Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

Volume 36, Issue 4
April 2012
ISSN 0378-4266

ELSEVIER

Journal of BANKING & FINANCE

Editorial Board
Editor:
I. Mathur
Assistant Editor:
T. Bartlett
Advisory Board:
G.M. Constantinides
R. Engle
K.R. French
C.M. James
F. Moshirian
R.W. Roll
Associate Editors:
L.F. Ackert
G. Alexander
C. Almeida
T. Bali
J.A. Batten
P.-H. Chou
D.W. Diamond
F. Fiordelisi
K. Giesecke
M. Goergen
M. Gordy
A. Guariglia
A. Guettler
C.R. Harvey
T. Hens
D.B. Humphrey
R. Ibragimov
V. Ivashina
B. Jacobsen
S.A. Johnson
G.A. Karolyi
M. Levy
E. Liljeblom
B. Marshall
J. Mencia
A. Metrick
S. Ongena
M.A. Peterson
R.G. Rajan
L. Renneboog
P. Roosenboom
K. Rydqvist
V. Salas-Fumás
O. Scaillet
A. Subrahmanyam
B.M. Tabak
A.V. Thakor
P. Tufano
S. Van Nieuwerburgh
L. Wall
Y.P. Yafeh
V. Zakamouline

CONTENTS

(Contents continued on outside back cover)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Journal of Banking & Finance 36 (2012) 934–956

Contents lists available at SciVerse ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

A careful re-examination of seasonality in international stock markets: Comment on sentiment and stock returns

Mark J. Kamstra [a], Lisa A. Kramer [b,*], Maurice D. Levi [c]

[a] Schulich School of Business, York University, 4700 Keele Street, Toronto, Ont., Canada M3J 1P3
[b] Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ont., Canada M5S 3E6
[c] Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC, Canada V6T 1Z2

A R T I C L E I N F O

Article history:
Received 12 October 2010
Accepted 8 October 2011
Available online 18 October 2011

JEL classification:
G10
G11
G12

Keywords:
Seasonal affective disorder
SAD
Seasonal depression
Stock market cycles
Return seasonality

A B S T R A C T

In questioning Kamstra, Kramer, and Levi's (2003) finding of an economically and statistically significant seasonal affective disorder (SAD) effect, Kelly and Meschke (2010) make errors of commission and omission. They misrepresent their empirical results, claiming that the SAD effect arises due to a "mechanically induced" effect that is non-existent, labeling the SAD effect a "turn-of-year" effect (when in fact their models and ours separately control for turn-of-year effects), and ignoring coefficient-estimate patterns that strongly support the SAD effect. Our analysis of their data shows, even using their low-power statistical tests, there is significant international evidence supporting the SAD effect. Employing modern, panel/time-series statistical methods strengthens the case dramatically. Additionally, Kelly and Meschke represent the finance, psychology, and medical literatures in misleading ways, describing some findings as opposite to those reported by the researchers themselves, and choosing selective quotes that could easily lead readers to a distorted understanding of these findings.

© 2011 Elsevier B.V. All rights reserved.

"The reports of my death are greatly exaggerated."
– Mark Twain

1. Introduction

It is testimony to the widespread interest in seasonality in equity returns, corresponding in strength and nature to the presence of SAD in markets at different latitudes and hemispheres, that researchers such as Kelly and Meschke (2010; hereafter KM) are drawn to investigate the phenomenon. We establish that the SAD effect survives and is even strengthened by KM's examination. We show, with KM's own data, that the SAD effect first documented by Kamstra, Kramer, and Levi (2003; henceforth KKL2003) is a robust, economically meaningful, and statistically significant feature of financial markets. We also show that in challenging KKL2003, KM take liberties with the data, the literature they cite, and the literature they choose not to cite.

Errors of commission and omission emerge on even casual inspection of their estimation techniques. Perhaps most pertinent, KM mislead readers by describing the SAD effect as a turn-of-the-year effect when in fact their model (and our model) controls explicitly for the turn of the year. KM also introduce a new specification (consisting of three variables to capture the SAD effect) and then test the significance of the three variables *one-at-a-time*, rather than performing a joint test with a (standard) *F*-test. As we show, joint tests strongly reject the null of no SAD effect, with their data and their model, but one-at-a-time tests are compromised by multicollinearity in their new three-variable specification, further misleading readers that there is no SAD effect. Further, KM do not explore joint tests of the SAD hypothesis *across their data series.* Instead they use single-series-at-a-time tests and ordinary least squares (OLS) estimation, and they ignore modern methods such as system-of-equations generalized method of moments (GMM). KM use heteroskedasticity-robust standard errors, when heteroskedasticity and autocorrelation consistent (HAC) standard errors with data-dependent window width selection techniques are appropriate. GMM and HAC standard errors, which are commonly employed, are powerful and robust techniques that allow precise estimation of parameters and standard errors even in the presence of autocorrelation and heteroskedasticity. GMM is the standard for

* Corresponding author. Tel.: +1 416 978 2496; fax: +1 416 971 3048.
E-mail addresses: mkamstra@yorku.ca (M.J. Kamstra), Lkramer@rotman.utoronto.ca (L.A. Kramer), maurice.levi@sauder.ubc.ca (M.D. Levi).

0378-4266/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2011.10.010

Author's personal copy

performing system-of-equations estimation with equity returns data. See Hodrick and Zhang (2001), Jagannathan and Wang (2007), Bekaert et al. (2009), and Albuquerque et al. (2009). Nonetheless, we find significant evidence of the SAD effect even using OLS methods such as seemingly unrelated regression with panel/time-series estimation.

As Hirshleifer and Shumway (2003) argue persuasively, joint tests using panel data are more powerful than one-at-a-time single equation tests. We acknowledge that in KKL2003 we did not exploit the full power of systems equation estimation, joint tests, or the most powerful HAC standard error estimates available. The aim was to soundly show that the SAD effect is large and easily statistically significant, and so we took a conservative testing approach. Since KM question the very existence of a SAD effect, it is appropriate for them to give the established result the benefit of the doubt, and use the most powerful tests available. When we perform panel/time-series estimation and joint tests on KM's data, exploiting GMM and HAC, we easily reject the null of no SAD effect. Had KM paid attention to the characteristics of the data, for instance that their own coefficient estimates are almost always the sign and magnitude predicted by SAD, they would have reached different conclusions. KM's own results, as inefficient as their test procedures are, strongly support the SAD hypothesis, but this support is obscured by their reporting conventions and introduction of spuriously correlated regressors, as we detail below.

We also note the selective choice of studies KM cite and their incomplete description of the large and growing body of research on the SAD effect. First, they paint a one-sided picture of the SAD literature in finance. An even-handed exposition would cite not only the papers that contest the SAD hypothesis, but also the growing list of supportive papers. They write, "While there is a large and growing literature that uses KKL2003 to motivate their research, several other studies are critical of the SAD hypothesis" (p. 1309). There is no mention or analysis of the particular papers that find support for the SAD hypothesis, in spite of the fact that in some cases those papers use virtually the same data KM consider but come to very different conclusions. Second, there are multiple instances in which KM mischaracterize several established results in the psychology literature. For instance, they claim there is "mixed" evidence that depression is associated with increased risk aversion when in fact the evidence is overwhelmingly supportive on this point. And third, they misrepresent several papers in the finance literature, for example implying that Goetzmann and Zhu (2005) overturn the relationship between length of day and investor behavior when in fact Goetzmann and Zhu do not study length of day (nor do they claim to). We elaborate on all of these shortcomings below.

We describe the statistical and econometric problems inherent in KM's analysis in Section 2. In Section 3 we highlight the errors and bias KM reveal in their discussion of the finance literature. In Section 4 we describe KM's errors in citing the psychology and medical literatures. In Section 5 we revisit the empirical analysis using methods that do not exhibit the econometric problems of KM's analysis; we report results based on various model specifications, including single-equation OLS as well as several panel/time-series models that exploit cross-market correlation. Finally, in our Appendix A we describe the problems inherent in KM's Appendix A.

2. Statistical/econometric problems

In this section we describe statistical problems inherent in KM's analysis. Because KM employ single-equation estimation techniques, our discussion in this section mostly refers to results based on these methods. In Section 5 we report on more powerful system-of-equations methods appropriate for the analysis of cross-correlated series such as we have here.

2.1. Mechanical inducement of statistical significance

In describing their concern with the model specification KKL2003 employ, KM write:

> To illustrate, consider if returns were quite large during winter but in fall no different from spring and summer. In a specification with a fall and a fall–winter dummy, the fall–winter dummy would capture the positive winter returns and implicitly attribute them to the entire period from fall to winter, ... Hence, the overlap between the two dummies would mechanically induce statistical significance where a properly specified model would find none. (p. 1309)

There are many problems with KM's illustration. First, while we challenge the validity of KM's illustration of a "mechanical effect" below, even if we accept the validity of their illustration, this "mechanical effect" disappears when one controls for the large winter returns (i.e., when one controls for a turn-of-the-year effect). That is, their illustration is based on a misspecified model we do not estimate, with two *overlapping dummy variables* and *no control for a turn-of-the-year effec*t. In our analyses, we always control for a turn-of-the-year effect (and we do not employ overlapping dummy variables). In extended analysis we describe below, we find strong evidence supporting the SAD effect, even when controlling for a turn-of-the-year effect in multiple ways. That is, evidence in support of the SAD effect is not an artifact of failing to control for a turn-of-the-year effect. KM's suggestion to the contrary is simply incorrect, as we show.

Second, properly specified *tests* are just as important as properly specified regression models. After controlling for a turn-of-the-year effect, a careful test of KKL2003's SAD hypothesis would explore the joint significance of the SAD variables, namely the fall dummy and the length of night variable. *F*-tests are appropriate when one has a joint hypothesis on coefficients in a regression, in particular a regression that controls separately for, in this case, a turn-of-the-year effect. *F*-tests can also have much greater power than one-at-a-time *t*-tests when the individual variables (the length of night variable and the fall dummy variable in this case) are overlapping and correlated – features of these variables that KM enthusiastically highlight. But KM employ one-at-a-time *t*-tests on these variables, never discussing the joint significance of the variables intended to capture the SAD effect. Although we did not report these joint tests in KKL2003, we did perform such tests (in the context of a model that properly controlled for the turn-of-the-year effect) and the tests indicate the individually negative significant fall dummy variable and positive significant length of night variable are strongly *jointly* significant. (Note that we provide these tests below.) That is, the individual significance of each variable is not "mechanically" induced by ignoring large positive returns around the turn of the year. Rather than performing such joint tests when they propose a re-examination of the SAD effect in their Section 6.1, KM instead introduce a new, more disaggregated specification of KKL2003's SAD model and perform one-at-a-time *t*-tests on this new, less parsimonious specification. They remark:

> In this section we show that the SAD interaction term does not differ materially from a fall–winter dummy and that the SAD effect is mechanically driven by a *de facto* overlapping dummy-variable specification and higher returns around the turn of the year. ...A simple way to test whether the overlap of the SAD and fall variables drives the significant results on the fall dummy is to split the SAD variable into fallSAD and

Author's personal copy

winSAD ... where fallSAD is a fall dummy interacted with normalized length of night, and winSAD is a winter dummy interacted with normalized length of night. If the original model (Eq. (1)) is correctly specified, decomposing the SAD interaction term into fallSAD and winSAD should not affect the results. In contrast, if the significance of the fall dummy is mechanically induced, splitting the SAD variable should eliminate the significance of the fall dummy. (p. 1317)

KM's new, more complicated model is not a simple way to investigate whether higher returns around the turn of the year are producing some sort of mechanical SAD effect. The *simple* way is to control for the turn of the year and test for the joint significance of the SAD specification variables. When we do this, we find a strong and statistically significant SAD effect, independent of the turn-of-the-year effect. This is clear from our model specification and tests in KKL2003 and elsewhere, specifications that include turn-of-the-year control variables.

Third, we note KM's faulty logic. Splitting a variable into two separate variables *certainly can* affect results. This is statistical analysis, and increasing the number of parameters to estimate the SAD effect from two to three impacts power, systematically degrading one's ability to find significant effects. We note that the KKL2003 specification and the KM specification (which are identical except KM split SAD into two halves by seasons: fallSAD and winSAD) lead to roughly the same number of indices displaying the expected (negative) sign on the fall variable (roughly three quarters of the indices), and the magnitude of these coefficients is virtually identical across the KKL2003 and KM specifications. What KM discover by splitting the SAD variable into two halves is that standard errors get larger with extra parameters being estimated. This does not in any way constitute a definitive test for a so-called mechanical effect.

Fourth, KM assert that KKL2003's length of night variable "does not differ materially from a fall–winter dummy" (p. 1317). They fail to support this assertion, and their analysis based on the fallSAD and winSAD variables they propose is unable to shed light on the assertion since neither fallSAD nor winSAD is a dummy variable. That is, they make an assertion with reference to one model (based on a fall/winter dummy, equal to 1 for months in the fall or winter) and test it using a different model (based on splitting the SAD variable into fallSAD and winSAD). To correct their omission, we now consider KM's Eq. (2), and building on this specification construct two models that permit an investigation of material differences between KKL2003's length of night variable and the fall/winter dummy variable KM refer to in their assertion. In KM's Eq. (2), we replace the fallSAD and winSAD variables with either (i) a fall/winter dummy variable, to form Eq. (1) below, or (ii) KKL2003's length of night variable, to form Eq. (2) below. In addition, because KM fixate on the possibility that a turn-of-the-year effect facilitates a spurious SAD result, we include in these models a dummy variable for the first month of the fiscal year, which we label MTax, soaking up that month's return altogether.[1] That is, our Eqs. (1) and (2) include a tax-loss-selling variable for the few days around the turn of the year and also incorporate a variable for the full month of the start of the tax year. Finally, we estimate a third specification, identical to KM's Eq. (2), but with the addition of the MTax dummy variable. Note that omitting MTax from Eqs. (1)–(3) does not qualitatively change our results, so that the inclusion of this variable is not in itself generating a new "mechanical effect." The models are as follows:

[1] We thank an anonymous referee for recommending the inclusion of this variable to address KM's concern that it is otherwise empirically difficult to distinguish between, for instance, a January effect and the SAD effect.

$$\begin{aligned} y_{i,t} = {} & \alpha_i + \rho_{1,i} y_{i,t-1} + \rho_{2,i} y_{i,t-2} + \beta_{i,\mathrm{Tax}} \mathrm{Tax}_{i,t} + \beta_{i,\mathrm{MTax}} \mathrm{MTax}_{i,t} \\ & + \beta_{i,\mathrm{Monday}} \mathrm{Monday}_t + \beta_{i,\mathrm{Fall}} \mathrm{Fall}_{i,t} + \beta_{i,\mathrm{FallWinter}} \mathrm{FallWinter}_{i,t} \\ & + \beta_{i,\mathrm{Temp}} \mathrm{Temp}_{i,t} + \beta_{i,\mathrm{Cloud}} \mathrm{Cloud}_{i,t} + \beta_{i,\mathrm{Rain}} \mathrm{Rain}_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{1}$$

$$\begin{aligned} y_{i,t} = {} & \alpha_i + \rho_{1,i} y_{i,t-1} + \rho_{2,i} y_{i,t-2} + \beta_{i,\mathrm{Tax}} \mathrm{Tax}_{i,t} + \beta_{i,\mathrm{MTax}} \mathrm{MTax}_{i,t} \\ & + \beta_{i,\mathrm{Monday}} \mathrm{Monday}_t + \beta_{i,\mathrm{Fall}} \mathrm{Fall}_{i,t} + \beta_{i,\mathrm{SAD}} \mathrm{SAD}_{i,t} \\ & + \beta_{i,\mathrm{Temp}} \mathrm{Temp}_{i,t} + \beta_{i,\mathrm{Cloud}} \mathrm{Cloud}_{i,t} + \beta_{i,\mathrm{Rain}} \mathrm{Rain}_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{2}$$

$$\begin{aligned} y_{i,t} = {} & \alpha_i + \rho_{1,i} y_{i,t-1} + \rho_{2,i} y_{i,t-2} + \beta_{i,\mathrm{Tax}} \mathrm{Tax}_{i,t} + \beta_{i,\mathrm{MTax}} \mathrm{MTax}_{i,t} \\ & + \beta_{i,\mathrm{Monday}} \mathrm{Monday}_t + \beta_{i,\mathrm{Fall}} \mathrm{Fall}_{i,t} + \beta_{i,\mathrm{FallSAD}} \mathrm{FallSAD}_{i,t} \\ & + \beta_{i,\mathrm{WinSAD}} \mathrm{WinSAD}_{i,t} + \beta_{i,\mathrm{Temp}} \mathrm{Temp}_{i,t} + \beta_{i,\mathrm{Cloud}} \mathrm{Cloud}_{i,t} \\ & + \beta_{i,\mathrm{Rain}} \mathrm{Rain}_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{3}$$

where $y_{i,t}$ is the return to country/index i at time t; $\mathrm{Tax}_{i,t}$ is is a dummy variable equal to one on the first five trading days and the last trading day of a country's fiscal year and zero otherwise; $\mathrm{MTax}_{i,t}$ is a dummy variable equal to one in the first month of country i's fiscal year and zero otherwise; Monday_t is a dummy variable for trading days on Mondays; $\mathrm{FallWinter}_{i,t}$ is a dummy variable equal to one between September 21 and March 20 for countries in the northern hemisphere, equal to one between March 21 and September 20 in the southern hemisphere, and zero otherwise; $\mathrm{Fall}_{i,t}$ is a dummy variable equal to one between September 21 and December 20 for countries in the northern hemisphere, equal to one between March 21 and June 20 in the southern hemisphere, and zero otherwise; $\mathrm{SAD}_{i,t}$ is the normalized length of night variable for country i (as defined by KKL2003); and $\mathrm{Temp}_{i,t}$, $\mathrm{Cloud}_{i,t}$, and $\mathrm{Rain}_{i,t}$ are KM's daily weather variables. Following KKL2003 and KM, we include two lags of the dependent variable in each model. We thank Professors Kelly and Meschke for kindly providing us with their data.

Table 1 contains estimation results for Eqs. (1)–(3) for the full set of indices KM investigate, with results for a single index per line. (Later we consider estimation results eliminating duplication in KM's set of indices, ensuring any given country is represented by no more than one index.) The SAD hypothesis does not imply one should necessarily find evidence of a SAD effect in exchanges close to the equator where the variation in daylight across the year is minimal, but it does imply we should see relatively more and stronger evidence of SAD as we consider indexes from exchanges increasingly far from the equator.[2] Thus we group the exchanges into two sets based on proximity to the equator: exchanges located in the tropics and subtropics (between the equator and 40° latitude, both north and south), and exchanges located above 40° (these are exclusively northern hemisphere exchanges). Panel A contains results for the exchanges above 40°, where we expect the strongest effects due to SAD, and Panel B contains results for the exchanges in the tropics and sub-tropics. We report only coefficient estimates for the variables of interest related to the SAD hypothesis, due to space constraints. We provide p-values for joint tests of significance of a given equation's SAD variables, and for Eq. (2) we also present the estimated returns due to SAD for the second month in each of fall and spring (results are similar for all three models). At the bottom of each panel we also present the average for each coefficient estimate across each of the two latitude groupings and average p-values. We also indicate the proportion of indices in each grouping with a p-value below the 5% and 10% cutoff levels for testing the significance of a given SAD coefficient. All individual coefficients

[2] The SAD hypothesis rests on SAD prevalence estimates from medical research. Several studies, including Lam (1998) and Magnusson's (2000) survey, report that SAD is more prevalent at higher latitudes. Some studies suggest 40° latitude as a meaningful cutoff above which SAD is predominant; see, for instance, Morrissey et al. (1996, p. 584). Further, several studies report that SAD symptoms remit when patients relocate close to the equator (Lam, 1998, for instance) and that symptoms are milder close to the equator (Rosenthal et al., 1984, for example).

Author's personal copy

significant at the 10% level or better are indicated in bold font, and all tests are two-sided.

By comparing and contrasting regression results from estimating Eqs. (1) and (2), we can *directly* investigate KM's assertion that the length of night variable "does not differ materially from a fall–winter dummy" (p. 1317). Consider first the over-40° exchanges, shown in Panel A of Table 1. A joint test of the statistical significance of KKL2003's SAD and fall dummy variables in Eq. (2), using KM's data, show that at the 5% (10%) level 55% (68%) of the series demonstrate a statistically significant joint effect. A joint test of statistical significance of the KM fall/winter dummy variable plus a fall dummy variable in Eq. (1), using KM's data, show that at the 5% (10%) level 26% (35%) of the series demonstrate a statistically significant joint effect. That is, a simple fall/winter dummy variable specification fares markedly worse that the KKL2003 SAD specification. The difference in the performance of Eqs. (1) and (2) is even more marked if we consider a *t*-test on the KKL2003 SAD variable versus a *t*-test on the KM fall/winter dummy variable. A *t*-test on the SAD variable in Eq. (2) shows that at the 5% (10%) level 58% (71%) of the series demonstrate a statistically significant SAD coefficient whereas a *t*-test on the KM fall/winter variable in Eq. (1) shows that at the 5% (10%) level 10% (13%) of the series demonstrate a statistically significant fall/winter coefficient. Since we control for the turn-of-the-year effect with two separate variables, a tax year variable and a dummy for the first month of the fiscal year, the sign and significance of the SAD variable is *not* driven by large positive returns from the turn-of-the-year (contrary to KM's claim). Finally, a *t*-test on the fall dummy variable in Eq. (2) shows that at the 5% (10%) level approximately 61% (74%) of the series demonstrate a statistically significant fall effect whereas a *t*-test on the fall dummy variable in the KM overlapping dummy variable model (Eq. (1)) shows that at the 5% (10%) level 48% (61%) of the series demonstrate a statistically significant fall effect. Altogether, the KKL2003 fall dummy and length of night variables perform better in capturing seasonality than the overlapping dummy variable specification KM have in mind. This directly refutes KM's claim that "the SAD effect is mechanically driven by a *de facto* overlapping dummy-variable specification and higher returns around the turn of the year" (p. 1317). The KKL2003 specification does *not* behave like an overlapping dummy variable model, nor does the return around the turn-of-the-year inflate the significance of the SAD effect, as we *explicitly* control for this return in KKL2003's specification and in the expanded specifications above.

We are also puzzled by KM's exclusive reliance on one-at-a-time *t*-tests on the significance of the fall dummy, fallSAD, and winSAD in estimating their Eq. (2), reported in their Section 6.1. By design, the SAD effect in KM's Eq. (2) is captured by *all three variables*, and given the overlapping nature of the variables in their model, one-at-a-time *t*-tests may be insignificant due to multicollinearity. Joint tests are appropriate in this context. In our Eq. (3) we re-estimate KM's SAD model incorporating the MTax dummy variable. The results are in the last set of columns of Table 1. For now, we restrict our attention to results for exchanges above 40° latitude, reported in Panel A. Joint tests of statistical significance on the fallSAD, winSAD, and fall dummy variables in Eq. (3), using KM's data, show that at the 5% (10%) level 68% (68%) of the series exhibit a statistically significant joint SAD effect. This is strong support for the SAD hypothesis. Note that Eq. (3) controls explicitly for a turn-of-the-year effect with two different variables, thus the joint significance of these variables is *not* mechanically driven by higher returns around the turn of the year, refuting KM's claims to the contrary.

2.2. Statistical power

KM utilize a very large panel and largely overlapping time-series dataset of returns, a dataset which exhibits cross-sectional covariance, heteroskedasticity, and autocorrelation. They analyze their data series one-at-a-time with OLS, and they employ MacKinnon and White (1985) heteroskedasticity-consistent errors. KKL2003 apply similar techniques to return series over a much longer time span, and these series have long non-overlapping sub-periods across countries. Further, KKL2003 consider only countries where the incidence of SAD is well-documented to be high, in locations far from the equator. KM consider a great many equatorial countries and exchanges, in which the prevalence of SAD is low or non-existent and in which one ought not to expect to find a SAD effect in stock returns (consistent with Dowling and Lucey's, 2008, findings). We find that it is KM's shorter samples, their use of data from many exchanges located near the equator, and their use of single-equation-at-a-time estimations that lead to the reduced incidence of significant SAD effects in their results. The compromised power of KM's tests is evidenced most clearly by reference to the fact that KM's coefficient estimates throughout their tables are the same order of magnitude as those KKL2003 report, reproduced in KM's Table 2, Panel A. KM's coefficients are less statistically significant because of the larger standard errors associated with the shorter samples. Also contributing to the scarcity of significance in KM's analysis is their decision not to report rejections at the 10% level of significance, an unusual choice relative to the financial economics literature and one they elect not to mention as a source of difference between their findings and KKL2003's. Altogether, these three features of KM's analysis lead them to conclude that the SAD effect is "insignificant" in their extended set of country-indices. As we discuss below, even using the shorter samples and the extended range of countries located close to the equator, we find strong evidence of SAD based on panel/time-series estimation, controlling for cross-sectional correlation of returns.

KM further undermine support for SAD effects in a variety of ways beyond even their use of short samples, data from equatorial countries, and aggressive significance cutoff values (that is, ignoring significance at the 10% level). Results in their Table 8 are based on one-at-a-time regressions using no more than 1 year of data to estimate an *annual* cyclicality. Even taking account of their assurance that they require at least 125 days of data in each year, testing for an annual effect using a year or less of data is questionable. In Table 6 KM present results based on KM's Eq. (3), a specification consisting of a fall dummy variable in combination with six monthly dummies (defined over the period September 21 to March 21, each interacted with length of night). This evidence is even more problematic than that they report in their Table 8, as it consumes 7 degrees of freedom instead of 2 to estimate a single effect. Further, all of these monthly length-of-night variables are strongly collinear with the weather variables, impacting signs and mechanically *reducing* the significance of parameter estimates. We note this with some irony, in light of KM's assertion that KKL2003's specification "mechanically" *induces* an effect.

KM also compromise statistical power by neglecting to use joint tests on these multiple-variable SAD specifications. In untabulated joint tests on the six length-of-night variables plus the fall dummy variable from KM's Eq. (3) using KM's own data, we find that at the 5% (10%) level of significance 35% (39%) of the northern (above 40° latitude) series demonstrate a statistically significant SAD effect. If we restrict ourselves to the fall dummy variable and just the three length-of-night variables from the fall half of the period the SAD_t variable covers (October, November, and December, covering September 21st to December 21st) to avoid the turn of the year altogether, using KM's data we still find over 23% (32%) of the northern (above 40°) series reject the null hypothesis of no SAD effect at the 5% (10%) level. These are *particularly* weak tests, performed one series at a time on heavily over-parameterized models, yet again the tests show strong evidence of SAD effects across many indices

Author's personal copy

Table 1
Key coefficient estimates, *p*-values, and magnitudes for Eqs. (1)–(3).

Country, index, date range, latitude	Eq. (1) Fall	Eq. (1) Fall/winter	Eq. (1) Joint effect *p*-value	Eq. (2) Fall	Eq. (2) SAD	Eq. (2) Joint effect *p*-value	Eq. (2) Fall magnitude SAD effect[a]	Eq. (2) Winter magnitude SAD effect[a]	Eq. (3) Fall	Eq. (3) FallSAD	Eq. (3) WinSAD	Eq. (3) Joint effect *p*-value
Panel A: Exchanges located above 40° latitude												
Finland, HEX General Index, 1987–2008, 64°N	.004	–.009	.995	.013	–.004	.954	.000	–.017	.105	–.019	.012	.646
Iceland, OMX Iceland All Share, 1993–2008, 64°N	**–.195**	.059	**.088**	**–.205**	.015	.145	–.163	.055	–.217	.017	**.013**	.130
Norway, Total Market, 1980–2008, 62°N	**–.111**	–.012	**.046**	**–.180**	**.022**	**.012**	–.122	.076	**–.153**	.017	**.028**	**.017**
Sweden, Veckans Affärer, 1982–2003, 59°N	**–.115**	.090	.131	**–.117**	.023	.205	–.066	.067	–.067	.012	**.034**	.124
Sweden, OMX Affärsvärldens Generalinde, 1980–2008, 59°N	**–.098**	.087	.134	**–.103**	**.023**	.161	–.051	.068	–.036	.008	**.038**	**.042**
Denmark, Copenhagen KFX, 1989–2008, 56°N	–.040	.018	.750	–.077	.022	.344	–.033	.058	.033	–.005	**.050**	**.027**
Ireland, Total Market, 1973–2008, 53°N	**–.162**	**.177**	**.000**	**–.124**	**.039**	**.001**	–.057	.089	–.035	.012	**.050**	**.000**
Netherlands, AEX Index, 1973–2008, 52°N	–.046	–.001	.416	**–.080**	.018	.257	–.050	.039	–.073	.016	.019	.375
UK, FTSE100, 1984–2008, 51°N	–.040	.059	.467	–.029	.015	.553	–.005	.031	.001	.005	.019	.639
UK, Total Market, 1973–2008, 51°N	**–.084**	**.086**	**.042**	**–.075**	**.027**	**.028**	–.031	.058	–.027	.011	**.034**	**.026**
Belgium, Banque Bruxelles Lambert 30, 1973–2008, 50°N	**–.063**	.012	.121	**–.091**	**.022**	**.045**	–.058	.044	–.064	.012	**.029**	**.038**
Germany, DAX100, 1988–1998, 50°N	**–.147**	.102	.123	**–.213**	**.081**	**.019**	–.085	.166	–.146	.059	**.102**	**.010**
Austria, ATX50, 1973–2008, 47°N	**–.083**	.038	**.030**	**–.104**	**.031**	**.008**	–.062	.057	–.053	.012	**.050**	**.001**
Switzerland, Total Market, 1973–2008, 47°N	–.010	–.015	.756	–.042	.017	.463	–.019	.031	–.005	.003	**.030**	.199
France, Total Market, 1973–2008, 46°N	**–.084**	.034	**.100**	**–.093**	.021	**.083**	–.065	.037	–.026	–.005	**.042**	**.015**
Canada, TSX300, 1973–2008, 43°N	–.054	.038	.373	**–.115**	**.065**	**.001**	–.037	.102	–.098	**.058**	**.072**	**.001**
Italy, Total Market, 1973–2008, 42°N	**–.096**	.049	.102	**–.107**	.034	**.075**	–.068	.051	–.052	.009	**.057**	**.034**
US, DJIA 1948–2008, 41°N	–.025	.039	.358	–.034	**.028**	**.068**	–.003	.041	–.038	**.030**	**.026**	.144
US, EW AMEX ex-Div, 1962–2008, 41°N	**–.084**	.037	**.000**	**–.102**	**.035**	**.000**	–.064	.050	**–.057**	.013	**.053**	**.000**
US, EW NYSE ex-Div, 1948–2008, 41°N	**–.046**	.030	.123	**–.069**	**.035**	**.002**	–.031	.051	**–.067**	**.034**	**.036**	**.004**
US, EW NASDAQ ex-Div, 1972–2008, 41°N	**–.081**	.029	**.008**	**–.107**	**.037**	**.000**	–.066	.054	–.060	.015	**.056**	**.000**
US, S&P500, 1948–2008, 41°N	–.009	.015	.854	–.027	**.023**	.170	–.002	.034	–.040	.029	.018	.314
US, VW AMEX ex-Div, 1962–2008, 41°N	**–.054**	.019	.105	**–.096**	**.046**	**.000**	–.046	.066	–.065	.031	**.058**	**.000**
US, VW NYSE ex-Div, 1948–2008, 41°N	–.013	.017	.785	–.032	**.024**	.107	–.005	.035	–.040	.028	**.021**	.211
US, VW NASDAQ ex-Div, 1972–2008, 41°N	–.041	.020	.648	**–.099**	**.059**	**.018**	–.034	.086	–.080	.050	**.067**	**.027**
US, EW AMEX ex-Div, 1962–2007, 41°N	**–.087**	**.050**	**.000**	**–.093**	**.032**	**.000**	–.059	.046	–.037	.005	**.055**	**.000**
US, EW NYSE, 1948–2007, 41°N	**–.045**	**.042**	**.078**	**–.055**	**.031**	**.009**	–.021	.045	–.047	**.027**	**.035**	**.016**
US, EW NASDAQ, 1972–2007, 41°N	**–.080**	.044	**.006**	**–.092**	**.033**	**.002**	–.056	.048	–.031	.004	**.057**	**.000**
US, VW AMEX, 1962–2007, 41°N	**–.056**	.034	.105	**–.088**	**.045**	**.001**	–.039	.065	–.047	.025	**.062**	**.000**
US, VW NYSE, 1948–2007, 41°N	–.013	.029	.491	–.023	**.024**	**.096**	.003	.034	–.025	.025	**.022**	.195
US, VW NASDAQ, 1972–2007, 41°N	–.045	.042	.589	**–.088**	**.058**	**.021**	–.024	.084	–.045	.038	**.075**	**.019**
Average value	–.068	.041	.285	–.089	.032	.124	–.046	.057	–.051	.043	.019	.105
Proportion of p-values ⩽ .05/.10	.48/.61	.10/.13	.26/.35	.61/.74	.58/.71	.55/.68			.00/.10	.03/.13	.68/.87	.68/.68
Panel B: Exchanges located in the tropical and sub-tropical latitudes (40° and below)												
Spain, Madrid SE General, 1974–2008, 40°N	–.060	.029	.393	–.063	.020	.397	–.042	.028	–.030	.003	.034	.356
China, Total Market, 1991–2008, 40°N	–.113	–.001	.638	–.029	–.093	.344	–.135	–.126	.095	–.151	–.041	.434
Greece, Total Market, 1988–2008, 39°N	**–.138**	–.017	**.062**	**–.220**	.063	**.024**	–.157	.085	**–.275**	.092	.040	**.059**
Turkey, Total Market, 1988–2008, 39°N	.153	–.006	.268	.087	.054	.216	.141	.073	.126	.033	.070	.332
Korea, Korea South Composite (KOSPI), 1975–2008, 37°N	.065	–.063	.591	.043	–.015	.806	.028	–.019	–.013	.013	–.049	.751
Japan, Nikkei 225, 1953–2008, 36°N	–.013	–.047	.395	–.049	.013	.464	–.038	.015	–.046	.011	.014	.662
Jordan, Amman SE Financial Market, 1988–2008, 31°N	.111	–.087	.262	.077	–.030	.434	.055	–.028	.030	.005	–.060	.567
Mexico, FTSE Mexico Index, 1987–2008, 23°N	–.133	.051	.312	**–.171**	.132	.122	–.109	.084	–.132	.088	.168	.189
Taiwan, Taiwan Weighted, 1973–2001, 23°N	**–.166**	.034	**.068**	**–.221**	.132	**.038**	–.160	.082	–.176	.085	.177	**.067**
Hong Kong, Total Market, 1973–2008, 22°N	**–.144**	.118	.121	**–.155**	**.201**	**.015**	–.066	.119	–.019	.050	**.283**	**.003**
India, National Index (100), 1989–2008, 20°N	–.097	.033	.520	**–.164**	**.216**	**.079**	–.079	.114	–.149	.197	**.228**	.159
Thailand, Stock Exchange of Thailand Index, 1975–2008, 15°N	–.031	.012	.861	–.032	.033	.843	–.024	.012	.042	–.104	.167	.414
Philippines, P.S.E. Composite Index, 1986–2008, 13°N	–.021	.007	.965	–.070	.194	.497	–.032	.056	.019	–.008	**.443**	.226
Sri Lanka, All Share, 1985–2008, 7°N	–.013	.023	.920	–.017	.196	.682	–.009	.018	–.004	.116	.252	.858
Malaysia, Composite, 1980–2008, 2°N	–.066	**.100**	.145	.020	.232	.859	.002	–.015	.040	.571	.093	.928
Singapore, Total Market, 1973–2008, 1°N	**–.079**	.054	.220	–.058	–.307	.404	–.027	.029	–.069	–.421	–.299	.609

Author's personal copy

Indonesia, Jakarta Composite Index, 1983–2008, 5°S	**.109**	**−.126**	**.053**	.078	**−.247**	.129	.017	−.070	−.030	.087	**−.289**	.169
South Africa, Total Market, 1973–2008, 26°S	−.013	.014	.946	−.049	**.091**	.113	.023	.091	−.097	**.128**	**.083**	.178
Australia, Total Market, 1973–2008, 34°S	.003	.031	.617	.016	.006	.712	.022	.008	**.098**	−.041	.032	.189
Australia, All Ordinaries, 1980–2008, 34°S	.010	.040	.316	.034	.001	.389	.035	.002	**.119**	−.047	.028	**.098**
New Zealand, Capital 40, 1996–2004, 37°S	−.001	−.083	.412	−.070	.020	.487	−.047	.030	−.060	.015	.021	.696
New Zealand, FTSE New Zealand Index, 1996–2008, 37°S	−.033	−.059	.257	−.074	.008	.332	−.066	.012	−.063	.002	.009	.529
Average value	−.030	.003	.425	−.049	.042	.381	−.030	.027	−.027	.064	.033	.385
Proportion of p-values ⩽ .05/.10	.05/.23	.09/.09	.00/.14	.23/.23	.09/.18	.14/.18			.09/.14	.00/.05	.09/.23	.05/.18

Notes: We produce regression results with one-at-a-time OLS and MacKinnon–White (1985) heteroskedastity-consistent HC3 standard errors. Bolded coefficient values and *p*-values are significant at the 10% level or better. Data are from Patrick Kelly and Felix Meschke.

[a] We use October for calculating the magnitude of the SAD effect in the fall, and we use February for calculating the magnitude of the SAD effect in the winter.

and countries, and SAD effects in the fall season *alone* as well, with *no* "mechanical" overlap into the winter. The above percentages are based on estimating KM's Eq. (3) using seasonally unadjusted weather data. If instead we use *seasonally adjusted* weather data, as Hirshleifer and Shumway (2003) do, we find even stronger results: 42% (45%) of the above-40° series demonstrate a statistically significant joint SAD effect (in a model that uses seven variables to capture the SAD effect) at the 5% (10%) level.

KM additionally reduce the power of their inference by placing little emphasis on *patterns* that emerge in the data *across latitudes*. (Recall that the SAD hypothesis has different predictions for exchanges located close to versus far from the equator.) Panel B of Table 1 contains estimation results for exchanges in the sub-tropics and tropics, allowing for easy comparison to the more northerly exchanges. Recall (based on results in Panel A) that for Eqs. (2) and (3), 68% of the indices in countries located above 40° latitude display joint significance at the 10% level. In contrast, only 18% of the indices from countries in the sub-tropics and tropics display joint significance at the 10% level. For Eq. (2) we also report in Table 1 the impact of SAD on returns across the seasons. The average impact of SAD on returns in the month of October is −4.6% for indices from countries located above 40° and −3.0% for indices from countries in the sub-tropics and tropics.[3] The average impact of SAD on returns in February (having controlled separately for the turn-of-the-year effect as specified in Eq. (2)) is 5.7% for indices from countries located above 40° and 2.7% for indices in sub-tropical and tropical countries.[4,5] This relationship between the impact of SAD and the latitude of the exchange is consistent with the SAD hypothesis predictions. Furthermore, consider the coefficient signs for the most northerly exchanges for all three equations, shown in Panel A of Table 1. Virtually every coefficient value for the fall dummy variable across Eqs. (1)–(3) has the expected (negative) sign, and similarly the SAD variables across the three equations display the expected (positive) signed coefficient values for virtually every index. In Panel B (the sub-tropics and tropics), only about half of the indices have the expected signs on the fall and SAD variables. The relationship between the signs of coefficients and the latitude of the exchange supplements the relationship between the latitude of the exchange and both economic magnitude and statistical significance of the coefficient estimates. Overall, these results strongly support the SAD hypothesis.

2.3. Assessing signs and magnitudes

In addition to KM disregarding the fact that their coefficient estimates match the magnitudes KKL2003 report, so that KM's results hinge on inflating standard error estimates, KM also disregard the signs of their estimates, which are overwhelmingly consistent with the SAD hypothesis. Consider the results in KM's Table 7, where KM explore the onset/recovery variable, the current best measure of the timing of SAD in the population. (We discuss the SAD onset/recovery variable more fully in Section 5.) Over 80% of KM's onset/recovery coefficient estimates have the expected sign. This is in spite of KM having included in their data many likely uninformative markets (including many tropical exchanges, which are so close to the equator that they experience very little seasonal variation in light) and having placed them on equal footing with markets in large, industrialized, non-equatorial countries with

[3] We calculate this impact by multiplying the October value of $Fall_{it}$ (1) by the Eq. (2) coefficient value $\beta_{i,Fall}$.

[4] We calculate this impact by multiplying the February value of SAD_{it} by the Eq. (2) coefficient value $\beta_{i,SAD}$.

[5] The Hong Kong index contains many stocks cross-listed from the London exchange (FTSE). Hence, including the Hong Kong index in the tropics group inflates the magnitude of the SAD impact for that region. We comment on this point more fully at the end of this section.

Author's personal copy

broad-based economies. When Hirshleifer and Shumway (2003) consider similar patterns of signs in their analysis of the impact of cloud cover on returns (70% and 96% negatively signed cloud cover coefficients, depending on the model specification), they investigate statistical significance using *panel/time-series models*. When we return to this issue in Section 5 below using panel/time-series models, we find a strongly statistically significant pattern of SAD onset/recovery coefficient-estimate signs, even accounting for contemporaneous correlation across exchanges.

2.4. Use of prevalence statistics from KM's Appendix A

In their Table 4, KM list correlations between various variables, including "Prevalence of SAD." KM's SAD prevalence statistics are from a set of studies they list in their Appendix A, statistics which they claim represent "[p]revalence of seasonal affective disorder (from) general population studies." (p. 1324) Even a casual examination of these so-called "general population" studies leaves the reader skeptical of the use of these correlations to measure characteristics of general populations. For instance, one study KM use to form the average prevalence rate for Canada, Magnusson and Axelsson (1993), is titled "The prevalence of seasonal affective disorder is low among descendants of Icelandic emigrants in Canada." That study examines why people of Icelandic ancestry are far *less* susceptible to SAD than others. Magnusson and Axelsson purposefully set out to measure a sub-population of Canada because of differences between that sub-population and the general population with regard to SAD prevalence. We believe that study, among others KM reference, is not accurately described as a general population study of SAD.

Even setting aside concerns regarding the generality of these studies, there are additional well-documented problems that compromise cross-country comparisons of studies that report *general population* SAD prevalence rates. Nonetheless, in cases where KM consider multiple statistics for a given country, they average them assuming they are directly comparable. This causes particular problems in drawing inferences from KM's Table 4. For a detailed discussion of these issues, please see our Appendix A.

2.5. Other empirical issues

KM write: "If they are predictable *ex ante*, stock return seasonals constitute an important challenge to the efficient market hypothesis because rational traders should be able to exploit them for large economic gains" (p. 1308). In fact, market efficiency does not preclude predictable returns. Predictable returns resulting from swings in *risk aversion* may arise not only due to SAD, but also as a consequence of habit persistence in consumption or even as a reaction to a recession. (See Fama and French, 1989, and Campbell and Cochrane, 1999, for instance.) Predictable swings in risk aversion arising from habit persistence or recessions lead necessarily to predictable returns. Yet these returns, though predictable, are not exploitable. Similarly, the notion that SAD could lead to predictable returns is well within the bounds of efficient markets.

Additionally, any study that considers a large number of stock return indices must contend with the difficulty that arises in properly accounting for the unique characteristics of each exchange. For instance, the Hong Kong index behaves much like London's FTSE index with respect to SAD effects. This surprises KM (see p. 1316), but the Hong Kong index includes many stocks that are cross-listed on the London exchange, making Hong Kong returns behave very similarly to the London returns. One might reasonably wonder: how many other countries in KM's sample exhibit these sorts of relationships? Does cross-listing fully explain the similarity between the Hong Kong index and the FTSE index (and similarities between other exchanges that have cross-listed stocks)? KM do not address this issue.[6] Furthermore, KM refer to the length of night variable as "highly persistent." The length of night is deterministic, not persistent.

2.6. Eliminating duplication from the set of countries KM consider

As noted above, there is duplication in data KM consider. Specifically, they use multiple indices for Sweden, the UK, the US, Australia, and New Zealand. Here we consider the impact on the magnitudes and proportions reported in Table 1, eliminating the duplication. Table 2 contains coefficient estimates, *p*-values (for individual tests and joint tests), economic magnitudes, and proportions of statistics significant at the 5% and 10% levels of significance based on considering only one index (the longest available total market index) for each country with duplicate data in KM's sample.[7] (Note that for the US we employ an equal-weighted total market index instead of one of the US series KM employ for two reasons. First, it is a total market index, encompassing all of the securities captured in the individual series KM employ. Second, it is an *equal-weighted* index, which necessarily places relatively more weight on smaller, riskier stocks that are more likely to exhibit seasonally varying returns due to time-varying investor risk aversion.) The average coefficient values, proportion of significant statistics, and average economic values in Table 2 are virtually identical to those reported in Table 1, and where different, almost always show *stronger* support for a SAD effect. For instance, in Table 2 the average impact of SAD on returns in October is −5.8% for indices from countries located above 40° (versus −4.6% in Table 1). In Table 2 the average impact of SAD on returns in February (having controlled separately for the turn-of-the-year effect as specified in Eq. (2)) is 6.0% for indices from countries located above 40° (versus 5.7% in Table 1). And for Eq. (3), KM's model, 75% of the indices in countries above 40° display joint significance of the SAD coefficients at the 10% level in Table 2 (versus 68% in Table 1). Thus our earlier conclusions do not rest on the duplication in KM's sample.

3. The finance literature

KM make errors and commit omissions in describing results from the finance literature.

3.1. Omitted survey of the literature on SAD and financial markets

KM do not cite many of the papers that document the influence of SAD on financial markets, nor do they even broadly sketch their findings. This omission denies the reader insight as to why there is a large and growing SAD literature. The evidence supporting the impact SAD has on financial markets includes the following papers omitted from KM's discussion: Dolvin et al. (2009) and Lo and Wu (2008) who study analysts' stock earnings forecasts, Dolvin and Pyles (2007) and Kliger et al. (2010) who investigate the underpricing of initial public offerings, Pyles (2009) who considers returns to real estate investment trusts, Kaplanski and Levy (2009) who study the influence of SAD on volatility through the VIX, and Kliger and Levy (2008) who study the influence of SAD on investors' probability weighting functions. All find evidence consistent with the influence of SAD on markets. Furthermore, in contrast to KM's assertion that "a *previously untested* implication of the SAD model is that the

[6] We thank an anonymous referee for raising the point that the special cross-listing relationship between the Hong Kong index and the FTSE may apply in additional cases.

[7] We retain the following indices among countries with duplicates: the Australian Total Market index, the FTSE New Zealand Index, the Swedish OMX Affärsvärldens Generalinde, the UK Total Market index, and the US equal-weighted total market (NYSE, NASDAQ, Amex) index, including distributions, obtained from CRSP.

Author's personal copy

Table 2
Key coefficient estimates, *p*-values, and magnitudes for Eqs. (1)–(3), based on a non-duplicative set of indices.

Country, index, date range, latitude	Eq. (1) Fall	Eq. (1) Fall/winter	Eq. (1) Joint effect *p*-value	Eq. (2) Fall	Eq. (2) SAD	Eq. (2) Joint effect *p*-value	Eq. (2) Fall magnitude SAD effect[a]	Eq. (2) Winter magnitude SAD effect[a]	Eq. (3) Fall	Eq. (3) FallSAD	Eq. (3) WinSAD	Eq. (3) Joint effect *p*-value
Panel A: Exchanges located above 40° latitude												
Finland, HEX General Index, 1987–2008, 64°N	.004	–.009	.995	.013	–.004	.954	.000	–.017	.105	–.019	.012	.646
Iceland, OMX Iceland All Share, 1993–2008, 64°N	**–.195**	.059	**.088**	**–.205**	.015	.145	–.163	.055	–.217	.017	**.013**	.130
Norway, Total Market, 1980–2008, 62°N	**–.111**	–.012	**.046**	**–.180**	**.022**	**.012**	–.122	.076	**–.153**	.017	**.028**	**.017**
Sweden, OMX Affärsvärldens Generalinde, 1980–2008, 59°N	**–.098**	.087	.134	**–.103**	**.023**	.161	–.051	.068	–.036	.008	**.038**	**.042**
Denmark, Copenhagen KFX, 1989–2008, 56°N	–.040	.018	.750	–.077	.022	.344	–.033	.058	.033	–.005	**.050**	**.027**
Ireland, Total Market, 1973–2008, 53°N	**–.162**	**.177**	**.000**	**–.124**	**.039**	**.001**	–.057	.089	–.035	.012	**.050**	**.000**
Netherlands, AEX Index, 1973–2008, 52°N	–.046	–.001	.416	**–.080**	.018	.257	–.050	.039	–.073	.016	.019	.375
UK, Total Market, 1973–2008, 51°N	**–.084**	**.086**	**.042**	**–.075**	**.027**	**.028**	–.031	.058	–.027	.011	**.034**	**.026**
Belgium, Banque Bruxelles Lambert 30, 1973–2008, 50°N	**–.063**	.012	.121	**–.091**	**.022**	**.045**	–.058	.044	–.064	.012	**.029**	**.038**
Germany, DAX100, 1988–1998, 50°N	**–.147**	.102	.123	**–.213**	**.081**	**.019**	–.085	.166	–.146	.059	**.102**	**.010**
Austria, ATX50, 1973–2008, 47°N	**–.083**	.038	**.030**	**–.104**	**.031**	**.008**	–.062	.057	–.053	.012	**.050**	**.001**
Switzerland, Total Market, 1973–2008, 47°N	–.010	–.015	.756	–.042	.017	.463	–.019	.031	–.005	.003	**.030**	.199
France, Total Market, 1973–2008, 46°N	**–.084**	.034	**.100**	**–.093**	.021	**.083**	–.065	.037	–.026	–.005	**.042**	**.015**
Canada, TSX300, 1973–2008, 43°N	–.054	.038	.373	**–.115**	**.065**	**.001**	–.037	.102	–.098	**.058**	**.072**	**.001**
Italy, Total Market, 1973–2008, 42°N	**–.096**	.049	.102	**–.107**	.034	**.075**	–.068	.051	–.052	.009	**.057**	**.034**
US, EW CRSP Total Market, 1948–2008, 41°N	**–.051**	.032	**.029**	**–.070**	**.033**	**.000**	–.034	.048	**–.052**	**.024**	**.040**	**.000**
Average value	–.083	.043	.257	–.104	.029	.162	–.058	.060	–.056	.042	.014	.098
Proportion of p-values ⩽ .05/.10	.50/.69	.13/.13	.31/.44	.63/.81	.38/.56	.50/.63			.00/.13.	.06/.13	.69/.88	.75/.75
Panel B: Exchanges located in the tropical and sub-tropical latitudes (40° and below)												
Spain, Madrid SE General, 1974–2008, 40°N	–.060	.029	.393	–.063	.020	.397	–.042	.028	–.030	.003	.034	.356
China, Total Market, 1991–2008, 40°N	–.113	–.001	.638	–.029	–.093	.344	–.135	–.126	.095	–.151	–.041	.434
Greece, Total Market, 1988–2008, 39°N	**–.138**	–.017	**.062**	**–.220**	.063	**.024**	–.157	.085	**–.275**	.092	.040	**.059**
Turkey, Total Market, 1988–2008, 39°N	.153	–.006	.268	.087	.054	.216	.141	.073	.126	.033	.070	.332
Korea, Korea South Composite (KOSPI), 1975–2008, 37°N	.065	–.063	.591	.043	–.015	.806	.028	–.019	–.013	.013	–.049	.751
Japan, Nikkei 225, 1953–2008, 36°N	–.013	–.047	.395	–.049	.013	.464	–.038	.015	–.046	.011	.014	.662
Jordan, Amman SE Financial Market, 1988–2008, 31°N	.111	–.087	.262	.077	–.030	.434	.055	–.028	.030	.005	–.060	.567
Mexico, FTSE Mexico Index, 1987–2008, 23°N	–.133	.051	.312	**–.171**	.132	.122	–.109	.084	–.132	.088	.168	.189
Taiwan, Taiwan Weighted, 1973–2001, 23°N	**–.166**	.034	**.068**	**–.221**	.132	**.038**	–.160	.082	–.176	.085	.177	**.067**
Hong Kong, Total Market, 1973–2008, 22°N	**–.144**	.118	.121	**–.155**	**.201**	**.015**	–.066	.119	–.019	.050	**.283**	**.003**
India, National Index (100), 1989–2008, 20°N	–.097	.033	.520	**–.164**	**.216**	**.079**	–.079	.114	–.149	.197	**.228**	.159
Thailand, Stock Exchange of Thailand Index, 1975–2008, 15°N	–.031	.012	.861	–.032	.033	.843	–.024	.012	.042	–.104	.167	.414
Philippines, P.S.E. Composite Index, 1986–2008, 13°N	–.021	.007	.965	–.070	.194	.497	–.032	.056	.019	–.008	**.443**	.226
Sri Lanka, All Share, 1985–2008, 7°N	–.013	.023	.920	–.017	.196	.682	–.009	.018	–.004	.116	.252	.858
Malaysia, Composite, 1980–2008, 2°N	–.066	**.100**	.145	.020	.232	.859	.002	–.015	.040	.571	.093	.928
Singapore, Total Market, 1973–2008, 1°N	**–.079**	.054	.220	–.058	–.307	.404	–.027	.029	–.069	–.421	–.299	.609
Indonesia, Jakarta Composite Index, 1983–2008, 5°S	**.109**	**–.126**	**.053**	.078	**–.247**	.129	.017	–.070	–.030	.087	**–.289**	.169
South Africa, Total Market, 1973–2008, 26°S	–.013	.014	.946	–.049	**.091**	.113	.023	.091	–.097	**.128**	**.083**	.178
Australia, Total Market, 1973–2008, 34°S	.003	.031	.617	.016	.006	.712	.022	.008	**.098**	–.041	.032	.189
New Zealand, FTSE New Zealand Index, 1996–2008, 37°S	–.033	–.059	.257	–.074	.008	.332	–.066	.012	–.063	.002	.009	.529
Average value	–.034	.005	.431	–.053	.045	.376	–.033	.028	–.033	.068	.038	.384
Proportion of p-values ⩽ .05/.10	.05/.25	.10/.10	.00/.15	.25/.25	.10/.20	.15/.20			.05/.10	.00/.05	.10/.25	.05/.15

Notes: We produce regression results with one-at-a-time OLS and MacKinnon–White (1985) heteroskedastity-consistent HC3 standard errors. Bolded coefficient values and *p*-values are significant at the 10% level or better. Data are from Patrick Kelly and Felix Meschke, unless otherwise noted. We exclude duplicate exchanges KM consider for a given country, which applies to Sweden, the US, the UK, Australia, and New Zealand. In each case, we employ the exchange with the longest time series: for Australia we employ the Total Market index, for New Zealand we use the FTSE New Zealand Index, for Sweden we use the OMX Affärsvärldens Generalinde, and for the UK we utilize the Total Market index. For the US we employ the equal-weighted total market (NYSE, NASDAQ, Amex) index, including distributions, obtained from CRSP. We employ this series for the US instead of one of the US series KM employ for two reasons. First, it is a total market index, encompassing all of the securities captured in the individual series KM employ. Second, it is an equal-weighted index, which necessarily places more weight on smaller, riskier stocks (relative to a value-weighted index) that are more likely to exhibit seasonally varying returns due to time-varying investor risk aversion.

[a] We use October for calculating the magnitude of the SAD effect in the fall, and we use February for calculating the magnitude of the SAD effect in the winter.

Author's personal copy

seasonal patterns in stock index returns are more pronounced in countries where SAD is more prevalent" (p. 1317; emphasis added), Dowling and Lucey (2008) enlarge the KKL2003 study to 37 countries and find strong SAD effects and evidence that the influence of SAD on markets increases with latitude.

From the above set of papers, the one most closely related to KM is by Dowling and Lucey (2008). How do Dowling and Lucey reach such different conclusions than KM? First, Dowling and Lucey group countries into those close to the equator versus those that are further away from the equator. Second, Dowling and Lucey employ maximum likelihood estimation and a GARCH specification. Both of these points may contribute to minor differences. However, the most significant difference is the assessment of results. Dowling and Lucey consider statistical significance, of course, but they adopt a technique Hirshleifer and Shumway (2003) use to quantify the significance of cloudy days. Specifically, Dowling and Lucey consider the proportion of coefficients that have the expected sign. Like KM, Dowling and Lucey document a striking preponderance of correctly signed coefficient estimates. Dowling and Lucey then also evaluate whether the significance of the SAD effect is greater for countries more distant from the equator. They find that it is. Further, in addition to studying total market indices, they consider relatively riskier small-capitalization indices. If the SAD hypothesis is correct, these riskier indices should display a more pronounced SAD effect. Dowling and Lucey find they do. Much of this same evidence appears in KM's own tables; unfortunately they seem to be unaware of it.

Other papers that study the potential influence of SAD on financial markets include several of our own. Kamstra, Kramer, and Levi (2011; cited by KM) document seasonality in US Treasury returns that is consistent with the time-varying risk aversion hypothesis. In that paper we test a large number of alternative explanations to the SAD hypothesis, including cross-market hedging, time-varying sentiment (employing both the Baker and Wurgler (2007), sentiment measure and the Michigan consumer sentiment index), macroeconomic cyclicality, and others, and find that none is capable of eliminating the statistical evidence in support of the SAD hypothesis. Kamstra et al. (2011c) investigate the flow of funds between safe and risky categories of US, Canadian, and Australian mutual funds and find both raw flows and flows after controlling for other factors (advertising, trend-chasing, capital gains impacts, etc.) display net flows out of risky funds and into safe funds in fall, with the patterns reversing in the winter, consistent with the SAD hypothesis. Garrett et al. (2005) explore time-varying risk aversion in an equilibrium asset pricing model which allows the price of risk to vary through the seasons, finding evidence consistent with the SAD hypothesis. DeGennaro et al. (2008) study bid-ask spreads, and find, among other things, that market makers quote wider spreads during periods of increased risk aversion. Kamstra et al. (2011b) explore a theoretical asset pricing model to determine the degree of seasonality in model parameters necessary to generate observed seasonal patterns in risky and risk-free security returns. They conclude that the necessary values of risk aversion and the intertemporal elasticity of substitution across the seasons are within standard acceptable norms.

Overall, KM neglect to mention supportive evidence from markets for various security types, based on a range of metrics (including volatility, spreads, quantities of fund flows, etc.), using data from a range of locations around the world, with much of the work conducted by researchers independent of KKL2003. A thorough discussion would have included these findings.

3.2. Discussion of the literature relevant to Jacobsen and Marquering

Two papers to which KM note they "owe the greatest intellectual debt" (p. 1309), Jacobsen and Marquering (2008, 2009), require particular attention. Jacobsen and Marquering (2008; hereafter JM2008) question the evidence supporting the SAD hypothesis and Jacobsen and Marquering (2009; hereafter JM2009) respond to Kamstra, Kramer, and Levi's (2009) comment regarding JM2008. A fair-handed treatment would require at least a passing reference to the Kamstra et al. (2009) comment, which did, after all, prompt the writing of the reply, JM2009. Because KM neglect to cite the comment, we offer a brief overview. In Kamstra et al. (2009), we state that we are unable to replicate JM2008's findings, even after corresponding with Professor Jacobsen. Note that in their response, JM2009 concede that their results could not be replicated with the data and methods they describe in JM2008, and they admit to data manipulation mistakes which they claim account for our inability to replicate their findings. After following the revised instructions in JM2009, we still find a strong SAD effect in the data and cannot replicate the JM2009 findings, in addition to not being able to replicate the JM2008 results. There is nothing in JM2008 and JM2009 that causes us to reconsider the SAD hypothesis, in particular in light of Professors Jacobsen and Marquering's own admission of data errors.

3.3. Misrepresentation of Goetzmann and Zhu's findings

Careful framing of the results in the literature is important in any study, but it is particularly crucial when authors challenge established work. KM distort Goetzmann and Zhu's (2005) findings in writing (on p. 1309):

> Goetzmann and Zhu (2005) examines investor trading activity in five major US cities from January 1991 to November 1996 and concludes that their 'results offer little support for the argument that investor behaviour is influenced by seasonality in the length of daytime hours'.

An isolated reading of this extracted sentence leaves the reader with the impression that Goetzmann and Zhu produce results showing that the length of day does not impact investor behavior. The full sentence from Goetzmann and Zhu is:

> Given the obvious seasonality in overall sky cover (see Fig. 1), our results offer little support for the argument that investor behavior is influenced by seasonality in the length of daytime hours. (p. 566)

That is, Goetzmann and Zhu speculate that seasonality in cloud cover (which they show is not likely related to investor behavior) is similar enough to the length of day that their results carry over to length of day. While we find the opinions of Goetzmann and Zhu interesting, we note that they provide no results relating the length of day to investor behavior, nor do they claim to.

Evaluation of the relationship between sky cover and SAD should be based on research that investigates the topic. Keller et al. (2005) investigate a wide range of environmental factors and find sky cover is unrelated to SAD. They also find the single most important determinant of SAD onset and recovery is the length of day. This has been shown in several studies, some of which are summarized by Young (2001):

> Existing research suggests that weather variables are not a factor in the basic aetiology of SAD. Weather may affect how individuals with SAD may feel, but these effects tend to be idiosyncratic to the individual. The experiences of those with SAD may be similar to the weather complaints of non-SAD individuals, may exacerbate SAD symptoms, or may be exacerbated by SAD symptoms. (p. 172)

Clearly, sky cover and SAD are not systematically related to one another.

Author's personal copy

3.4. Misrepresentation of Grinblatt and Keloharju's findings

KM write:

> With day-by-day records of investors' psychological states, their trading and their portfolio holdings, we could directly examine how seasonal changes in depression translate into seasonal variations in portfolio holdings. A good example for this line of inquiry is Grinblatt and Keloharju (2008), which examines how sensation seeking and overconfidence affect the tendency of investors to trade stocks. They do so by matching results from a psychological test given to all Finnish males by the Finnish Armed Forces with portfolios and trading records (1995–2002) of all household investors domiciled in Finland. (p. 1310)

This summary could easily lead one to conclude Grinblatt and Keloharju (2008) do several things they do not do (and do not purport to do). First, it might appear that Grinblatt and Keloharju have records on investors' day-by-day psychological states. They do not. Second, it might appear that Grinblatt and Keloharju study depression. They do not. Third, it might appear that Grinblatt and Keloharju find a relationship between an individual-based measure of risk taking (sensation seeking) and investor trading. They do not. Grinblatt and Keloharju develop a sensation-seeking proxy using the number of speeding ticket convictions an individual accumulates over time.

While on one hand KM critique us for failing to employ data (including daily investor mood) to properly test our hypotheses, they simultaneously imply that such data have been used in other studies, which is not the case. We agree that it would be wonderful if such data were available, however the data are not currently at our disposal, if they exist at all. Only data over very short recent time spans, unsuitable for analysis of an annual seasonal effect such as that implied by SAD, are currently being exploited. See, for instance, Bollen et al. (2011) and Gilbert and Karahalios (2010), who employ daily mood proxies extracted from online social media (e.g., Twitter and LiveJournal). In the future, researchers may try to employ these daily mood proxies to further explore behavioral effects. We have to caution, however, that daily "mood" does not necessarily capture clinical depression due to SAD.

4. The psychology and medical literatures

4.1. The timing of seasonal depression

We agree with KM that KKL2003 need not have cited Palinkas et al. (1996) and Palinkas and Houseal (2000). Those citations are superfluous to the economic argument; the hypothesis regarding the timing of depression due to SAD and its influence on stock returns does not depend on what happens to people who spend winter in Antarctica.[8] The best way to model the timing of SAD symptoms is based on studies that document the timing of SAD symptoms. We were unaware when we wrote KKL2003 that such statistics were available, thus the KKL2003 SAD measure is based on the instrument most closely tied by clinical research to SAD symptoms, length of night. Kamstra et al. (2011a) have since developed an improved measure of the timing of SAD symptoms based on Lam's (1998) clinical study of SAD patients.[9]

In Section 6.3 their study, KM employ Kamstra, Kramer, and Levi's (2011) improved measure of the timing of SAD based on clinically observed SAD symptoms, so clearly they were aware of its existence at the time they wrote their paper. Nonetheless, KM obfuscate what is well-established about the timing of SAD symptoms with reference to a study by Kasper et al. (1989b), which is based on telephone interviews with 416 randomly selected residents of Montgomery County, Maryland.[10] The Kasper et al. (1989b) survey was conducted by dialing random phone numbers in the month of November and asking subjects to recall their normal experiences. The interview question of greatest interest to KM asked participants to recall the month of the year during which they normally "feel worst". KM extrapolate Kasper et al.'s (1989b) survey findings about the annual timing of participants "feeling worst" in an attempt to infer the timing of *depression* symptoms among individuals who suffer from SAD. As Lam and Levitt (1999, pp. 37–38) note in reference to the questionnaire Kasper et al. (1989b) employ, the timing of feeling worst need not correspond to when individuals are depressed. Unfortunately, Kasper et al. did not measure the time of year at which study participants were *depressed*.

Shortly after responding to the question about the timing of feeling worst, participants in the Kasper et al. (1989b) study were asked which environmental conditions contributed to their own personal notion of feeling worst. The various sources of participants' mood deterioration appear in Kasper et al. (1989b)'s Table 3. In the total sample of 416 individuals, the most often stated cause of "feeling worst" is humid days (77% of subjects), a weather condition that to the best of our knowledge has never been associated with SAD by any study, perhaps because it is not typically encountered during the fall or winter seasons. (Even in the sub-sample of 180 participants who appear to have "winter-type" mood variation, 88% report humid days as a contributing factor, again, the most popular reason chosen in that group.) Back to the full sample, participants mention hot weather more frequently than cold weather as a cause of "feeling worst" (48% versus 43%), and they mention high pollen count more frequently than short days (51% versus 47%). The surprising list of causes for feeling worst, and the fact that the primary accepted cause of seasonal depression (short days) ranks so low on participants' lists of causes, should be a red flag for anyone attempting to extrapolate the Kasper et al. (1989b) findings to the timing of seasonal depression in absence of diagnostic information about depression among individuals.

KM make much of the fact that Kasper et al. (1989b) state that the peak month during which their survey participants report "feel worst" is February. Even if this coincides with the point in time when individuals experience depression at its worst (and since the Kasper et al. survey does not test for depression at any point in time, one has no way of knowing whether it does), this would not necessarily be at odds with the SAD hypothesis. We know from research on individuals who suffer from SAD (e.g., Lam, 1998; Young et al., 1997) that people experience onset

[8] Instead, the economic argument rests on the simple premise that seasonal depression causes seasonally varying risk aversion. During autumn, some investors experience a reduced appetite for risky securities. Distaste for risky securities drives stock returns lower during autumn: associated higher expected returns are required to persuade investors to hold risky securities being shunned those who experience seasonally varying risk aversion. Investor appetite for risk remains reduced until it rebounds sometime in the new year.

[9] There exist other clinical studies that document the timing of SAD symptoms, including Young et al. (1997). We base our measure on data from the Lam (1998) study because, unlike data from other clinical studies, the Lam dataset details the timing of both symptom onset and symptom recovery. Our measure is qualitatively identical if we combine data from the Lam and Young et al. studies.

[10] KM list in their references a different "Kasper et al. (1989)" paper than the one they describe in their paper (and we discuss in this paper). They describe the Kasper et al. (1989b) study of residents of Montgomery County (and include statistics from that study in their Appendix A) but their references list only the unrelated Kasper et al. (1989a) study.

Author's personal copy

and recovery from SAD at different times.[11] That is, there is a *flow* (or rate) of individuals experiencing their onset of seasonal depression, and as more and more people succumb, the *stock* (or level) of people suffering increases. The peak in the *flow* of onset, according to medical research, is around October. We speculate that the timing of the peak flow may correspond with the peak timing of the negative impact on markets, since an investor who experiences a surge in depression, and hence risk aversion, in the early fall will at some point have sufficiently rebalanced his/her portfolio and hence will no longer be the marginal investor (even if, say, the intensity of that investor's depression continues to strengthen and peaks well after having rebalanced toward a safer portfolio allocation). We have not tested the conjecture that the timing of peak flow of onset coincides with the timing of peak impact on markets, though it appears broadly consistent with the seasonal pattern of risky returns. As people continue to succumb to seasonal depression during the fall, they may continue to shift their portfolio allocations, but the flow of people succumbing will be lower in late fall than in early fall. The peak in the *level* (number) of people suffering from SAD, according to medical research, is around winter solstice. With little variation in daylight around December 21, clinicians find that few people are either succumbing to or recovering from SAD at that point in the year.[12] After winter solstice, individuals recover on different schedules. The peak in the *flow* of people *recovering* from SAD, again according to medical research, is around March. Clinical evidence indicates that a small fraction of people begin recovering as early as January, but the peak point of recovery coincides approximately with the time of year when hours of daylight change most rapidly – the positive point of inflection in the annual daylight cycle – around March 21. People may not resume their previous risky holdings until they are fully recovered or they may resume them gradually as they recover; again, we do not have trader-based evidence to pinpoint the timing definitively. The differences between the *stock* and *flow* of people experiencing seasonal depression is precisely what led to the development of the Kamstra et al. (2011a) SAD measure based on clinical data. This SAD onset/recovery measure reflects the *flow* of people succumbing to and recovering from SAD through the seasons.[13]

Overall, KM mischaracterize established findings on the timing of depression due to SAD.

4.2. Depression versus induced sadness

In discussing the evidence from the psychology literature on whether depressed people are more risk averse in their financial decisions, KM misrepresent the literature. They write "the evidence here is quite mixed" (p. 1310). In fact the evidence is very clear. Studies of depressed people are virtually uniform in finding that depression is significantly associated with greater risk aversion. See, for instance, Zuckerman (1979, 1984, 1994), Pietromonaco and Rook (1987), Carton et al. (1992), Carton et al. (1995), and Smoski et al. (2008).

We suspect KM's mistaken impression regarding the strong link between depression and increased risk aversion is based on their confusion of studies that examine the serious medical state of depression with studies that take *non-depressed* subjects and induce a temporary state of mild sadness, for instance by having them view a brief film clip. KM cite studies from these two streams of the literature as though they were directly comparable.

Unlike studies of depressed people, studies that induce temporary sadness in healthy (non-depressed) people using movie clips have mixed findings. Many, such as Raghunathan and Pham (1999), find people induced into a sad mood tend to select risky over safe choices. Others, such as Yuen and Lee (2003), find induced sadness leads to safe choices over risky alternatives. Still others, such as Leith and Baumeister (1996), find induced sadness does not impact risk taking. Clearly there is ambiguity in the induced-sadness literature, but it is inappropriate to hyperextend these results to tarnish the established link between *depression* and increased risk aversion.

We are always clear in studying the connection between SAD and financial markets that we are interested in mood states that can persist for months at a time: clinical depression and related sub-clinical depression of a seasonal nature, not fleeting sadness that may arise from watching an excerpt of a tear-jerker film. For our purposes, perhaps the most important difference between induced sadness and clinical depression is the fact that sadness induced by researchers in an experiment is transient and short-lived by construction, whereas depression is a relatively stable mood state. This difference may help explain why experimentally induced sadness and depression have such different implications for risk aversion. The studies of depressed people find those individuals exhibit behavior consistent with what is known as the affect infusion model (AIM; see Forgas, 1995). Under AIM, people in a negative mood focus on negative cues disproportionately, which alters their subjective probability assessments and leads to less risk-taking behavior. In contrast, the studies of people in a temporarily induced mood state mostly find that such people exhibit behavior that is perhaps more consistent with an analogue to the mood maintenance hypothesis (MMH; see, for instance, Isen et al., 1988). MMH focuses on the behavior of people in a positive mood, but if one is willing to accept that the influence of temporary negative mood on risk taking is opposite to the influence of temporary positive mood on risk taking, then MMH implies that people in a negative mood should increase risk-taking in hopes of achieving a positive outcome to jolt them out of their temporary funk. Such behavior makes sense only if one believes their current negative mood state is transient and easily overcome. That is, MMH is not a plausible description of the behavior of clinically depressed individuals (nor did the authors who developed MMH intend it to be). Depressed individuals generally know from their difficult and often failed experiences with standard depression treatments such as psychotherapy, medication, and light exposure, that their depression is not easily cured. Hence it is understandable that they would not pin their hopes on the possibility that a positive outcome from a risky gamble would chase away their woes.

4.3. Mood and risk taking

We now turn to studies KM highlight in the context of questioning a connection between mood and willingness to take risks (especially financial risks). First, KM report that Hockey et al. (2000) find "no such association between mood and risk-taking behavior" (p. 1310). This directly contradicts Hockey et al.'s own statement: "The findings from the three studies show that the degree of risk taken in everyday decision making may be affected by variations in state mood" (p. 849). This important misrepresentation aside, Hockey et al.'s paper bears little relation to the key question pertaining to the association between *depression* and risk-taking.

[11] The timing of peak seasonal depression for any given individual could in principle fall at any point between SAD onset and recovery, and as we argue below, an individual's peak timing of depression need not have any bearing on the timing of that individual's seasonal portfolio reallocation or aggregate stock returns. Indeed, it would require formidable discipline for a SAD-sufferer to wait until they feel the very worst before reallocating assets.

[12] In spite of this, KM perplexingly claim that we assume "depressive effects subside and returns increase *immediately* following December 21" (p. 1311, emphasis added). We do not assume anything of the sort.

[13] A variant of that measure, "SAD incidence," reflects the *level* of people suffering at a given point in time, and is used in studying *level* quantities of economic variables (such as bid-ask spreads) in contrast to rate of change (*flow*) variables such as returns. See DeGennaro et al. (2008) for details on the SAD incidence variable.

Author's personal copy

Hockey et al.'s Study 1 and Study 2 did not include any standard depression questionnaires or depression diagnostic criteria; instead they asked participants how they felt over the past few hours. Their Study 3 examined induced fatigue and also did not employ standard measures of depression. Hence Hockey et al.'s study offers no insight into the connection between *depression* and risk taking.

Second, KM describe Clark et al.'s (2001) findings as failing to find a link between mood and risk-taking behavior (on p. 1310). Clark et al. induce "depression" using music. While such a procedure may induce temporary feelings of mild sadness, we know of no evidence that music could cause someone to become depressed. Clark et al.'s lack of meaningful findings with respect to the link between depression and risk taking may therefore be a consequence of experimental design.

Third, KM cite Morse's (1998) study that fails to find a relation between sensation seeking (a measure of risk-taking tendency) and actual investment choices. A serious errors-in-variables problem plagues Morse's study, related to the arbitrarily-assigned risk levels of the investment categories. For example, stamp collections are deemed by Morse as riskier than real estate but safer than financial derivatives, with no rationale for the ranking. Mutual funds are deemed as riskier than corporate bonds and safer than common stock, with no recognition of the fact that there are many varieties of mutual funds, each with its own risk characteristics. For instance some mutual funds invest exclusively in securities that are far safer than corporate bonds, some focus on aggressively risky securities that surely exceed the risk of a diversified portfolio of common stock, and so on. This feature of the Morse study lowers the power of statistical tests.

Fourth, KM inaccurately describe papers on risk-taking among individuals who are not depressed and who have not had their moods manipulated (i.e. studies that focus on healthy individuals who are in their natural mood state): Horvath and Zuckerman (1993), Eisenberg et al. (1998), and once again Grinblatt and Keloharju (2008). We describe each problem in turn.

In describing the findings of Horvath and Zuckerman (1993), KM write (on p. 1310):

> [Horvath and Zuckerman examine] the cross-sectional relationship between sensation seeking and four risk factors among college students. They find that sensation seeking 'was negatively and significantly correlated with own risk appraisal for all of the risk areas except financial risk' (p. 45) and argue: 'Financial risk taking may be a special type not as highly related to the more general sensation seeking trait' (p. 49).

KM create the misimpression that Horvath and Zuckerman fail to find the expected relationship. In fact, inspection of Horvath and Zuckerman's paper reveals that the quotes KM highlight refer to a special case in their Table 2, where they report a coefficient on financial risk, *though the expected negative sign*, is not significant at the 1% level or better. (It is unclear whether the expected negative estimate was significant at other conventional levels of significance because Horvath and Zuckerman reported significance using only the 1% level, but the important point is that the sign on financial risk was negative, as it was for other types of risk, and as Horvath and Zuckerman expected.) Further, this result pertains to an individual's own risk appraisal, *not behavior*. Overlooked by KM, when Horvath and Zuckerman examine behavior, they find that people who score high on sensation seeking tests are *significantly more likely to engage in risky financial activities*. Since risk-taking *behavior* is rather relevant to the debate at hand, we are perplexed that KM would choose to neglect reporting this important finding from a paper they cite. Describing this result, Horvath and Zuckerman write, "Positive correlations were found between sensation seeking and the [individuals'] reports of engaging in risky activities of all four types" (p. 46). Lest KM rest their entire case on Horvath and Zuckerman's finding that a properly signed coefficient failed to exhibit significance at the 1% level, we note that several other studies find statistically significant evidence that healthy people (in their natural mood state) who score high on sensation-seeking tests make riskier financial decisions. See, for instance, Harlow and Brown (1990), Wong and Carducci (1991), and Tokunaga (1993), among others.

Fifth, KM quote (on p. 1310) Eisenberg et al. (1998)'s statement "more depressive symptoms and more anxiety went with less tendency to act." Again, this is a selective quote, and KM should have provided full contextual information to avoid creating a false impression. The quote refers to one of two experiments Eisbenberg et al. conducted on healthy (*non-depressed*) individuals. In describing their full set of results, Eisenberg et al. explain that depressive symptoms are correlated with general passivity in one of their experiments, but not the other. Lest researchers be tempted to extrapolate their findings beyond the healthy samples they study, Eisenberg et al. caution readers: "perhaps further studies are warranted using people with greater degrees of depressive symptoms." We note as well that KM elect not to mention a very relevant finding from the Eisenberg et al. study. In both experiments, Eisenberg et al. find depression is correlated with reduced risk taking (significantly so in one of the experiments, at better than the 0.1% level).

Sixth, regarding their statement that "depression may not spur sufferers to act" (p. 1310), KM claim Grinblatt and Keloharju (2008) find "investors with lower sensation-seeking propensity are less likely to trade" (p. 1310). What Grinblatt and Keloharju actually say is that their proxy for sensation seeking (number of speeding tickets) "is less related to the decision of whether to trade at all and more related to the decision of how much to trade" (p. 553). That is, they find number of speeding ticket convictions is related to an individual's number of trades and trading turnover, neither of which is relevant to the willingness of depressed individuals to act or not to act, contrary to KM's claims. A study that does bear on depressed individuals' decision to act is that by Eisenberg et al. (1998). They conduct experiments in which individuals differing in degree of depression make a series of choices between pairs of risky and safe alternatives, including some of a financial nature. By setting choices such that in some cases the risky option is the default (not requiring action) and in other cases the safe option is the default, the researchers distinguish risk aversion from passivity, finding depressive symptoms correlate with risk aversion.

Seventh, KM report "no time-series evidence... relates *changes* in depression to changes in risk aversion" (p. 1310). While KM could not have known when they wrote their article, Kramer and Weber (forthcoming) provide time-series and cross-sectional evidence relating changes in depression to changes in risk aversion. Kramer and Weber study hundreds of individuals, some of whom suffer from SAD and some of whom do not, at three points in time over a year. They find the SAD-sufferers exhibit changes in their level of depression over time and changes in their financial risk aversion time. Specifically, those who suffer from SAD are significantly more depressed and significantly more financially risk averse in fall/winter than in summer.

In sum, KM's selective discussion of the literature obfuscates the accepted fact that depression is associated with increased risk aversion. It also muddles the fact that standard risk-taking measures are significantly associated with financial risk taking in healthy (non-depressed) individuals.

5. Empirical analysis revisited

KM utilize a very large panel/time-series dataset which exhibits cross-sectional covariance, heteroskedasticity, and autocorrelation. Yet they analyze their data series one-at-a-time with OLS, and they employ MacKinnon and White (1985) heteroskedasticity-consistent errors. We prefer in this context to exploit the power and efficiency

Author's personal copy

of system-of-equations estimation and standard errors robust to both autocorrelation and heteroskedasticity. Hirshleifer and Shumway (2003), for instance, estimate a city-specific fixed-effects panel model, incorporating the contemporaneous correlation of residuals, as well as being robust to autocorrelation and heteroskedasticity. This greatly enhances the power of their tests. Indeed, Hirshleifer and Shumway explicitly state that they use a panel approach to "[increase their] power to detect an effect" (p. 1014). We use panel estimation in recent work, including the analysis of SAD effects in Treasury returns (Kamstra et al., 2011a) and the consideration of the influence of SAD on US, Canadian, and Australian mutual fund flows (Kamstra et al., 2011c). By not using panel/times-series econometric techniques, KM necessarily compromise power. This problem is most clearly evident by reference to the fact that KM's coefficient estimates throughout their tables are the same order of magnitude as those KKL2003 report, reproduced in Panel A of KM's Table 2. It is KM's inefficient estimation, combined with their decision not to report rejections at the 10% (two-sided) level of significance, that enable KM to conclude the SAD effect is "insignificant" in their extended set of country-indices.

The concern KM raise regarding a turn-of-the-year effect, although specious, is nonetheless handled directly by use of an improved SAD measure developed by Kamstra et al. (2011a). This variable, SAD onset/recovery, is based directly on clinical diagnoses of SAD. The value of the SAD onset/recovery variable, which reflects the change in the number of SAD-affected individuals, is nearly zero during December and January and thus it cannot influence returns during those months. For this reason, and since the SAD onset/recovery variable closely captures what we intend to measure (the timing of symptoms experienced by those who suffer from SAD), we restrict our remaining analysis of the statistical significance of the cross-country SAD effect in stock returns using the SAD onset/recovery variable in a panel/time-series context.

First we form time series of weekly returns for each index KM consider (using KM's data unless stated otherwise.) We employ weekly data to avoid missing daily return observations in the cross-section across our indices, since different countries have different holidays. We fill in missing weather data with index-by-index average week-of-the-year values over the entire available weather dataset for a given index. The weekly return models we explore all include the SAD onset/recovery variable and a tax-loss-selling dummy variable. We discuss results using both seasonally adjusted weather data and KM's seasonally unadjusted weather data. We employ OLS with MacKinnon and White (1985) standard errors, seemingly unrelated regression (SUR) panel estimation with MacKinnon and White standard errors, and GMM estimation with HAC standard errors.[14,15] We present two main sets of results, reporting results based on five different estimation techniques within each set.

For our first set of results, we consider individual exchanges that have the longest span of available (non-missing) return data from both hemispheres and over a range of latitudes. These include indices from Australia, Ireland, the UK, Canada, Italy, France, Switzerland, Belgium, Austria, Netherlands, Japan, and the US. The sample period spans July 1973 to December 2008. For our second set of results, we consider all countries at once by forming new indices as the average returns of exchanges in groupings by latitude.[16] We form five sets of indices, with each index consisting of the equal-weighted average of the non-missing returns for the exchanges in that index at a point in time (thus, at the very beginning and end of the data span covered, there are fewer exchanges included in the average than in the middle of the range of the data; this permits a long span for these indices, from December 1972 to December 2008). The first index is based on exchanges in the southern-hemisphere tropics and sub-tropics (from the equator to 40°S),[17] the second index is based on exchanges in the northern-hemisphere tropics and sub-tropics (from the equator to 40°N),[18] the third index is the US only, the fourth index is based on exchanges in the latitude range above 40°N and below 50°N (excluding the US),[19] and the fifth index is based on exchanges at or above 50°N.[20] In all cases we estimate models of this form:

$$\begin{aligned} y_{i,t} = {} & \alpha_i + \rho_{1,i} y_{i,t-1} + \rho_{2,i} y_{i,t-2} + \beta_{i,\mathrm{Tax}} \mathrm{Tax}_{i,t} \\ & + \beta_{i,\mathrm{OnsetRecovery}} \mathrm{OnsetRecovery}_{i,t} + \beta_{i,\mathrm{Temp}} \mathrm{Temp}_{i,t} \\ & + \beta_{i,\mathrm{Cloud}} \mathrm{Cloud}_{i,t} + \beta_{i,\mathrm{Rain}} \mathrm{Rain}_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{4}$$

where $y_{i,t}$ is the return to index i for week t; $\mathrm{Tax}_{i,t}$ is the tax year dummy equal to 1 if week t immediately precedes/follows the fiscal year end; $\mathrm{OnsetRecovery}_{i,t}$ is Kamstra, Kramer, and Levi's (2011) SAD onset/recovery variable[21] averaged for week t; and $\mathrm{Temp}_{i,t}$, $\mathrm{Cloud}_{i,t}$, and $\mathrm{Rain}_{i,t}$ are KM's daily weather variables averaged for week t (or, in some cases, KM's weather variables deseasonalized, following Hirshleifer and Shumway (2003), as we describe below).

For each set of results, we estimate Eq. (4) with five variations labeled Models 1–5. Model 1 uses single-equation OLS which does not account for cross-sectional correlations in market returns (this is the technique KM adopt), Models 2 and 3 make use of SUR (an OLS-based panel/time-series estimation technique which *does* account for cross-sectional correlation). Models 4 and 5 use system-of-equations GMM estimation, which also accounts for cross-sectional correlation. For the three estimations using OLS techniques (single

[14] To calculate the HAC standard errors we follow Newey and West (1994) and use the Bartlett kernel and an automatic bandwidth parameter (autocovariance lags) equal to the integer value of $4(T/100)^{2/9}$. Ferson and Foerster (1994) study GMM in the context of systems of return portfolios, performing Monte Carlo experiments using systems of equations with as many as 14 assets and as many as 720 observations. Those Monte Carlo experiments show conventional GMM estimation exhibits very little bias, slightly under-rejecting at the 10% and 5% levels (9.5% and 4.5%, respectively) and slightly over-rejecting at the 2% and 1% levels (2.2% and 1.1%, respectively).

[15] The moment conditions include orthogonality between the regressors and the errors. For the weather variables, we form weekly average temperature, rainfall, and cloud cover variables across all indices, by hemisphere, to use as instruments. We do not use the individual weather series for each country as this would lead to a large number of instruments and would tend to produce test statistics which reject the null more often than nominal size would indicate. As poor choice of instruments can bias results, we perform the Hansen (1982) goodness-of-fit test of over-identifying moment restrictions, and we perform robustness checks using bootstrap-corrected probability values.

[16] In cases where KM use multiple series (Sweden, the US, the UK, Australia, and New Zealand), we use one index per country: always a total market index with the longest time series, as noted in Section 2. Again, for the US we use the equal-weighted total market (NYSE, NASDAQ, Amex) index, including distributions, obtained for CRSP, since it includes all of the securities captured in the series KM employ and since it places relatively more weight on the smaller, riskier stocks that are more likely to exhibit seasonally varying returns due to time-varying risk aversion.

[17] The southern-hemisphere tropics and sub-tropics index includes Indonesia, South Africa, Australia (Total Market), and New Zealand (FTSE New Zealand Index). There are no exchanges located above 40°S latitude.

[18] The northern tropics and northern sub-tropics region includes Singapore, Malaysia, Sri Lanka, the Philippines, Thailand, India, Hong Kong, Mexico, Taiwan, Jordan, Japan, Korea, Greece, Turkey, Spain, and China.

[19] The non-US markets between 40°N and 50°N include Italy, Canada, France, Austria, and Switzerland.

[20] The markets at or above 50°N include Belgium, Germany, the UK (Total Market), the Netherlands, Ireland, Denmark, Sweden (OMX Affärsvärldens Generalinde), Norway, Finland, and Iceland.

[21] To construct the SAD onset/recovery variable for the southern hemisphere we simply shift the northern-hemisphere SAD onset/recovery variable by six months. As we discuss below, this simple method may be inadequate for capturing the true timing of SAD onset/recovery in the southern hemisphere, but given current data limitations it is the best available approximation. Further, this measure of onset/recovery is constructed with North American clinical data, and so it may be an imperfect reflection of onset/recovery for other regions in general.

Author's personal copy

equation and SUR), like KKL2003 and KM we adopt a two-lag autocorrelation specification and find this specification absorbs much if not all the autocorrelation in our weekly data, and we use MacKinnon and White (1985) standard errors. We find that including additional lags up to 5 makes little or no difference to our results. For the two estimations that make use of GMM, we do not include the lagged autoregressive terms $y_{i,t-1}$ and $y_{i,t-2}$, though unreported results from models that include lags are similar. Further, the first SUR and GMM models, Models 2 and 4, have identical model specifications as the OLS model, Model 1, while the second SUR and GMM models, Models 3 and 5, impose a restriction that the coefficients on each weather variable are constant across indices; that is $\beta_{i,\text{Temp}}=\beta_{\text{Temp}}$, $\beta_{i,\text{Cloud}}=\beta_{\text{Cloud}}$, and $\beta_{i,\text{Rain}}=\beta_{\text{Rain}}$ for all i.[22]. The second GMM estimation, Model 5, additionally restricts the SAD variable coefficient to be constant across indices.

5.1. Results based on series for which we have long time series

Table 3 contains regression results for exchanges with the longest span of available (non-missing) return data, altogether five columns of results for Australia, Ireland, the UK, Canada, Italy, France, Switzerland, Belgium, Austria, Netherlands, Japan, and the US. (Again, regarding duplicate series from KM's sample, we use the CRSP equal-weighted total market index with dividends for the US and the longest available series for Sweden, New Zealand, Australia, and the UK.) For this set of regressions we employ seasonally adjusted weather data, as do Hirshleifer and Shumway (2003). Results using seasonally unadjusted data are similar.[23] For the sake of brevity, we omit coefficient estimates for the intercept and autoregressive terms.

Panels A through D of Table 3 contain parameter estimates, standard errors, and results of tests of statistical significance for $\beta_{i,\text{OnsetRecovery}}$, $\beta_{i,\text{Temp}}$, $\beta_{i,\text{Cloud}}$, and $\beta_{i,\text{Rain}}$ from Eq. (4). Panel A of Table 3 contains SAD onset/recovery coefficient estimates. The first line of Panel A contains results for the restricted coefficient estimate (only applicable to Model 5), indicated by "Restricted" in the first column and with "SAD" in the second column indicating that the coefficient estimate is for the SAD onset/recovery variable. The remaining lines of Panel A contain unrestricted coefficient estimates for the SAD onset/recovery variable, for each model and for each index (with the first column identifying the country of the index). Standard errors appear in parentheses below coefficient estimates. Estimates that are significant at the 10%, two-sided level are indicated by a single asterisk, 5% significance is indicated by two asterisks, and 1% significance is indicated by three asterisks. (In contrast, KM report significance based only on 5% and 1% significance levels, denoted by one and two asterisks respectively.) We report coefficient estimates for temperature in Panel B, cloud cover in Panel C, and rainfall in Panel D. Panel E contains summary statistics, including the number of equations in the model (one for each index/country), the number of model parameters (including intercepts and autoregressive terms), and the number of weekly observations.[24] For the GMM estimation we report the degrees of freedom (over-identifying restrictions), the criterion value, the number of moment conditions per equation, and the test of over-identifying restrictions (a model specification test).[25] We also present joint tests, where appropriate, on model coefficients, first that the SAD onset/recovery coefficients are jointly zero across countries, second that the SAD onset/recovery coefficients are equal across indices, and third that each of the weather variable coefficients are equal across indices. The SAD onset/recovery coefficients are arranged by latitude, with highest northern latitude first (Ireland), to lowest northern latitude (Japan), then southern latitude (Australia). Finally we present estimates of the economic impact of SAD on returns.

If the SAD hypothesis is supported by the data, we should see that the onset of SAD is associated with lower returns and recovery from SAD is associated with higher returns, and we should see increasing absolute coefficient magnitudes as we consider markets increasingly north of the equator. This is generally what we find, with the largest magnitude estimate being either Ireland or the Netherlands in every model estimation (these are the most northern countries in Table 3), and Japan displaying the lowest, or one of the lowest, absolute magnitude estimate in each case. Switzerland is also among the smallest magnitude SAD effects, but the coefficient estimate on Switzerland's SAD onset/recovery variable is insignificant in each model. We report the economic impact of returns due to SAD in Panels F and G.[26] We find the economic impact of SAD on the US returns is comparable to that reported by KKL2003 using their cruder SAD measure. For the US, the impact of SAD on returns from January through June is roughly between 1.8% and 2.5% for the models with unrestricted SAD onset/recovery coefficients (depending on the model) and for July through December the impact is roughly equal and opposite (the SAD onset/recovery variable imposes symmetry), roughly between −1.8% and −2.5%. The magnitude of the effect for Ireland and the Netherlands is roughly twice as large, 3.5% to 5% and −3.5% to −5%. We discuss the economic significance of the SAD effect for the restricted model below.

Our non-joint tests for significance of the SAD effect from Model 1 (the OLS single-equation-at-a-time model) are similar to those from KM's Eq. (2), although we perform these tests on shorter series and on weekly rather than daily data. Only three of the return series of the 12 shown for Model 1 in Panel A of Table 3 are individually significant at the 5% level or better: the Netherlands, Ireland, and Austria, which are the same three that KM find statistically significant out of this group of 12 in their Table 7. Notably, however, Panel E of Table 3 includes an F-test for joint significance (i.e., difference from 0) of the 12 SAD coefficients. This test shows significance of the SAD effect across the 12 indices at the 0.1% level. Further, when we employ SUR estimation (Model 2), we find *eight* of the 12 SAD onset/recovery coefficients in Panel A of Table 3 are individually significant at the 5%, two-sided level, and nine are significant at the 10%, two-sided level. Clearly, making use of SUR, a panel/time-series method, greatly improves the ability to find a SAD effect, index-by-index, relative to simple OLS one-equation-at-a-time estimation as performed by KM. In Model 2, the joint test of no SAD effect is still strongly rejected, at the 0.8% level (versus the 0.1% level based on the OLS

[22] We test and fail to reject each restriction.

[23] The only notable difference between the results based on seasonally adjusted and seasonally unadjusted weather data is that the SAD onset/recovery coefficient estimates are somewhat less significant (albeit still significant at conventional levels and still economically strong and typically monotonically increasing with latitude) based on seasonally unadjusted data. This is likely because the weather data, in particular temperature, are strongly correlated with length of day, as is the SAD onset/recovery variable, resulting in multicollinearity and reducing power.

[24] A very small number of weeks are lost due to one index or another having a full-week closing of the exchange, typically the last week of December. We exclude these observations from the panel estimation by deleting the observations; we also explore including them with weight 0 in our parameter estimation criterion function (primarily for lag formation issues). Both approaches lead to nearly identical results.

[25] Hansen (1982), Staiger and Stock (1997), and Stock and Wright (2000) detail conditions sufficient for consistency and asymptotically normality of GMM estimation and show that the optimized value of the objective function produced by GMM is asymptotically distributed as a chi-square, providing a goodness-of-fit test of the model.

[26] We calculate the economic impact by multiplying the SAD onset/recovery coefficient value by the value of the onset/recovery variable, week by week, and cumulating the implied SAD return using continuous compounding.

Author's personal copy

estimation; correcting for the lack of independence across series with SUR can lower the significance of joint tests relative to an OLS test imposing independence). We reject equality of the SAD onset/recovery coefficients for Model 2, at the 0.5% level, whereas Model 1's low-power OLS single-equation-at-a-time estimation is unable to detect significant differences between parameter estimates.

A test that the weather coefficients are identical across indices fails to reject the null of equality for both Models 1 and 2, as shown in Panel E of Table 3. As a result we estimate Model 3 restricting the weather coefficients to be the same across indices. This more tightly parameterized model (with 63 rather than 96 parameters to estimate) indicates even stronger rejections of the null of no SAD effect. The estimation of Model 4, with GMM, confirms the OLS and SUR results with most of the SAD onset/recovery coefficients individually significant (see Panel A of Table 3) and the joint tests again reject the null of no effect, now at the 0.1% level (see Panel E of Table 3). The GMM model specification test on the over-identifying restrictions does not reject (again, see Panel E).

The GMM estimation of Model 4, like Models 1 through 3, strongly rejects the null that the SAD onset/recovery coefficient estimates are constant across indices; the formal test is in Panel E of Table 3. Further, inspection of the SAD coefficients in Panel A, Models 1 through 4, indicates the magnitude of the coefficients increases with latitude, consistent with the SAD hypothesis. Model 5 permits a test of the SAD effect by imposing this pattern, as follows. For Model 5, we rescale the onset/recovery variable, imposing a larger magnitude SAD effect at higher latitudes.[27] With this model we can test whether the rescaled SAD variable is statistically significant in a one degree of freedom *t*-test, which is a more powerful test if the restriction is approximately correct.

As we show in Panel E of Table 3, the specification test on the over-identifying restrictions for Model 5 does not reject the null of correct specification of this restricted model, and the test for no SAD effect is rejected with a *p*-value of 0.2% on the restricted SAD onset/recovery coefficient. The significance of the weather variables varies by estimation approach (see Panels B–D of Table 3), though using weekly data compromises the ability of the model to pick up high-frequency weather effects. When we aggregate the data from daily to weekly, blurring weather effects, the SAD effect is virtually unaffected, and in other research we find this result extends to the use of monthly data. The robustness of the SAD effect to data frequency supports the view that the SAD variable is not merely capturing a weather effect.

We note that the SAD onset/recovery variable has an unexpected sign for Australia. We can confirm, however, KM's result that estimating KKL2003's two-variable SAD and fall dummy model generally produces the predicted signs and occasional statistical significance for southern hemisphere countries, which is consistent with the SAD hypothesis but inconsistent with the evidence from models employing SAD onset/recovery. Recall that to construct the onset/recovery variable for the southern hemisphere we simply shift by 6 months the northern-hemisphere SAD onset/recovery variable (which is based on SAD prevalence estimated on samples from North America). The inconsistency in Australian results based on the onset/recovery variable versus the two-variable SAD specification suggests that the 6-months-shifted northern hemisphere SAD onset/recovery variable may be a poor proxy for the southern hemisphere, particularly in light of research suggesting that SAD symptoms are weak or non-existent in countries close to the equator (see Footnote 2). Unfortunately, we have available no reliable estimates of the precise timing of SAD onset and recovery in southern hemisphere countries.

5.2. Results for indices of countries grouped by latitude

The next set of results we present are for the aggregated indices, grouped by latitude into five groups: the average weekly return for exchanges in the southern-hemisphere tropics and sub-tropics (from the equator to 40°S), exchanges in the northern-hemisphere tropics and sub-tropics (from the equator to 40°N), for the US, exchanges in the latitude range above 40°N and below 50°N (excluding the US), and exchanges at or above 50°N. Again, we use the same data as KM with the exception of the US for which we use the total market return including dividends, equal-weighted. For this set of regressions we use seasonally adjusted weather data as suggested by Hirshleifer and Shumway (2003). Results using seasonally unadjusted data are similar.

Table 4 contains results, using structure similar to that of Table 3. Generally we see that the impact of SAD onset/recovery on returns increases almost monotonically with latitude, with small and generally statistically *in*significant coefficient estimates below roughly 0.4 in magnitude for the tropical and sub-tropical indices. The SAD onset/recovery coefficient is negative for northern tropical and sub-tropical countries, and it is positive for southern tropical and sub-tropical countries. As noted earlier, SAD should not necessarily be playing a role in exchanges close to the equator, where the daily number of hours of daylight varies little across the seasons, and a great many of these tropical and sub-tropical exchanges are very close to the equator. The coefficient on the equal-weighted US return is significant at the 10% (two-sided) level or better, is larger (up to double the magnitude) than seen for exchanges closer to the equator depending on the estimation technique, and is negative as predicted. The indices in the region between 40°N and 50°N and in the region above 50°N exhibit very strongly significant coefficient estimates (rejecting the null of no effect at the 1% level or better), and both regions' SAD onset/recovery coefficient estimates are typically larger than the US estimate, again depending on whether we employ OLS, SUR, or GMM estimation. The economic impact of SAD on the US and northern 40s and 50s exchanges (see Panels F and G of Table 4) is comparable to that reported by Kamstra et al. (2003). For example, the northern 40s and 50s countries exhibit impacts of as much as plus and minus 350 basis points, depending on the model. Finally, the joint test of no SAD and the test that the SAD coefficients are identical are rejected for each model estimation (see Panel E of Table 4). The test that the weather coefficients are identical across indices is not rejected nor is the GMM model specification test.

The restricted GMM estimation (Model 5) in Table 4 employs a rescaled onset/recovery variable, identical to that used for Table 3. The GMM test of over-identifying restrictions fails to reject this model (see Panel E of Table 4), the SAD coefficient is strongly statistically significant at conventional levels of significance (Panels A and E), and the economic impact from SAD closely matches that produced by the other estimations (Panels F and G), although just as we see for Model 5 in Table 3, the im-

[27] We first form a scaling factor for each exchange equal to the exchange latitude divided by 41 (the latitude in New York City). This produces a value of 1 for the US, values less than one for tropical and subtropical exchanges, e.g., 0.9 for Japan, and values greater than one for high latitude exchanges, e.g., 1.2 for Germany. We then scale the onset/recovery variable for each exchange and index by multiplying the onset/recovery variable by the scaling factor.

Author's personal copy

pact is more muted than that estimated by the unrestricted models.

In untabulated robustness checks, we conduct many alternative estimations to ensure the system-of-equation results are not fragile. We consider all of the individual indices used in our five aggregated indices, now separately. This reduces the overlapping data period substantially, to January 1993 through November 2003. Naturally this adversely affects the power of joint tests, as does including many tropical and sub-tropical indices. Importantly, however, the joint tests for the regressions which account for contemporaneous correlation, SUR and GMM, still reject the null of no SAD at the 0.1% level. We also consider restricting the system-of-equation analysis to exchanges outside of the tropics and sub-tropics, since medical research indicates SAD is most predominant at latitudes above 40°. In considering only the exchanges located above 40°, the data span the last week of 1979 through to December 2008, and we find strong, statistically significant joint SAD effects.[28] Next, we restrict our attention to tropical exchanges exclusively, and we find that none of our estimation techniques indicate statistically significant joint SAD effects. This null result is consistent with the medical literature and the SAD hypothesis. Finally, we utilize bootstrap resampling techniques to determine the data-adjusted significance of the SAD effect in the GMM models. We employ the block bootstrap technique of Politis and Romano (1994), using blocks of data of random length, distributed according to the geometric distribution with mean block length b. The parameter b is chosen so that block length is data-dependent. We set $b = N^{1/3}$, where N = sample size.[29] Thus we preserve the cross-sectional correlation and heteroskedasticity structure of the indices, and the block structure controls for autocorrelation. We resample 500 times to estimate the bootstrapped probability. We confirm the statistical significance of the SAD effect is qualitatively unchanged, for instance finding a bootstrapped p-value significant at better than the 1% level for Model 5 in Tables 1 and 2.

6. Conclusion

In response to the growing literature about the role of SAD in financial markets since the effect was first documented, we have refined our analysis, introducing a SAD measure based directly on the clinical timing of SAD symptoms. We have also considered the possibility that SAD has an influence in other financial market contexts including Treasury security returns and the flows of funds between mutual funds in different risk classes. In the meantime, new evidence has emerged, with Kramer and Weber (forthcoming) showing seasonal variation in depression and financial risk aversion at an individual level. We have, for several years, discussed with Professors Kelly and Meschke their evidence that we strongly believe is *in favor* of the SAD hypothesis. We hope that this public airing of the evidence will help in clarifying the real issues and will lead others to conclude that the SAD effect is alive and well. While it is impossible to prove that mere chance did not generate the return patterns we (and KM) document in stock indices around the world, we have attempted to show convincingly that there is no basis to KM's claim that the statistical evidence in support of the SAD effect arises due to a mechanical effect.

In their conclusions, KM take a very bold stand:

> [T]hese studies tend to skip three important steps by not examining whether the event-induced mood change actually affects investor perception of financial risk or return, whether such a change in perception manifests itself in trading behavior, and whether these sentiment-based trades impact stock prices. It is essential to carefully scrutinize these links for one to conclude that sentiment affects security prices. (p. 1324)

However, many prominent studies conduct valid inference using methods analogous to ours, observing correlation between market returns and factors thought to be indicative of investor mood or investor preferences.[30] Asset prices are essentially an aggregated indicator of investors' revealed preferences, and drawing inferences based on revealed preference has a long, storied tradition in economics. We do not agree with KM that it is essential in the early stages of identifying an effect to simultaneously scrutinize all the links they propose. A more practical and conventional course of action for testing a hypothesis that the mood of investors impacts market prices is first to ensure that the hypothesis is based on an *economic* argument (such as the hypothesis that investor mood impacts investor risk aversion which affects investor willingness to hold different risk classes of assets and in turn impacts asset returns) rather than an ad hoc argument, and second to determine whether there is sufficient empirical evidence consistent with the hypothesis. Should convincing empirical evidence exist, then deeper questions should be explored, but by proposing arbitrary hurdles, KM serve only to stifle legitimate scientific inquiry.

Acknowledgements

We thank Ian Garrett, Mark Weber, Ike Mathur, and an anonymous referee for helpful comments. This research was completed in part while Kamstra and Kramer were Visiting Scholars at the Stanford University Department of Psychology. They thank members of the university community for their hospitality. We are grateful to the Social Sciences Research Council of Canada and the Canadian Securities Institute Research Foundation for generous financial support. Any remaining errors are our own.

Appendix A. Observations regarding KM's Appendix A

KM catalogue international SAD prevalence rates based on a set of studies they cite, and they use these rates in some of their

[28] Note that the consideration of latitudes above 40° latitude excludes southern hemisphere countries since only northern hemisphere exchanges exist at latitudes higher than 40°. We employ the following indices located above 40°N: Ireland, the UK, Canada, Italy, France, Switzerland, Belgium, Austria, Netherlands, an equal-weighted average index of Nordic exchanges (rather than these countries separately), and the US. The Nordic countries include Iceland, Denmark, Sweden, Norway, and Finland, and to this group we add Germany to form an index of "Nordic" countries. We form this Nordic index because the German, Swedish, and Icelandic series are short (the German and Swedish series KM employ end in 2003, and the Iceland series does not begin until 1993) and using these series separately would shorten the usable time series for this panel estimation considerably. We set this index equal to the average of non-missing returns from these countries. Excluding Iceland, Germany, and Sweden entirely and using Denmark, Norway, and Finland to form a separate index return series leads to similar results.

[29] We explore several settings for mean block length. The setting based on $N^{1/3}$ leads to a mean block length of approximately 13 observations in our sample, which is a fairly long block length for weekly return data. White (2000) remarks that a mean block length of 10 for daily data is appropriate given the weak autocorrelation of returns. This would translate to a mean block length of 2 for our weekly data. Hence we explore using mean block length of 2, 5, and 13. Our results are virtually identical in all cases.

[30] Examples include Hirshleifer and Shumway (2003), Baker and Wurgler (2007), and Edmans et al. (2007), among many others.

Author's personal copy

Table 3
Analysis of series for which we have long time series.

Country	Parameter or Statistic	Model 1 OLS	Model 2 SUR	Model 3 SUR	Model 4 GMM	Model 5 GMM	Country	Parameter or Statistic	Model 1 OLS	Model 2 SUR	Model 3 SUR	Model 4 GMM	Model 5 GMM
Panel A: SAD onset/recovery coefficient estimates							*Panel B: Temperature coefficient estimates*						
Restricted	SAD					−.303***	Restricted	Temperature			−.005	−.014***	−.014***
						(.099)					(.006)	(.003)	(.003)
Ireland	SAD	−.872**	−1.141***	−1.206***	−.867***		Ireland	Temperature	−.020	−.009			
		(.380)	(.380)	(.343)	(.262)				(.018)	(.018)			
Netherlands	SAD	−.906***	−1.281***	−1.221***	−.848***		Netherlands	Temperature	−.018	−.004			
		(.349)	(.350)	(.324)	(.228)				(.012)	(.012)			
UK	SAD	−.268	−.636*	−.744**	−.408*		UK	Temperature	−.025**	−.016			
		(.346)	(.346)	(.324)	(.232)				(.012)	(.012)			
Belgium	SAD	−.467	−.686**	−.829***	−.590***		Belgium	Temperature	−.027**	−.019*			
		(.312)	(.311)	(.292)	(.223)				(.011)	(.011)			
Austria	SAD	−.549**	−.675**	−.765***	−.661***		Austria	Temperature	−.023***	−.014*			
		(.269)	(.269)	(.269)	(.254)				(.008)	(.008)			
Switzerland	SAD	−.208	−.441	−.400	−.226		Switzerland	Temperature	−.012	−.001			
		(.296)	(.296)	(.289)	(.189)				(.009)	(.009)			
France	SAD	−.505	−.800**	−.903***	−.532**		France	Temperature	−.028***	−.015			
		(.340)	(.344)	(.326)	(.259)				(.011)	(.011)			
Canada	SAD	−.479	−.625**	−.536*	−.321		Canada	Temperature	−.003	−.000			
		(.295)	(.296)	(.304)	(.212)				(.006)	(.006)			
Italy	SAD	−.467	−.645	−.905**	−.627**		Italy	Temperature	−.033**	−.027**			
		(.424)	(.425)	(.375)	(.312)				(.013)	(.014)			
US	SAD	−.442	−.635**	−.570**	−.460**		US	Temperature	−.006	−.002			
		(.286)	(.288)	(.288)	(.225)				(.006)	(.006)			
Japan	SAD	−.362	−.603	−.608*	−.112		Japan	Temperature	−.014	−.006			
		(.384)	(.384)	(.359)	(.275)				(.010)	(.010)			
Australia	SAD	.482*	.615**	.622**	.644***		Australia	Temperature	.010	−.008			
		(.278)	(.279)	(.254)	(.235)				(.015)	(.015)			
Panel C: Cloud cover coefficient estimates							*Panel D: Rainfall coefficient estimates*						
Restricted	Cloud Cover			−.033**	−0.015	−0.008	Restricted	Rainfall			−.000	−.018*	−.016*
				(.016)	(.016)	(.015)					(.006)	(.009)	(.009)
Ireland	Cloud Cover	0.024	0.006				Ireland	Rainfall	−.025	−.036*			
		(.063)	(.063)						(.021)	(.021)			
Netherlands	Cloud Cover	−.039	−.046				Netherlands	Rainfall	.003	.029			
		(.040)	(.041)						(.021)	(.023)			
UK	CloudCover	.010	−.058				UK	Rainfall	−.028*	−.006			
		(.050)	(.050)						(.016)	(.015)			
Belgium	CloudCover	−.055	−.063*				Belgium	Rainfall	−.028	−.009			
		(.038)	(.037)						(.027)	(.020)			
Austria	CloudCover	−.105***	−.072*				Austria	Rainfall	−.008	−.017			
		(.041)	(.040)						(.014)	(.016)			
Switzerland	CloudCover	−.032	−.009				Switzerland	Rainfall	−.029	−.010			
		(.040)	(.039)						(.024)	(.018)			
France	CloudCover	−.043	−.038				France	Rainfall	−.003	.010			
		(.040)	(.041)						(.018)	(.016)			
Canada	CloudCover	.016	−.014				Canada	Rainfall	−.014	−.014			
		(.048)	(.048)						(.023)	(.022)			
Italy	CloudCover	−.038	−.049				Italy	Rainfall	.036*	.043**			
		(.054)	(.053)						(.019)	(.019)			
US	CloudCover	−.042	−.043				US	Rainfall	−.008	.004			
		(.035)	(.035)						(.023)	(.023)			
Japan	CloudCover	−.024	−.015				Japan	Rainfall	−.016	−.004			
		(.039)	(.039)						(.016)	(.016)			
Australia	CloudCover	−.045	−.015				Australia	Rainfall	0.010	.006			
		(.040)	(.040)						(.017)	(.016)			

Author's personal copy

	Model 1 OLS	Model 2 SUR	Model 3 SUR	Model 4 GMM	Model 5 GMM
Panel E: Summary statistics					
Number of observations	1841	1841	1841	1841	1841
Number of equations	12	12	12	12	12
Number of parameters	96	96	63	39	28
GMM criterion value				111.05	114.90
Degrees of freedom				117	128
Number of restrictions				13	13
Over-identification test (p-value)				.638	.790
Test of SAD equal across indices (p-value)	.192	.005***	.000***	.001***	N/A
Test of SAD equal to 0 (p-value)	.001***	.008***	.000***	.001***	.002***
Test of weather coefficient estimates equal across indices (p-value)	.598	.421	N/A	N/A	N/A
Panel F: Economic impact due to SAD onset					
Ireland	−3.482	−4.535	−4.786	−3.465	−3.300
Netherlands	−3.615	−5.076	−4.844	−3.387	−2.634
UK	−1.081	−2.550	−2.981	−1.643	−2.092
Belgium	−1.881	−2.749	−3.315	−2.368	−1.653
Austria	−2.207	−2.705	−3.061	−2.651	−.790
Switzerland	−.841	−1.776	−1.613	−.913	−.790
Canada	−1.928	−2.506	−2.153	−1.296	−.272
France	−2.031	−3.199	−3.605	−2.140	−.611
Italy	−1.878	−2.589	−3.613	−2.516	−.205
US	−1.780	−2.549	−2.288	−1.853	−.154
Japan	−1.458	−2.420	−2.440	−.454	−.032
Australia	1.942	2.471	2.496	2.584	−.016
Panel G: Economic impact due to SAD recovery					
Ireland	3.495	4.599	4.865	3.478	3.307
Netherlands	3.634	5.174	4.927	3.397	2.623
UK	1.061	2.537	2.978	1.621	2.072
Belgium	1.860	2.740	3.323	2.352	1.631
Austria	2.188	2.695	3.060	2.640	.773
Switzerland	.823	1.754	1.591	.894	.773
Canada	1.907	2.493	2.135	1.274	.265
France	2.011	3.203	3.623	2.121	.597
Italy	1.857	2.576	3.631	2.502	.200
US	1.758	2.536	2.271	1.831	.150
Japan	1.436	2.405	2.426	.443	.031
Australia	−1.920	−2.455	−2.481	−2.571	.016

Notes: We consider the same indexes as KM (with the exception of the US series, which is composed of the total US equal-weighted return for AMEX, NASDAQ, and NYSE-listed equities, including dividends obtained from CRSP). Of those series, here we report results based on the series that have the longest span of available (non-missing) data, spanning July 1973 through to December 2008. We estimate Eq. (1) using various estimation techniques. Model 1 is single-equation OLS, Model 2 is based on SUR estimation, Model 3 is based on SUR estimation with constraints on weather coefficients across countries, Model 4 is based on GMM estimation, and Model 5 is based on GMM estimation with restrictions on weather and SAD onset/recovery coefficients across countries. Standard errors appear in parentheses. Models 1, 2, and 3 use MacKinnon–White (1985) heteroskedastity-consistent HC3 standard errors, and Models 4 and 5 use Newey–West (1994) standard errors.

* Significance at the 10% level, based on two-sided tests.

** Significance at the 5% level, based on two-sided tests.

*** Significance at the 1% level, based on two-sided tests.

Author's personal copy

Table 4
Analysis of indices by latitude.

Country		Model 1 OLS	Model 2 SUR	Model 3 SUR	Model 4 GMM	Model 5 GMM
Panel A: SAD onset/recovery coefficient estimates						
Restricted	SAD					−.429*** (.100)
North 50s	SAD	−.577** (.231)	−.766*** (.231)	−.768*** (.231)	−.860*** (.238)	
North 40s	SAD	−.637*** (.210)	−.746*** (.212)	−.747*** (.212)	−.759*** (.216)	
US	SAD	−.458* (.240)	−.568** (.241)	−.566** (.240)	−.768*** (.257)	
North Tropics and Sub-Tropics	SAD	−.232 (.229)	−.364 (.230)	−.351 (.229)	−.457* (.258)	
Southern Hemisphere	SAD	.254 (.230)	.359 (.231)	.361 (.230)	.235 (.253)	
Panel B: Temperature coefficient estimates						
Restricted	Temperature			.009 (.012)	−.003 (.014)	−.003 (.014)
North 50s	Temperature	−.019 (.019)	.009 (.019)			
North 40s	Temperature	−.021 (.020)	.001 (.020)			
US	Temperature	.014 (.015)	.009 (.016)			
North Tropics and Sub-Tropics	Temperature	.019 (.051)	.028 (.051)			
Southern Hemisphere	Temperature	.044 (.054)	.038 (.054)			
Panel C: Cloud cover coefficient estimates						
Restricted	Cloud Cover			−.041 (.026)	.019 (.031)	.018 (.030)
North 50s	Cloud Cover	−.005 (.058)	−.047 (.058)			
North 40s	Cloud Cover	−.073 (.055)	−.071 (.055)			
US	Cloud Cover	−.044 (.036)	−.026 (.036)			
North Tropics and Sub-Tropics	Cloud Cover	−.199 (.125)	−.186 (.126)			
Southern Hemisphere	Cloud Cover	−.075 (.069)	.011 (.069)			
Panel D: Rainfall coefficient estimates						
Restricted	Rainfall			−.002 (.005)	−.005 (.004)	−.005 (.004)
North 50s	Rainfall	−.055** (.024)	−.030 (.023)			
North 40s	Rainfall	−.017 (.028)	.021 (.029)			
US	Rainfall	−.009 (.023)	−.002 (.023)			
North Tropics and Sub-Tropics	Rainfall	.016 (.030)	.002 (.030)			
Southern Hemisphere	Rainfall	−.002 (.006)	−.002 (.006)			
Panel E: Summary statistics						
Number of observations		1874	1874	1874	1869	1869
Number of equations		5	5	5	5	5
Number of parameters		40	40	28	18	14
GMM criterion value					55.222	57.202
Degrees of freedom					47	51
Number of restrictions					13	13
Over-identification test (p-value)					.192	.256
Test of SAD equal to 0 (p-value)		.001***	.008***	.007***	.003***	.000***
Test of SAD equal across indices (p-value)		.037**	.014**	.011**	.052*	N/A
Test of weather coefficient estimates equal across indices (p-value)		.757	.989	N/A	N/A	N/A
Panel F: Economic impact due to SAD onset						
North 50s		−2.319	−3.068	−3.075	−3.438	−2.358
North 40s		−2.558	−2.989	−2.992	−3.040	−1.896
US		−1.842	−2.282	−2.274	−3.075	−1.729
North Tropics and Sub-Tropics		−.939	−1.468	−1.417	−1.840	−1.026
Southern Hemisphere		1.025	1.447	1.455	.952	−1.233
Panel G: Economic impact due to SAD recovery						
North 50s		2.305	3.071	3.079	3.454	2.345

Author's personal copy

Table 4 (*continued*)

Country	Model 1 OLS	Model 2 SUR	Model 3 SUR	Model 4 GMM	Model 5 GMM
North40s	2.548	2.990	2.993	3.042	1.877
US	1.823	2.268	2.260	3.078	1.709
North Tropics and Sub-Tropics	.922	1.448	1.396	1.821	1.007
Southern Hemisphere	−1.006	−1.426	−1.433	−.934	1.185

Notes: We consider index data for countries grouped according to their latitude. The groupings are: Southern Hemisphere, the US, North Tropics and Sub-Tropics, North 40s, and North 50s. The Southern Hemisphere grouping includes southern hemisphere countries in the tropics or sub-tropics (from the equator to 40°S): Indonesia, South Africa, Australia (Total Market), and New Zealand (FTSE New Zealand Index). North Tropics and Sub-Tropics grouping includes Singapore, Malaysia, Sri Lanka, the Philippines, Thailand, India, Hong Kong, Mexico, Taiwan, Jordan, Japan, Korea, Greece, Turkey, Spain, and China. The North 40s includes northern hemisphere countries at latitudes above 40°N and below 50°N: Italy, Canada, France, Austria, and Switzerland. The North 50s includes countries at or above 50°N: Belgium, Germany, the UK (Total Market), the Netherlands, Ireland, Denmark, Sweden (OMX Affärsvärldens Generalinde), Norway, Finland, and Iceland. The US series is composed of the total US equal-weighted return for AMEX, NASDAQ, and NYSE-listed equities, including dividends obtained from CRSP. The data span December 1972 through to December 2008. We estimate Eq. (1) using various estimation techniques. Model 1 is single-equation OLS, Model 2 is based on SUR estimation, Model 3 is based on SUR estimation with constraints on weather coefficients across countries, Model 4 is based on GMM estimation, and Model 5 is based on GMM estimation with restrictions on weather and SAD onset/recovery coefficients across countries. Standard errors appear in parentheses. Models 1, 2, and 3 use MacKinnon–White (1985) heteroskedastity-consistent HC3 standard errors, and Models 4 and 5 use Newey–West (1994) standard errors.

* Significance at the 10% level, based on two-sided tests.

** Significance at the 5% level, based on two-sided tests.

*** Significance at the 1% level, based on two-sided tests.

empirical analyses.[31,32] There are several caveats that one must keep in mind in interpreting and especially in comparing these data. For example, cross-country comparisons are severely compromised by sample selection biases, differences in the way SAD is identified among study participants, and failure of most of the studies to use methods designed to accurately identify clinical depression. We detail these problems, and several others, below.

First, KM state that all data reported in their Appendix A are "from general population studies" (p. 1325). In fact, several datasets they employ exhibit well-documented selection biases. KM describe one study as having sampled Canadians in the province of Manitoba (Magnusson and Axelsson, 1993), but this was far from a representative sample. In that study, the researchers ensured their sample contained *only* individuals whose Icelandic ancestry could be traced back at least as far as 1840. This sample was assembled *specifically because* people of Icelandic origin have demonstrated some resistance to SAD relative to people of other ancestry.[33] Thus, results from the Magnusson and Axelsson study are not representative of the general population of Canada or even the province of Manitoba. Other studies KM cite as representative of the general population use samples drawn from a workplace, such as Eagles et al. (1996), Hedge and Woodson (1996), Ito et al. (1992), and Ozaki et al. (1995). A study by Mersch (2001) (which KM cite in a different context) lists issues that can arise in conducting a study of SAD prevalence based on workplace samples. For instance, employees may be concerned about the confidentiality of their responses but may feel pressured to participate nonetheless (and this may compromise their responses) and absenteeism due to SAD may bias the sample. Similar problems likely apply to the studies KM cite that use student samples, including Han et al. (2000a, 2000b), Soriano et al. (2007), Lee et al. (2005), and Lee et al. (2006). Authors of several of the studies KM cite in their Appendix A are careful to note, themselves, that their samples are not representative of the general population. For instance Eagles et al. (1996) indicate their sample is "not a random sample of the population" (p. 132) and Soriano et al. (2007) write "[m]edical students likely do not represent the Romanian population as a whole" (p. 877).

All of the studies KM cite in their Appendix A use the seasonal pattern assessment questionnaire (SPAQ) method of identifying SAD, developed by Rosenthal et al. (1987). This method has some limitations. Lam and Levitt (1999, pp. 37–38) summarize these limitations as follows (note that they reference various versions of the manual psychiatric professionals consult in diagnosing mental illnesses, the Diagnostic and Statistical Manual of Mental Illness, DSM):

> (1) The SPAQ includes only four symptoms (appetite/weight, mood, sleep, energy) of the nine symptoms required to make a DSM-III-R or DSM-IV diagnosis of major depressive disorder. (2) The SPAQ does not directly assess impaired function that may result from each of these symptoms. (3) The SPAQ does not distinguish symptoms that might result from medical or physical conditions or drugs. (4) The SPAQ does not determine the number of major depressive episodes that the individual may have experienced in the past, nor their relationship to one another or to the seasons. For a diagnosis of SAD, DSM-III-R requires that three such episodes have occurred, two in consecutive years, and the DSM-IV requires two episodes in the past 2 years. (5) The SPAQ does not determine whether episodes were followed by complete remissions. (6) The month(s) in which mood is "best" or "worst" is (are) reported, but not when mood may be "depressed" or "high" or "normal."

[31] Note that the set of studies KM consider is not comprehensive. Omitted studies that document SAD prevalence rates are too numerous to list here.

[32] There are serious inconsistencies in KM's Appendix A. The percentage of the Magnusson and Stefansson (1993) Iceland sample that exhibit SAD is 3.8%, not 2.8% as KM report. The statistics KM attribute to Han et al. (2000a) come from Han et al. (2000b), and vice versa. We are unable to find several papers KM cite (Broman and Hetta, 1998; Hagfors et al., 1995; Konradsen, 1995, "Mersch et al., 1995," and Wirz-Justice et al., 1992). We requested copies of these papers from Professors Kelly and Meschke, and they reported that they too were unable to find them. The "Mersch et al. (1995)" study, which KM cite as having been published in the journal Acta Neuropsychiatrica, appears not to exist. In its place, KM may have meant to cite Mersch et al. (1999). In replying to our query, Professor Meschke reported that the data KM state as coming from the unavailable Wirz-Justice et al. (1992) abstract are from Wirz-Justice et al. (2003). He also reported that the second line of Finnish data in KM's Appendix A, which they attribute to Hagfors et al. (1995), are from Hagfors et al. (1992). (We have also been unable to find the Hagfors et al. (1992) study.) Obviously we are unable to comment on the validity of methods or prevalence statistics for the papers KM cite that we are unable to find. Thus, when we refer to "all" the studies KM cite, we are referring to all that we can access. Additional errors in KM's references are as follows: In citing the source of data from a Maryland study, KM cite "Kasper et al. (1989)" (in the Appendix and elsewhere in the paper). Their list of references indicates that study is by Kasper et al. (1989a). In fact the Maryland study is described in a different 1989 paper in the same journal by a different set of authors, Kasper et al. (1989b). A paper KM cite as "Elbi (2002)" on p. 1325 was written by multiple authors, Elbi et al. (2002). KM refer to one of the authors of the book in which the Mersch (2001) study appears as Patonen, but that author's surname is Partonen.

[33] Magnusson and Axelsson (1993) write "It was suggested that there may have been a genetic selection within the Icelandic population that has helped it to adapt to the long arctic winter. If this is correct, one might expect to find relatively low prevalence rates of SAD . . . among people of Icelandic descent living outside Iceland" (p. 947).

Author's personal copy

In light of these points, Lam and Levitt (1999) suggest that when using of SPAQ, researchers use supplemental diagnostic methods to identify current depressive episodes and use prospective studies to document the seasonality of the condition by sampling across seasons.

Further, studies that use SPAQ responses to identify individuals who suffer from SAD differ in the types of responses they accept as indicative of SAD, and this poses a problem for cross-study comparisons. Referring to SPAQ, Booker and Hellekson (1992) write "Differing threshold values for global seasonality score have been used" (p. 1177). Magnusson (2000) writes "[a] major problem with epidemiological SPAQ studies is that different cut-off points have been used for SPAQ classifications of SAD and S-SAD. Most studies include [a seasonality score] of 11 or greater as one of the criteria for SAD, thus 11 seems to be the standard... However, others have used [a seasonality score] of 10 as a cut-off point" (p. 181). Both cut-off values are variously adopted in the studies KM cite. For instance, Magnusson and Stefansson (1993) employ a cut-off value of 11, while Kasper et al. (1989b) use 10. Differences such as these influence prevalence rate estimates and necessarily complicate comparisons across studies.

In reference to yet another problem, Magnusson (2000) states:

> There are two additional SPAQ criteria for SAD, namely that the person has to feel worst during one of the winter months, and that the seasonal changes are experienced as a problem at least to a moderate degree. Researchers have not been consistent when defining which months are 'winter months', but I suggest that November, December, January and February should be the standard. Moreover, some studies have added 'feeling worst' in one of the summer months as an exclusion criterion. (p. 181)

The studies KM list in Appendix A differ in how they define winter, which further complicates comparisons across studies. Some require that subjects report feeling worst in one of December, January, or February (e.g. Soriano et al. (2007), Hedge and Woodson (1996)). Others require that subjects report feeling worst in January or February (e.g., Rosen et al. (1990), Muscettola et al. (1995), Kasper et al. (1989b)), and so on. Additionally complicating matters, some studies modify the SPAQ even more drastically than just redefining the timing of winter. Morrissey et al. (1996) employ SPAQ but arbitrarily elect not to refer to the four seasons of the year, instead modifying the seasons to suit the local northern Australian climate by using two seasons: wet (November through March) and dry (April through October). Unique peculiarities of individual studies such as these severely compromise KM's cross-study comparisons.

Some SAD studies apply stringent depression diagnostics; these are a small minority of the studies KM list in their Appendix A. For example, Booker and Hellekson (1992) use DSM-III-R criteria for diagnosing SAD, which was state-of-the-art at the time. But diagnostic criteria are updated as the medical profession's understanding of the condition changes. This complicates direct data comparisons across studies. Blazer et al. (1998) show that estimates of the prevalence of SAD vary considerably depending on the criteria employed. Rosen et al. (1990) find subjects score higher on the seasonality scale when interviewed in person, and Mersch (2001) observes that SPAQ-based studies using telephone interviews identify more cases of SAD than SPAQ-based mail surveys conducted by the same researchers studying the same geographic area.

Some authors offer possible reasons for differences in prevalence rates across countries. For instance, Ozaki et al. (1995) remark that differences between countries might reflect either "methodological differences or genuine differences in prevalence" (p. 1227) and that "Japanese subjects may be reluctant to complain about their problems, or American subjects might tend to exaggerate theirs" (p. 1227). Ozaki et al. caution against generalizing on the basis of their findings, even within country: "Our ability to generalize from the present findings is limited because the subjects studied were not representative of the Japanese population as a whole" (p. 1226). One must keep such points in mind when different statistics arise from geographically similar locations: the Ozaki et al. study finds less than 1% of their Japanese sample experiences SAD while Lee et al. (2005, 2006) find more than 10% of their Korean samples experience SAD. Yet KM fail to qualify their cross-country and cross-study comparisons in any way.

The season when individuals participate in a survey can influence both sample composition and subjects' responses (due to differences in memory recall regarding the timing of the depressive experience; see, for instance, Kasper et al., 1989b, and Magnusson, 2000). This factor varies markedly across the studies KM cite. For example, Morrissey et al. (1996) mailed surveys in September but do not report how long they waited for responses, Rastad et al. (2005) mailed surveys to residents in the first week of December and waited until the end of February for responses to be mailed back, Eagles et al. (1996) mailed questionnaires in early December and reminded potential participants once if a response had not been received within 4 weeks, Muscettola et al. (1995) mailed in early January and waited until early March, Booker and Hellekson (1992) collected their data from mid-January through mid-March, and Magnusson and Stefansson (1993) mailed questionnaires out in spring and sent up to three follow-up reminders. These differences complicate cross-study comparisons.

Some of the studies KM cite consider very few subjects. For instance, Srisurapanont and Intaprasert (1999) study 92 people. Haggarty et al. (2002) study 88.[34] Some of the cited researchers themselves caution against generalizing on the basis of their findings: "Since these findings are based on a small sample, we must be cautious about drawing conclusions" (Rosen et al., 1990, p. 133, in reference to the subset of participants they interviewed in person).

A further technical issue that complicates comparisons across studies is that there is variation across studies in the length of the window during which individuals need to experience onset or recovery to be labeled as suffering from SAD. Referencing a Japanese study that is not included in KM's Appendix A, Takahashi et al. (1991) write:

> In the DSM-IIIR criteria, the window of onset and remission from the depressive episode is set at a particular 60 days of the year in individuals. We therefore examined the variation in onset time of the depressive episode in each patient, as shown in Table 1. Only 41% (or 40%) of all SAD patients examined met the criteria, suggesting that a 60 day window is insufficient to identify the disorder. If we had followed the DSM-IIIR criteria for seasonal pattern strictly, more than half of the SAD patients in Japan would have been overlooked. (p. 71)

Note that according to Lam and Levitt (1999), the DSM-III-R 60-day window criterion, which many researchers view as overly restrictive, is no longer part of the current DSM-IV criteria. Among the handful of studies KM cite that use DSM criteria, all use the restrictive DSM-III-R 60-day window criterion.

Not all studies KM cite in their Appendix A study winter SAD. For example, Soriano et al. (2007) examine the influence of air conditioning on incidence of reverse (summer) SAD.

---

[34] KM report their Appendix A estimates to two decimal places. The original studies rarely provide that number of significant digits; e.g., Rosen et al. (1990) report the rate of SAD in NH as 11% but KM report this as 11.00%.

Author's personal copy

In sum, there are large differences in the methodologies underlying the studies KM cite in their Appendix A. These differences seriously compromise cross-country comparisons. Furthermore, employing the SAD prevalence statistics from these studies as explanatory variables in regression analysis, as KM do in their Table 4, leads to an errors-in-variables problem that can invalidate inference. Lam and Levitt (1999) suggest that researchers should supplement the use of SAD identification methods with the use of diagnostic methods to identify depression and that one should ideally document seasonality by sampling at multiple points during the year. The clinical study of SAD is still in a relative state of infancy, with the condition having first been formally identified by Rosenthal et al. (1984). As the medical profession's understanding of the condition improves, and as more researchers adopt robust methodologies for documenting SAD prevalence around the world, the type of comparisons KM attempt to make may become feasible. In the meantime, we believe the best estimate of the influence of SAD on equity returns around the world is the SAD onset/recovery variable based on data from Lam (1998), perhaps scaled by latitude for application to markets outside of North America.

References

Albuquerque, R., Bauer, G.H., Schneider, M., 2009. Global private information in international equity markets. Journal of Financial Economics 94 (1), 18–46.

Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. Journal of Economic Perspectives 21 (2), 129–151.

Bekaert, G., Engstrom, E., Xing, Y., 2009. Risk, uncertainty, and asset prices. Journal of Financial Economics 91 (1), 59–82.

Blazer, D.G., Kessler, R.C., Swartz, M.S., 1998. Epidemiology of recurrent major and minor depression with a seasonal pattern: the National Comorbidity Survey. British Journal of Psychiatry 172, 164–167.

Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. Journal of Computational Science 2 (1), 1–8.

Booker, J.M., Hellekson, C.J., 1992. Prevalence of seasonal affective disorder in Alaska. American Journal of Psychiatry 149, 1176–1182.

Broman, J.E., Hetta, J., 1998. Prevalence of seasonal affective disorders and related symptoms at two latitudes in Sweden. Society for Light Treatment and Biological Rhythms Abstracts 10, 228.

Campbell, J., Cochrane, J., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107 (2), 205–241.

Carton, S., Jouvent, R., Bungener, C., Wildlöcher, D., 1992. Sensation seeking and depressive mood. Personality and Individual Differences 13, 843–849.

Carton, S., Morand, P., Bungenera, C., Jouvent, R., 1995. Sensation-seeking and emotional disturbances in depression: relationship and evolution. Journal of Affective Disorders 34, 219–225.

Clark, L., Iversen, S., Goodwin, G., 2001. The influence of positive and negative mood states on risk taking, verbal fluency, and salivary cortisol. Journal of Affective Disorders 63, 179–187.

DeGennaro, R.P., Kamstra, M.J., Kramer, L.A., 2008. Does Risk Aversion Vary During the Year? Evidence from Bid-Ask Spreads. Working Paper, University of Toronto.

Dolvin, S., Pyles, M., 2007. Seasonal affective disorder and the pricing of IPOs. Review of Accounting and Finance 6 (2), 214–228.

Dolvin, S., Pyles, M., Wu, Q., 2009. Analysts get SAD too: the effect of seasonal affective disorder on stock analysts' earnings estimates. Journal of Behavioral Finance 10 (4), 214–225.

Dowling, M., Lucey, B.M., 2008. Robust global mood influences in equity pricing. Journal of Multinational Financial Management 18, 145–164.

Eagles, J.M., Gavin, M., Andrew, J., Boshier, F.J., 1996. Seasonal affective disorder among psychiatric nurses in Aberdeen. Journal of Affective Disorders 37, 129–135.

Edmans, A., Garcia, D., Norli, Ø., 2007. Sports sentiment and stock returns. Journal of Finance 62 (4), 1967–1998.

Eisenberg, A., Baron, J., Seligman, M., 1998. Individual Differences in Risk Aversion and Anxiety. Working Paper, University of Pennsylvania.

Elbi, H., Noyan, A., Korukoğlu, S., Ünal, S., Bekaroğlu, M., Oğuzhanoğlu, N., Türköz, N., Abay, E., Kumbasar, H., Yurdakul, S., 2002. Seasonal affective disorder in eight groups in Turkey: a cross-national perspective. Journal of Affective Disorders 70, 77–84.

Fama, E., French, K., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23–49.

Ferson, W.E., Foerster, S., 1994. Finite sample properties of generalized methods of moments in tests of conditional asset pricing models. Journal of Financial Economics 36, 29–55.

Forgas, J.P., 1995. Mood and judgment: the affect infusion model (AIM). Psychological Bulletin 117 (1), 39–66.

Garrett, I., Kamstra, M.J., Kramer, L.A., 2005. Winter blues and time variation in the price of risk. Journal of Empirical Finance 12 (2), 291–316.

Gilbert, E., Karahalios, K., 2010. Widespread worry and the stock market. In: 4th International AAAI Conference on Weblogs and Social Media (ICWSM).

Goetzmann, W.N., Zhu, N., 2005. Rain or shine: where is the weather effect? European Financial Management 5, 559–578.

Grinblatt, M., Keloharju, M., 2008. Sensation seeking, overconfidence, and trading activity. Journal of Finance 64, 549–578.

Hagfors, C., Koskela, K., Tikkanen, J., 1992. Seasonal affective disorder (SAD) in Finland: an epidemiological study. Society for Light Treatment and Biological Rhythms Abstract 4, 24.

Hagfors, C., Thorell, L., Arned, M., 1995. Seasonality in Finland and Sweden, an epidemiologic study, preliminary results. Society for Light Treatment and Biological Rhythms Abstracts 7, 22.

Haggarty, J.M., Cernovsky, Z., Husni, M., Minor, P., Kermeen, P., Merskey, H., 2002. Seasonal affective disorder in an Arctic community. Acta Psychiatrica Scandinavica 105, 378–384.

Han, L., Wang, K., Cheng, Y., Du, Z., Rosenthal, N., Primeau, F., 2000a. Summer and winter patterns of seasonality in Chinese college students: a replication. Comprehensive Psychiatry 41, 57–62.

Han, L., Wang, K., Du, Z., Chen, Y., Simons, J., Rosenthal, N., 2000b. Seasonal variation in mood and behavior among Chinese medical students. American Journal of Psychiatry 157, 133–135.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1084.

Harlow, W.V., Brown, K.C., 1990. Understanding and assessing financial risk tolerance: a biological perspective. Financial Analysts Journal 6 (6), 50–80.

Hedge, A.L., Woodson, H., 1996. Prevalence of seasonal changes in mood and behavior during the winter months in central Texas. Psychiatry Research 62, 265–271.

Hirshleifer, D., Shumway, T., 2003. Good day sunshine: stock returns and the weather. Journal of Finance 58, 1009–1032.

Hockey, G.R.J., Maule, A.J., Clough, P.J., Bdzola, L., 2000. Effects of negative mood states on risk in everyday decision making. Cognition and Emotion 14, 823–855.

Hodrick, R.J., Zhang, X., 2001. Evaluating the specification errors of asset pricing models. Journal of Financial Economics 62, 327–376.

Horvath, P., Zuckerman, M., 1993. Sensation seeking, risk appraisal and risky behavior. Personality and Individual Differences 14, 41–51.

Isen, A.M., Nygren, T.E., Ashby, F.G., 1988. Influence of positive affect on the subjective utility of gains and losses: it is just not worth the risk. Journal of Personality and Social Psychology 55, 710–717.

Ito, A., Ichihara, M., Hisanaga, N., Ono, Y., Kayukawa, Y., Ohta, T., Okada, T., Ozaki, N., 1992. Prevalence of seasonal mood changes in low latitude area: seasonal pattern assessment questionnaire score of Quezon city workers. Japanese Journal of Psychiatry and Neurology 46, 249.

Jacobsen, B., Marquering, W., 2008. Is it the weather? A comment on studies linking weather and stock market behaviour. Journal of Banking and Finance 32, 526–540.

Jacobsen, B., Marquering, W., 2009. Is it the weather? Response. Journal of Banking and Finance 33, 583–587.

Jagannathan, R., Wang, Z., 2007. Lazy investors, discretionary consumption, and the cross-section of stock returns. Journal of Finance 62 (4), 1623–1661.

Kamstra, M.J., Kramer, L.A., Levi, M.D., 2003. Winter blues: a SAD stock market cycle. American Economic Review 93 (1), 324–343.

Kamstra, M.J., Kramer, L.A., Levi, M.D., 2009. Is it the weather? Comment. Journal of Banking and Finance 33, 578–582.

Kamstra, M.J., Kramer, L.A., Levi, M.D., 2011a. Seasonal Variation in Treasury Returns. Working Paper, University of Toronto. Available on SSRN: <http://ssrn.com/abstract=1076644>.

Kamstra, M.J., Kramer, L.A., Levi, M.D., Wang, T., 2011b. Seasonally Varying Preferences: Foundations for an Empirical Regularity. Working Paper, University of Toronto.

Kamstra, M.J., Kramer, L.A., Levi, M.D., Wermers, R., 2011c. Seasonal Asset Allocation: Evidence from Mutual Fund Flows. Working Paper, University of Maryland. Available on SSRN: <http://ssrn.com/abstract=1907904>.

Kaplanski, G., Levy, H., 2009. Seasonality in Perceived Risk: A Sentiment Effect. Working Paper, The Hebrew University of Jerusalem. Available on SSRN: <http://ssrn.com/abstract=1116180>.

Kasper, S., Rogers, S.L., Yancey, A., Schulz, P.M., Skwerer, R.G., Rosenthal, N.E., 1989a. Phototherapy in individuals with and without subsyndromal seasonal affective disorder. Archives of General Psychiatry 46 (9), 837–844.

Kasper, S., Wehr, T.A., Bartko, J.J., Gwst, P.A., Rosenthal, N.E., 1989b. Epidemiological findings of seasonal changes in mood and behavior: a telephone survey of Montgomery County, Maryland. Archives of General Psychiatry 46 (9), 823–833.

Kelly, P.J., Meschke, F., 2010. Sentiment and stock returns: the SAD anomaly revisited. Journal of Banking and Finance 34, 1308–1326.

Keller, M.C., Fredrickson, B.L., Ybarra, O., Côté, S., Johnson, K., Mikels, J., Conway, A., Wager, T., 2005. A warm heart and a clear head: the contingent effects of weather on mood and cognition. Psychological Science 16 (9), 724–731.

Kliger, D., Levy, O., 2008. Mood impacts on probability weighting functions: "large-gamble" evidence. Journal of Socio-Economics 37, 1397–1411.

Kliger, D., Gurevich, G., Haim, A., 2010. When Chronobiology Met Economics – Seasonal Affective Impact on the Demand for IPOs. Working Paper, University of Haifa, Available at SSRN: <http://ssrn.com/abstract=1502566>.

Author's personal copy

Konradsen, H., 1995. Årstidsavhengig affektiv forstyrrelse: En emperisk undersøkelse av forekomst blant 6300 unge voksne fra 58 til 703N. Working Paper, University of Tromsø.

Kramer, L.A., Weber, J.M., forthcoming. This is your portfolio on winter: seasonal affective disorder and risk aversion in financial decision-making. Social Psychological and Personality Science.

Lam, R.W., 1998. Seasonal affective disorder: diagnosis and management. Primary Care Psychiatry 4, 63–74.

Lam, R.W., Levitt, A.J., 1999. Canadian Consensus Guidelines for the Treatment of Seasonal Affective Disorder. Clinical & Academic Publishing, Vancouver.

Leith, K., Baumeister, R., 1996. Why do bad moods increase self-defeating behavior? Emotion, risk taking, and self regulation. Journal of Personality and Social Psychology 71 (6), 1250–1267.

Lee, H., Sung, S., Han, C., Kim, Y., Kim, S., Lee, M., 2005. G-protein β_3 subunit C825T polymorphism tends to be associated with seasonal variation in young male college students. Neuropsychobiology 52, 135–139.

Lee, H., Sung, S., Lim, S., Paik, J., Kim, L., 2006. Seasonality associated with the serotonin 2A receptor – 1438 A/G polymorphism. Journal of Affective Disorders 95, 145–148.

Lo, K., Wu, S.S., 2008. The Impact of Seasonal Affective Disorder on Financial Analysts and Equity Market Returns, Working Paper, University of British Columbia.

MacKinnon, J.G., White, H., 1985. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. Journal of Econometrics 29 (3), 305–325.

Magnusson, A., 2000. An overview of epidemiological studies on seasonal affective disorder. Acta Psychiatrica Scandinavica 101, 176–184.

Magnusson, A., Axelsson, J., 1993. The prevalence of seasonal affective disorder is low among descendants of Icelandic emigrants in Canada. Archives of General Psychiatry 50, 947–951.

Magnusson, A., Stefansson, J.G., 1993. Prevalence of seasonal affective disorder in Iceland. Archives of General Psychiatry 50, 941–946.

Mersch, P.P., Middendorp, H.M., Bouhuys, A.L., Beersma, D.G.M., van den Hoofdakker, R.H., 1999. The prevalence of seasonal affective disorder in the Netherlands: a prospective and retrospective study of seasonal mood variation in the general population. Biological Psychiatry 45, 1013–1022.

Mersch, P., 2001. Prevalence from population surveys. In: Partonen, T., Magnusson, A. (Eds.), Seasonal Affective Disorder: Practice and Research. Oxford University Press, Oxford.

Morrissey, S.A., Raggatt, P.T., James, B., Rogers, J., 1996. Seasonal affective disorder: some epidemiological findings from a tropical climate. Australian and New Zealand Journal of Psychiatry 30, 579–586.

Morse, W., 1998. Risk taking in personal investments. Journal of Business and Psychology 13 (2), 281–288.

Muscettola, G., Barbato, G., Ficca, G., Beatrice, M., Puca, M., Aguglia, E., Amati, A., 1995. Seasonality of mood in Italy: role of latitude and sociocultural factors. Journal of Affective Disorders 33, 135–139.

Newey, W.K., West, K.D., 1994. Automatic lag selection in covariance matrix estimation. Review of Economic Studies 61, 631–653.

Ozaki, N., Ono, Y., Ito, A., Rosenthal, N., 1995. Prevalence of seasonal difficulties in mood and behavior among Japanese civil servants. American Journal of Psychiatry 152, 1225–1227.

Palinkas, L., Houseal, M., 2000. Stages of change in mood and behavior during a winter in Antarctica. Environment and Behavior 32, 128–141.

Palinkas, L., Houseal, M., Rosenthal, N., 1996. Subsyndromal seasonal affective disorder in Antarctica. Journal of Nervous and Mental Disease 184, 530–534.

Pietromonaco, P.R., Rook, K.S., 1987. Decision style in depression: the contribution of perceived risks versus benefits. Journal of Personality and Social Psychology 52 (2), 399–408.

Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. Journal of the American Statistical Association 89 (428), 1303–1313.

Pyles, M.K., 2009. The influence of seasonal depression on equity returns: further evidence from real estate investment trusts. Quarterly Journal of Finance and Accounting 48 (2), 63–83.

Raghunathan, R., Pham, M.T., 1999. All negative moods are not equal: motivational influences of anxiety and sadness on decision making. Organizational Behavior and Human Decision Processes 79 (1), 56–77.

Rastad, C., Sjoden, P., Ulfberg, J., 2005. High prevalence of self-reported winter depression in a Swedish county. Psychiatry and Clinical Neurosciences 59, 666–675.

Rosen, L.N., Targum, S.D., Terman, M., Bryant, M.J., Hoffman, H., Kasper, S.F., Hamovit, J.R., Docherty, J.P., Welch, B., Rosenthal, N.E., 1990. Prevalence of seasonal affective disorder at four latitudes. Psychiatry Research 31, 131–144.

Rosenthal, N.E., Sack, D.A., Gillin, J.C., Lewy, A.J., Goodwin, F.K., Davenport, Y., Mueller, P., Newsome, D.A., Wehr, T.A., 1984. Seasonal affective disorder: a description of the syndrome and preliminary findings with light therapy. Archives of General Psychiatry 41, 72–80.

Rosenthal, N.E., Genhart, M., Sack, D.A., Skwerer, R.G., Wehr, T.A., 1987. Seasonal affective disorder: relevance for treatment and research of bulimia. In: Hudson, J.I., Pope, H.G. (Eds.), Psychobiology of Bulimia. American Psychiatric Press, Washington, DC.

Smoski, M., Lynch, T., Rosenthal, Z., Cheavens, J., Chapman, A., Krishnan, R., 2008. Decision-making and risk aversion among depressive adults. Journal of Behavior Therapy and Experimental Psychiatry 39, 567–576.

Soriano, J., Ciupagea, C., Rohan, K., Neculai, D., Yousufi, S., Guzman, A., Postolache, T., 2007. Seasonal variations in mood and behavior in Romanian postgraduate students. The Scientific World Journal 7, 870–879.

Srisurapanont, M., Intaprasert, S., 1999. Seasonal variations in mood and behaviour: epidemiological findings in the north tropics. Journal of Affective Disorders 54, 97–99.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557–586.

Stock, J., Wright, J., 2000. GMM with weak identification. Econometrica 68 (5), 1055–1096.

Takahashi, K., Asano, Y., Kohsaka, M., Okawa, M., Sasaki, M., Honda, Y., Higuchi, T., Yamazaki, J., Ishizuka, Y., Kawaguchi, K., Ohta, T., Hanada, K., Sugita, Y., Maeda, K., Nagayama, H., Kotorii, T., Egashira, K., Takahashi, S., 1991. Multi-center study of seasonal affective disorders in Japan: a preliminary report. Journal of Affective Disorders 21, 57–65.

Tokunaga, H., 1993. The use and abuse of consumer credit: application of psychological theory and research. Journal of Economic Psychology 14, 285–316.

White, H., 2000. A reality check for data snooping. Econometrica 68 (5), 1097–1126.

Wirz-Justice, A., Kräuchi, K., Graw, P., Schulman, J., Wirz, H., 1992. Seasonality in Switzerland: an epidemiological survey. Society for Light Treatment and Biological Rhythms Abstracts 4, 33.

Wirz-Justice, A., Graw, P., Kräuchi, K., Wacker, H.R., 2003. Seasonality in affective disorders in Switzerland. Acta Psychiatrica Scandinavia 108, 92–95.

Wong, A., Carducci, B., 1991. Sensation seeking and financial risk taking in everyday money matters. Journal of Business and Psychology 5, 525–530.

Young, M.A., Meaden, P.M., Fogg, L.F., Cherin, E.A., Eastman, C.I., 1997. Which environmental variables are related to the onset of seasonal affective disorder? Journal of Abnormal Psychology 106 (4), 554–562.

Young, M.A., 2001. Weather. In: Partonen, T., Magnusson, A. (Eds.), Seasonal Affective Disorder: Practice and Research. Oxford University Press, Oxford.

Yuen, K.S.L., Lee, T.M.C., 2003. Could mood state affect risk-taking decisions? Journal of Affective Disorders 75, 11–18.

Zuckerman, M., 1979. Sensation Seeking: Beyond the Optimal Level of Arousal. Lawrence Erlbaum Associates, Hillsdale.

Zuckerman, M., 1984. Sensation seeking: a comparative approach to a human trait. Behavioral and Brain Sciences 7, 413–471.

Zuckerman, M., 1994. Behavioral Expression and Biosocial Bases of Sensation Seeking. Cambridge University Press, Cambridge.