The Effects of Total Sleep Deprivation on Bayesian Updating*

<table>
<thead>
<tr>
<th>David L. Dickinson</th>
<th>Sean P.A. Drummond</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dept. of Economics</td>
<td>Dept. of Psychiatry</td>
</tr>
<tr>
<td>Appalachian State University</td>
<td>Laboratory for Sleep and Behavioral Neuroscience</td>
</tr>
<tr>
<td></td>
<td>University of California—San Diego</td>
</tr>
<tr>
<td></td>
<td>Psychology Service</td>
</tr>
<tr>
<td></td>
<td>Veterans Affairs San Diego Healthcare System</td>
</tr>
</tbody>
</table>

ABSTRACT

Recent evidence suggests that nearly 25% of U.S. adults (47 million) suffer from some level of sleep deprivation. The impact of this sleep deprivation on the U.S. economy includes direct medical expenses related to sleep deprivation and related disorders, the cost of accidents, and the cost of reduced worker productivity. Sleep research has examined the effects of sleep deprivation on a number of performance measures, but the effects of sleep deprivation on decision-making under uncertainty are largely unknown. In this article, subjects perform a decision task (Grether, 1980) in both a well-rested and experimentally sleep-deprived state. We have two main results: 1) final choice accuracy is unaffected by sleep deprivation, and 2) the estimated decision model used is significantly different when sleep-deprived compared to well-rested. Following sleep deprivation, subjects weigh all sources of information to a lesser degree, but they also do not display the tendency to over-weight new information in forming beliefs when well-rested. Because the altered decision process still maintains decision accuracy, it suggests that increased accident and error rates attributed to reduced sleep in modern society may result more from a decline in auxiliary functions (e.g., slowed reaction time, reduced motor skills) or other components of decision making, rather than the inability to process new information.

JEL Key Words: Bayes Rule, Uncertainty, Information, Experiments, Sleep.
JEL Codes: D81, D83, C91

*The authors wish to thank the U.S. National Academy of Sciences and the Japanese Society for the Promotion of Sciences for sponsoring the Japanese-American Beckman Frontiers of Science Symposium, which led to this research collaboration. Helpful comments were provided by Rob Oxoby, participants at the Economic Science Association meetings, and seminar participants at the U.S. Air Force Academy and Appalachian State University. The authors are particularly grateful to Henry J. Orff and Benjamin S. McKenna for their work on experiment administration and data management. This study was funded in part by the US Department of the Army award #DAMD17-02-1-0201 and NIH M01 RR00827. The US Army Medical Research Acquisition Activity is the awarding and administering acquisition office. The content herein does not necessarily reflect the position or policy of the Government, and no official endorsement should be inferred.
A large volume of evidence suggests that individuals in industrialized nations are becoming increasingly sleep-deprived. According to a recent poll conducted by the *National Sleep Foundation*, the average American adult slept less than 7 hours per night in 2005. The nightly average was 7.5 hours in 1975 and 9 hours per night in 1910 (Coren, 1996). This trend towards less and less sleep has significant implications given the known effects of sleep deprivation: decreased motor and cognitive performance, reduced vigilance and reaction time, worsened mood, and reduced ability to think flexibly (Pilcher and Huffcutt, 1996; Harrison and Horne, 1999; Harrison and Horne, 2000). Indeed, even 7 hours of sleep per night leads to significantly diminished cognitive performance relative to 8 or 9 hours (Van Dongen, et al, 2003; Belenky, et al., 2003). Nearly 50 million Americans, close to 25% of all adults, are estimated to suffer from some level of sleep deprivation. Sixty percent of adults surveyed reported driving while drowsy, while 37% reported falling asleep or nodding off at some point while driving.¹ Estimates of the cost of lost U.S. worker productivity caused by sleep deprivation vary, but a conservative estimate—based on a 4% reduction in productivity for sleep-deprived working adults—is over $40 billion dollars annually (Stoller, 1997).

It may be tempting to argue that reduced hourly productivity in labor markets, for example, is offset by the longer work hours made possible through sleep reduction. However, Biddle and Hammermesh (1990) find that sleep reduction resulting from higher wages results in increased waking leisure time rather than increased work time. Furthermore, the full impact of sleep loss in the workplace includes a doubled rate of accidents (Melamed and Oksenberg, 2002), increased risk of fatal accidents (Akerstedt et al, 2002), increased absenteeism (Phillips et al., 1991; Kupperman et al., 1995), greater

¹ This data is reported by the *National Sleep Foundation*, and can be accessed at [www.sleepfoundation.org](http://www.sleepfoundation.org).
medical morbidity and related costs of such (Drake et al., 2004), and even slower career advancement (Johnson and Spinweber, 1983).

Many occupations promote a culture of sleep deprivation, whether it be the use of shift work in factories or hospitals, or the need to alter sleep schedules to monitor real-time foreign financial market activity. Certain professions that give rise to more significant sleep deprivation as a matter of routine—emergency personnel, medical residents, military personnel, long-haul truck drivers—are also those where impaired functioning can put others’ lives at risk. A study of long-haul truck drivers in Canada and the U.S. (Mitler, et al. 1997) found that they averaged only about 5 hours of sleep per night. A recent study of first- and second-year medical residency students found that two-thirds reported sleeping an average of six or less hours per night (Baldwin, et al. 2004). A smaller fraction (20%) averaged five hours of sleep a night, and such residents were more likely to report, among other things, having made significant medical errors. Weinger and Ancoli-Israel (2002) concluded that sleep deprivation significantly impairs doctors’ performance, thereby impacting patient safety, in part due to poor decisions made by sleep deprived physicians.

Accidents linked to sleepiness are, of course, not limited to the workplace (e.g., home, driving). Leger (1994) estimated the costs of accidents across numerous settings (work, home, driving, public accidents) to be $43-$56 billion, in 1988 dollars. Sleep deprivation has also been considered at least partially responsible for several major historical disasters, including the Space Shuttle Challenger explosion, the Exxon Valdez oil spill, and the Chernobyl Nuclear plant explosion (Coren, 1996). In sum, the impact of
sleep deprivation in the workplace and on society as a whole, while difficult to measure precisely, is massive.

This paper reports results from a laboratory study that examines the information processing abilities of subjects in a well-rested versus an experimentally sleep-deprived state. Much of the existing sleep deprivation research examines subject performance on sustained attention, mathematical and/or verbal tasks, such as simple reaction time tasks, arithmetic processing, grammatical reasoning, or verbal learning. Examinations of flexible thinking, strategy updating, and risk assessment are relatively new to sleep research (see references in Harrison and Horne, 2000). In particular, there are only a few studies examining decision making in the context of sleep deprivation. These studies have typically used very complex tasks involving many decisions that together produce an end result that is evaluated as “correct” or not. Such designs do not allow the investigators to identify specific aspects of decision making that may be altered by sleep deprivation.

The current study aims to examine the effects of sleep deprivation on one specific aspect of decisions familiar to economists: Bayesian updating. We examine differences in subjects’ propensity to incorporate new information as they update prior probabilities to form posterior (subjective) probability estimates. A Bayes rule experiment is administered to subjects both well-rested and after 22-25 hours ($\mu=22.72, \sigma=.60$) of total sleep deprivation (TSD). For comparison to existing economics research, we utilize the Bayes rule experiment presented in Grether (1980). His results indicate that, contrary to Bayesian updating, subjects tend to overweight new evidence relative to prior odds when forming subjective beliefs. The result, further confirmed in Grether (1992), is largely due
to subjects’ tendencies to utilize a ‘representativeness’ heuristic in cases where new sample information looks representative of one population versus another (see, e.g., Kahneman and Tversky, 1972).²

The results from our pooled sample (well-rested and sleep-deprived data) are quite similar to those in Grether (1980). However, we provide statistical evidence that the same decision model parameter estimates do not fit the well-rested and TSD condition subsamples. The decision model used by subjects following TSD is more consistent with the use of Bayes rule than the estimated well-rested decision model. This suggests that the heuristic typically used (i.e., more weight on evidence than on odds) is abandoned when one is sleep deprived. Instead, subjects may adopt a distinct heuristic that relies on equal (but reduced) weights on each parameter. Put differently, sleep deprivation may reduce a tendency to “overthink” the problem.

The outcome measure we analyze reveals no significant difference in final-choice accuracy comparing the well-rested and TSD data. This indicates that the distinct decision process we estimate following TSD results in maintained decision quality. There is, however, some evidence that decision model error terms have higher variance in the TSD subsample, which implies somewhat less consistent behavior under TSD.

Because information updating is a fundamental component of decision making under uncertainty, this unique examination of Bayesian updating following sleep deprivation is relevant to a large variety of behavioral applications. Our results are also important because they indicate that certain components of decision-making are necessarily impaired following TSD. The empirical data on increased accidents/errors

² Grether (1992) indicates that the representativeness heuristic is used when available, and it is available a high proportion of the time in his earlier (1980) design. When not available as often in the (1992) experimental design, overweighting of new evidence is not borne out as a more general result.
due to sleep deprivation may ultimately result from impaired functioning in areas other than one’s ability to process new information.

2. Background

An examination of Bayesian updating under sleep deprivation contributes significantly to both the literatures in economics and sleep. Sleep research indirectly points towards failed information assimilation under sleep deprivation (e.g., increased hesitance and reduced focus among sleep-deprived junior doctors in Goldman et al, 1972, and increased stereotyping of responses in Harrison and Horne 1997, 1998). However, direct evidence on decision making under uncertainty and information updating following TSD is needed, and Harrison and Horne (2000) recognize the lack of sleep deprivation research on specific decision models. We believe that research from economics can bring something to bear on this subject. Bayes rule is a fundamental decision model of belief revision and decision-making under uncertainty, and it has application to a variety of contexts. The relevance of this research to economists stems from our desire to understand decision-making behavior, and the evidence indicates that a good portion of decision-makers are, in their typical state, sleep-deprived to some degree. So, any identified effects of sleep deprivation on the decision making process or behavioral responses relevant to economic decision-making models highlight the importance of understanding an individual’s sleep-deprived state when conducting behavioral analysis. Such effects would also identify a previously uncontrolled potential confound in experimental data sets (for example, certain subject populations, such as students or those
employed as shift workers, may include relatively more sleep-deprived subjects than other populations).

Only a small amount of economics research has examined sleep. Biddle and Hammermesh (1990) incorporate labor productivity effects of sleep in a theoretical model of the allocation of time. Their empirical results from a variety of sources lead them to conclude that increased wages reduce sleep (more so for men than women), while increasing waking leisure time. In other words, additional waking hours generated by reducing one’s sleep need not imply additional hours in the workplace. This is consistent with the aggregate evidence on sleep reduction in many industrialized countries with rising wages, and it implies that sleep deprivation may be an inevitable byproduct of wage growth in a society.

Kamstra et al. (2000) examine the effects of daylight saving time changes on financial market returns. Interestingly, stock market returns drop both after losing an hour (Spring) and gaining an hour (Fall). This suggests that even a minor disruption of one’s internal (biological) circadian rhythm can affect behavior and decisions, even when sleep is not necessarily lost. Saunders (1993) finds that mood swings due to weather fluctuations can have a significant impact on stock prices. Because sleep deprivation has been found to worsen mood even more than it worsens cognitive or motor performance (Pilcher and Huffcutt, 1996), some of the effects of sleep deprivation in our economy may be difficult to measure. As a whole, sleep is largely unexplored by economists, and the limited studies in this area have yet to examine sleep-loss effects on specific decision models. This paper is a significant first step towards understanding the effects of sleep reduction on an important and fundamental decision-making process.
Sleep deprivation can be either partial or total, where total sleep-deprivation implies no sleep at all during a given day(s) (i.e., one or more 24 hour periods). Intuition might suggest that total sleep deprivation impairs functioning more than partial sleep deprivation. If this were true then one might not feel as concerned about the average partially sleep deprived adult—a college student studying all night for an exam would be the exception. However, existing research indicates that there are just as many reasons to be concerned about the effects of partial sleep deprivation. Van Dongen et al. (2003) found that chronic partial sleep deprivation of 4 or 6 hours per night for as few as six consecutive nights resulted in significant deficits on cognitive performance. In fact, the deficits were equivalent to those from up to two nights of total sleep deprivation experienced by a separate treatment group. In other words, chronic partial sleep deprivation can cause performance deficits on cognitive tasks equivalent to those from 1-2 nights of zero sleep. And yet, partially sleep deprived subjects did not (subjectively) report feeling as sleepy as TSD subjects. Pilcher and Huffcutt (1996) also find that the average partial sleep deprivation study included in their meta-analysis reported evidence of significant performance and mood effects, and they note that these partial sleep deprivation effects have perhaps been underestimated in some narrative reviews of the sleep literature.³

3) The Experiments

³ In this paper, we focus on the effects of TSD on information processing, which are relatively understudied compared to cognitive tasks of various sorts. There is another family of effects of sleep deprivation, which includes decreased glucose metabolism, increased risk of obesity, and decreased release of growth hormone, among others.
As noted, the experiments replicate the Grether (1980) design for a hand-run Bayes rule decision task. Two bingo cages are each filled with six colored balls: Cage A is filled with four green and two red balls, and Cage B is filled with three red and three green balls. Six draws, with replacement, are to be made from one of the cages. Each subject was informed of a ‘prior’ probability of using Cage A in terms of a die roll. For example, a 1/3 prior odds of Cage A was implemented by informing the subject that Cage A would be used if the die roll was 1-2 (3-6 implied use of Cage B). Subjects did not see the actual die roll, but its result tells the experimenter from which Cage to make the six draws from behind an opaque divider. The subject was shown each draw from the bingo cage, and after six draws was asked to indicate whether the balls were drawn from Cage A or B. A correct cage response resulted in payment of $12, whereas an incorrect response paid $2.

The procedure—choose the cage, draw a sample of six balls, subject indicates cage used—was repeated six times, and subjects were informed that only one of these times from each administration (i.e., well-rested and sleep deprived) would count for payment as determined by a random draw after completion of both administrations (so subjects did not know their accuracy or winnings until after all testing was complete). The design was balanced across prior A odds of 1/3, 1/2, and 2/3.\(^4\) Note that because compensation is higher when correctly indicating the cage, it is incentive compatible to indicate Cage A only if subjects perceive the (posterior) probability of Cage A to be greater than 50%.

If subjects use Bayes rule in their cage choice, they will form a posterior probability that Cage A is used from the particular sample of green/red balls drawn—

\(^4\) One implementation accidentally utilized one instance each of the prior odds of 1/6 and 5/6.
\[ P(\text{Cage A} \mid \# \text{green balls}) = \frac{P_A \cdot P(\# \text{green balls} \mid \text{Cage A})}{P_A \cdot P(\# \text{green balls} \mid \text{Cage A}) + P_B \cdot P(\# \text{green balls} \mid \text{Cage B})}, \]

where \( P_A \) is the prior odds of the cage \( i \) being used. That is, the new sample information is used to update the prior probability of Cage A. Table 1 shows the Bayesian updated posterior odds of Cage A in this Grether (1980) design. If subjects use a representativeness heuristic to make their choice of which cage is used, then a sample draw of three or four green balls out of six will induce a choice of Cage B or A, respectively, simply because the sample drawn looks like the population of one of the cages. One can see from Table 1 that the use of the representativeness heuristic can lead to an incorrect cage choice. For example, the posterior probabilities indicate that Cage A is more likely when \( P_A=2/3 \) and three green balls are drawn, but this sample looks like the Cage B population. Similarly, when \( P_A=1/3 \) and four green balls are drawn, Bayesian updating would lead one to indicate that Cage B was used—the posterior probability of Cage A is less than 1/2. The design is balanced so that, on average, the proportion of representative samples should not bias accuracy in favor of (or against) Bayesian choices.

A total of 24 subjects were administered the Bayes rules experiment as part of their participation in a total sleep deprivation study, which involved a stay of 6 consecutive nights and days in the Laboratory for Sleep and Chronobiology at the University of California-San Diego.\(^5\) Subjects were compensated several hundred dollars for the entire stay at the lab, but it was made clear to the subjects that these experiments afforded the opportunity to earn additional payments that were unrelated to their fixed compensation. Lab staff generally indicated that the subjects were more engaged in these

\(^5\) Though the sample size is small, multiple subject trials create a panel of \( N=144 \) well-rested and \( N=144 \) TSD observations. A small number of total subjects is quite common in sleep-deprivation studies, because of the screening criteria, the requirement that subjects stay in the sleep lab several days, and the total compensation per subject for a TSD experiment (often several hundred dollars per subject).
Bayes rule experiments than in other cognitive task experiments in which they participated during their lab stay, and so the extra compensation appeared salient to the subjects. Subjects were tested on various cognitive dimensions during their entire lab stay, with testing occurring approximately every two hours. This basic Bayes rule experiment was performed twice by each subject (so, they had the opportunity to earn $12 twice); once in a well-rested state, and once after 22-24 hours of total sleep deprivation. Each administration of the Bayes rule experiment lasted approximately thirty minutes.

Screening criteria for this study only allowed subjects who were right-handed, healthy, and considered ‘normal’ sleepers—those who had a consistent sleep-wake schedules that included 7-9 hours in bed each night. Subjects are indirectly monitored for one week prior to reporting to the sleep lab by keeping a sleep journal and wearing an actigraph.6 Because we motivated the relevance of this research by indicating how common it is to not be a normal sleeper, one may question the external validity of using only normal sleepers. As experimentalists, however, we face the usual trade-off of internal control versus external validity in conducting a sleep deprivation study. Only by using otherwise normal sleepers can we be confident that subjects were not sleep deprived when they entered the lab and engaged in the “well-rested” test administration. During the TSD treatment, subjects were not allowed any sleep, not allowed stimulants of any sort, and they were under constant supervision by lab staff to ensure no sleep during

---

6 The actigraph measures wrist movement as a proxy of gross motor activity. This movement, in turn, is used to determine sleep and wake. These data verify that subjects are engaged in normal sleep patterns prior to their lab stay and are not partially sleep deprived at the beginning of the experiment. The complete list of experimental inclusion/exclusion criteria is fairly standard for sleep deprivation research, and they are available on request.
this time. Figure 1 describes the basic timeline of the subjects’ lab stay relative to their participation in these decision experiments.

In a more recent paper, Grether (1992) notes that there are limits to what can be gleaned from the data using his simpler 1980 design. Because the design favors generating samples that are representative of Cage A or B, we are somewhat limited in our ability to generalize towards instances in which new information is not necessarily representative. On the other hand, we chose the more simple design in order to present subjects the most straightforward decision task that involved prior and new-sample information. As stated above, this design also provides an efficient evaluation of the use of a representative heuristic compared to a Bayes rule in subjects’ decision making. The dichotomous choice of Cage A or B does not allow us to infer strength of belief (i.e., 55% versus 95% certain that the balls came from Cage A), as does Grether (1992) in a modified design. However, given the known debilitating effects of sleep deprivation on vigilance, we felt this was a reasonable trade-off in design choice in order to be more assured that subjects understood the decision task, even after total sleep deprivation.

The particular placement of our Bayes rule task during the subjects’ lab stay implies that all subjects complete their second Bayes rule decision task in their sleep deprived state.7 As such, one might be concerned that subject learning may be generating some of the data. To explore this possibility, the Bayes rule experiment was also administered to an additional twelve control subjects (N=144 total observations). These

---

7 Due to an un-planned deviation from the sleep lab protocol for these experiments, one subject was administered the Bayes rule task under the TSD treatment first, in which case the coding of the TSD dummy variable distinguishes this one subject from the others. Ideally, the ordering of the TSD and well-rested administration would be counter-balanced but, as described above, the surrounding evidence does not indicate that subject learning is generating the TSD treatment effects due to our addition of control subject data.
control subjects performed the Bayes rule decision task twice, at the same 22-24 hour interval, but the control subjects were well-rested in both instances. Decision model estimates for these control subjects (see Appendix) find no significant difference in the weight placed on the evidence during the second Bayes rule experiment—contrary to the main finding in the TSD data.

In other words, we find no evidence that the differences in decision-making we report in the next section are due to subject learning across the two administrations of the experiment. In addition, subject learning would imply that choice accuracy should be higher the second Bayes rule experiment, but it is not. Or, learning might imply that a particular empirical model should better fit the data as choices converge to a particular set of model parameters—Grether (1980) finds this among experienced subjects, for example. Our results also show that this is not the case. We are therefore confident in attributing the second trial effects to the TSD treatment.

4) Results

Our subjects ranged from 18 and 39 years of age (μ=23.83, σ=5.37), and each gave voluntary consent to participate in the total sleep deprivation study. Because each Bayes rule experiment involves 6 trials of the choice task, the total number of observations generated by our 24 subjects is N=288 (N=144 in the well-rested state and N=144 in the sleep deprived state). The econometric estimations reported in this section account for the potential non-independence of decisions of a given subject across trials as a subject-specific random effect, but our results are robust to error-term specification.
Table 1 shows the Bayesian posterior probabilities of Cage A, which imply posterior odds of either Cage A or B being more likely. For example, the posterior probability of Cage A of .584 indicates a posterior odds of Cage A of approximately 1.40:1. Certain prior odds and sample draws imply a relatively easier choice for the subject in the sense that the posterior odds of the more likely cage are quite high (e.g., if \( P_A = 1/3 \) and only one green ball is drawn, the posterior odds of the more likely cage (Cage B in this case) are about 11:1. The bold cells in Table 1, for example, highlight the sample possibilities for Cage A that lead to the most difficult choices among all possibilities, but average choice difficulty is similar across treatments and all choices are analyzed below.\(^8\)

Table 2 shows the main econometric results. Here, following Grether (1980) for comparison, we estimate the following decision model:

\[
Y_{it}^* = \alpha + \beta_1 \ln LR(A)_t + \beta_2 \ln \left( \frac{P_A}{1-P_A} \right)_t + \mu_t + \epsilon_{it}
\]

where \( Y_{it}^* \) is the subject \( i \)'s subjective log odds in favor of Cage A in trial \( t \), \( LR(A)_t \) is the likelihood ratio for Cage A, and \( \left( \frac{P_A}{1-P_A} \right)_t \) is the prior odds ratio for Cage A. The dichotomous variable \( Y_{it} \) is observed equal to 1 if \( Y_{it}^* \geq 0 \), and so we estimate (1) using a random effects probit estimation. Grether (1980) estimates logit results for this model, and does not account for subject-specific random effects, and so our econometric specifications are similar but not identical. A Bayes rule hypothesis amounts to testing

\(^8\) Overall, the average difficulty of the entire well-rested and TSD subsamples were 3.3 and 3.1, respectively, in terms of posterior odds of the more likely cage. Thus, overall choice difficulty was quite similar across the well-rested and TSD subsamples, as one would predict given the balanced design.
jointly whether $\alpha=0$, $\beta_1=\beta_2>0$, while the representativeness heuristic would be supported if $\beta_1>\beta_2\geq 0$. In other words, a Bayesian subject will weight the evidence and the prior odds equally, while a subject who uses the representativeness heuristic would place more weight on the evidence than the prior odds of cage A. The basic findings of Grether (1980), who estimates a version of (1) as a logit model, support the representativeness heuristic hypothesis. That is, $\beta_1>\beta_2\geq 0$ for most of his subject groups, indicating that subjects overweight the evidence (i.e., the likelihood ratio) relative to the prior odds.

Table 2 shows random effects probit estimates of model (1) for our subjects, though our results are robust to estimation of a fixed effects model as well. A key result of this paper is found by examining whether or not the same model parameters $(\alpha$, $\beta_1$, and $\beta_2)$ apply to the well-rested and TSD data subsamples (i.e., a test for a structural break in the data resulting from the TSD treatment). Using the likelihood ratio test on the restricted model of pooled data and the unrestricted models of the separate TSD=1 and TSD=0 subsamples, we reject the null hypothesis that a single set of model parameters applies to both sets of data (the chi-squared statistic=13.36—significant at the p=.01 level for the test of three restrictions). Thus, the results indicate a structural change in the decision model following TSD, and so we next turn our focus to the model estimates for the separate well-rested and TSD subsamples. As noted earlier, results from additional control subjects do not support the hypothesis that the differences in the well-

---

9 For comparison to Grether’s (1980) logit estimations, we also perform a logit estimation of the model similar to (1) above, but without the random effects error-term specification. The pooled results that Grether reports for his financially motivated subjects yield the estimated model $Y_i = -.11 + 2.25 \ln LR(A)_{it} + 1.82 P_{Ait}(1-P_{Ait})$, where $\alpha$, $\beta_1$, and $\beta_2$ are statistically significant. In estimating the same logit model for our pooled data, the results are $Y_i = .04 + 2.26 \ln LR(A)_{it} + 1.95 P_{Ait}(1-P_{Ait})$, with $\beta_1$ and $\beta_2$ being statistically significant (p=.00). So, our results are quite comparable to those reported in Grether (1980), and logit estimations of any of the models in this section are consistent with the results we find in the probit estimations that we report. The results we find are also similar for a fixed effects specification (logit and/or fixed effects estimation results available from the authors on request).
rested and TSD data are due to subject learning. The supporting evidence from these control subjects is given in detail in the Appendix.

One difference that stands out in Table 2 is that well-rested subjects place more weight on the evidence than the prior odds. This difference is statistically significant using the chi-squared test for the restriction that $\beta_1 = \beta_2$ ($p=.06$). When subjects are well-rested, the estimated decision model replicates this key result from Grether (1980) using the same basic experimental design. Following TSD, however, there is no significant difference in the weight the subjects place on the prior odds versus the sample evidence ($p=.91$). TSD reduces the decision-weight placed on the evidence relative to the prior odds. Ironically, the decision model following TSD is consistent with the Bayes rule hypothesis, because TSD apparently eliminates the overweighting that well-rested subjects tend to place on the evidence. In all cases, the models do a reasonably good job of predicting the Cage A and Cage B choices of the subjects, correctly predicting their choice between 83% and 85% of the time.\textsuperscript{10}

In addition to the coefficient estimates, the estimated marginal effects are shown in Table 2 for interpretability. Consider the marginal effect on the log odds term $\ln\left(\frac{P_A}{1-P_A}\right)$. With our particular experimental parameterization, this term increases by about one when comparing $P_A=1/3$ to $P_A=2/3$. So, the marginal effects of .54 and .34 for the well-rested and TSD data, respectively, imply that this increase in prior odds makes

\textsuperscript{10} An alternative model that Grether (1980) estimates includes dummy terms for samples that are representative of either Cage A or B. Our key results appear to hold under this alternative empirical model, although the model failed to converge properly for the well-rested subsample of data. Nevertheless, relative to the pooled data, the TSD sample estimates for weight placed on the prior odds and the evidence are both less that those estimated for the pooled data, and significant in both cases. Some evidence for use of the representativeness heuristic is found more specifically in this alternative estimation, though it is only significant for the case when the sample looks like Cage B—subject are then significantly less likely to choose Cage A. That is, following TSD subjects still use the representativeness heuristic to some extent.
subjects 54 percentage points more likely to choose Cage A when well-rested, but only 34 percentage points more likely to choose Cage A when sleep-deprived. This difference between marginal effects on the log odds terms may not be statistically significant, however. Consider an alternative formulation for the pooled data set with a dummy variable for TSD=1, along with interaction terms

\[ Y_{it}^* = \alpha + \beta_1 \ln LR(A)_i + \beta_2 \ln \left( \frac{P_A}{1-P_A} \right)_i + \beta_3 \cdot TSD + \beta_4 (\ln LR(A)_i \cdot TSD) + \beta_5 \left( \ln \left( \frac{P_A}{1-P_A} \right)_i \cdot TSD \right) + \mu_i + \varepsilon_{it} \]

(2)

We estimate this random effects probit specification to allow a more direct parameter estimate comparisons. The results are:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \alpha )</th>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( \beta_3 )</th>
<th>( \beta_4 )</th>
<th>( \beta_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal effect</td>
<td>.07</td>
<td>.91</td>
<td>.58</td>
<td>-.09</td>
<td>-.56</td>
<td>-.24</td>
</tr>
<tr>
<td>p-value (two-tailed test)</td>
<td>.64</td>
<td>.00***</td>
<td>.00***</td>
<td>.61</td>
<td>.00***</td>
<td>.23</td>
</tr>
</tbody>
</table>

These estimates are consistent with Table 2 results in showing that the tendency to significantly overweight the evidence \((\beta_1 > \beta_2)\) is mitigated following TSD \((\beta_4 < 0)\). The coefficient on \( \beta_5 < 0 \) is in the direction indicating that TSD significantly reduces the weight one places on the prior odds, but the estimate is not statistically significant.\(^{11}\)

The marginal effect on the evidence, \( \ln LR(A) \), term in Table 2 represents the marginal change to the probability of choosing cage A for a one-unit change in the log-

\(^{11}\) For a similarly estimated logit model, \( \beta_3 \) significance is at p=.12. We also examine the relative difficulty of the different choices subjects would make, as proxied by the Bayesian posterior-odds of the more likely choice—higher odds represent an easier choice. As expected, we find that more difficult choices reduce the likelihood that subjects pick the correct cage. However, the TSD treatment does not significantly affect subject choice-accuracy, neither in general—noted earlier—nor for varying difficulty levels of choice, relative to when subjects are well-rested. These results are available from the authors on request.
likelihood ratio for Cage A. For this design, a sample of two green and four red balls, for example, generates a likelihood ratio of -1.05, while a sample of three green and three red balls generates $\ln LR(A) = -.353$, which is an increase in $\ln LR(A)$ of about .70. The estimated marginal effect for well-rested subjects implies that this change in $\ln LR(A)$ from drawing one additional green ball would make subjects 57 percentage points more likely to choose Cage A. For sleep-deprived subjects the comparative marginal effect is only about 25 percentage points. Of course, this does not take into account the fact that ‘representative’ samples may affect decisions independent of their effect on the likelihood ratio, but it is clear that these effects are behaviorally, as well as statistically, significant. The different sample draws in our experiment created a range of likelihood ratios from $\ln LRA(A) = -2.50$ for the case where six red balls were drawn, to $\ln LRA(A) = 1.73$ for the case where six green balls were drawn, though the extreme draws were rare.

These differences in the parameter estimates for the decision models when comparing subjects well-rested and following TSD are significant given that they indicate that TSD causes subjects to place a decreased decision-weight on both new evidence and on prior odds. The estimated effect is significant in the case of the likelihood ratio (i.e., the evidence), and the effect is robust to model specification (compare sub-sample estimates in Table 2 with estimates of model (2)). We also estimate that TSD reduces the decision-weight that subjects place on the prior odds, though the effect is not as large in magnitude and did not reach statistical significance.

It is intriguing, however, that the accuracy of the subjects’ choices is no worse following TSD than when well-rested, on average. For all N=144 observations of both
well-rested and TSD data, subjects indicated the correct cage 67-68% of the time. Our model estimates indicate that less weight is placed on both prior odds and new information following TSD. To the extent that weighting both sources of information increases accuracy, TSD should therefore reduce choice accuracy. However, TSD also eliminates the significant (non-Bayesian) overweighting of the new information, which should increase choice accuracy. Actual choice accuracy may be biased if the more likely Cage, based on Bayesian updated probabilities, is often not the actual Cage (by random chance). However, further examination of subject choices indicate that they coincide with the more likely Bayesian event 85% and 84% of the time when subjects are well-rested and following TSD, respectively. Thus, we find no evidence that TSD affects the quality of final choices in this environment. If TSD causes subjects to abandon a common decision heuristic (like paying too much attention to new information) and resort to more well-engrained habits, then this altered decision process is not necessarily “worse” if the abandoned heuristic is itself imperfect.

For our control subjects, we find that choice accuracy drops for the second administration of the task (76% to 61% accuracy). However, subject choices coincide with the more likely Bayesian event 85% and 86% of the time for first and second administration of the task, respectively. This suggests the drop in actual choice accuracy during on the second administration of the task for control subjects is an artifact of a disproportionate number of Cage outcomes that did not coincide with the statistical evidence. Thus, our evidence overall is consistent with recent research on sleep.

---

12 Choices and accuracy are not consistent with random decisions. In the well-rested subsample, the actual Cage A frequency is 54.2%, and subjects chose Cage A 52.8% of the time (actual accuracy was 68.1%). In the TSD subsample, Cage A frequency was 43.8%, and Cage A choice occurred 46.5% of the time (67.4% accuracy).
deprivation that has found that the underlying cognitive process may be quite different following TSD even though task performance is unaffected (Drummond et al., 2000). The different parameter estimates of our decision model are likely an important first clue to the type of cognitive process change that results from TSD.

An examination of the residuals from estimating (1) indicate that the TSD sample yields somewhat higher-variance residuals, though the difference is not statistically significant (two-sample F-test for variance, \( p=.20 \)). This may suggest that choices following TSD are not as convergent upon the decision model in (1) as when well-rested. Though our residuals-variance result is statistically insignificant, it is similar to Grether’s (1980) finding of less consistent behavior for inexperienced subjects. Our lack of significance may be due to our limited sample size, but the result is consistent with results in the sleep literature that indicate increased variability and statistical variance under TSD.

It is also worth noting that our result of unaffected choice accuracy in the Bayes rule experiment following sleep deprivation only implies that subjects are equally accurate in assessing the likelihood of being in state A versus state B. This does not imply that a TSD subject is as adept at dealing with any further ramifications of being in one state versus the other. This latter consideration will also be a function of TSD effects on other factors like vigilance, reaction time, and motor skills.

5. Discussion

The topic of sleep deprivation is virtually unexplored in research on economic decision models. Because of the evidence indicating that, as a society, we are more sleep
deprived at present than in any previous generation, the implications this has on decision-making under uncertainty across many environments are worth exploring. Not only is the impact of sleep-deprivation significant to an economy (e.g., lost worker productivity), but any adverse effects of sleep-deprivation take on increased significance in certain susceptible labor markets when one considers the public health/safety ramifications (e.g., medical residency, long-haul truck driving or piloting, the military). Because recent sleep research indicates that performance of selected tasks may be just as affected under chronic partial sleep deprivation as under total sleep deprivation (Van Dongen et al., 2003), the effects of sleep deprivation on decision-making are not likely to be limited to only the short-term totally sleep deprived individual.

This paper examines the effects of sleep deprivation on a particular type of decision-making that is of interest to decision scientists, in general, and is unexplored by sleep researchers. We administer a Bayes rule decision experiment to subjects in experimentally controlled well-rested and sleep-deprived states. Because the general population does not exactly fit either of these experimentally induced states, the results can be viewed as indicative of the decision processes of a given individual when approaching either the well-rested or TSD state. This decision experiment provides a fundamental look at how subjects process and filter information in uncertain choice environments. That is, a Bayesian subject is assumed to update a prior belief with new information on a situation in order to form a posterior belief of event occurrence. So, the experiment examines a basic decision model that may serve as a building block for many more complicated decision environments of interest to economists, among others.
We find that, following sleep deprivation, subjects no longer overweight new information in forming subjective probability estimates.\textsuperscript{13} Ironically, when subjects no longer employ the usual heuristic used to make decisions (i.e., focus more on the evidence), their behavior becomes more consistent with Bayes rule than when those same subjects are well-rested. There is also some indication that, following TSD, subjects reduce the decision weight placed on prior odds (i.e., prior information), although this is somewhat less conclusive. In terms of the experimental outcome measure, we find that choice accuracy is statistically equivalent when well-rested versus sleep-deprived.

For the experimental economist, these results indicate that there may be an important confound in laboratory data for certain types of experiments. For example, neuro-economists who compete with other scientists for the use of scanning equipment may conduct experiments at abnormally late evening hours when subjects would be especially tired. Another example is if experiments are conducted during exam week, when student subjects might be functioning on less sleep than normal.\textsuperscript{14} Sleep deprivation may be an unidentified confound in the behavioral (and neural) data generated in such circumstances.

The result is also significant in today’s modern sleep-deprived society. Though there is ample documentation of the detrimental effects of TSD on cognitive and motor skills, and certain decision tasks, the present results (showing intact decision performance through an adapted decision model) suggest that not all components of decision-making are necessarily harmed by TSD. This is an important finding because it suggests that

\textsuperscript{13} Grether (1980) finds this overweighting of the evidence among a typical sample of student subjects in his design that we replicate.

\textsuperscript{14} If students, on average, are more sleep deprived than the general population as survey data suggests, then the data from any experiments using student subjects will contain sleep deprivation related confounds.
increased accident or error rates attributed to sleep deprivation result more from auxiliary function impairment (e.g., reduced vigilance, reaction times, or short-term memory) and/or other aspects of decision making than the ability to integrate new information.

Our experiment involves an unavoidable risky decision environment. Further research is needed to examine potential changes in risk preference during sleep deprivation. For example, if an individual is less likely to avoid a risky decision environment, when the opportunity to sort oneself out of the decision exists, then sleep deprivation may lead individuals to choose more risky decision environments, on average. Though we find error rates to be unaffected by TSD, the cost of each error may be higher in a riskier scenario. This has interesting implications for, among others, military personnel choosing to engage or not engage in a riskier outcome scenario, or a physician choosing between two courses of surgical action.

Because we find that subject decision accuracy in the Bayes rule experiment is unaffected by TSD, the finding of significant differences in estimated decision models for subjects whether well-rested or TSD merits further exploration. Even when sleep deprivation might not affect some behavioral outcome measures, there is still much to understand about how underlying decision processes might be altered. Drummond et al. (2000) is an intriguing study that shows how versatile the brain can be under adversity. In their study, recognition memory on a verbal learning task showed no significant change as a result of a TSD treatment, though the pattern of brain activity was different following sleep deprivation. Specifically, the subjects’ parietal lobes, especially in the left hemisphere, came ‘on-line’ after total sleep deprivation. Because the parietal lobes are related to performance, their activation after TSD compensated for any decreased
performance resulting from deficits in other brain regions. Others have reported similar increases in brain activation and resultant intact performance during TSD on a variety of tasks (Drummond et al., 2001, 2004, 2005; Portas et al., 1998; Chee and Choo, 2004), and Stricker et al. (2006) have reported changes in the neural networks that perform a given task after sleep deprivation. Hsu et al., (2005) examine decision-making under uncertainty in a neuroeconomics experiment, and they suggest a multi-regional neural system for evaluating uncertainty.

Though we examine behavioral outcomes in this paper, the evidence we find in support of distinct decision-weights across sleep-states may be a clue indicating neural activation differences in information-updating environments. For example, the ventrolateral prefrontal cortex has been implicated in the neural process of integrating new contingencies (Paulus, et al, 2004). Our finding that subjects decrease the decision weight placed on new evidence following TSD might indicate decreased activation of the ventrolateral prefrontal cortex. Studies of the Bayes rule decision task we use are absent in the neuroimaging literature, but research on other tasks suggests that compensatory activation may occur in the parietal lobes following TSD, thus allowing performance to maintain intact (e.g., Drummond et al., 2000, 2001, 2005). Together, these results indicate that our understanding of decision-making following sleep deprivation is incomplete at best, and more exploration is needed even in cases where individuals apparently retain functional ability.

Our results are an important step towards furthering our understanding of how the brain processes information and how it may react to adversity. Some emerging neuroeconomics research suggests that certain brain regions form a neural system for
evaluating uncertain decision environments (Hsu et al, 2005). Other neuroimaging
studies support the hypothesis of compensatory brain activation (e.g., Drummond et al.,
altered decision-model following TSD may be an initial indication of compensatory
neural activity following TSD that we intend to explore further.
Table 1: Posterior probabilities of Cage A

<table>
<thead>
<tr>
<th>Prior probability of Cage A</th>
<th>Number of Green Balls Drawn</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2/3</td>
<td>.149</td>
<td>.260</td>
<td><strong>.413</strong></td>
<td><strong>.584</strong></td>
<td>.737</td>
<td>.849</td>
</tr>
<tr>
<td>1/2</td>
<td>.081</td>
<td>.149</td>
<td>.260</td>
<td><strong>.413</strong></td>
<td><strong>.584</strong></td>
<td>.737</td>
</tr>
<tr>
<td>1/3</td>
<td>.042</td>
<td>.081</td>
<td>.149</td>
<td>.260</td>
<td><strong>.413</strong></td>
<td><strong>.584</strong></td>
</tr>
</tbody>
</table>

Table replicated from Grether (1980) Table 1. Bold cells represent “difficult choices” of approximately equal posterior odds of the more likely Cage (i.e., choices of approximately equal difficulty for subjects).

Table 2: Probit estimates of $Y^*_h$ model

(random effects specification. p-values given in parenthesis)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pooled (N=288)</th>
<th>Well-rested (N=144)</th>
<th>Sleep-deprived (N=144)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff. marg. effect</td>
<td>Coeff. marg. effect</td>
<td>Coeff. marg. effect</td>
</tr>
<tr>
<td>Constant</td>
<td>.03 (.83) .01 (.83)</td>
<td>.13 (.42) .05 (.42)</td>
<td>-.08 (.68) -.03 (.68)</td>
</tr>
<tr>
<td>lnLR(A)</td>
<td>1.27 (.00)*** .48 (.00)***</td>
<td>2.20 (.00)*** .81 (.00)***</td>
<td>1.01 (.00)*** .36 (.00)***</td>
</tr>
<tr>
<td>ln(P_A/(1-P_A))</td>
<td>1.10 (.00)*** .42 (.00)***</td>
<td>1.46 (.00)*** .54 (.00)***</td>
<td>.97 (.00)*** .34 (.00)***</td>
</tr>
<tr>
<td>% correctly predicted</td>
<td>84.38%</td>
<td>85.42%</td>
<td>83.33%</td>
</tr>
</tbody>
</table>
FIGURE 1
A week in the sleep lab: time-line

Note: Some subjects stayed in the lab one less day and participated in a one-night TSD study. Our examination of TSD effects after one night of TSD allows us to combined subjects from different sleep studies, whether or not they participated in a one or two night TSD lab stay.
REFERENCES


Appendix: Control Subject Data

The experimental protocol was administered to an additional twelve subjects, who were well-rested for both the first and second administration of the Bayes rule experiment—well-rested was verified using similar measures as for the sleep deprivation subjects, and average subject age was similar to non-control subjects ($\mu=24.12$ years old, $\sigma=4.278$). The results of estimation equation (1) from the text for the sample of $N=144$ Bayes rule decisions are shown in Table A1 below.

Table A1: Probit estimates of $Y^*$ model for
CONTROL SUBJECTS
(random effects specification. p-values given in parenthesis)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pooled (N=144)</th>
<th>marg. effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.02 (.88)</td>
<td>.01 (.88)</td>
</tr>
<tr>
<td>$\ln \text{LR}(A)$</td>
<td>1.62 (.00)***</td>
<td>.63 (.00)**</td>
</tr>
<tr>
<td>$\ln \left( \frac{P_A}{1-P_A} \right)$</td>
<td>1.47 (.00)***</td>
<td>.58 (.00)***</td>
</tr>
</tbody>
</table>

% correctly predicted 85.41%

As can be seen, results are similar to those from the main data set, except that the estimated weights on evidence and prior odds are somewhat higher. Estimation up to an unknown scale parameter, however, prohibits a direct comparison across models. A test for structural change in the data across the first- and second-administration fails to reject the null hypothesis that the same parameter estimates apply to both the subsamples of first- and second-administration of the experiment (the Likelihood Ratio statistic is 1.772 compared with the 90% critical value of 6.25 for the $X^2$ statistic for $n=3$ restrictions). This contrasts with the results from the well-rested and TSD subsamples of the main data. Control subjects who are not sleep-deprived for the second administration of the task fail to display a significant difference in the estimated decision model across the two administrations of the task. In other words, the differences found in the main data appear to be a result of the sleep deprivation treatment as opposed to learning from first to second administration of the experiment.

These results are robust to alternative examinations of the control subject data as well. For example, we might also estimate a model similar to (2) in the text. That is, the pooled control-subject data is analyzed with dummy variables for second-administration of the experiment, with interaction terms that allow for the second-administration effects to potentially differ with respect to decision weights on evidence and prior odds. Thus, we estimate
\[ Y_{it}^* = \alpha + \beta_1 \ln LR(A) + \beta_2 \ln \left( \frac{P_A}{1 - P_A} \right)_t + \beta_3 \times 2\text{ndAdmin}_{it} + \]

\[ \beta_4 (\ln LR(A) \times 2\text{ndAdmin}_{it}) + \beta_5 \ln \left( \frac{P_A}{1 - P_A} \right)_t \times 2\text{ndAdmin}_{it} + \mu_t + \epsilon_{it} \]

(1A)

The results are

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \alpha )</th>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( \beta_3 )</th>
<th>( \beta_4 )</th>
<th>( \beta_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal effect</td>
<td>-.004</td>
<td>.55</td>
<td>.60</td>
<td>-.30</td>
<td>.27</td>
<td>.01</td>
</tr>
<tr>
<td>p-value (two-tailed test)</td>
<td>.96</td>
<td>.00***</td>
<td>.00***</td>
<td>.30</td>
<td>.22</td>
<td>.97</td>
</tr>
</tbody>
</table>

The only significant variables are the prior odds and the evidence (we fail to reject the null hypothesis that \( \beta_1 = \beta_2, p=.75 \)). The second administration of the Bayes rule task does not significantly affect the likelihood of choosing Cage A. The results are also unchanged if one considers a model of the entire pooled data set (main data and control subject data), with dummy variables for TSD, 2nd administration of the task for control subjects, and interaction terms. The only significant variables remain the \( \ln LR(A) \), \( \ln \left( \frac{P_A}{1 - P_A} \right)_t \), and the interaction of TSD and \( \ln LR(A) \).

In short, the data indicate that there are no significant differences in the control subjects’ decision model from one day to the next. This evidence is in support of our conclusions that the sleep deprivation treatment, not learning, is generating the behavioral differences we estimate in the main text. The data are also consistent with the hypothesis that there is compensatory effort engaged following sleep deprivation that helps maintain choice accuracy, though the mechanism involved cannot be fully explored in the current data.

In short, the data indicate that there are no significant differences in the control subjects’ decision model from one day to the next. This evidence is in support of our conclusions that the sleep deprivation treatment, not learning, is generating the behavioral differences we estimate in the main text. The data are also consistent with the hypothesis that there is compensatory effort engaged following sleep deprivation that helps maintain choice accuracy, though the mechanism involved cannot be fully explored in the current data.