Econometrics Journal (2004), volume **7**, pp. 528–549.

Testing for duration dependence in economic cycles

JONATHAN OHN*, LARRY W. TAYLOR[†§] AND ADRIAN PAGAN[‡]

**Department of Business Administration, Wagner College, One Campus Road, Staten Island, NY 10301, USA*

[†]College of Business and Economics, Rauch Business Center, Lehigh University, Bethlehem, PA 18015-3117, USA

[‡]Economics Program, Research School of Social Sciences, The Australian National University, Canberra ACT 0200, Australia

Received: October 2002

Summary In this paper, we discuss discrete-time tests for duration dependence. Two of our test statistics are new to the econometrics literature, and we make an important distinction between the discrete and continuous time frameworks. We then test for duration dependence in business and stock market cycles, and compare our results for business cycles with those of Diebold and Rudebusch (1990, 1991). Our null hypothesis is that once an expansion or contraction has exceeded some minimum duration, the probability of a turning point is independent of its age—a proposition that dates back to Fisher (1925) and McCulloch (1975).

Keywords: Duration dependence, Discrete time, Business cycles, Stock market cycles.

1. INTRODUCTION

Survival analysis has been applied to study events such as the length of unemployment spells, wars, marriages, lifetimes of firms, birth intervals, time until the adoption of a new technological innovation and business cycles. One of the main themes has been to examine the question of whether the probability of exiting the state of interest depends upon how long one has spent in it. If it does, we say that there is duration dependence. There is quite a bit of empirical work on duration dependence in the business cycle, largely motivated by the question of whether it is possible to predict the termination of a boom or a recession. Fisher (1925) was one of the first investigators to consider this question, raising the issue of whether the probability of exiting any phase of the cycle is just a constant, as might be expected to happen when the series underlying the business cycle was not serially correlated.

Casual observation from the experience of the 1980s and 1990s might lead us to infer that long expansions may be inclined to continue in that state, while long contractions are very likely to terminate. Contrary evidence might be the long and frustrating Great Depression. One might even hold to McCulloch's (1975) view, as summarized by Niemira (1991), that, once an expansion

[§]Corresponding author.

or contraction has exceeded its historical minimum duration, the probability of a turning point is indeed independent of its age. Thus findings, such as those by Diebold and Rudebusch (1990, 1991), that there is evidence of duration dependence in U.S. business cycles, have attracted quite a bit of attention and are often cited.

Economic cycles are normally described by binary random variables taking the values of unity and zero, with unity indicating a state of expansion (say) and zero as a state of contraction. Hence this paper discusses tests for duration dependence that might be applied to the analysis of such binary data, either in its original form or after some aggregation. We will refer to the states distinguished by the binary outcomes for S_t as *phases* and the sample of data available to us might be $S_t,\ t = 1, \ldots, T$. Often, however, it comes in an aggregated form as the time spent in each phase, i.e. as *duration data*. Thus, if we divide the T observations into n phases the duration of time spent in the ith phase will be designated as X_i. We can then formally define any *duration dependence* within a given phase in one of two ways. In the first we focus upon the continuation probability $\Pr(S_t = j|S_{t-1} = j)$ and ask if this probability provides a complete description of the process within the phase. If so, the process will be first-order Markov and we will have *duration independence*. When data are discrete, the time spent in the jth phase up to $t - 1$ is just the sum of past S_t, so that it is clear that duration dependence is a statement that $\Pr(S_t = j|S_{t-1} = j) \neq \Pr(S_t = j|S_{t-1} = j, S_{t-2} = j, \ldots, S_0 = j)$. In the second approach the implications of duration dependence for the density of the X_i are derived and then a comparison would be made of the density of the data on X_i with that expected under duration independence.

In Section 2, we present three tests for duration dependence based on using either the S_t or the X_i. In Section 3, we then apply the tests to the analysis of U.S. business and stock market cycles, while in Section 4 we summarize our findings.

2. TESTING FOR DURATION DEPENDENCE

2.1. Some basic considerations

Consider a random sample of n observations $(X_1, X_2, \ldots, X_n)$ from a continuous distribution F, such that $F(a) = 0$ for $a < 0$. We assume that the X_i are duration data and they represent the time spent in one of two phases. An example of the latter would be an expansion or contraction of the business cycle. Let the density function of the random variable underlying the duration data be $f(x)$. Then the hazard rate function is defined as

$$h(x) = f(x)/G(x), \tag{1}$$

where $h(x)$ is the hazard (or failure) rate and $G(x) = \mathrm{P}(X \geq x)$ is known as the survival function. For a small Δ, $h(x)\Delta$ is the probability that the expansion will terminate during the interval $(x, x + \Delta)$ given it has lasted until time x. If there is to be no duration dependence, then the hazard rate must be constant, regardless of the duration of time spent in the phase. Consequently, the hazard rate does not depend on x, i.e.

$$H_0 : h(x) = \theta \quad \text{for some } \theta > 0 \text{ and all } x > 0. \tag{2}$$

For a continuous random variable there is only one density for $f(x)$ that satisfies the constant-hazard assumption in *continuous time*, namely, the *Exponential density.* Given this fact it is possible to write an alternative null hypothesis, namely, $H_0 : F$ is Exponential, and this is the key to many tests for duration dependence.

However, the duration data that are available on economic cycles is invariably discrete since statistics on economic variables are collected at discrete intervals of time, e.g. business cycle duration data come at intervals that are no shorter than a month. Although stock market data are potentially available at much shorter intervals than a month, the cycles that are of interest tend be on a monthly frequency and so we will use such data in the analysis that follows.

With discretely measured duration data the density of X_i is a geometric density when there is duration independence. For a geometric density $P(X = x) = (1 - p)^x p$ for $0 \leq x < \infty$. Define the hazard function as

$$h(x) = P(X = x)/P(X \geq s). \tag{3}$$

It is clear that this becomes $h(x) = p$ in the case of a geometric density, i.e. there is a constant hazard. The first two moments of the geometric density are $E(X) = (1 - p)/p$ and $V(X) = (1 - p)/p^2$ and this leads to the equality $V(X) - [E(X)]^2 - E(X) = 0$, a restriction that is in contrast to the relation between the first and second moments that holds for an exponentially distributed random variable, namely, $V(X) - E(X)^2 = 0$. It follows that one must exercise some care when using moments for testing for duration dependence with discretely measured data. If one used the moment relations that come with the exponential rather than those appropriate to the geometric density, the test statistics would be incorrectly centred, although the error may be small if p is either large or small.[1]

Another difference to the continuous time case is that, in discrete time, we can focus directly on the probability that a phase terminates within a given hour, day, week, month or year—depending on our unit of measurement.[2] For the empirical examples that follow in Section 3, we will measure the contractions and expansions in business and stock market cycles on a monthly basis. We will say that expansions exhibit a constant hazard if the hazard rate remains constant from month to month, making it independent of the duration of the expansion. As we have mentioned above, there is only one distribution that satisfies the constant-hazard assumption in discrete time, namely, the *Geometric density*. Therefore, we can alternatively write the null hypothesis as, H_0 : F is Geometric.

The first set of tests for duration dependence check whether the implied density is compatible with the data, i.e. the question asked is whether the sample X_i is compatible with $f(x)$ being of a particular form. We might term these *consistency tests,* as they check to see if the data are consistent with the null hypothesis. Exactly how one should check for consistency varies a good deal. A *weak-form* test might involve checking if the relations between selected moments of the random variable X implied by a geometric density are satisfied. A *strong-form* test would compare some nonparametric estimate of a density from the data with the hypothesized parametric form, e.g. the geometric. These would be *direct* tests of the hypothesis. Equivalent tests are available by examining hypotheses that are derived from any one-to-one transformation of the density function. Thus, in our context, it is natural to ask if the hazard function is a constant, rather than that the density is geometric. These consistency tests can be very useful when it comes to a graphical display of what the data says about duration dependence, e.g. a plot of a nonparametric estimate of $h(x)$ may indicate whether an assumption of constancy is a reasonable one. Moreover, they can

[1]When p is small or large the square of $E(X)$ should be close to $V(X)$.

[2]In continuous time, the probability that a phase terminates at a specific point in time is zero, while in discrete time we can mass the probability at integer values. As such, in the discrete case the hazard function, $h(x)$, is technically not a 'rate'.

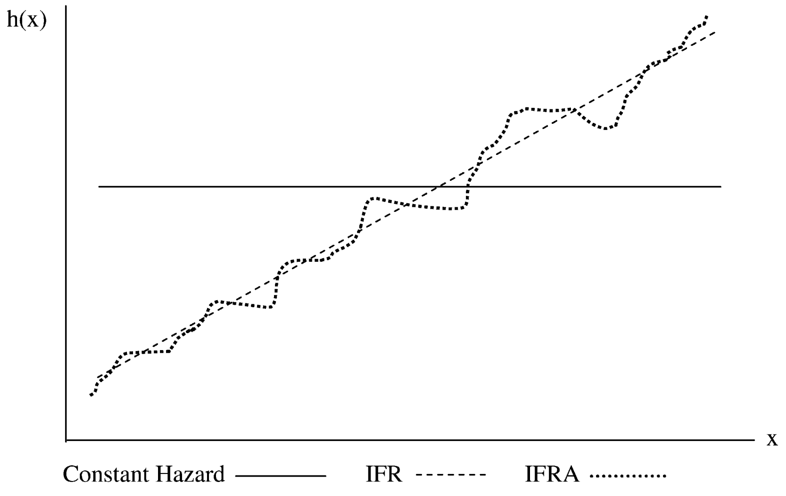

Figure 1. Hazard functions for various life distributions.

be very informative about why a test statistic rejects the null hypothesis. The latter information is useful in assessing the appropriate response to any rejection.

2.2. Alternative hypotheses

Most consistency tests are sensitive to specific directions of departure from the null hypothesis. Indeed, we may employ selected tests to search for departures in a given direction. This leads one to ask which types of duration dependence we might expect in phases and what the distributions might look like for those alternatives. In most cases they are motivated by physical analogues and are intended to suggest what functions of the data one might focus upon in the event that this particular type of dependence was present. From the discussion in Hollander and Wolfe (1999), the two foremost alternatives are:

(1) *Increasing (Decreasing) Failure Rate: IFR (DFR)*;
(2) *Increasing (Decreasing) Failure Rate Average: IFRA(DFRA).*

To illustrate the various alternatives, Figure 1 provides examples of hazard functions from IFR and IFRA *life distributions*.[3] Because there is a one-to-one relationship between the hazard function and the probability density function (PDF) a comparison of hazard rates is generally a natural way of analyzing the nature of the probabilities of failure, more so than a comparison of density functions. This is particularly true given the nature of the underlying null hypothesis.

In Figure 1, the constant hazard corresponds to the exponential density. In the constant hazard case, expansions (say) are such that new ones are no more (or less) likely to terminate than mature ones. In contrast, if expansions are IFR, then the hazard (or failure) rate is never decreasing, and our illustrated IFR hazard function implies an ever more likely chance of termination (or mortality). Yet, this is not the only type of hazard function that has a tendency to rise. While our depicted

[3]A so-called life distribution is one that places all of its probability on nonnegative values, and the term 'life distribution' is coined from the study of mortality, where many of these distributions were first employed.

IFRA hazard function has periods of decline, it is clear that IFRA exhibits the same overall upward trend as IFR. In particular, IFRA initially increases at a faster rate than IFR, whereas there are periods where IFRA is lower than IFR. One example would be if the hazard rate for an IFRA expansion fluctuates due to seasonal variation, even though it increases on average over the life of the expansion. It might also be the case that a recession could initially have a decreasing hazard rate for a short period, but then exhibit an increasing hazard rate over the majority of the phase due to (say) federal intervention. Thus policy might be late to perceive the recession but, once recognized, effective action is taken.

It is easy to see that any IFR distribution is also an IFRA distribution, although the converse is not true. While the IFRA class is fundamentally like the IFR class, the former relaxes the strict assumptions associated with IFR. A completely analogous situation holds for DFR and DFRA distributions.

The geometric and exponential densities are central to the study of duration analysis. If we find that durations do *not* follow the geometric density, then there must be some predictability in the length of the durations and the timing of turning points. It is the presence or absence of such predictability that defines *duration independence*. *Positive duration dependence* means that (say) an expansion is more likely to end as its duration increases and so corresponds to the IFR and IFRA classes of life distributions. The opposite pattern, *negative duration dependence*, means that the given state (either a contraction or expansion) tends to persist, and so corresponds to the DFR and DFRA classes of life distributions.

As mentioned earlier the geometric and exponential densities exhibit relations between the first and second moments. Thus, for the exponential density $V(X) = E(X)^2$. On the one hand, when the average length of expansions is *greater* than the standard deviation, this is taken to be evidence of *positive* duration. On the other hand, if the average length of expansions is *less* than the standard deviation, this is evidence of *negative* duration dependence.

2.3. Weak-form tests in discrete time

The basic test employed under this heading is that coming from the moment condition implied by the geometric density, i.e. $V(X) - [E(X)]^2 - E(X) - \gamma = 0$. One can test if $\gamma = 0$ using the GMM of γ from this moment condition. This was done by Mudambi and Taylor (1995). They determined whether $\mathrm{MT} = [(1/T)\sum(x_i - \bar{x})^2] - \bar{x}^2 - \bar{x}$ was significantly different from zero. We expect MT to be especially sensitive to IFRA alternatives since it is completely analogous to the continuous time test based on $V(X) - (E(X))^2 = 0$ designed for such alternatives.[4] While the MT test statistic is asymptotically centred at zero and can be standardized so that it is asymptotically $N(0,1)$, Mudambi and Taylor (1995) found that its distribution is highly skewed in finite samples, and thus it was necessary to use simulations to obtain finite-sample critical values.

In spite of the skewed distribution, a primary advantage of MT (once standardized) is that it is asymptotically pivotal, i.e. asymptotically it does not depend on unknown parameters. For asymptotically pivotal statistics, the bootstrapped critical values are generally more accurate than those based on first-order asymptotic theory. Horowitz (2001) provides a highly readable account of why it is generally desirable to use pivotal statistics when bootstrapping.

[4]See Lee *et al*. (1980) for more details on this.

It is also possible to implement a simple regression-based test that is closely related to MT.[5] We first define a state variable S_t that is assigned unity if the observed index is a month of expansion, and zero for a contraction. For constant-hazard expansions and contractions, S_t is a Markov process, and Hamilton (1989) shows that it can be written as an AR(1):

$$S_t = c_0 + c_1 S_{t-1} + \eta_t, \tag{4}$$

where η_t is a disturbance term for which $E_{t-1}(\eta_t) = 0$. Moreover, $c_0 = p_{1|0} = 1 - p_{0|0}$ and $c_1 = p_{1|1} + p_{0|0} - 1$, where $p_{a|b}$ is the conditional probability of moving to state 'a' from state 'b'. Consequently, duration dependence could be thought of as shifts in c_0 and/or c_1, with the shifts being related to how long one has been in a given state.

Consider the regression equation:

$$S_t = c_0 + c_1 S_{t-1} + c_2 S_{t-1} d_{t-1} + \text{error}, \tag{5}$$

where d_t is the number of consecutive months (i.e. 'duration') spent in an expansion up and through time t and the error term has a density such as to make realizations of S_t either zero or unity. A simple variable-addition test of duration dependence in expansions is therefore available by testing the null hypothesis $H_0 : c_2 = 0$.[6] In general, we would expect the actual relation between S_t and d_{t-1} to be a nonlinear one, as in the Durland and McCurdy (1994) logistic formulation. However, since we are performing a diagnostic test for duration dependence rather than trying to estimate its form, the linear model at (5) is a suitable vehicle for doing that, although the power of the test might be affected.[7]

We term this the SB test to indicate that it is based upon the states S_t rather than durations. In the Appendix, we show that the SB test based upon $\hat{c}_2$ is effectively checking if $V(X) - [E(X)]^2 + E(X)$ is zero. At first sight this seems to be inconsistent with the moment implications of a geometric density given earlier, but the resolution of the difference comes from noting that, by the definition of the binary indicator, the duration of any expansion or contraction must be at least one period. Hence, it is not possible to get a duration of zero periods when using the S_t. This suggests that we should check the relationship between the moments when the geometric density is left-censored at unity. Now the censored probability function will be $\mathrm{P}(X = x) = (1 - p)^{x-1} p$ for $1 \le x < \infty$ and $E(X) = 1/p$, $V(X) = (1 - p)/p^2$. Thus, in the censored geometric case, the relation between moments is $V(X) - E(X)^2 + \mathrm{E}(X) = 0$, agreeing with what SB tests.

The SB test has a number of attractions. First, it focuses directly upon the conditional probabilities and what influences them. Second, it involves a regression and so it is generally easy to explain the outcome to nonspecialists. Third, it can be used to examine prediction issues. Finally, since the parameters can be recursively estimated we can study how duration dependence might have changed over time.

Some care, however, must be used in implementing the SB test. If one constructs the test with a sample consisting of *both* contractions and expansions, the disturbance in the regression

[5]This material was originally presented as part of the third author's Walras-Bowley Lecture 'Bulls and Bears: A Tale of Two States', delivered to the Summer Meetings of the North American chapter of the Econometric Society in Montreal in 1998.

[6]For contractions one replaces S_t by $(1 - S_t)$ and d_t will need to be the duration of contractions.

[7]Durland and McCurdy tested for duration dependence in a latent state process that was taken as driving the growth rate in GDP. Designating these latent states by z_t they test for variation in the transition probabilities. However, it is rarely the case that $z_t = S_t$. Thus even if there is no duration dependence in the z_t process there can be in the S_t. In contrast, duration dependence in the latent states would almost always imply duration dependence in S_t.

at equations (4) or (5) is conditionally heteroskedastic. In fact, $V(\eta_t\ |S_{t-1} = 0) = p_{1|0}\ (1 - p_{1|0})$ and $V(\eta_t|S_{t-1} = 1) = p_{1|1}(1 - p_{1|1})$. One then needs to account for that when forming the test, and it is unclear whether the robust measures that are used in most regression packages will work as well with the binary random variables in the SB regression. We can eliminate this problem by operating separately on expansions and contractions since the disturbance is then homoskedastic.

As an example, consider the following string for S_t: 1,1,1,0,0,1,1,1,0,0,1,1,0. We are in a period of expansion for observations 1–3, 6–8 and 11–12. The string for S_{t-1} is as follows: ?, 1,1,1,0,0,1,1,1,0,0,1,1. Arranging S_t, S_{t-1} and d_{t-1} in matrix form gives:

S_t	S_{t-1}	d_{t-1}
1	?	0
1	1	1
1	1	2
0	1	3
0	0	0
1	0	0
1	1	1
1	1	2
0	1	3
0	0	0
1	0	0
1	1	1
0	1	2

For expansions only, our (half-cycle) data matrix is as follows:

S_t	S_{t-1}	d_{t-1}
1	1	1
1	1	2
0	1	3
1	1	1
1	1	2
0	1	3
1	1	1
0	1	2

It is clear that we will lose some expansion points. We ignore the first observation as well as any observations that are part of incomplete (i.e. censored) spells. For business cycles, the censored observations will invariably be those towards the end of the data matrix. Our resulting econometric sample size is denoted as T, with $T - n$ points for which $S_t = 1$ and 'n' turning points for which $S_t = 0$.

Having taken the subsample for which $S_{t-1} = 1$, we thus employ the simplified regression equation:

$$S_t = b_0 + b_1 d_{t-1} + \text{disturbance}, \tag{6}$$

where, again, d_t is the number of consecutive months (i.e. 'duration') spent in an expansion up and through time t. The test of duration dependence in expansions is obtained by testing the null hypothesis $H_0 : b_1 = 0$. An important consideration is that the standard t-test from this regression is asymptotically pivotal, and thus it is this test statistic that we wish to bootstrap.

It is also the case that the degree of censoring used in constructing the states can create difficulties. It is well known that the NBER business cycle dates have a minimum phase restriction of two quarters and it is equally clear that equivalent dates for other cycles almost always involve some such constraint. There are two ways of dealing with this problem. First, if there is a two-quarter minimum to the phase restriction, then it can be shown that the states must obey the following model—see Harding and Pagan (2003):

$$S_t = c_0 + c_1 S_{t-1} + c_2 S_{t-2} + c_3 S_{t-1} S_{t-2} + \text{disturbance}, \tag{7}$$

and so the duration dependence test would involve trying to add on $S_{t-1}d_{t-1}$to this regression rather than to the first-order one at (5). Alternatively we can subtract unity from each of the observed durations and then implement the SB test. The two-quarter minimum phase is then automatically satisfied and we can implement the test at (5). So, for example, if we observe a phase lasting 8 quarters, we transform the observation to 7 quarters and likewise for the other observations.

2.4. Strong-form density-based tests in discrete time

An obvious choice for testing if the empirical density of the X_i is a geometric density is the chi-square goodness-of-fit test. The test statistic is $\chi^2 = \sum_{j=1}^{K} [(O_j - E_j)^2 / E_j]$, where O_j is the observed number of elements in the jth bin and E_j is the expected number of elements in the jth bin under the geometric density. Using simulated critical values based on 5,000 replications, Diebold and Rudebusch (1991) used χ^2 to shed some light on duration dependence in business cycles. They varied the number of bins (K) from 2 to 5 in order to provide a sensitivity analysis. As constructed, χ^2 is asymptotically pivotal. Yet, while the size of χ^2 is properly controlled, construction of the test in this manner may result in low power. With this in mind, it is notable that, except for pre-war expansions, Diebold and Rudebusch (1991) find only weak evidence of duration dependence.

Our approach to bin selection is somewhat different. A well-known rule-of-thumb is that the expected frequency (E_j) should be at least 5 (or perhaps 6) for all bins. To be on the safe side we use 6 rather than 5.[8] Since ideally the bootstrapped critical values provide a higher-order approximation to the first-order asymptotic critical values, it appears reasonable to keep to this rule. In keeping with the rule, it is also possible to use the asymptotic critical values for comparison.

As normally practiced, suppose that $E_1, \ldots, E_K$ are ordered such that E_1 corresponds to the bin with the lowest values for the realizations of the random variable. E_K then corresponds to the

[8]The rule-of-thumb concerning the chi-square test dates back many years and is discussed by Hoel (1954). He states that experience and theoretical investigations indicate that the goodness-of-fit test based on the chi-square approximation is satisfactory when the number of cells and expected frequency within each cell is at least 5. The expected frequency should be somewhat larger than 5 when the number of cells is less than 5.

Table 1. Business Cycle Summary Statistics.

Sample	Mean duration	Standard deviation	Sample size
	Expansions		
Entire sample	34.6	21.8	31
Entire sample, excluding wars	28.9	15.3	26
Post-WWII	48.6	28.9	9
Post-WWII, excluding wars	40.9	22.3	7
Pre-WWII	26.5	10.7	21
Pre-WWII, excluding wars	24.5	9.2	19
	Contractions		
Entire sample	18.1	12.5	30
Post-WWII	10.7	3.4	9
Pre-WWII	21.2	13.6	21

bin with the highest values for the realizations. For example, E_1 may correspond to the interval '[0, 3]' which includes all values from 0 to 3, inclusive, while E_K corresponds to the open interval '>15'. Due to the small sample sizes encountered in macroeconomic data, we take a conservative approach and construct our intervals such that $E_1, \ldots, E_{K-1}$ approach the value 6 from the right. If the residual bin has an expected frequency less than 5, $E_K < 5$, we combine bins K and $K-1$. Simulated critical values are constructed in accordance with this rule. Of course, the construction of the bins is dependent on the estimated hazard rate and thus our χ^2 test is not truly pivotal. Still, this way of constructing the test seems to work well in practice.

3. EMPIRICAL APPLICATIONS

3.1. Business cycles

To compare and contrast our procedures with earlier work, we replicate the studies of Diebold and Rudebusch (1990, 1991) to determine whether there exist any major discrepancies. Diebold and Rudebusch employ continuous time tests in their 1990 paper and the discrete-time chi-square test in their 1991 paper. As stated above, in this paper we implement the chi-square test in a slightly different manner.

As have others, we consider various so-called minimum phase durations, denoted as τ_0. McCulloch (1975) set τ_0 to be the historical minimum duration, while Diebold and Rudebusch (1990) and Mudambi and Taylor (1991, 1995) set τ_0 to be *at most* the historical minimum. The argument here is that the uncertainty associated with the precise timing of turning points (especially for macroeconomic time series) calls for the examination of various minimum phase durations.[9] But, of course, simply allowing the minimum phase to vary does not overcome the stringent nature of the geometric density. We refer to imposing a minimum phase jointly with the assumption of duration independence as the Markov hypothesis.

[9]Epstein (1960) defines the two-parameter Exponential p.d.f. as $f(t; \theta, \tau_0) = 1/\theta \exp(-(t-\tau_0)/\theta)$ for $t \geq \tau_0 \geq 0$, and $f(t; \theta, \tau_0) = 0$ elsewhere. Here, X is two-parameter Geometric with $P(X = i) = (1-p)^{\tau_0 - 1} p$ $(i = \tau_0 + 1, \ldots)$. To impose a given minimum phase, τ_0, note that $Y = X - \tau_0$ also follows the geometric density. The SB test is designed for $\tau_0 = 1$, and thus we only transform the data for $\tau_0 > 1$.

Table 2. Tests for Constant-Hazard Contractions.

Statistic	All ($N = 30$)	Pre-war ($N = 21$)	Post-war ($N = 9$)
SB ($\tau_0 = 4$)	0.5778	0.3544	0.0402+
SB ($\tau_0 = 5$)	0.8624	0.5172	0.1038
SB ($\tau_0 = 6$)	0.7954	0.7286	0.2804
MT ($\tau_0 = 4$)	0.6230	0.3940	0.0376+
MT ($\tau_0 = 5$)	0.9004	0.5638	0.1080
MT ($\tau_0 = 6$)	0.7850	0.7740	0.3188

The finite-sample two-tailed p-values are obtained through simulation. A plus sign (+) indicates statistically significant positive duration dependence.

Table 3. Tests for Constant-Hazard Expansions.

Statistic	Entire sample ($N = 31$)	Entire sample no wars ($N = 26$)	Post-WWII ($N = 9$)	Post-WWII no wars ($N = 7$)	Pre-WWII ($N = 21$)	Pre-WWII no wars ($N = 19$)
SB ($\tau_0 = 8$)	0.3224	0.1424	0.4148	0.3998	0.0104+	0.0096+
SB ($\tau_0 = 9$)	0.4378	0.2256	0.4628	0.4516	0.0180+	0.0194+
SB ($\tau_0 = 10$)	0.5822	0.3426	0.5166	0.5118	0.0332+	0.0400+
MT ($\tau_0 = 8$)	0.3474	0.2168	0.4798	0.5360	0.0108+	0.0280+
MT ($\tau_0 = 9$)	0.4672	0.3166	0.5318	0.6002	0.0196+	0.0556+
MT ($\tau_0 = 10$)	0.6042	0.4494	0.5938	0.6662	0.0408+	0.1018

The finite-sample two-tailed p-values are obtained through simulation. A plus sign (+) indicates statistically significant positive duration dependence.

Summary statistics for the business cycle data are taken from Diebold and Rudebusch (1990) and replicated in Table 1, while our tests for duration dependence are found in Tables 2–5. To construct the p-values, we generate 10,000 samples of size N from a geometric distribution with the parametric parameter p fixed at the maximum likelihood estimator for an assumed value of τ_0. The p-value for a given statistic is then obtained from the 10,000 ordered realized values.[10]

Consider first the weak-form tests found in Tables 2 and 3. As expected from our theoretical arguments, SB and MT yield similar conclusions. On the one hand we find some evidence that post-war contractions and pre-war expansions fall into the IFRA class of alternatives. On the other hand, there is no evidence of duration dependence in post-war expansions, though we note that the extremely small post-war sample sizes are likely to result in tests with low power. Our findings so far corroborate those of Diebold and Rudebusch (1990).

In contrast, the results based on the strong-form chi-square test presented in Tables 4 and 5 differ a bit from those of Diebold and Rudebusch (1991). Not only do we find strong evidence for duration dependence in pre-war expansions, but there is also strong evidence that pre-war

[10]This parametric bootstrap approach for generating finite-sample critical values (i.e. conditioning on the estimated value of p) was recommended by Sargan and Bhargava (1983) when they tested for a random walk in the errors from a regression equation.

Table 4. Chi-Square Tests for Constant-Hazard Contractions (One-Tailed *p*-Values Are Reported in Parentheses).

Minimum phase								
$\tau_0 = 4$			$\tau_0 = 5$			$\tau_0 = 6$		
Interval	*E*	*O*	Interval	*E*	*O*	Interval	*E*	*O*
Entire sample ($N = 30$)								
[0, 3]	7.206	2	[0, 3]	7.663	6	[0, 2]	6.374	5
[4, 8]	6.625	8	[4, 8]	6.888	8	[3, 6]	6.444	5
[9, 15]	6.171	13	[9, 15]	6.229	9	[7, 12]	6.526	13
>15	9.998	7	>15	9.220	7	>12	10.656	7
χ^2 (2) = 12.50 (0.0023)			χ^2 (2) = 3.34 (0.2411)			χ^2 (2) = 8.30 (0.0173)		
Pre-WWII ($N = 21$)								
[0, 5]	6.028	2	[0, 5]	6.328	3	[0, 5]	6.659	3
[6, 15]	6.453	12	[6, 14]	6.104	11	[6, 14]	6.248	11
>15	8.519	7	>15	8.568	7	>15	8.093	7
χ^2 (1) = 7.73 (0.0042)			χ^2 (1) = 5.97 (0.0155)			χ^2 (1) = 5.77 (0.0164)		
Post-WWII ($N = 9$)								
[0, 7]	6.058	7	[0, 6]	6.115	7	[0, 5]	6.007	7
>7	2.942	2	>6	2.885	2	>5	2.993	2

From the geometric distribution, *E* is the expected number of observations that lie in the stated interval. From the sample, *O* is the observed number in the interval. The finite-sample one-tailed *p*-values are obtained through simulation. The chi-square test is not applied to post-WWII contractions due to the small sample size. The 5% critical value for a chi-square distribution with 2 (1) degree(s) of freedom is 5.99 (3.84).

contractions are duration dependent.[11] The clustering effect of the observed frequencies suggests *positive* duration dependence. As an example, consider pre-war contractions and a minimum phase of 4. The observed frequency for the interval [6, 15] is 12 while the expected frequency is a mere 6.453. The observed frequency for the interval [0, 5] is only 2 while the expected frequency is 6.028. Thus, relative to a geometric density, there are fewer short-term contractions than expected and more intermediate-term contractions than expected. Diebold and Rudebusch (1991) detected only some evidence of duration dependence for contractions when they applied the chi-square test by varying the number of bins from 2 to 5. Their results were strongest for the pre-war period, but, overall, they concluded that the evidence was weak.

To further investigate the nature of the hazard functions, we plot them for pre- and post-war contractions and expansions (including wars) in Figures 2 and 3. The nonparametric hazard rates are computed by the life-table method of Cutler and Ederer (1958) using the computer package LIMDEP 7.0. An approximate (but intuitive) explanation of the procedure is as follows. We first place the contractions in ascending order, and then construct the hazard rate at time '*t*' as the ratio of the number of contractions terminating *at* month '*t*' over the number of contractions lasting for *at least* '*t*' months. That is, we use the sample information to estimate $P(X = t)/P(X \geq t)$. A contraction that terminates is said to exit the sample, while those contractions lasting at least

[11] Except for a minimum phase of 5 months, the full-sample results on contractions also indicate duration dependence. A post-war analysis was not attempted since the expected frequency <5 in the last bin.

Table 5. Chi-Square Tests for Constant-Hazard Expansions (One-Tailed *p*-Values Are Reported in Parentheses).

Minimum phase								
$\tau_0 = 8$			$\tau_0 = 9$			$\tau_0 = 10$		
Interval	*E*	*O*	Interval	*E*	*O*	Interval	*E*	*O*
Entire sample ($N = 31$)								
[0, 5]	6.155	3	[0, 5]	6.365	3	[0, 5]	6.589	3
[6, 13]	6.349	6	[6, 13]	6.501	9	[6, 13]	6.660	9
[14, 24]	6.169	8	[14, 24]	6.235	6	[14, 24]	6.297	7
[25, 43]	6.211	10	[25, 43]	6.152	9	[25, 43]	6.080	8
>43	6.116	4	>43	5.747	4	>43	5.374	4
χ^2 (3) = 5.22 (0.1653)			χ^2 (3) = 4.60 (0.2234)			χ^2 (3) = 3.81 (0.3067)		
Entire sample, no wars ($N = 26$)								
[0, 5]	6.352	3	[0, 5]	6.620	3	[0, 5]	6.911	3
[6, 13]	6.124	6	[6, 13]	6.282	9	[6, 13]	6.446	9
[14, 26]	6.153	10	[14, 26]	6.168	7	[14, 26]	6.170	9
[27, 63]	6.061	6	[27, 68]	6.044	6	[27, 77]	6.005	5
>63	6.310	1	>68	5.886	1	>78	5.468	0
χ^2 (3) = 8.65 (0.0132)			χ^2 (3) = 7.32 (0.0277)			χ^2 (3) = 10.16 (0.0057)		
Post-WWII ($N = 9$)								
[0, 45]	6.065	6	[0, 44]	6.074	6	[0, 42]	6.007	6
>45	2.935	3	>45	2.926	3	>42	2.993	3
Post-WWII, no wars ($N = 7$)								
[0, 64]	6.002	6	[0, 62]	6.001	6	[0, 61]	6.030	6
>64	0.998	1	>63	0.999	1	>61	0.970	1
Pre-WWII ($N = 21$)								
[0, 6]	6.480	2	[0, 6]	6.775	2	[0, 5]	6.254	2
[7, 17]	6.389	10	[7, 18]	6.071	10	[6, 14]	6.069	10
>17	8.131	9	>18	8.154	9	>14	8.677	9
χ^2 (1) = 5.23 (0.0252)			χ^2 (1) = 6.00 (0.0138)			χ^2 (1) = 5.45 (0.0231)		
Pre-WWII, no wars ($N = 19$)								
[0, 6]	6.406	2	[0, 6]	6.726	2	[0, 5]	6.257	2
[7, 18]	6.371	10	[7, 17]	6.097	10	[6, 15]	6.195	10
>18	6.224	7	>18	6.177	7	>16	6.548	7
χ^2 (1) = 5.19 (0.0284)			χ^2 (1) = 5.93 (0.0155)			χ^2 (1) = 5.27 (0.0257)		

From the geometric distribution, *E* is the expected number of observations that lie in the stated interval. From the sample, *O* is the observed number in the interval. The finite-sample one-tailed *p*-values are obtained through simulation. The chi-square test is not applied to post-WWII expansions due to the small sample sizes. The 5% critical value for a chi-square distribution with 3 (1) degree(s) of freedom is 7.81 (3.84).

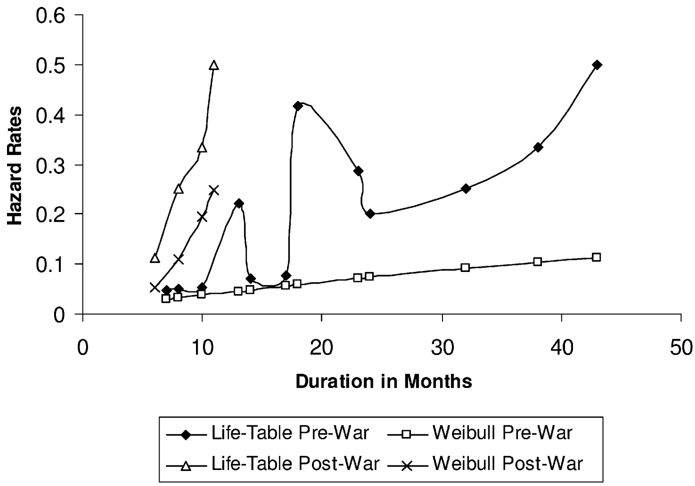

Figure 2. Contractions.

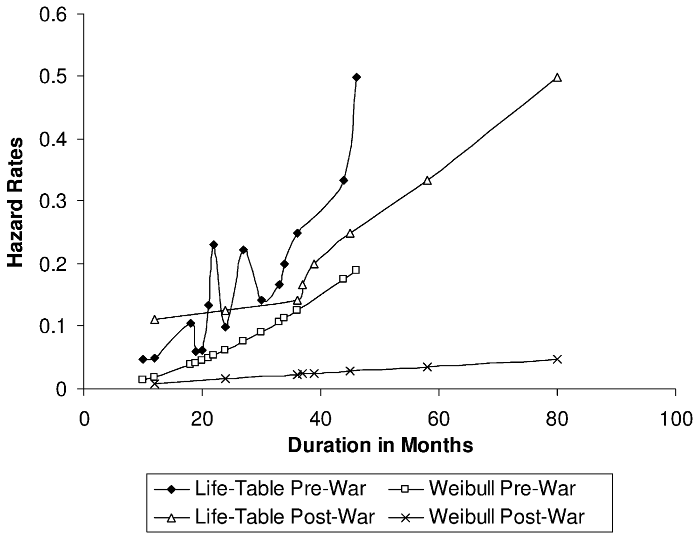

Figure 3. Expansions.

'*t*' months are said to be still at risk. Of course, the pool of contractions still *at risk* decreases with '*t*'. This implies that the effective sample size for estimating the hazard rates for relatively long contractions is less than for short contractions. While each of our empirical hazard functions appear to increase (on average) in Figures 2 and 3, the formal statistical tests given in Tables 2–5 are necessary to avoid spurious conclusions from inspecting the graphs alone.

Yet, in conjunction with our statistical tests, the graphs are powerful in their message concerning duration dependence for expansions and contractions. The hazard function for post-war contractions clearly appears to fall within the IFR class of alternatives, never decreasing with '*t*'. On the other hand, pre-war contractions appear to fall more within the IFRA class (but not the smaller IFR class) since the hazard function increases, but only on average. For either of the pre- or post-war periods, new contractions appear less prone to terminate than mature contractions, so that the mean residual lives of mature contractions are less than the mean residual lives of young contractions. To summarize, from the information gleaned from Figure 2 it is not surprising that our formal statistical tests reject duration independence, with stronger evidence found for post-war contractions.

From Figure 3, our story for expansions is somewhat reversed. Pre-war expansions clearly appear to fall within the IFRA class of alternatives, while the hazard rates for post-war expansions are rather flat for the shorter durations.[12] From the graph, one might conclude that post-war expansions have a relatively flat hazard function, at least for the first 3 years or so. Moreover, our graphical analysis concurs with our formal statistical analysis. There is strong evidence of positive duration dependence for pre-war expansions, but very little in the post-war years.

For comparison, we also graph the parametric Weibull hazard rates. Sichel (1991) suggests that using parametric methods to test for duration dependence will result in more powerful tests for duration dependence, but yet it is clear from Figures 2 and 3 that the Weibull model is insufficient to capture the richness of the IFRA pre-war contractions and expansions. Further, our nonparametric discrete-time MT and SB tests are expected to be especially sensitive to IFRA alternatives, while the chi-square test is expected to perform well provided that we observe the well-known rule of thumb that each cell should have an expected frequency of about 6. It is thus hard to see why using the continuous-time Weibull model would be superior in the context of a time series measured at regular discrete intervals for which the hazard may be highly irregular. But business cycle data have these characteristics.

Attempts to apply more flexible continuous-time methods have met with mixed success. Diebold *et al.* (1993), for example, apply their nonlinear exponential linear model to business cycle data that mostly duplicates the results from the Weibull model. Zuehlke (2003) then applies the nonlinear model of Mudholkar *et al.* (1996) that nests the Weibull model.[13] Yet, the hazard function for the Mudholkar model allows only for hazards that are monotonically increasing, monotonically decreasing, U-shaped or inverted U-shaped. None of these shapes adequately describes the hazard functions for pre-war contractions and expansions. Using the Mudholkar model, however, Zuehlke (2003) does find evidence of duration dependence in pre-war contractions just as we do. But yet the Mudholkar model does not improve upon the Weibull model for post-war contractions, though this is not so surprising given the life-table estimates from Figure 2. Finally, Zuehlke (2003) only finds evidence of positive duration dependence in post-war expansions once the sample is extended through 2001.

3.2. Stock market cycles

Based on the reference dates from Pagan and Sossounov (2002), we next examine the duration dependence of U.S. bull and bear markets. The *largest* possible value of τ_0 is 4 for bull markets since this is the *smallest* observed duration of a bull market. The largest possible value of τ_0 is 3 for bear markets since this is the smallest observed duration of a bear market.

[12]Recall that for the shorter durations, the effective sample sizes are larger for estimating the hazard rates.

[13]Zuehlke (2003) notes that Diebold *et al.* (1993) *simply* subtracts the minimum phase from each duration to account for the censoring and then uses the unconditional density, whereas, Sichel (1991) uses the conditional density to correct for censoring. Yet, these approaches are equivalent as is clear from examining Epstein's (1960) two-parameter Exponential PDF.

Table 6. Bull and Bear Market Reference Dates and Durations.

Trough	Peak	Bear	Bull	Trough	Peak	Bear	Bull
July 1837	Oct. 1838	NA	15	Dec. 1914	Nov. 1916	27	23
Dec. 1839	Nov. 1840	14	11	Dec. 1917	July 1919	13	19
Feb. 1843	Jan. 1846	27	35	Aug. 1921	Mar. 1923	25	19
Feb. 1847	Sept. 1847	13	7	Oct. 1923	Sept. 1929	7	71
Dec. 1848	Jan. 1853	15	49	June 1932	Feb. 1934	33	20
Feb. 1855	Aug. 1855	25	6	Mar. 1935	Feb. 1937	13	23
Nov. 1857	April 1858	27	5	April 1938	Oct. 1939	14	**18**
Aug. 1859	Nov. 1860	16	15	May 1942	May 1946	**31**	**48**
Aug. 1861	April 1864	9	32	Feb. 1948	June 1948	21	4
April 1865	Nov. 1866	12	19	June 1949	Dec. 1952	12	42
May 1867	April 1872	6	59	Aug. 1953	July 1956	8	35
Nov. 1873	April 1875	19	17	Dec. 1957	July 1959	17	19
June 1877	June 1881	26	48	Oct. 1960	Dec. 1961	15	14
Jan. 1885	May 1887	43	28	June 1962	Jan. 1966	6	43
June 1888	May 1890	13	23	Sept. 1966	Nov. 1968	8	26
Dec. 1890	Aug. 1892	7	20	June 1970	April 1971	19	10
Aug. 1893	April 1894	12	8	Nov. 1971	Dec. 1972	7	13
Mar. 1895	Sept. 1895	11	6	Sept. 1974	Dec. 1976	21	27
Aug. 1896	Sept. 1897	11	13	Feb. 1978	Nov. 1980	14	33
April 1898	April 1899	7	12	July 1982	June 1983	20	11
Sept. 1900	Sept. 1902	17	24	May 1984	Aug. 1987	11	39
Oct. 1903	Sept. 1906	13	35	Nov. 1987	May 1990	3	30
Nov. 1907	Dec. 1909	14	25	Oct. 1990	Jan. 1994	5	39
July 1910	Sept. 1912	7	26	June 1994	?	5	NA

Bull and Bear phases during World War II are given in boldface.

Our reference period is July 1837–June 1994. Reference dates for the stock market data are presented in Table 6, with summary statistics presented in Table 7. Observe that for all samples the mean duration is greater than the standard deviation.

This suggests that both bull and bear phases display positive duration dependence. Consider, for instance, the weak-form tests found in Tables 8 and 9. There is very strong statistical evidence of positive duration dependence in bear markets regardless of the sample period. For bull markets, we again reject in favour of positive duration dependence quite often, and every test rejects the constant-hazard assumption for the full sample and post-war period. Positive duration dependence implies that new bull phases are more robust to failure than more mature bull markets, and this is also consistent with a clustering of the observed duration of bull markets. The clustering effect will be revisited below.

There is less evidence of duration dependence in pre-war bull markets, and in fact, SB rejects only when $\tau_0 = 2$. While positive duration dependence may indeed be a feature of the pre-war period bull markets, the evidence is rather inconclusive when based solely on the results in Table 9.

Table 7. Stock Market Cycle Summary Statistics.

Sample	Mean duration	Standard deviation	Sample size
	Bull market		
Entire sample	24.8	14.9	47
Post-WWII	25.7	13.0	15
Pre-WWII	23.8	15.8	30
	Bear market		
Entire sample	15.3	8.5	47
Post-WWII	12.0	6.3	16
Pre-WWII	16.5	8.8	30

Table 8. Tests for Constant-Hazard Bear Markets.

Statistic	All ($N = 47$)	Pre-war ($N = 30$)	Post-war ($N = 16$)
SB ($\tau_0 = 1$)	0.0002+	0.0012+	0.0096+
SB ($\tau_0 = 2$)	0.0008+	0.0032+	0.0304+
SB ($\tau_0 = 3$)	0.0026+	0.0138+	0.0822+
MT ($\tau_0 = 1$)	0.0004+	0.0022+	0.0016+
MT ($\tau_0 = 2$)	0.0024+	0.0054+	0.0112+
MT ($\tau_0 = 3$)	0.0094+	0.0194+	0.0418+

The two-tailed p-values are obtained through simulation. A plus sign (+) indicates statistically significant positive duration dependence.

Table 9. Tests for Constant-Hazard Bull Markets.

Statistic	All ($N = 47$)	Pre-war ($N = 30$)	Post-war ($N = 15$)
SB ($\tau_0 = 2$)	0.0020+	0.0868+	0.0096+
SB ($\tau_0 = 3$)	0.0050+	0.1456	0.0126+
SB ($\tau_0 = 4$)	0.0140+	0.2378	0.0224+
MT ($\tau_0 = 2$)	0.0042+	0.1084	0.0042+
MT ($\tau_0 = 3$)	0.0108+	0.1768	0.0074+
MT ($\tau_0 = 4$)	0.0272+	0.2704	0.0130+

The two-tailed p-values are obtained through simulation. A plus sign (+) indicates statistically significant positive duration dependence.

To further investigate the nature of the hazard functions, we construct the chi-square goodness-of-fit tests. The results are found in Tables 10 and 11.[14] From Table 10, the evidence is overwhelming that the full-sample and pre-war bear markets exhibit positive duration dependence. The clustering effect of such dependence is very evident in both samples. For example, for the full-sample period when $\tau_0 = 3$, we have an expected frequency of 6.802 in the closed interval [0, 1] and 6.753 in the closed interval [8, 11]. Yet, the respective observed frequencies are 1 and 15.

[14]An analysis of Post-WWII was not attempted since the expected frequency was less than 5 in the last bin.

Table 10. Chi-Square Tests for Constant-Hazard Bear Markets (One-Tailed p-Values Are Reported in Parentheses).

Minimum phase								
$\tau_0 = 1$			$\tau_0 = 2$			$\tau_0 = 3$		
Interval	E	O	Interval	E	O	Interval	E	O
Entire sample ($N = 47$)								
[0, 2]	8.626	1	[0, 1]	6.344	1	[0, 1]	6.802	1
[3, 5]	7.043	4	[2, 4]	7.947	4	[2, 4]	8.403	9
[6, 9]	7.422	8	[5, 7]	6.393	8	[5, 7]	6.646	3
[10, 14]	6.856	17	[8, 11]	6.624	11	[8, 11]	6.753	15
[15, 21]	6.428	8	[12, 17]	6.946	11	[12, 17]	6.887	8
>21	10.625	9	[18, 26]	6.108	9	[18, 27]	6.242	8
–	–	–	>26	6.638	3	>27	5.267	3
χ^2 (4) = 23.74 (0.0000)			χ^2 (5) = 15.48 (0.0086)			χ^2 (5) = 18.72 (0.0025)		
Pre-WWII ($N = 30$)								
[0, 3]	6.638	0	[0, 3]	7.024	0	[0, 3]	7.458	1
[4, 8]	6.272	6	[4, 8]	6.515	6	[4, 8]	6.772	7
[9, 15]	6.058	14	[9, 15]	6.140	15	[9, 15]	6.207	13
>15	11.032	10	>15	10.321	9	>15	9.563	9
χ^2 (2) = 17.16 (0.0003)			χ^2 (2) = 20.02 (0.0001)			χ^2 (2) = 13.07 (0.0016)		
Post-WWII ($N = 16$)								
[0, 5]	6.507	4	[0, 4]	6.065	4	[0, 4]	6.552	5
>5	9.493	12	>4	9.935	12	>4	9.448	11

From the geometric distribution, E is the expected number of observations that lie in the stated interval. From the sample, O is the observed number in the interval. The finite-sample one-tailed p-values are obtained through simulation. The chi-square test is not applied to post-WWII bear markets due to the small sample size. The 5% critical value for a chi-square distribution with 5 degrees of freedom is 11.07, while the 5% critical value for a chi-square distribution with 4 (2) degrees of freedom is 9.49 (5.99).

Relative to the geometric density, there are too few short bear phases and too many intermediate ones. This fact is also reflected in Table 7 where the mean duration of the bear markets is almost twice the standard deviation.

For the bull markets, we reject the constant-hazard assumption in the pre-war period for imposed minimum phases of either 3 or 4. This is most interesting since neither SB nor MT reject the null geometric density for the pre-war period when $\tau_0 > 2$. Coupling the results presented in Table 9 with those from Table 11, there appears to be strong evidence that bull markets exhibit positive duration dependence for both the pre- and post-WWII sample periods.

We also investigate the nature of the duration dependence by using graphs. Figures 4 and 5 plot the hazards for the pre- and post-war bear and bull markets. For the post-war period, bull markets appear to fall within the strict IFR class of alternatives since the hazard function appears to continuously rise, while bear markets fall into the IFRA class since the hazard function rises but only on average. For the pre-war period, both bull and bear markets exhibit IFRA behaviour, but again, the Weibull hazard fails to reveal the complexity of the IFRA case. Based upon either

Table 11. Chi-Square Tests for Constant-Hazard Bull Markets (One-Tailed p-Values Are Reported in Parentheses).

Minimum phase								
$\tau_0 = 2$			$\tau_0 = 3$			$\tau_0 = 4$		
Interval	E	O	Interval	E	O	Interval	E	O
Entire sample ($N = 47$)								
[0, 3]	7.415	2	[0, 3]	7.719	4	[0, 2]	6.176	4
[4, 7]	6.245	4	[4, 7]	6.451	3	[3, 6]	6.991	3
[8, 12]	6.440	7	[8, 12]	6.595	8	[7, 11]	7.080	8
[13, 18]	6.108	10	[13, 18]	6.190	8	[12, 17]	6.569	8
[19, 26]	6.043	9	[19, 26]	6.043	9	[18, 25]	6.321	9
[27, 39]	6.308	8	[27, 39]	6.186	9	[26, 38]	6.334	9
>39	8.441	7	>39	7.816	6	>38	7.529	6
χ^2 (5) = 9.44 (0.0918)			χ^2 (5) = 7.61 (0.1840)			χ^2 (5) = 6.04 (0.3189)		
Pre-WWII ($N = 30$)								
[0, 4]	6.027	3	[0, 4]	6.278	4	[0, 4]	6.551	5
[5, 11]	6.460	5	[5, 11]	6.646	4	[5, 10]	6.002	3
[12, 21]	6.330	11	[12, 21]	6.399	12	[11, 19]	6.249	11
>21	11.183	11	>21	10.677	10	[20, 35]	6.108	7
–	–	–	–	–	–	>35	5.090	4
χ^2 (2) = 5.30 (0.1030)			χ^2 (2) = 6.82 (0.0449)			χ^2 (3) = 5.84 (0.0765)		
Post-WWII ($N = 15$)								
[0, 12]	6.235	5	[0, 11]	6.058	5	[0, 11]	6.264	5
>12	8.765	10	>11	8.942	10	>11	8.736	10

From the geometric distribution, E is the expected number of observations that lie in the stated interval. From the sample, O is the observed number in the interval. The finite-sample one-tailed p-values are obtained through simulation. The chi-square test is not applied to post-WWII bull markets due to the small sample size. The 5% critical value for a chi-square distribution with 5 degrees of freedom is 11.07, while the 5% critical value for a chi-square distribution with 3 (2) degrees of freedom is 7.81 (5.99).

the parametric or nonparametric graphical evidence, however, we conclude that there is positive duration dependence for both bear and bull markets.

4. SUMMARY

We have discussed discrete-time tests for duration dependence that may be applied to the analysis of many types of economic data. Our weak form tests concentrate on moment conditions, and we examine two simple asymptotically pivotal test statistics whose finite-sample p-values are obtained though bootstrapping. The first of these tests is derived within the generalized method of moments framework and the second is based on regression analysis. Our strong-form chi-square test directly tests against a geometric density. We pay special attention to the rule-of-thumb that each bin should have an expected frequency of about 6, and we report results based on both asymptotic theory and bootstrapping.

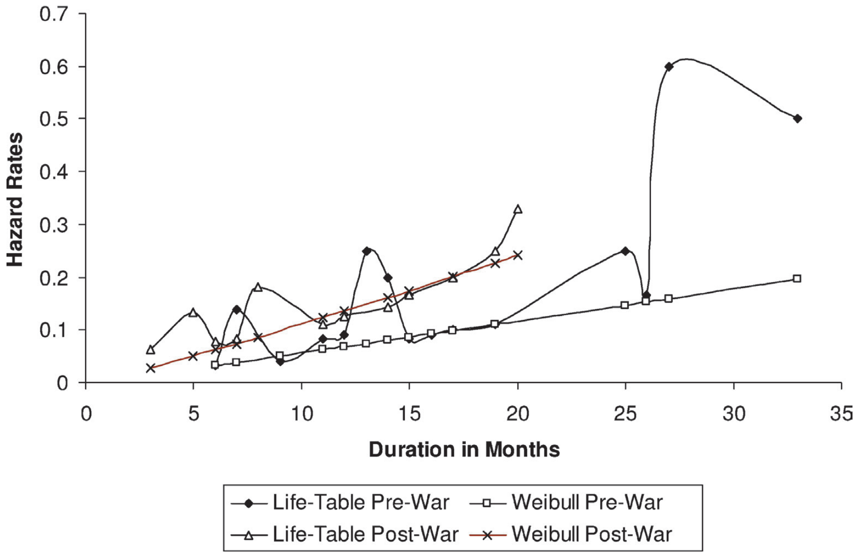

Figure 4. Bear markets.

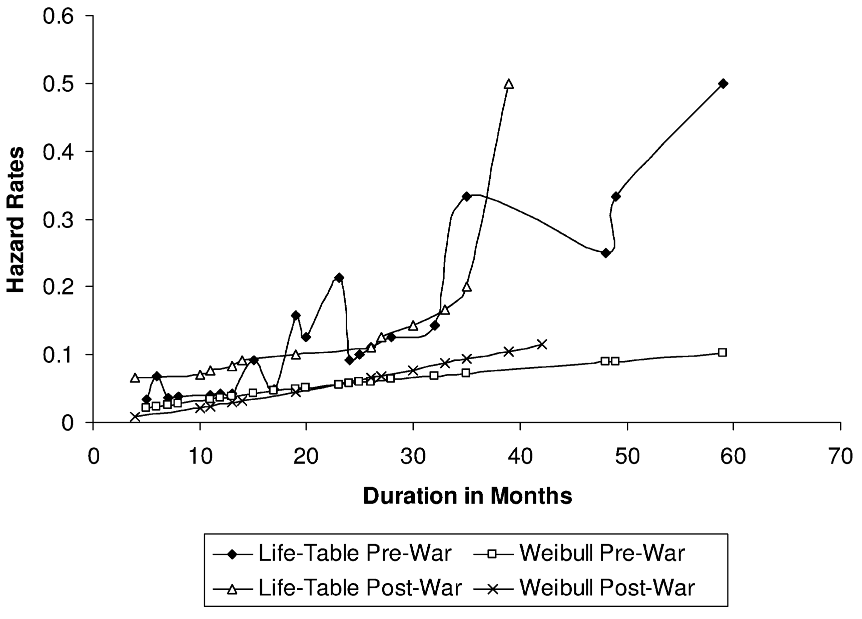

Figure 5. Bull markets.

Our first empirical application was to revisit the work of Diebold and Rudebusch (1990, 1991) on business cycles. In addition to finding evidence of an increasing hazard in post-war contractions (as did Diebold and Rudebusch (1990)), we also find some evidence of positive duration dependence in pre-war contractions. Positive duration dependence implies that new contractions are more robust to failure than more mature contractions.

Based on graphical analysis, the hazard function for post-war contractions clearly appears to fall within the IFR class of alternatives, never decreasing with 't'. On the other hand, pre-war contractions appear to fall more within the IFRA class (but not the smaller IFR class) since the hazard function increases, but only on average. Pre-war expansions clearly appear to fall within the IFRA class of alternatives, while the hazard rates for post-war expansions are rather flat for the shorter durations. So, there is strong evidence of positive duration dependence for pre-war expansions, but very little in the post-war years. In particular, we have no strong statistical evidence to support the notion that post-war expansions are duration dependent, though this may be due to our small sample size.

We next examined the duration dependence of U.S. bull and bear markets. From graphical analysis and formal statistical reasoning, both bull and bear phases display positive duration dependence. For the post-WWII period, bull markets appear to fall within the strict IFR class of alternatives since the hazard function appears to rise continuously, while bear markets fall into the IFRA class since the hazard function rises, but only on average. For the pre-war period, both bull and bear markets appear to fall within the IFRA class.

Our work follows that of Diebold and Rudebusch (1990) in that we have tested a version of the Markov hypothesis based upon the two-parameter geometric density. That hypothesis states that, once an expansion or contraction has exceeded some minimum value, the probability of a turning point is independent of its age. Suppose, for example, that it is correct to impose a minimum phase of 6 months. If so, then the hazard function (or the hazard rate in discrete time) must be *identically zero* for months 1–6 and some nonzero value, p, thereafter. In other words, the phase cannot terminate in less than 6 months, but has a constant hazard rate thereafter. Only under this Markov hypothesis is it appropriate to truncate durations at 6 months, since in using the filter we transform the durations to follow the geometric density. Note, in particular, that it is not true that we begin with durations that follow the geometric density, but rather we obtain the said density only after we impose the minimum phase of 6 months.

To reiterate, we find that subtracting a minimum phase from the observed phases is typically insufficient to eliminate the duration dependence in business and stock-market phases. This is not so surprising. For business cycles the growth rate of Gross Domestic Product is a reasonable proxy for the implicit series used by the NBER committee to date the phases. While it is necessary to account for any censoring prior to conducting the duration analysis, the observed serial correlation in GDP should result in duration dependence. Failure to reject the null hypothesis for (say) post-war expansions may be due to the small sample size.

ACKNOWLEDGEMENTS

We would like to thank William Lastrapes and seminar participants at the NSF/NBER Time Series Conference 2002, the University of Georgia and Lehigh University for helpful comments and suggestions. The comments of two referees and Neil Shephard have also been extremely useful in the rewriting of an earlier version of the paper. The second author wishes to acknowledge partial financial support from the Arthur F. Searing Fund.

REFERENCES

Cutler, S. and F. Ederer (1958). Maximum utilization of the life table in analyzing survival. *Journal of Chronic Disorders 8*, 699–712.

Diebold, F. X. and G. D. Rudebusch (1990). A nonparametric investigation of duration dependence in the American business cycle. *Journal of Political Economy 98*, 596–616.

Diebold, F. X. and G. D. Rudebusch (1991). Turning point prediction with the composite leading index: A real-time analysis. In K. Lahiri and G. H. Moore (eds.). *Leading Economic Indicators: New Approaches and Forecasting Records*, pp. 231–256.

Diebold, F. X., G. D. Rudebusch and D. E. Sichel (1993). Further evidence on business cycle duration dependence. In J. H. Stock and M. W. Watson (eds.). *Business Cycles, Indicators, and Forecasting*, pp. 87–116.

Durland, J. M. and T. H. McCurdy (1994). Duration dependent transitions in a markov model of U.S. GNP growth. *Journal of Business and Economic Statistics 12*, 279–88.

Epstein, B. (1960). Tests for the validity of the assumption that the underlying distribution of life is exponential, I. *Technometrics 2*, 83–101.

Fisher, I. (1925). Our unstable dollar and the so-called business cycle. *Journal of the American Statistical Association 20*, 179–202.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*, 357–84.

Harding, D. and A. R. Pagan (2003). A comparison of two business cycle dating methods. *Journal of Economic Dynamics and Control 27*, 1681–90.

Hoel, P. G. (1954). *Introduction to Mathematical Statistics*. New York: John Wiley and Sons.

Hollander, M. and D. A. Wolfe (1999). Life distributions and survival analysis. *Nonparametric Statistical Methods*, pp. 495–765. New York, NY: John Wiley and Sons.

Horowitz, J. (2001). The bootstrap in econometrics. In J. J. Heckman and E. E. Leamer (eds.). *Handbook of Econometrics*, vol. 5, pp. 3159–228.

Lee, S. C. S., C. Locke and J. D. Spurrier (1980). On a class of tests of exponentiality. *Technometrics 22*, 547–54.

LIMDEP (1995), Version 7.0, Econometric Software, Inc.

McCulloch, J. H. (1975). The Monte Carlo cycle in business cycle activity. *Economic Inquiry 13*, 303–21.

Mudambi, R. and L. W. Taylor (1991). A nonparametric investigation of duration dependence in the American Business Cycle: A note. *Journal of Political Economy 99*, 654–6.

Mudambi, R. and L. W. Taylor (1995). Some nonparametric tests for duration dependence. *Journal of Applied Statistics 22*, 163–77.

Mudholkar, G. S., D. K. Srivastava and G. D. Kollia (1996). A generalization of the weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association* 91, 1575–83.

Niemira, M. P. (1991). International application of Neftci's probability approach. In K. Lahiri and G. H. Moore (eds.). *Leading Economic Indicators: New Approaches and Forecasting Records*, pp. 91–108.

Pagan, A. R. and K. A. Sossounov (2002). A simple framework for analyzing bull and bear markets. *Journal of Applied Econometrics 18*, 23–46.

Sargan, J. D. and A. Bhargava (1983). Testing residuals from least squares regression for being generated by the Gaussian random walk. *Econometrica 51*, 153–74.

Sichel, D. E. (1991). Business cycle duration dependence: A parametric approach. *Review of Economics and Statistics 73*, 254–60.

Zuehlke, T. W. (2003). Business cycle duration dependence reconsidered. *Journal of Business and Economic Statistics 21*, 564–9.

APPENDIX: AN ANALYTICAL FORM FOR THE ESTIMATE OF c_2 IN THE SB TEST

Consider the SB regression equation:

$$S_t = c_0 + c_1 S_{t-1} + c_2 S_{t-1} d_{t-1} + \text{disturbance}, \tag{A.1}$$

where S_t is the state variable assigned unity if the observed index is a month of expansion (and zero for a contraction) and d_t is the number of consecutive months (i.e. the 'duration') spent in an expansion up and through time t. It is possible to find an analytic form for the estimate of c_2. To do so, first recognize that the intercept in (A.1) can be eliminated by mean correcting all the variables. Let the sample means of S_{t-1} and $S_{t-1}d_{t-1}$ be $\overline{S}$ and $\overline{SD}$, respectively.[15] Then,

$$\frac{1}{T}\sum_{t=1}^{T} S_{t-1}S_t = \overline{S} - \frac{n}{T}, \tag{A.2}$$

where T is the econometric sample size and and 'n' is the number of expansions, and

$$\overline{SD} = \frac{1}{T}\sum_{t=1}^{T} S_{t-1}d_{t-1} = \frac{1}{T}\sum_{i=1}^{n}\frac{(x_i+1)x_i}{2}, \tag{A.3}$$

where x_{i} is the length of the ith expansion. Finally,

$$\frac{1}{T}\sum_{t=1}^{T} S_{t-1}S_t d_{t-1} = \frac{1}{T}\sum_{i=1}^{n}\frac{(x_i-1)x_i}{2} = \overline{SD} - \overline{S}. \tag{A.4}$$

Let the (2,1) element in the inverse of the cross product matrix of the regression of $S_t - \overline{S}$ on $S_{t-1} - \overline{S}$ and $S_{t-1}d_{t-1} - \overline{SD}$ (scaled by $1/T$) be 'a' and the (2,2) element be 'b'. Then,

$$\hat{c}_2 = a\left(\overline{S}(1-\overline{S}) - \frac{n}{T}\right) + b(\overline{SD}(1-\overline{S}) - \overline{S}). \tag{A.5}$$

We can determine that

$$a = -\Delta^{-1}\overline{SD}(1-\overline{S}),\ b = \Delta^{-1}\overline{S}(1-\overline{S}), \tag{A.6}$$

where Δ is the determinant of the scaled cross product matrix.

Putting all these elements together we get,

$$\begin{aligned}\hat{c}_2 &= -\Delta^{-1}\overline{SD}(1-\overline{S})\left\{\overline{S}(1-\overline{S}) - \tfrac{n}{T}\right\} + \Delta^{-1}\overline{S}(1-\overline{S})\{\overline{SD}(1-\overline{S}) - \overline{S}\}\\ &= \Delta^{-1}\tfrac{n}{T}(1-\overline{S})\left\{\overline{SD} - \tfrac{T}{n}\overline{S}^2\right\}\\ &= \Delta^{-1}\tfrac{n}{T}(1-\overline{S})\left\{\tfrac{1}{T}\sum_{i=1}^{n}\tfrac{(x_i+1)x_i}{2} - \tfrac{T}{n}(\tfrac{n}{T}\bar{x})^2\right\}\\ &= \Delta^{-1}\tfrac{n}{T}(1-\overline{S})\tfrac{n}{2T}\{\hat{\sigma}^2 - \bar{x}^2 + \bar{x}\}.\end{aligned} \tag{A.7}$$

Since the factors of proportionality are irrelevant to the test statistic, the SB test based upon $\hat{c}_2$ is therefore effectively checking if $V(X) - [E(X)]^2 + E(X)$ is zero.

[15] The sample means of S_{t} and S_{t-1} are not exactly the same but we will ignore such end effects.