circulation, in the gasoline mark-up example

State	Posterior mean	Posterior standard deviation	Posterior probability of negative circulation
2	−0.0015	0.0030	0.6986
3	−0.0198	0.0051	1.0000
4	−0.0029	0.0048	0.7310
5	−0.0031	0.0030	0.8685
Total	−0.0272	0.0091	0.9985

some evidence of negative circulation through states 2, 4 and 5. Results not reported suggest that the strong evidence in favor of negative total circulation is robust to reasonable perturbations of the bin limits, and to the number of bins, as long as there are at least four.

This evidence is consistent with previous observations that prices tend to rise sharply and decline gradually. It also suggests that total circulation measures a type of time reversibility that is empirically relevant.

The second empirical example investigates time irreversibility of investment growth in the U.S.A. We have 205 quarterly observations, from the first quarter of 1947 to the first

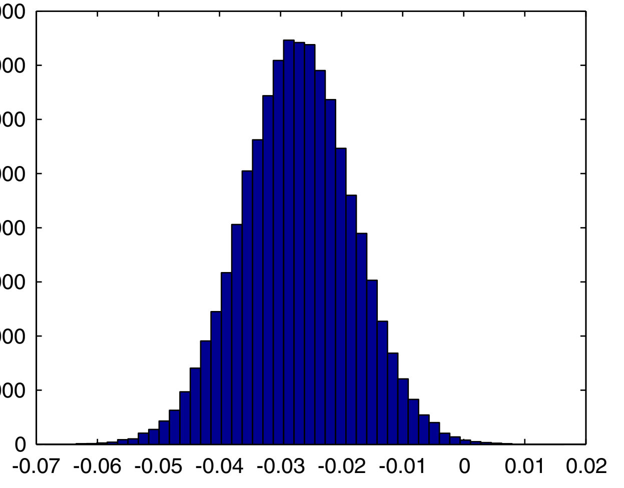

Fig. 5. Posterior histogram of total circulation for gasoline mark-up example.

quarter of 2004, of real gross private domestic investment[5] I_t. We construct the series $g_t \equiv \log(I_t/I_{t-1})$ of investment growth. This is the same series used in Clements and Krolzig (2003) in their investigation of business cycle asymmetries, updated to the first quarter of 2004.

We use one of the models in Clements and Krolzig (2003), originally described in Hamilton (1989), for g_t. There is a latent stationary regular Markov state sequence $\{s_t\}$ on the state set $\{1, \ldots, m\}$, with transition matrix P. The process $\{g_t\}$, conditional on $\{s_t\}$, is given by

$$g_t - \mu_{s_t} = \sum_{i=1}^{p} \phi_i(g_{t-i} - \mu_{s_{t-1}}) + \varepsilon_t, \tag{11}$$

$$\varepsilon_t \sim \text{ i.i.d. N}(0, h^{-1}). \tag{12}$$

The quantities $\mu \equiv (\mu_1, \ldots, \mu_m)$, h, P and $\phi \equiv (\phi_1, \ldots, \phi_p)$ are unknown parameters.

Following Clements and Krolzig, we set $m = 3$ and $p = 4$. We complete the model with a prior on μ, h, P and ϕ where the parameters are independent. μ has a truncated multivariate normal distribution. The untruncated distribution has independent μ_2, $\mu_3 - \mu_2$ and $\mu_2 - \mu_1$, with means[6] 0.01, 0.04 and 0.04 and standard deviations 0.02, 0.04 and 0.04. Truncation is to the set $\{\mu : \mu_1 < \mu_2 < \mu_3\}$. The prior for h has $0.0004 \cdot h \sim \chi^2(1)$. The rows of P are independent, with the following Dirichlet distributions:

$$(P_{11}, P_{12}, P_{13}) \sim \text{Di}(2, 2, 1), \tag{13}$$

$$(P_{21}, P_{22}, P_{23}) \sim \text{Di}(1, 3, 1), \tag{14}$$

[5] I_t is in billions of chained 2000 dollars, seasonally adjusted, at an annual rate. The source is the FRED II (Federal Reserve Economic Data) database (www.stls.ftb.org/fred/data/gdp.html) and the series ID is GPDIC96.

[6] $\mu_2 = 0.01$ means an average 1% growth, at an annual rate with continuous compounding, in the moderate growth rate state.

$$(P_{31}, P_{32}, P_{33}) \sim \mathrm{Di}(1, 2, 2). \tag{15}$$

ϕ has a truncated multivariate normal distribution. The untruncated distribution has independent ϕ_i, with $\phi_1 \sim \mathrm{N}(0, 1)$ and $\phi_i \sim \mathrm{N}(0, 0.25)$ for $i \in \{2, 3, 4\}$. Truncation is to the stationary region.

We note that $y_t - \mu_{s_t}$ is stationary and Gaussian and therefore time reversible, so any time irreversibility of y_t must come from s_t. We also note that the density for P is invariant to the relabelling of states 1 and 3, so that the prior is neutral with respect to the direction of cycles in the L matrix.

Using the BACC software, we generate a posterior sample of 100,000 draws of all parameters. The software implements the simulation methods described in Chib (1996), in which $\{s_t\}_1^T$ is drawn as a single Gibbs block. We then construct posterior samples for π, Π, X and L. Table 5 shows the posterior mean and standard deviation of the elements of P, π, Π, X and L. Fig. 6 displays a histogram for the posterior sample of total circulation. The posterior probability of negative total circulation is 0.7519.

Testing time reversibility would be desirable. One possible test in a Bayesian context involves comparing, using Bayes factors, two models for $\{s_t\}$: one with a prior marginal density $f(X)$ and $L = 0$ with probability 1 and the other with the *same* density $f(X)$ and a prior conditional density $f(L|X)$. Such a test is conceptually simple, and since both models imply exactly the same prior distributions for π, the state persistence P_{ii} in every state and the total flow $\Pi_{ij} + \Pi_{ji}$ between any two states (recall Result 3.2), the test is very suitable. However, usual simulation methods require evaluation of the prior $f(X, L)$, and evaluation of the normalization factor of $f(L|X)$, which depends on X, is not always easy, even for uniform $f(L|X)$. For this reason, the issue of testing for time reversibility is left to another paper, currently under preparation.

5. Conclusions

We have introduced the parameterization of a stationary regular finite-state Markov chain by its joint probability matrix Π, proposed a decomposition of Π into its reversible

Table 5
Posterior mean and standard deviation of the elements of P, π, Π, X and L in the investment growth example

P	0.2953 (0.1253)	0.2864 (0.1702)	0.4182 (0.1627)
	0.0442 (0.0297)	0.9020 (0.0882)	0.0538 (0.0805)
	0.2104 (0.1163)	0.3061 (0.1449)	0.4835 (0.1493)
π	0.0876 (0.0335)	0.7676 (0.1004)	0.1448 (0.0875)
Π	0.0273 (0.0196)	0.0245 (0.0171)	0.0358 (0.0184)
	0.0328 (0.0181)	0.6989 (0.1312)	0.0359 (0.0423)
	0.0275 (0.0180)	0.0442 (0.0372)	0.0732 (0.0618)
X	0.0273 (0.0196)	0.0286 (0.0158)	0.0316 (0.0164)
		0.6989 (0.1312)	0.0400 (0.0391)
			0.0732 (0.0618)
L	0 (0)	−0.0041 (0.0079)	0.0041 (0.0079)
		0 (0)	−0.0041 (0.0079)
			0 (0)

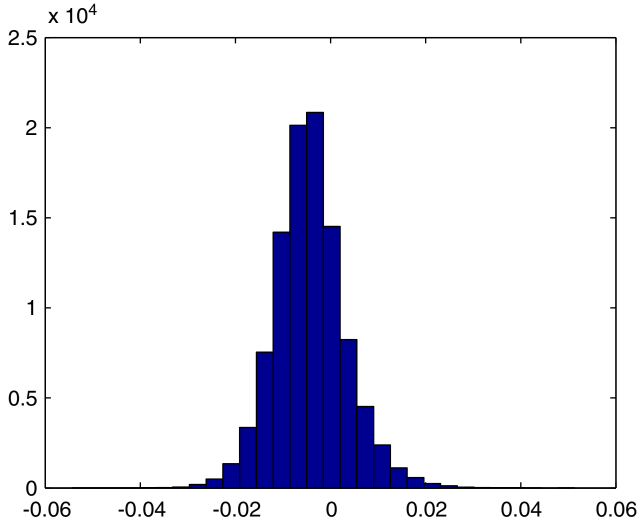

Fig. 6. Posterior histogram of total circulation for the investment growth example.

part X and irreversible part L, and suggested an index, total circulation, describing a certain kind of irreversibility in chains whose states are naturally ordered. Empirical examples illustrate the use of these quantities, for both directly observed and latent chains. The results suggest that total circulation is an empirically relevant quantity.

In these examples, uncertainty about the dynamics of the Markov chain is expressed as a distribution over the Markov transition matrix P. Providing instead a marginal prior for X and a conditional prior for $L|X$ would have many advantages:

1. A prior for X is easy to elicit. For example, one can choose a Dirichlet distribution (or a mixture of such) for the following vector:

$$\theta_X \equiv (X_{11}, 2X_{12}, \ldots, 2X_{1m}, X_{22}, 2X_{23}, \ldots, 2X_{2m}, \ldots, X_{mm}). \tag{16}$$

2. The implied prior for π depends only on the prior for X (recall Result 3.2) and the fact that π is a linear function of X ($\pi = \iota' X$)[7] means that first and second moments of π may be easy derived from the first two moments of X. The mean is particularly transparent: $\mathrm{E}[\pi] = \iota'\mathrm{E}[X]$. Consider the difficulty of finding moments of π given a prior on P.
3. Testing for reversibility of a stationary regular finite-state Markov chain $\{s_t\}$ can be very disciplined, as we have seen.
4. We can choose truncated priors for $L|X$ to impose restrictions such as the direction of total circulation C or a common sign on the circulation through all interior states. We can also choose priors that are neutral about the direction of circulation without resorting to the extreme form of symmetry used in this paper.

However, elicitation of $L|X$ is not easy. Even with a flat prior, computing the normalization factor for $f(L|X)$, which we seem to need[8] for posterior simulation since

[7] ι is an $m \times 1$ vector of ones.

[8] We have a paper in preparation which shows that we do not in fact need this normalization factor.

it depends on X, is no simple feat. A paper in preparation shows how we can elicit priors on $L|X$ and therefore realize the above advantages.

Acknowledgments

We thank FQRSC (Fonds québécois de la recherche sur la société et la culture) for financial support, Simon Leblond for excellent research assistance, and Andrew Eckert for providing us with data. John Geweke, two anonymous referees and an associate editor provided many helpful comments. Computations reported in this paper were undertaken using the Bayesian Analysis, Computation and Communication (BACC) software described in Geweke (1999) and McCausland (2004b). Additional software for computations reported in this paper is available from the author. We alone are responsible for errors.

References

Chen, Y.-T., Kuan, C.-M., 2001. Time irreversibility and EGARCH effect in US stock index returns. In: Proceedings of the International Conference on Modelling and Forecasting Financial Volatility.

Chen, Y.-T., Chou, R.Y., Kuan, C.-M., 2000. Testing time reversibility without moment restrictions. Journal of Econometrics 95, 199–218.

Chib, S., 1996. Calculating posterior distributions and modal estimates in Markov mixture models. Journal of Econometrics 75, 79–97.

Clements, M.P., Krolzig, H.-M., 2003. Business cycle asymmetries: characterization and testing based on Markov-switching autoregressions. Journal of Business and Economic Statistics 21, 196–211.

Eckert, A., 2002. Retail price cycles and response asymmetry. Canadian Journal of Economics 35, 52–77.

Eckert, A., 2003. Retail price cycles and presence of small firms. International Journal of Industrial Organization 21, 151–170.

Fong, W.M., 2003. Time reversibility tests of volume-volatility dynamics for stock returns. Economics Letters 81, 39–45.

Geweke, J., 1999. Using simulation methods for Bayesian econometric models: inference, development, and communication. Econometric Reviews 18, 1–126.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357–384.

Hinich, M.J., Rothman, R., 1998. A frequency domain test of time reversibility. Macroeconomic Dynamics 2, 72–88.

Iosifescu, M., 1980. Finite Markov Processes and Their Applications. Wiley, New York, NY.

Maskin, E., Tirole, J., 1988. A theory of dynamic oligopoly II: price competition, kinked demand curves and Edgeworth cycles. Econometrica 56, 571–599.

McCausland, W.J., 2004a. Time reversibility of stationary regular finite state Markov chains. Cahiers de recherche du Département de sciences économiques. Université de Montréal, no. 2004-06.

McCausland, W.J., 2004b. Using the BACC software for Bayesian inference. Journal of Computational Economics 23, 201–218.

Noel, M., 2003. Edgeworth price cycles, cost-based pricing and sticky pricing in retail gasoline markets. Working Paper, University of California, San Diego.

Peng, L., Xiong, W., 2003. Time to digest and volatility dynamics. Working Paper, Princeton University.

Ramsey, J.B., Rothman, P., 1996. Time irreversibility and business cycle asymmetry. Journal of Money, Credit, and Banking 28, 3–20.

Robinson, P.M., 1991. Consistent nonparametric entropy-based testing. Review of Economic Studies 58, 437–453.

Scott, S.L., 2002. Bayesian methods for hidden Markov models: recursive computing in the 21st century. Journal of the American Statistical Association 97, 337–351.

Sims, C.A., 2003. Implications of rational inattention. Journal of Monetary Economics 50, 665–690.