Losing Sleep at the Market: The Daylight Saving Anomaly: Reply

By Mark J. Kamstra, Lisa A. Kramer, and Maurice D. Levi*

J. Michael Pinegar (2002) revisits the issue of daylight-saving-time changes impacting financial markets, suggesting that the evidence we provided in Kamstra et al. (2000; henceforth KKL) is nothing to lose sleep over. His claim is that further robustness checks including new statistical tests and removing so-called "outliers" undermine the evidence in favor of a daylight-saving effect. However, he largely ignores the international evidence compiled by KKL, presenting tabular evidence only for U.S. indices. In addition and more importantly, some tests he performs are invalid, he systematically understates the statistical significance of the valid tests he performs, and he misrepresents our findings. In this comment we do not update our sample, but rather use the original data presented in KKL so as to make our point without changing the goalposts. We do note, however, that two of the more recent daylight-saving weekends (beyond Pinegar's elongated sample) witnessed very large negative returns on many indices, as large as −5 percent, and we wonder if Pinegar would argue that these observations should also be excluded as "outliers." We will leave it to the reader to decide.

One of the main features of Pinegar's comment provides a useful starting point for our reply. He plots the cumulative distribution functions (c.d.f.'s) of weekend returns for the NYSE equal-weighted index to investigate whether there is a significant difference between daylight-saving weekend returns and regular weekend returns. In doing so he draws our attention to the tails of the distributions. Although he focuses on only the relatively fatter *lower* tail of the c.d.f. for returns on daylight-saving weekends, his plot also inadvertently highlights the fact that daylight-saving weekends are unusual for their relative dearth of positive returns. This point is noteworthy because it reveals that the daylight-saving effect arises not only due to the extreme negative observations in the lower tail, but also because of a lack of positive return observations. That is, on daylight-saving weekends, the entire distribution of returns shifts to the left. Not only are there more negative returns on those weekends, but there are also fewer positive observations. (We thank Pinegar for leading us to this insight.) We complete the exercise Pinegar began by presenting similar plots of returns for the remaining indices and countries in our sample (including some countries that experience daylight-saving-time changes on different dates relative to the United States, as outlined in KKL). Our findings considerably strengthen the case for a daylight-saving effect.

Figures 1–3 present cumulative distribution functions of returns for each of the indices we covered in KKL. In each plot, the line indicated with solid dots is the c.d.f. for regular weekend returns exclusive of daylight-saving returns, and the line indicated with hollow dots is the c.d.f. for daylight-saving weekend returns. Figure 1 presents c.d.f. plots for the NYSE value-weighted and equal-weighted indices (data from the Center for Research in Security Prices)[1] and for the S&P 500 index, 1928–1966 and post-1966. Figure 2 presents plots for the NASDAQ and AMEX value-weighted and equal-weighted indices, and Figure 3 presents plots for the TSE

* Kamstra: Research Department, Finance, Atlanta Federal Reserve Bank, 1000 Peachtree Street N.E., Atlanta, GA 30309; (e-mail: Mark.Kamstra@atl.frb.org); Kramer: Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, Canada M5S 3E6 (e-mail: Lkramer@chass.utoronto.ca); Levi: Faculty of Commerce and Business Administration, University of British Columbia, 2053 Main Mall, Vancouver, B.C., Canada V6T 1Z2 (e-mail: maurice.levi@commerce.ubc.ca). The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

[1] These returns do not include distributions. Our results strengthen when all distributions are included in the daily returns.

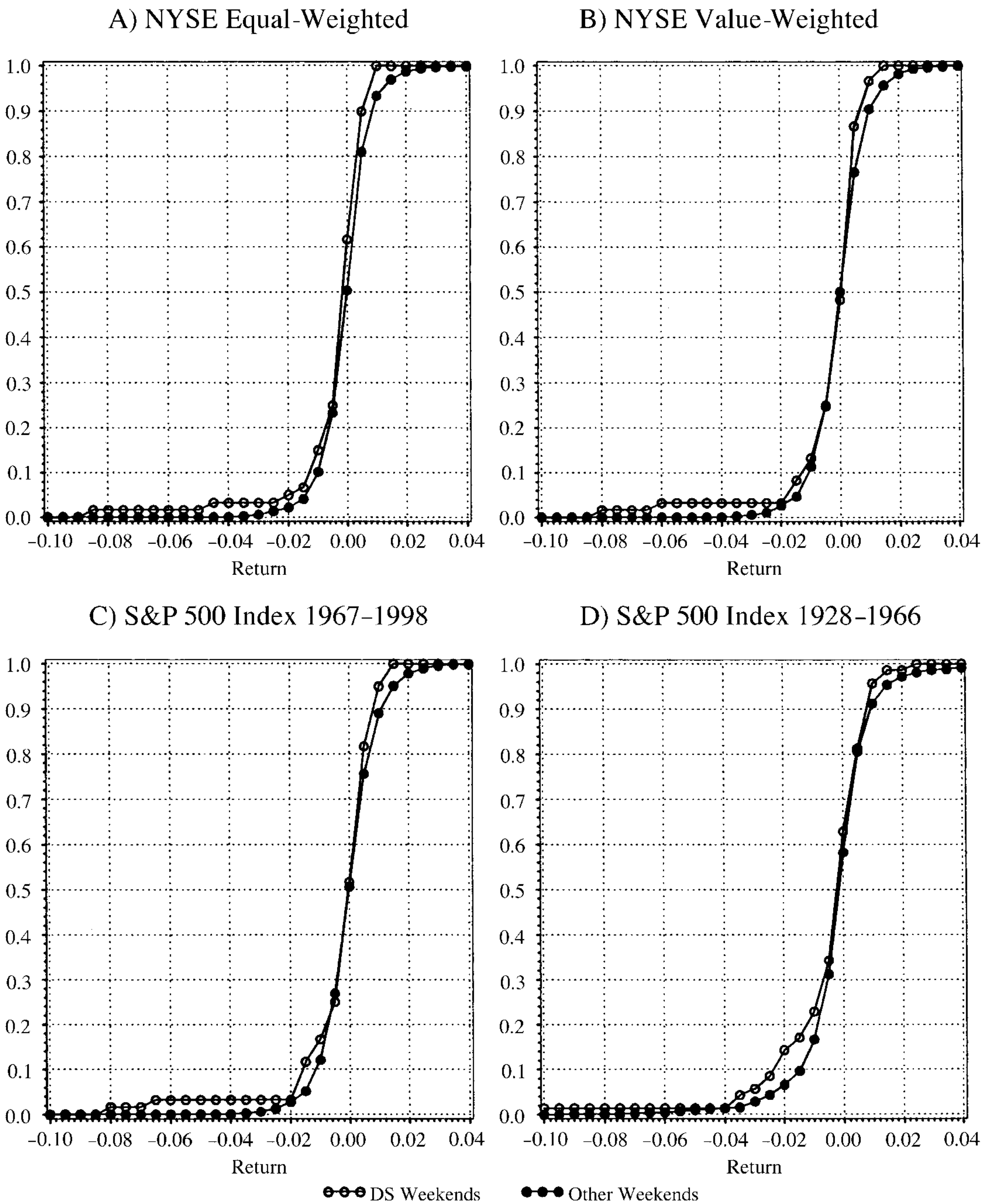

FIGURE 1. CUMULATIVE DISTRIBUTION FUNCTIONS FOR DAYLIGHT-SAVING WEEKEND RETURNS AND OTHER WEEKEND RETURNS FOR VARIOUS U.S. INDICES, SPECIFICALLY, THE NYSE EQUAL-WEIGHTED AND VALUE-WEIGHTED INDICES (BOTH 1 JANUARY 1967–31 DECEMBER 1997) AND THE S&P 500 INDEX (1 JANUARY 1967–31 DECEMBER 1998 AND 1 JANUARY 1928–31 DECEMBER 1966)

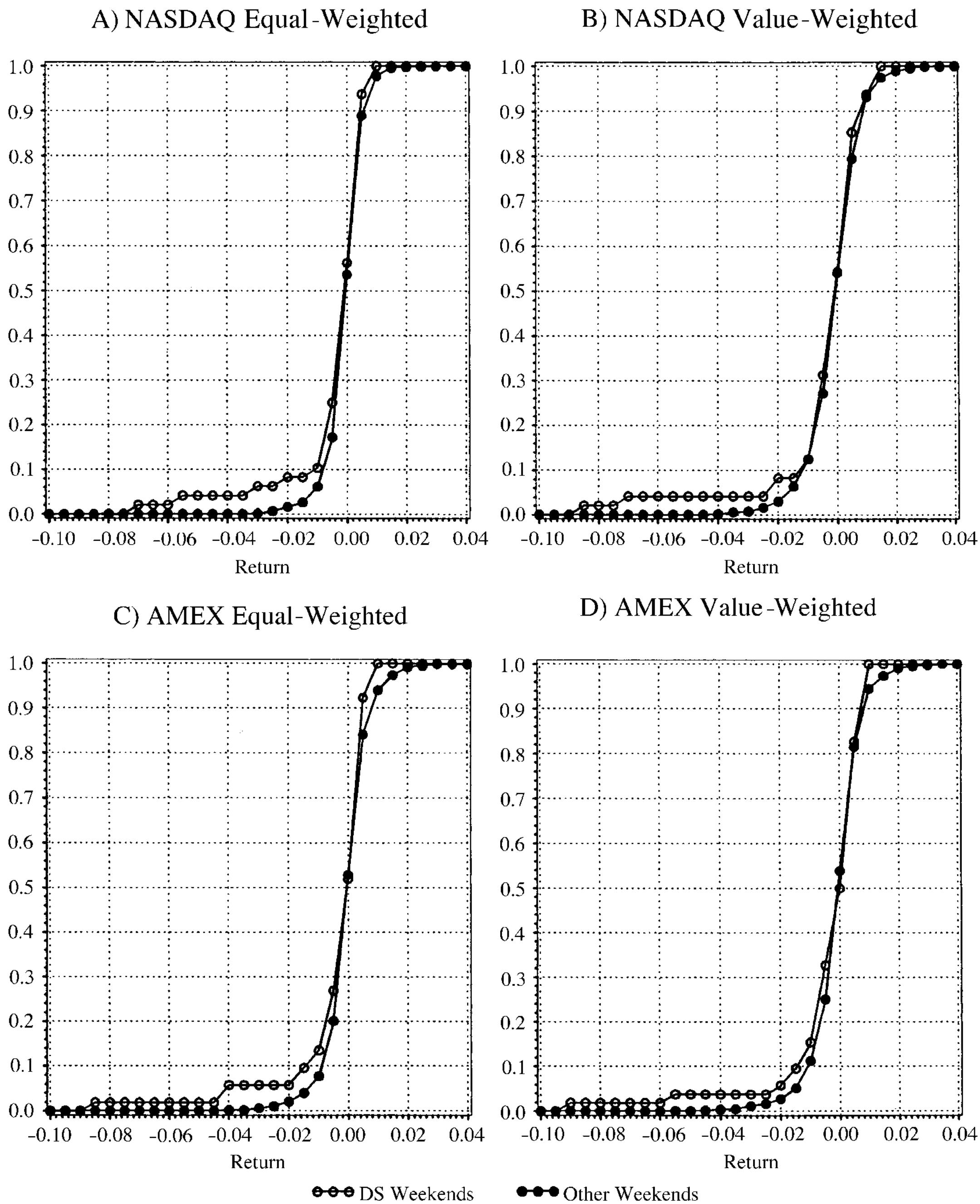

FIGURE 2. CUMULATIVE DISTRIBUTION FUNCTIONS FOR DAYLIGHT-SAVING WEEKEND RETURNS AND OTHER WEEKEND RETURNS FOR VARIOUS U.S. INDICES, SPECIFICALLY, THE NASDAQ EQUAL-WEIGHTED AND VALUE-WEIGHTED INDICES AND THE AMEX EQUAL-WEIGHTED AND VALUE-WEIGHTED INDICES (ALL 1 JANUARY 1967–31 DECEMBER 1997)

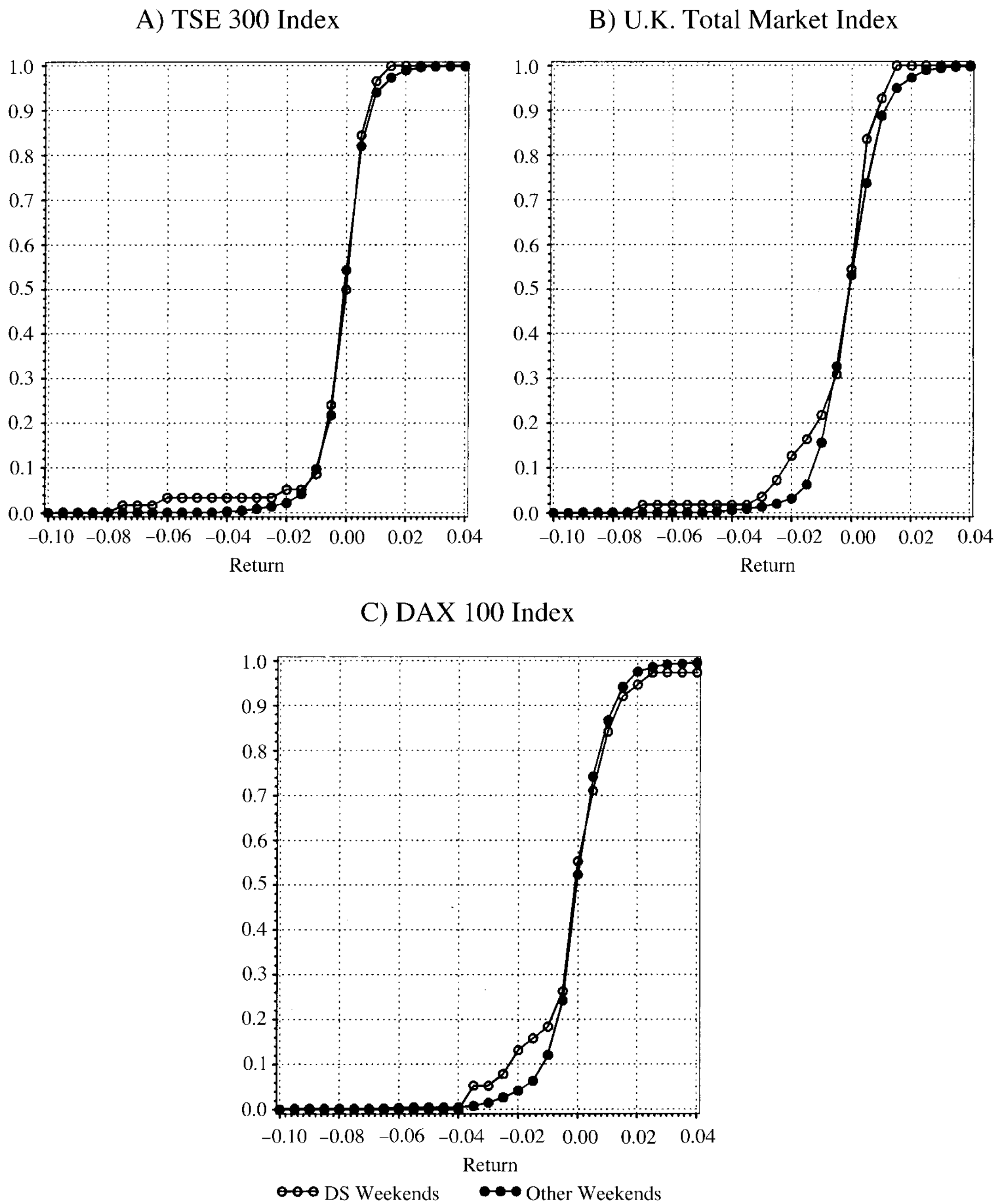

FIGURE 3. CUMULATIVE DISTRIBUTION FUNCTIONS FOR DAYLIGHT-SAVING WEEKEND RETURNS AND OTHER WEEKEND RETURNS FOR VARIOUS INTERNATIONAL INDICES, SPECIFICALLY, CANADA'S TSE INDEX (1 JANUARY 1969–17 DECEMBER 1998), THE U.K. TOTAL MARKET INDEX (1 JANUARY 1969–18 DECEMBER 1998), AND GERMANY'S DAX 100 INDEX (1 JANUARY 1973–18 DECEMBER 1998)

300 (Toronto), the U.K. Total Market, and the DAX 100 (Germany).[2]

If differences between the distributions of daylight-saving and non-daylight-saving weekends were due only to a few negative outliers, as Pinegar suggests, we would observe largely overlapping c.d.f.'s with a lower tail for the daylight-saving weekend returns falling to the left of that for regular weekend returns. In each plot we do indeed see a remarkable downward shift (to the left) in the lower tail of negative returns for daylight-saving weekends relative to regular weekends. This is true across countries that do not share the same day for the daylight saving time change (avoiding international contagion concerns), as well as across value-weighted and equal-weighted indices. However, more notably, we find virtually without exception that the upper tail of positive returns (returns larger than zero, not just large positive returns) for the daylight-saving weekend returns c.d.f. is also shifted to the left relative to the c.d.f. for regular weekend returns. This is most striking for the NYSE equal-weighted and value-weighted, the S&P 500 pre- and post-1966, the AMEX equal- and value-weighted, and the U.K. Total Market indices, but is also true of the TSE 300 and the NASDAQ equal- and value-weighted indices. The DAX 100 index is the only index for which this pattern does not hold, and it was the only index for which KKL did not find statistically significant evidence of a daylight-saving effect. The differences are in not just one tail but *both* tails of the c.d.f.'s for regular and daylight-saving weekend returns, suggesting the daylight-saving effect is not simply an artifact of a few negative outliers.

It should be noted that removal of the two observations denoted as outliers by Professor Pinegar does not affect the positive tail of the c.d.f. plots for returns. If the daylight-saving effect documented by KKL was caused by these observations, then removing them would yield largely overlapping c.d.f.'s for daylight-saving weekend returns and non-daylight-saving weekend returns. In fact, when we produce plots excluding the two observations in question, we find first that the positive tails remain markedly to the left of the c.d.f. for regular weekend returns, and second that the daylight-saving weekend returns maintain a larger negative mean than regular weekend returns. Thus, we feel that the evidence strongly supports the prominence of the daylight-saving effect. We see not only a downward shift in the lower tail of the returns distribution, but also a downward shift in the positive tail of the returns distribution, that is, a relative lack of "good days" in the market over 30 years of daylight-saving weekends. The daylight-saving effect is not simply driven by a few outliers, but is a systematic effect which shifts the entire distribution of returns.

Pinegar goes on to present test results he claims indicate a lack of statistically significant difference between regular weekend and daylight-saving weekend returns distributions, even though each daylight-saving weekend return entails an economically significant loss of roughly $30 billion in the United States alone. We now consider his employment of the Kolmogorov-Smirnov (KS) test to highlight the flaws in his statistical methodology. The notoriously weak KS test is unable to distinguish two data distributions unless a large number of observations are available for each of the two samples.[3] In the case we consider here, at most 62 observations are available for the daylight-saving return distribution, a very small sample for this test. Second, the KS test assumes independent and identically distributed observations. Stock returns are well documented to be autocorrelated and to exhibit autoregressive conditional heteroscedasticity, both strong forms of dependence over time. While one might make a case for independence of daylight-saving observations that occur six months apart, returns that occur within a week of each other, like those occurring after regular weekends, are most certainly strongly dependent. Thus use of the KS test employed by Pinegar is invalid, and inferences based on the test are not reliable.

Pinegar's other main claim is that the daylight-saving effect is made insignificant or

[2] As in the original Kamstra et al. (2000) paper, all of the time series considered are post-1966 (with the exception of the S&P 500 series which starts in 1928), with various start and end dates. For details see the footnotes to the figures and the original KKL paper.

[3] See, for example, E. L. Lehmann (1985 p. 105) for a discussion of the low power of the KS test.

marginally significant when adjustments are made for heteroscedasticity and/or when the large sample is accounted for with Bayesian sample-size adjustments. With regard to the significance of the daylight-saving effect, we note that Pinegar understated the significance by reporting two-sided tests, rather than one-sided tests of the null of no effect. We were not testing that the time change caused investors to bid prices up, but rather that they bid prices down, a one-sided alternative based both on our theory of sleep desynchronosis and the fact that weekend returns are historically negative to start with. Virtually all the joint fall and spring daylight-saving effect and the fall daylight-saving effect test statistics reported in Pinegar's table 1 are significant at the 10-percent level against the appropriate one-sided alternative, and half are significant at the 5-percent level or better. His ordinary least-squares (OLS) *t* statistics are significant for the joint fall and spring daylight-saving effect, and for the fall daylight-saving effect at the 1-percent level against the appropriate one-sided alternative. This is true for each and every U.S. index he reports, as well as our international data (with the exception of the DAX 100). The spring effect is weaker than the fall effect, significant in only one case, but even so it is typically 3–4 times the magnitude of the regular weekend negative return, a huge difference of considerable economic import, and a magnitude repeated across the several countries we investigated.

Regarding the use of a Bayesian adjustment for large sample sizes, Pinegar does not indicate why one might want to adopt this procedure. A common motivation for using the Bayesian adjustment is the perceived "problem" that any null hypothesis can be rejected with sufficient observations. Of course this is not literally true, but the source of concern is the fact that the variance of statistical estimators falls as the sample size increases. As the variance falls, ever smaller deviations from the null become increasingly statistically significant, even if the deviations are economically irrelevant. We should note that the deviations we find with regard to the daylight-saving effect are fairly large and, we would argue, economically relevant, amounting to roughly $60 billion in losses on average annually in the United States alone. But do we find this small variance effect with our data? The variance of the estimator of the regular-weekend dummy coefficient is indeed small. For the NYSE equal-weighted index[4] the standard deviation of the regular-weekend dummy coefficient estimate is a scant 0.00019. This compares to the coefficient estimate of roughly −0.001, a negative 0.1-percent return on average on weekends, yielding a large *t* statistic of close to −4. This estimate is effectively based on roughly 1,500 weekend observations, a large number without question. The standard deviation of the daylight-saving dummy coefficient is an order of magnitude larger, at 0.001, a larger variance because, in part, there are only approximately 60 daylight-saving weekends. This is not a large sample at all. Although we have more than 8,000 observations in the NYSE sample, for example, only about 1,500 of them contain any information about weekends, and only about 60 of them contain any information about daylight-saving weekends. The appropriate adjustment, even for a Bayesian, would have to take into account this feature of the data. Should a sample-size adjustment use the sample size of 60, 1,500, or 8,000? To us, 8,000 seems excessive for the appropriate sample size given that we have only approximately 1,500 regular weekends and 60 daylight-saving weekends.

We also believe that the Bayesian procedure administered by Pinegar is essentially ad hoc, as it can be manipulated virtually to an arbitrary extent with appropriate choice of priors and posteriors. This is not a new observation, and even proponents of prior odds analysis admit to "some arbitrary aspects" (see e.g., Robert A. Connolly, 1991 p. 57). Donald W. K. Andrews (1994) provides a correspondence between classical hypothesis tests and Bayesian posterior odds ratios, showing that for certain choices of priors the Bayesian posterior odds ratio test is large-sample equivalent to classical tests. Our question, then, is what priors are reasonable? The tests Pinegar performs turn a classical test which was significant at the 5-percent level into a test favoring the null over the alternative 364 to 1!

[4] See Kamstra et al. (2000) for details on coefficient and standard error estimates.

We turn now to Pinegar's discussion of excluding outliers from the set of daylight-saving weekend returns. Three of the 20 largest percentage declines in the S&P 500 over the past 80 years took place immediately after a fall daylight-saving-time change. While we do not believe daylight-saving-time changes *cause* market crashes, we do believe that daylight-saving-time changes affect the degree of market fluctuations. Hence, it may be no coincidence that some of the largest market fluctuations on record occurred within close proximity of a daylight-saving-time change. One of these daylight-saving observations Pinegar considers to be potentially excludable happened to occur "a few weeks following" a market crash. We speculate that severe downturns are more likely following daylight-saving weekends, and we argue that the data support this contention.

Pinegar remarks that we find no daylight-saving anomaly in our German data. This is inaccurate. While we find no statistical significance at conventional levels, we do find that the magnitude of the effect in Germany relative to the regular weekend effect is at least as large as in the rest of our sample, roughly six times the regular weekend effect. Lack of significance here might reflect the relatively limited number of daylight-saving-time changes affecting this market.

Our final remark pertains to Pinegar's statement that "sleep desynchronosis may contribute to the so-called 'day-of-the-week' effect on non-daylight-saving Mondays also" (p. 1256), an observation we communicated to him in private correspondence. He also points out that the crash of 1987 occurred on a "*non-daylight-saving*" Monday (italics his) which actually supports our conjecture that regular Monday effects may be a function of sleep desynchronosis. It is interesting to see Pinegar admit the potential role of sleep desynchronosis not only on daylight-saving weekends, but on all weekends. The evidence clearly supports the argument in our original paper that daylight-saving-time changes are indeed something to lose sleep over.

REFERENCES

Andrews, Donald W. K. "The Large Sample Correspondence between Classical Hypothesis Tests and Bayesian Posterior Odds Tests." *Econometrica*, September 1994, *62*(5), pp. 1207–32.

Connolly, Robert A. "A Posterior Odds Analysis of the Weekend Effect." *Journal of Econometrics*, July–August 1991, *49*(1–2), pp. 51–104.

Kamstra, Mark J.; Kramer, Lisa A. and Levi, Maurice D. "Losing Sleep at the Market: The Daylight Saving Anomaly." *American Economic Review*, September 2000, *90*(4), pp. 1005–11.

Lehmann, E. L. *Nonparametrics: Statistical methods based on ranks.* San Francisco, CA: Holden-Day, 1985.

Pinegar, J. Michael. "Losing Sleep at the Market: The Daylight Savings Anomaly: Comment." *American Economic Review*, September 2002, *92*(4), pp. 1251–56.