Tax Compliance and Psychic Costs: Behavioral Experimental Evidence Using a Physiological Marker

Working Paper No. 2012-19

CREMA Gellertstrasse 18 CH - 4052 Basel www.crema-research.ch

Tax Compliance and Psychic Costs: Behavioral Experimental Evidence Using a Physiological Marker

Uwe Dulleck[1], Jonas Fooken[1], Cameron Newton[2], Andrea Ristl[3], Markus Schaffner[1], and Benno Torgler[1,4,5*]

Abstract

Although paying taxes is a key element in a well-functioning civilized society, the understanding of why people pay taxes is still limited. What current evidence shows is that, given relatively low audit probabilities and penalties in case of tax evasion, compliance levels are higher than would be predicted by traditional economics-of-crime models. Models emphasizing that taxpayers make strategic, financially motivated compliance decisions, seemingly assume an overly restrictive view of human nature. Law abidance may be more accurately explained by social norms, a concept that has gained growing importance as a facet in better understanding the tax compliance puzzle. This study analyzes the relation between psychic cost arising from breaking social norms and tax compliance using a heart rate variability (HRV) measure that captures the psychobiological or neural equivalents of psychic costs (e.g., feelings of guilt or shame) that may arise from the contemplation of real or imagined actions and produce immediate consequential physiologic discomfort. Specifically, this nonintrusive HRV measurement method obtains information on activity in two branches of the autonomous nervous system (ANS), the excitatory sympathetic nervous system and the inhibitory parasympathetic system. Using time-frequency analysis of the (interpolated) heart rate signal, it identifies the level of activity (power) at different velocities of change (frequencies), whose LF (low frequency) to HF (high frequency band) ratio can be used as an index of sympathovagal balance or psychic stress. Our results, based on a large set of observations in a laboratory setting, provide empirical evidence of a positive correlation between psychic stress and tax compliance and thus underscore the importance of moral sentiment in the tax compliance context.

JEL Classification: H26, H41, K42, D31, D63, C91
Keywords: tax compliance, psychic costs, stress, tax morale, cooperation, heart rate variability, biomarkers, experiment

---

[1] Queensland Behavioural Economics Group (QuBE), School of Economics and Finance, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4001, Australia. [2] School of Management, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4001, Australia. [3] AUTONOM TALENT® Consulting GmbH, Mariahilfer Straße 54/15, 1070 Vienna, Austria. [4] CREMA – Center for Research in Economics, Management and the Arts, Gellertstrasse 18, CH-4052 Basel, Switzerland. [5] EBS Universität für Wirtschaft und Recht, EBS Business School, Rheingaustraße 1, 65375 Oestrich-Winkel, Germany. This work was supported by the Australian Research Council (ARC), Linkage Grant LP0884074 and Future Fellowship FT110100463. We thank Brian Erard, Jonathan Feinstein, Richard Jefferson, and the participants of several seminars (University of New South Wales, Deakin University, University of Technology, Sydney, and Uppsala University) and a conference in Muenster for very helpful comments.

[*] Correspondence should be addressed to: Benno Torgler (benno.torgler@qut.edu.au).

I. INTRODUCTION

Over the entrance to the U.S. Internal Revenue Service (IRS) the following words from Oliver Wendell Holmes are inscribed: “Taxes are what we pay for a civilized society.” In other words, taxes are the fuel that makes civilizations run and few civilizations have failed to impose them (Adams 1993). However, despite the crucial importance of taxation in citizens’ lives, understanding of why people pay taxes remains limited (see, e.g., Slemrod 1992, Torgler 2007, Alm, Martinez-Vazquez, and Torgler 2010, Konrad and Qari 2012). One possibility for encouraging tax payment is the use of a deterrence policy, a major component of the traditional economics-of-crime approach (Becker 1968, Allingham and Sandmo 1972). Nevertheless, this approach, although it has produced many useful insights, seems too narrow as a framework for fully understanding tax compliance because it assumes a rational individual who maximizes the expected utility of the tax evasion gamble by weighing the benefits of successful cheating against the prospect of detection and punishment. Qualitatively, this theory suggests that tax evasion is negatively correlated with the probability of detection and the degree of punishment. However, not only is the level of deterrence too low in many countries to explain the high degree of tax compliance observed, but even in the least compliant countries, a purely on this type of standard economic theory based approach can only explain the observed levels of tax evasion in terms of unrealistic preferences (Alm et al. 2010). Thus, the tax compliance decision must be affected in ways not fully captured by the basic economics-of-crime approach. The inability of such a model to explain compliance behavior is probably related to a variety of factors, including the normative social influence and social and cognitive dissonance widely theorized in the psychological literature and conceptualized on both social and individual levels (Asch 1952, Festinger 1957). Thus, many researchers (Graetz and Wilde 1985, Alm, McClelland and Schulze 1992) have argued that addressing social norms may help resolve the high tax compliance puzzle because such norms act as an alternative to law enforcement in channeling individual behavior. Specifically, the violation of social norms has consequences like internal sanctioning (e.g., guilt, remorse) or, in the case of detection, external legal and social sanctions such as prosecution, gossip or ostracism (Polinsky and Shavell 2000). From this perspective, the obligation to pay taxes is an accepted social norm that enables governments to impose and collect them. Internal sanctions, in particular, can be captured by level of psychic stress (e.g., guilt, shame or the related sentiments that can be defined as moral), although these experiences may vary across taxpayers according to the social norms internalized. Theoretically, it has been stressed that the *intention* of a false declaration generates anxiety, guilt or a reduction in the taxpayer's self-image or self-esteem, which in turn reduces the utility of tax evasion and thus the likelihood of tax evasion (Erard and Feinstein 1994). Hence, a dissonance may arise from the mere contemplation of real or imagined actions (e.g., choosing not to pay taxes) that may lead to the individual's experiencing psychic stress (anticipated uneasiness at breaking social norms) and its immediate consequence of physiological discomfort (Lewis 1982). This study therefore takes an innovative approach by measuring physiological indicators of stress in individuals contemplating such decisions.

As a marker for such psychological strain, we use measurements of heart rate variability (HRV) in an experiment that requires participants to make compliance choices. This approach is novel not only because evidence at the *behavioral* level is rare, but because this particular theoretical framework has been little empirically tested in the tax compliance context. Its use is, however, supported by survey evidence that the anticipated psychic stress associated with committing tax evasion serves as a much stronger compliance enforcement mechanism than the perceived threat of legal sanctions (Grasmick and Bursik 1990). Measuring heart rate variability has the strength of capturing psychobiological or neural equivalents of psychic strain (e.g., feelings of guilt or shame) that arise from the contemplation of real or imagined actions (Green and Paxton 2009). HRV thus allows exploration of immediate consequential physiologic reactions captured by changes in the heart rate – an established correlate to psychic strain (Dishman et al. 2000).

Nevertheless, not only has past research with such a focus been criticized for failure to measure psychic stress reliably (Andreoni, Erard, and Feinstein 1998), but the measurement of tax evasion and tax compliance in particular is problematic because obtaining honest information about tax compliance behavior is difficult. Even when data on tax evaders are available, such behavior could have been affected by specific circumstances that are difficult to control. Therefore, to circumvent the problem of acquiring honest answers on illegal behavior, we use a laboratory experimental approach. Such experimentation, however, has itself been criticized for lacking realism: choices in the artificial laboratory setting may not accurately reflect choices in "the outside world." We therefore try to increase the external validity of our method by expressing relevant factors like audit probability, penalty rate, and tax rate as real values. We also implement financial incentives (performance-dependent subject payments) in the hope of adding realism and increasing participants' motivation to act realistically (Alm, Jackson, and McKee 1992).

In addition, our investigation into the possible impact of psychic stress on tax compliance employs neuroscientific tools that have the potential to generate a better biological micro-foundation for compliance behavior. Most particularly, a focus on brain systems, heart rate, skin resistance, genes, neurons and neurotransmitters can reveal otherwise unobservable aspects of the individual decision-making process. For example, functional magnetic resonance imaging (fMRI) evidence suggests that mandatory tax-like transfers to charity elicit neural activities in areas linked to reward processing (Harbaugh, Mayr, and Burghart 2007). Thus, it is not surprising that there is an increasing interest among economists to integrate biological approaches using various techniques. Non-invasive procedures are in particular attractive to understand the biological bases of human decisions in a lab experimental setting (Dulleck et al. 2011). Coricelli et al. (2010) were the first who explored the link between physiological measures and tax evasion focusing on skin conductance responses (SCR). They show that SCR is correlated with self-reported emotional arousal and hedonic valence. They explored 48 subjects 30 times finding that an increased SCR measured before the decision is made (higher anticipated and anticipatory emotional arousal) is linked with a higher likelihood to evade and more tax evasion. Such a result is not consistent with the idea that the intention of tax evasion induces feelings such as guilt that reduces the utility of tax evasion and therefore the likelihood or the degree of tax evasion. However, they observe that a higher fine is related with negative feelings stressing that this could be due to regret and anger. In addition, learning that their picture is going to be disseminated is also correlated with more negative feelings which could be driven by shame. Not getting audited on the other hand generates positive feelings possibly due to a relief and joy of getting higher earnings.

In this study, we use the heart rate variability (HRV) measure to better understand the tax compliance phenomenon. HRV technology, because it is nonintrusive and requires only compact equipment, allows experimenters in line with SCR to design more complex environments (i.e., a large variety of settings and interpersonal contexts) and observe a larger group of participants/interactions in the laboratory than allowed by alternatives like fMRI. Thus, in our laboratory experiment, we were able to observe a substantially larger sample size than is possible in fMRI studies. Recently, HRV measures have been used to understand responses to social interactions. Falk et al. (2011) find evaluating data of 70 subjects that an unfair pay (actual share and discrepancy between and actual and appropriate share) is correlated with lower heart rate variability (higher stress) and therefore with adverse effects at the physiological level. Lange et al. (2011) find that introducing an experimental protocol with measuring heart rate promotes behavioral trust in trust games and reciprocal giving in the trustee possibly due to interpersonal (touch and communicating care) and intrapersonal mechanisms (arousal and self-awareness). On the other hand, Brandts and Garofalo (2012) were not able to find that HRV has a statistically significant effect on human performance (resolution of a task). They explored how people react to an audience observing that male contrary to female participants are strongly influenced by the gender of the audience. Men’s blood pressure as opposed to the HRV also reacted to the gender pairing in the decision task. Thus, these discussed studies suggest to explore in detail when and how the body reacts to systematically to social interactions and an array of problems and issues associated with understanding human nature and governing and managing our society.

II. HRV ANALYSIS

In general, HRV analysis is used to identify certain medical predispositions (Malik et al. 1996), as well as psychological, emotional and mental activities (Koelsch et al. 2007, Yang et al. 2007, Appelhans and Luecken 2006, Crone et al. 2004, 2005, McCraty et al. 1995). Specifically, the HRV measurement method obtains information on activity in two major parts of the ANS, the sympathetic and parasympathetic systems. The first (the fight-or-flight response) affects the heart rate indirectly through the sympathetic nerves and by releasing cell-stimulating hormones (mainly adrenaline) into the blood stream (neurotransmission of norepinephrine). The second (which is responsible for rest and relaxation) influences the heart directly through the vagal nerve's connection to specific "pacemaker" cells (vagal stimulation releases acetylcholine, which increases the pacemaker cells' K^+ conductance and regulates the heart rate) (Appelhans and Luecken 2006, Levy and Martin 1979). The oscillations in heart rate generated by these two autonomic system branches occur at different speeds or frequencies (Appelhans and Luecken 2006): those induced by the sympathetic system are of considerably longer duration (maximum effect after more than 5 seconds) than those induced by the parasympathetic system (maximum effect after less than 5 seconds) (Levy et al. 1970). This timing difference can be used to identify the extent of sympathetic and parasympathetic activity. Most particularly, in experimental environments that control (avoidance of) physical activity, eating and drinking, measuring HRV can generate indicators of the participants' psychological state (Berntson and Cacioppo 2008, Seong et al. 2004) by correlating the changes in balance between the sympathetic and parasympathetic systems and mental (rather than motor) activity in the anterior insular, dorsolateral prefrontal and anterior cingulate cortices. In fact, research has shown that individuals react to mental stress with either increased sympathetic and/or decreased parasympathetic activity (Berntson et al. 1994).

Thus, purpose of using heart rate variability (HRV), in the context of behavior under psychic stress, is to identify whether the sympathetic or parasympathetic system is more active during a given time interval. Not only does such activity reflect the physiological and mental processes in the body, but HRV measures are specific correlates of activity in particular brain regions. The use of HRV measures builds on the observation that heart beats are not independent events but a realization of a continuous charging and releasing of electrical potentials, referred to here as the heart rate signal. This signal is based on electrocardiogram (ECG) recording of heart beat strength (essentially the length of the so-called QRS complex) (*Figure A1* in the Appendix). To arrive at a continuous measure from consecutive QRS data points, the data is cubically interpolated to a 5 Hz signal (tachogram). Usually, at this stage the signal must also be adjusted for noise in the recording and misreading in the ECG. From this signal, various measures of HRV can be built, to describe how the heart rate changes or varies over different time intervals. These intervals can be long (i.e., to explore daytime variation) or very short (i.e., covering only minutes or even seconds).

The estimation of these PSDs is, however, not trivial and is dependent on the estimation technique used. The three most common time-frequency decomposition methods are (i) the short-term fast Fourier transformation (SFFT), (ii) the autoregressive spectral estimation (AR) and (iii) the smoothed pseudo Wigner-Ville distribution (SPWVD, a wavelet transformation). The Appendix provides a brief description of these techniques. Each has different properties with respect to accuracy and smoothness. All approaches must deal with two issues: (i) the sympathetic and vagal system are active alongside several other prominent systems, including those that regulate respiration, temperature and blood pressure (Hainsworth 2008) and (ii) there is a large degree of heterogeneity in heart rates and individual heart rate reactions. One solution is to use time-frequency analysis of the (interpolated) heart rate signal to identify the level of activity (power) given different velocities of change (frequencies). The most commonly used method for obtaining this information for the frequencies used is the smoothed pseudo Wigner-Ville distribution (SPWVD) wavelet transformation (Seong et al. 2004, Bianchi et al. 1993, Wiklund, Akay, and Niklasson 1997), a technique that until now has been applied primarily in medical research. Such prior application, however, has clearly shown that activity in the sympathetic system is reflected mainly by high spectral power in the low frequency band (LF [0.033 - 0.15 Hz]) and that in the parasympathetic system (and respiration) by high spectral power in the high frequency band (HF [0.15 - 0.4 Hz]) (Malik 2008). Moreover, despite some controversy over whether LF HRV is contaminated by parasympathetic influence, in general, the ratio of activity in the low frequency band to that in the high frequency band (i.e., the LF/HF ratio) can be used as an index of sympathovagal balance (Appelhans and Luecken 2006), which serves as a useful indicator of psychic stress. Daytime variation and long- to mid-term changes in the heart rate signal are of minor interest. To obtain accurate estimates, any slowly varying processes (waves with a length of more than 27 seconds [0.033 Hz]) must be eliminated (detrended) using a standard DWT wavelet filter (Wiklund et al. 1997).

III. EXPERIMENTAL SETTING

Our computer-based laboratory experiment was designed to replicate the structure of a voluntary income reporting tax event (see, e.g., Alm, McClelland, Schulze 1992, Alm, Jackson and McKee 1992, 1993), repeated over several rounds, in which participants received income, had to declare all of it to the tax authority and then had to pay taxes on the declared income while the after-tax amount went into wealth generation. However, they also faced the possibility of an audit that would detect and penalize evasion. Because the percentage of individual income tax declarations subjected to a thorough tax audit is quite small in most countries and the penalty for fraudulent evasion seldom exceeds more than the amount of unpaid taxes (Alm et al. 2010), the tax administration and the tax system were simulated in the experiment by defining (a) a tax rate of 30%, (b) an audit probability of 12.5% and (c) a tax penalty on tax evasion equal to the unpaid tax. No participant paid taxes on undeclared income, but each participant was faced with the possibility that any evasion would be detected and penalized. During the entire experiment, participant ECGs were recorded using a Holter Digital ECG recorder (see *Figure A2*), and from these measurements, an average LF/HF ratio was calculated for each income declaration period (from the moment the income declaration screen appeared to the moment in which the participants clicked the OK button after declaring their income[2]).

The experiment was conducted in April 2009 over a period of three weeks at the Queensland University of Technology, Brisbane, Australia. The participants were informed that the experiment would include the measurement of their heart rate and that they would not be allowed to eat or drink anything (except water) 90 minutes prior to the experiment. The experiment was reviewed by the Queensland University of Technology Faculty Research Ethics Advisory Board and confirmed to meet the requirements of the National Statement on Ethical Conduct in Human Research. The experiment lasted approximately 90 minutes. The participants were volunteers recruited primarily from first-year economics units (using a faculty-wide invitation email and an in-lecture advertisement). Their ages ranged between 17 and 51 years (M = 23) and 56% were female. All sessions were held in the afternoon to minimize the effect of daytime variation in heart rate. None of the participants had previous experience with such an experiment. Within each session, participants were randomly divided into groups of four (45 groups in total). They were welcomed to the

---

[2] See *Figure A4* for the screen.

experiment and received instructions and assistance on how to apply the heart rate monitors. Three single-use electrodes were placed on the participant's chest and connected to the heart rate monitor.

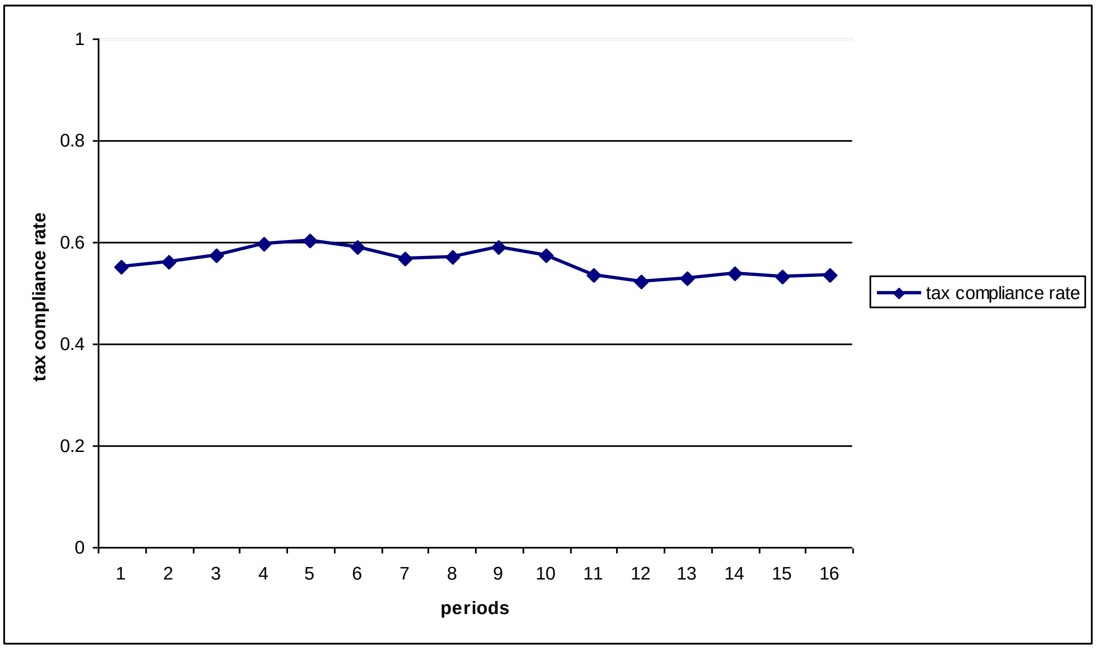
The experiment was programmed and conducted using z-Tree (Fischbacher 2007). On sitting down at the computer, participants received on-screen instructions for the tax evasion game (see *Figure A3*), which were also read aloud by a native English speaker. Communication was forbidden during the experiment. Before the experiment began, three test rounds were completed to familiarize participants with the experimental setup. During the experiment, participants were asked to report the income they wanted to declare (see *Figure A4*).

The experiment was conducted over 14 sessions that each involved 16 rounds of reporting. The 180 participants were divided into 45 groups of 4 participants, none of whom knew in advance when the experiment would end. *Figure 1* shows the mean tax compliance rate over all participants in a given round. Tax compliance over time is substantial, and the compliance rate is relatively stable throughout the 16 periods of the experiment. This stability implies that our experiment is unaffected by the previously reported empirical finding that repetition in public good experiments results in decreased cooperation, an observation that has intensified discussion on whether confusion and inexperience are driving forces for such results (Houser and Kurzban 2002, Ledyard 1995).

At the beginning of each round, participants received income and had to choose an amount to declare, paying taxes on the declared income but reserving the remainder for wealth generation. Based on their accumulated wealth during the experiment, each earned between 10 and 29 AUD (M = 20 AUD) and only the individual knew his or her actual income. Prior to the tax declaration event, participants in a group were divided equally into two income categories: exogenous or endogenous income distribution. In experiments using endogenous income distribution (84% of the cases), income distribution was based on individual performance on a cognitive skill test (50 questions, 12 minutes, see *Figure A5*).

In treatments that included public good (40 groups), this variable was introduced in a manner comparable to the literature on voluntary provision of public goods. The taxes of the individuals in a group were paid into a group fund (with and without the participants' own contribution), which was then redistributed to group members (in one treatment by a multiple of 2 before redistribution) in equal shares (see "transfer income" in the declaration screen in *Figure A4*). Amounts collected via audits were not added to the group fund. In those treatments that did not include public good, participants received nothing for their tax payments.

Figure 1: Mean Tax Compliance Rate Over Time

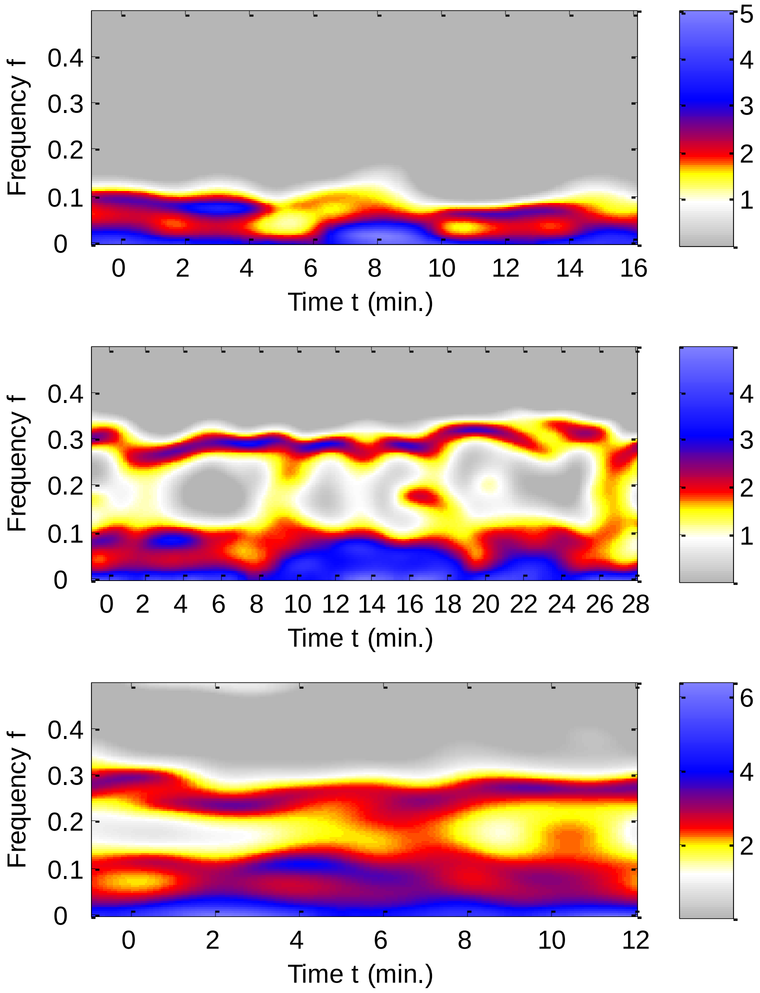


At the end of the experiment, basic socio-demographic characteristics and other variables were collected by questionnaire. In line with the literature, we measured tax morale using an item that asked participants whether cheating on taxes given the chance can always be justified, never be justified, or something in between (measured on a 10-point scale with two extreme points: 1 = “never justified,” 10 = “always justified”). Prior to the main experiment, we assessed the degree of risk aversion (*Figure A6*) using a menu of 10 safe/unsafe choices.

IV. EMPIRICAL RESULTS

To measure stress, we developed spectrograms of HRV frequency band activity for participants who exhibited distinct behavior during the 16 rounds of the experiment (see *Figure 1*). A state of stress was characterized by a lack of activation in the high frequency band, HF [0.15-0.4 Hz] but strong activation in the low frequency band, LF [0.033-0.15 Hz]. The resulting time-frequency distribution, denoted by color, assigns a different power for each frequency bin for each time interval and strength (i.e., amount of variance within a given frequency range). Overall, changes in visible HRV activation patterns correlate with substantially higher standard deviations in tax compliance rates. The upper participant has an overall tax compliance rate (ratio of reported to true income) of 0.94 (almost full compliance), indicating low variability in the tax compliance rate (SD = 0.130). This participant’s spectrogram depicts a state of stress characterized by a lack of activation in the high frequency band, HF [0.15-0.4 Hz] but strong activation in the low frequency band, LF [0.033-0.15 Hz], which also fluctuates over the course of the experiment. The middle participant has a lower tax compliance rate (0.45) and larger standard deviation (0.219), with more activation in the high frequency band indicating lower levels of stress. The third participant (lower image) has the lowest compliance rate (0.168; SD = 0.171).

Figure 1: Spectrograms of HRV Frequency Band Activity for Three Selected Cases



Notes: The time-frequency distribution was obtained using a pseudo smoothed Wigner-Ville distribution wavelet transformation to the detrended cubic interpolation of each participant’s heart rate signal. The measure used overlapping time windows of 300 seconds and frequency windows of 101 bins (total = 512 bins).

To explore the relation between psychological strain and tax compliance, we apply an ordinary least squares multiple linear regression. *Table 1* shows that our simple regression model in specification (1) gives us the following relation:

$$Compliance_{it} = constant + 0.025 * (LF/HF)_{it} + \beta \boldsymbol{X_{it}} + TD_t + GD_i + \varepsilon_{it}$$

where $Compliance_{it}$ is defined as the share of income declared. To better isolate the impact of the LF/HF ratio on compliance, we controlled for several factors (X_{it}), including the accumulated wealth at the time of the decision (to capture a participants' experimental history), being caught and punished in the previous round, cognitive skills, and other individual characteristics such as gender, age, risk aversion, and religiosity, which the literature reports as key driving forces in compliance (see Torgler 2006, 2007). Controlling for risk aversion ensures that a change in the LF/HF ratio cannot be explained by a reaction to the uncertain nature of the audit. We also controlled for experimental design factors (e.g., group dummies, GD_i) and time dynamics (time dummies, TD_t). ε_{it} denotes the error term.

The results indicate a highly statistically significant positive correlation between psychic stress and tax compliance (P = 0.007). A one-unit increase in the LF/HF ratio increases tax compliance by more than 2 percentage points. Thus, psychic stress does make a significant contribution to tax compliance beyond that of other factors. We can interpret these findings to mean that the decision to pay more taxes (and cheat less) is more likely when the experimental participant is exposed to psychic stress during the tax declaration decision. Moreover, moving from the lowest to the highest LF/HF value increases tax compliance by more than 17 percentage points.

Table 1: Tax Compliance and Psychic Stress

Dependent variable: tax compliance rate	**(1)**		**(2)**	
	Coeff.	t-stat.	Coeff.	t-stat.
Independent variables				
Psychic stress				
$(LF/HF)_t$	**0.025*** **	**2.68**	**0.024** **	**2.41**
Deterrence				
Being audited$_{(t-1)}$	0.001	0.02	0.093	2.26
Fine$_{(t-1)}$	-0.001***	-5.47	-0.002***	-9.8
Wealth				
Log(wealth)$_t$	-0.384***	-9.56	-0.087***	-2.41
Individual characteristics				
Cognitive skills	0.023***	11.72	0.004***	2.5
Age	0.018***	9.93	0.016***	9.26
Female	0.152***	7.19	0.185***	9.48
Church attendance	0.449***	8.49	0.267***	7.30
Risk aversion	0.001	0.40	-0.002	-0.42
Experimental factors				
Group dummies	YES			
Time dummies	YES	YES	YES	
Session dummies		YES	YES	
Redistribution factor 1			0.172***	2.65
Redistribution factor 2			0.388***	4.69
Redistribution rule change			-0.003	-0.19
Performance-based income distribution			0.041	1.07
Number of observations	2094		2094	
Prob > F	0.000		0.000	
R^2	0.398		0.372	

Notes: Dependent variable = the tax compliance rate as the ratio of reported to true income. Significance levels: * 0.05 < P < 0.10, ** 0.01< P < 0.05, *** P < 0.01. *Being audited*$_{(t-1)}$ = a dummy for having been audited in the previous period; *Fine*$_{(t-1)}$ = size of the fine in the previous period; and *Log(wealth)*$_t$ = accumulated wealth (negatively correlated with tax compliance). *Church attendance*: 1 = once a week or more than once a week, otherwise 0.

Instead of using group dummies we present in specification (2) results with treatment variables. The coefficient of our key variable, namely psychic stress, hardly changes. When the provision of public good was included, tax compliance increased (see the dummy variables *redistribution factor 1* and *redistribution factor 2*; *no provision of public good* = the reference group). The higher the redistribution factor, the higher the level of tax compliance. On the other hand, a *redistribution rule change* (1= group tax fund based on the contribution of the other 3 group participants, 0 = a pool based on the contribution of all 4 group participants) or a *performance-based income distribution* (1 = cognitive skill test performed, 0 = random allocation) has no implication on tax compliance.

Looking at further control variables we observe that women are more compliant than men. Older people are also more compliant, although our participants were relatively young and therefore the results should be treated with caution. Higher cognitive skills and a higher level of religiosity are also associated with higher tax compliance. The coefficient for risk aversion is not statistically significant. Prior audit and the size of the penalty paid did not improve tax compliance in the subsequent period. These results, however, cannot be compared to the available evidence in the tax compliance literature because this latter focuses primarily on fine rate changes and changes in the audit probability (Alm 1999), two variables that remained constant in our experimental design. The results are, however, comparable to experimental studies that use similar proxies (Torgler and Schaltegger 2005). Field evidence based on the Taxpayer Compliance Measurement Program also indicates that experiencing an audit has little effect on future reporting behavior (Erard 1992). Quite possibly, the negative audit experience (Frey 1997) may goad taxpayers into being even more evasive in future as retaliation against the tax agency (Andreoni et al. 1998). This assumption may also explain the negative relation to the penalty amount.

So far we have excluded participants that show inconsistencies in the degree of risk aversion (more than one switching point, see *Figure A6*). This was the case for 41 of the participants. However, the results remain robust when including these participants in the multivariate analysis controlling them with a dummy and controlling for the share of participants within a group that reported inconsistencies in their degree of risk aversion (coeff. of psychic stress (LF/HF) = 0.035, t = 3.09, N=2722). Thus, the quantitative effect of psychic stress is even larger as beforehand.

We next controlled for participant tax morale (justifiability of tax evasion) based on the assumption that those with high tax morale are unlikely to experience psychic stress. Our results indicate that, even though the impact of psychic stress remains stable (coeff. = 0.025; t = 2.63), tax morale and tax compliance (coeff. = 0.024; t = 5.88) are strongly correlated. Thus, the influence of psychic stress remains after controlling for individuals' tax morale. In addition, we explore what happens if we exclude those cases where people committed full tax evasion. The results remain robust (coeff. = 0.022; t = 2.34; N=1659). We also checked what happens if we exclude hardcore tax evaders, namely individuals always reported zero income. Four of the subjects can be classified in this category. Also here the results are stable (coeff. = 0.027; t = 2.89; N = 2030).

Similarly, we also explore whether we find a non-linear relationship between psychic costs and tax compliance. Adding $(LF/HF)_t^2$ in the two specifications reported in *Table 1* we observe that the coefficient $(LF/HF)_t^2$ is negative while $(LF/HF)_t$ remains positive (tax compliance is increasing at a decreasing rate with an increase in psychic costs). However, only in the first specification the coefficient for $(LF/HF)_t^2$ is statistically significant and only at the 10% level (t=−1.79). The coefficient value of $(LF/HF)_t$ is now larger, namely 0.040, while $(LF/HF)_t^2$ reports a value of -0.011 which indicates, ceteris paribus, that the turning point is not reached after full compliance.

Besides looking at the degree of tax compliance we also looked at the likelihood of tax evasion using a linear probability and a probit model. For specification (1) in *Table 1* we observe for the linear probability model a coefficient of -0.031 and a t-value of -2.64. For the probit model we receive a z-value of -2.70 with a marginal effect of -0.035 calculated at the multivariate point of means. Thus, an increase of psychic stress from the average by one unit reduces the likelihood of tax evasion by 3.5%.

Moreover, we examined whether individuals also experience psychic stress after committing an illegal act (remorse). This is also related to what Cooter and Ulen (2004, p. 464) call “Saturday Night Fever” waking up the next morning saying: “I can’t believe what I did last night!”. Our results reveal that there is indeed a negative correlation between tax compliance in a previous period (t-1) and psychic stress in a current one (t) using only group and time dummy variables as controls (coeff. tax compliance = −0.091, $t = -2.13$). We also explored whether the psychic stress experienced in a previous period (t-1) has an impact on tax compliance in the current round (t). The results are very similar: the coefficient is statistically significant at the 1% level with comparable quantitative effects. Adding both factors, namely psychic stress in t and $t-1$ together in the specification indicates that the current round is more important.

Lastly, to do further checks the robustness, we added controls for health factors such as alcohol consumption (drinks per week), smoking habits (“daily,” “less frequently,” “not at all”) and physical/sport activity (minutes per week) that are related to HRV. These additions, however, barely changed the impact of the LF/HF ratio on tax compliance ($P = 0.009$; coeff. = 0.024). Even when we added in the information that 33 of the 180 participants reported taking prescribed medication, our results remained robust ($P = 0.024$; coeff. = 0.021). We also explored a particular subsample of participants, 92 students who reported having or complaining of headache, symptoms of depression, or high blood pressure during the previous month. Of these, 14% reported symptoms of depression, 5% reported both depression and headache, 79% reported only headache, and 1% reported high blood pressure. In fact, focusing on this subgroup actually increased the correlation between psychic stress and tax compliance (coeff. = 0.046, *t* = 3.83, *n* = 1017). Controlling for these symptoms in the entire sample produced similar results (coeff. = 0.030, *t* = 3.16).

V. CONCLUSIONS

In general, these results imply that the analysis of tax evasion may benefit from including factors not traditionally viewed as essential policy tools. Contrary to Coricelli et al. (2010) and in line with Erard and Feinstein (1994) theoretical argument we find that the intention of a false declaration generates anxiety, guilt or a reduction in the taxpayer's self-image or self-esteem, which in turn reduces the utility of tax evasion and thus the level of tax evasion. Thus, based on our results it seems that potential emotions such as excitement due to the prospect of higher payoffs were outweighed by the experienced psychic costs. If such excitement is also measured with our HRV variable we would observe a lower bound estimate of the impact of psychic costs.

In general, valuable insights could be gleaned from a better understanding of the social and economic conditions under which paying taxes elicits more or less psychic stress. Admittedly, the psychic stress related to social norms of tax payment may also be directly affected by factors like perceived fairness of the tax system, tax administrators' treatment of taxpayers, the level of disapproval of government policies or general culture. In fact, there is evidence at the attitudinal level that these factors shape tax morale or social norms of compliance (Slemrod 1992, Torgler, 2007, Alm et al. 2010). At the behavioral level, however, little is known. Future research might also benefit from investigating the large set of other moral sentiments that might contribute to a more sophisticated biological micro-foundation for compliance behavior.

Moreover, further conclusions can be drawn from the present study. First, our results have also implication beyond tax compliance such as, e.g., contributions to the commons (Dietz, Ostrom, and Stern 2003) which underlie many public good activities. If our findings hold independent of the specific domain, then factors that increases awareness and valuation of the commons may help to raise contribution levels. More specifically, with respect to tax compliance as a form of contribution to the commons, our evidence shows that factors not traditionally viewed as essential policy tools may improve contribution rates. Hence, understanding the social and economic conditions under which contribution elicits more or less psychic strain may be a way to improve the effectiveness and fairness of tax systems.

Finally, the present study shows that HRV data can be used as a biomarker to observe decision behavior influenced by nonmonetary and noneconomic incentives (such as norms) underlying contributions to the commons or society in general. Not only can these data help to predict and understand participants' behavior, but carefully designed experiments that record the participants' HRV can help to identify situations, and their associated factors, in which nonstandard motivations like moral sentiment and social norms affect decision making.

REFERENCES

Adams, C., 1993. For Good and Evil: The Impact of Taxes on the Course of Civilization. Madison Books, London.

Asch, S. E. 1952. Social Psychology. Prentice Hall, New York.

Allingham, M.G., Sandmo, A., 1972. Income Tax Evasion: A Theoretical Analysis. Journal of Public Economics 1, 323–338.

Andreoni, J., Erard, B., Feinstein, J., 1998. Tax Compliance. Journal of Economic Literature 36, 818–860.
Alm, J., 1999. Tax Compliance and Administration, in: Hildreth, W.B., Richardson, J.A. (Eds.), Handbook on Taxation. Marcel Dekker, New York, pp. 741–768.

Alm, J., Jackson, B.R., McKee, M., 1992. Institutional Uncertainty and Taxpayer Compliance. The American Economic Review 82, 1018–1026.

Alm, J., Jackson, B.R., McKee, M., 1993. Fiscal Exchange, Collective Decision Institutions, and Tax Compliance. Journal of Economic Behavior & Organization 22, 285–303.

Alm, J., McClelland, G.H., Schulze, W.D., 1992. Why Do People Pay Taxes? Journal of Public Economics 48, 21–38.

Alm, J. J. Martinez-Vazquez, B. Torgler (Eds.) 2010. Developing Alternative Frameworks for Explaining Tax Compliance. Routledge, London.

Appelhans, B.M., Luecken, L.J., 2006. Heart Rate Variability as an Index of Regulated Emotional Responding. Review of General Psychology 10, 229–240.

Becker, G.S., 1968. Crime and Punishment: An Economic Approach. The Journal of Political Economy 76, 169–217.

Berntson, G.G., Cacioppo, J.T., 2008. Heart Rate Variability: Stress and Psychiatric Conditions, in: Camm, J.A., Malik, M. (Eds.), Dynamic Electrocardiography. John Wiley & Sons, Ltd, Oxford.

Berntson, G.G., Cacioppo, J.T., Binkley, P.F., Uchino, B.N., Quigley, K.S., Fieldstone, A., 1994. Autonomic cardiac Control. III. Psychological Stress and Cardiac Response in Autonomic Space as Revealed by Pharmacological Blockades. Psychophysiology 31, 599–608

Bianchi, A.M., Mainardi, L., Petrucci, E., Signorini, M.G., Mainardi, M., Cerutti, S., 1993. Time-Variant Power Spectrum Analysis for the Detection of Transient Episodes in HRV signal. Biomedical Engineering, IEEE Transactions on 40, 136–144.

Brandts, J., Garofalo, O., 2012. Gender Pairings and Accountability Effects, Journal of Economic Behavior & Organization 83, 31-41.

Cooter, R., Ulen, T., 2044. Law & Economics. Pearson Addison Wesley, Boston.

Coricelli, G., Joffily, M., Montmarquette, C., Villeval, M. C., 2010. Cheating, Emotions, and Rationality: An Experiment on Tax Evasion 13, 226-247.

Crone, E.A., Bunge, S.A., de Klerk, P., van der Molen, M.W., 2005. Cardiac Concomitants of Performance Monitoring: Context Dependence and Individual Differences. Cognitive Brain Research 23, 93–106.

Crone, E.A., Somsen, R.J.M., van Beek, B., van der Molen, M.W., 2004. Heart Rate and Skin Conductance Analysis of Antecendents and Consequences of Decision Making. Psychophysiology 41, 531–540.

Dietz, T., Ostrom, E., Stern, P. Govern the Commons Science 302, 1907-1912.

Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L., Blair, S. N., 2000. Heart Rate Variability, Trait Anxiety, and Perceived Stress Among Physically Fit Men and Women. International Journal of Psychophysiology 37, 121-133 (2000).

Dulleck, U., Ristl, A., Schaffner, M., Torgler, B., 2011. Heart Rate Variability, the Autonomous Nervous System, and Neuroeconomic Experiments, Journal of Neuroscience, Psychology, and Economics 4, 117-124.

Engelmann, D., Strobel, M., 2004. Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments. American Economic Review, American Economic Review 94, 857–869.

Erard, B., 1992. The Influence of Tax Audits on Reporting Behavior, in: Slemrod, J. (Ed.), Why People Pay Taxes. University of Michigan Press, Michigan, pp. 95–114.

Erard, B., Feinstein, J., 1994. The Role of Moral Sentiments and Audit Perceptions in Tax Compliance. Public Finance 49, 70–89.

Falk, A., Menrath, I., Verde, P. E., Siegrist, J., 2011. Cardiovascular Consequences of Unfair Pay, IZA Discussion Paper Series No. 5720, May 2011.

Festinger, L., 1957. A Theory of Cognitive Dissonance. Stanford University Press, Stanford.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. Experimental Economics 10, 171–178.

Frey, B. S., 1997. Not Just for Money: An Economic Theory of Personal Motivation. Edward Elgar, Cheltenham, UK.

Graetz, M.J., Wilde, L.L., 1985. The Economics of Tax Compliance: Fact and Fantasy. National Tax Journal 38, 355–363.

Grasmick, H.G., Bursik, R.J., 1990. Conscience, Significant Others, and Rational Choice: Extending the Deterrence Model. Law & Society Review 24, 837–861.

Green, J.D., Paxton, J.M., 2009. Patterns of Neural Activity Associated with Honest and Dishonest moral Decisions. Proceedings of the National Academy of Sciences of the United States of America 106, 12506–12511.

Hainsworth, R., 2008. Physiological Background of Heart Rate Variability, in: Camm, J.A., Malik, M. (Eds.), Dynamic Electrocardiography. John Wiley & Sons, Ltd, Oxford.

Harbaugh, W.T., Mayr, U., Burghart, D.R., 2007. Neural Responses to Taxation and Voluntary Giving Reveal Motives for Charitable Donations. Science 316, 1622–1625.

Holt, C.A., Laury, S.K., 2002. Risk Aversion and Incentive Effects. American Economic Review 92, 1644–1655.

Houser, D., Kurzban, R., 2002. Revisiting Kindness and Confusion in Public Goods Experiments. The American Economic Review 92, 1062–1069.

Koelsch, S., Remppis, A., Sammler, D., Jentschke, S., Mietchen, D., Fritz, T., Bonnemeier, H., Siebel, W.A., 2007. A Cardiac Signature of Emotionality. European Journal of Neuroscience 26, 3328–3338.

Konrad, K. A., Qari, S., 2012. The Last Refuge of a Scoundrel? Patriotism and Tax Compliance, Economica 79, 516-533.

Ledyard, J.O., 1995. Public Goods: A Survey of Experimental Research, in: Kagel, J.H., Roth, A.E. (Eds.), The Handbook of Experimental Economics. Princeton University Press, Princeton, pp. 111–194.

Levy, M.N., Martin, P.J., 1979. Neural Control of the Heart, in: Berne, R.M. (Ed.), Handbook of Physiology. American Physiological Society, Bethesda, pp. 581–620.

Levy, M.N., Martin, P.J., Iano, T., Zieske, H., 1970. Effects of Single Vagal Stimuli on Heart Rate and Atrioventricular Conduction. American Journal of Physiology 218, 1256–1262.

Lewis, A., 1982. The Psychology of Taxation. Martin Robertson, Oxford.

Malik, M., 2008. Standard Measurement of Heart Rate Variability, in: Camm, J.A., Malik, M. (Eds.), Dynamic Electrocardiography. John Wiley & Sons, Ltd, Oxford.

Malik, M., Bigger, T.J., Camm, J.A., Kleiger, R.E., Malliani, A., Moss, A.J., Schwartz, P.J., 1996. Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use. European Heart Journal 17, 354 –381.

McCraty, R., Atkinson, M., Tiller, W.A., Rein, G., Watkins, A.D., 1995. The Effects of Emotions on Short-Term Power Spectrum Analysis of Heart Rate Variability. American Journal of Cardiology 76, 1089–1093.

Polinsky, A.M., Shavell, S., 2000. The Economic Theory of Public Enforcement of Law. Journal of Economic Literature 38, 45–76.

Seong, H.M., Lee, J.S., Shin, T.M., Kim, W.S., Yoon, Y.R., 2004. The Analysis of Mental Stress Using Time-Frequency Distribution of Heart Rate Variability Signal, in: Engineering in Medicine and Biology Society, 2004. IEMBS ’04. 26th Annual International Conference of the IEEE. Presented at the Engineering in Medicine and Biology Society, IEMBS ’04. 26th Annual International Conference of the IEEE, pp. 283–285.

Slemrod, J., 1992. Tax Compliance and Enforcement, in: Slemrod, J. (Ed.), Why People Pay Taxes. University of Michigan Press, Michigan.

Spyers-Ashby, J.M., Bain, P.G., Roberts, S.J., 1998. A Comparison of Fast Fourier transform (FFT) and Autoregressive (AR) Spectral Estimation Techniques for the Analysis of Tremor Data. Journal of Neuroscience Methods 83, 35–43.

Torgler, B., 2007. Tax Compliance and Tax Morale: A Theoretical and Empirical Analysis. Edward Elgar Publishing, Cheltenham.

Torgler, B., Schaltegger, C. A., 2005. Tax Amnesties and Political Participation, Public Finance Review 33, 403-431.

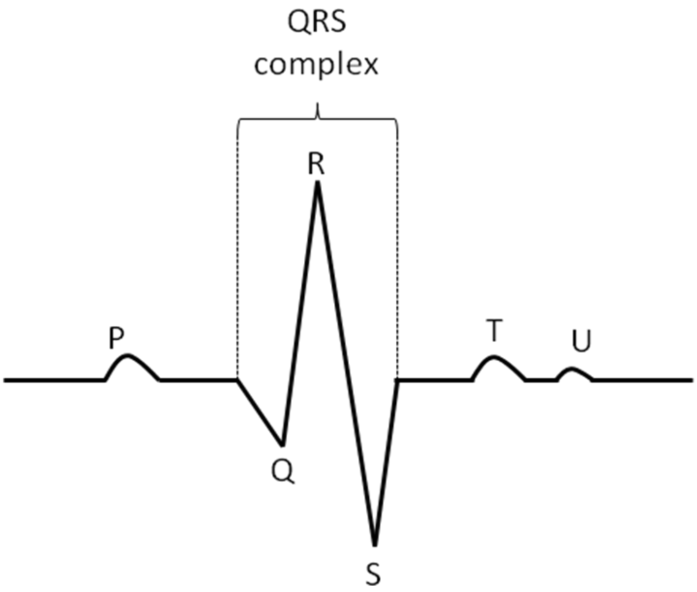
Van Lange, P. A. M., Finkenauer, C., Popma, A., van Vugt, M., 2011. Electrodes as Social Glue: Measuring Heart Rate Promotes Giving in the Trust Game, International Journal of Psychophysiology 80, 246-250.

Wiklund, U., Akay, M., Niklasson, U., 1997. Short-Term Analysis of Heart-Rate Variability of Adapted Wavelet Transforms. Engineering in Medicine and Biology Magazine, IEEE 16, 113–118.

Yang, T.T., Simmons, A.N., Matthews, S.C., Tapert, S.F., Bischoff-Grethe, A., Frank, G.K.W., Arce, E., Paulus, M.P., 2007. Increased Amygdala Activation is Related to Heart Rate During Emotion Processing in Adolescent Subjects. Neuroscience Letters 428, 109–114.

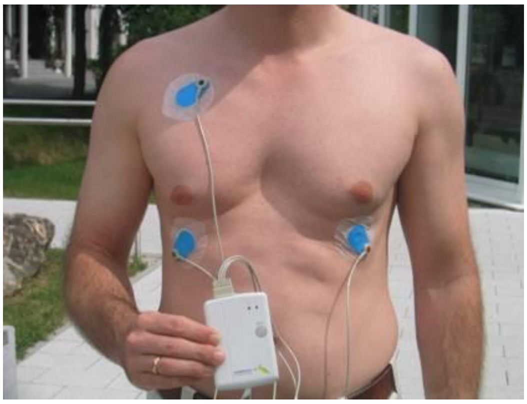
APPENDIX

Figure A1: ECG Signal



Notes: Medical and cardiology research understands the ECG signal of a heart beat as consisting of waves, and refers to the wavelength of the main signal as the QRS complex.

Figure A2: Heart Rate Measurement Equipment



Notes: The heart rate measurement equipment used was the nonintrusive Holter ECG Recorder AR12, a pocket-sized monitoring device (10x7x2.5cm) that connects to a human subject's chest through three conducting electrodes. This device not only records the ECG with a high sampling rate of 4096 Hz, it also records respiration levels and has a built-in QRS detection algorithm. The data recorded are stored on a CF disk and can be read using a proprietary software package or, alternatively, exported into a number of data formats (e.g., Matlab). Any irregularities in the ECG recording (assumedly random, and present in only 5% of observations) were excluded from the data analysis. To reduce the potential for such irregularities, we had previously tested the heart rate monitor application using the IrDA (Infrared Data Association) interface.

Figure A3: Instruction Screen

INSTRUCTION

This is an experiment in the economics of decision-making. You will have the possibility to earn money through your participation in this experiment. Please follow the instructions carefully and do not hesitate to raise your hand if you have a question. Please do not ask your question out loud; wait until an experiment administrator comes to you to answer your question privately. During this session, any communication between participants is forbidden. Please turn off your mobile phones.

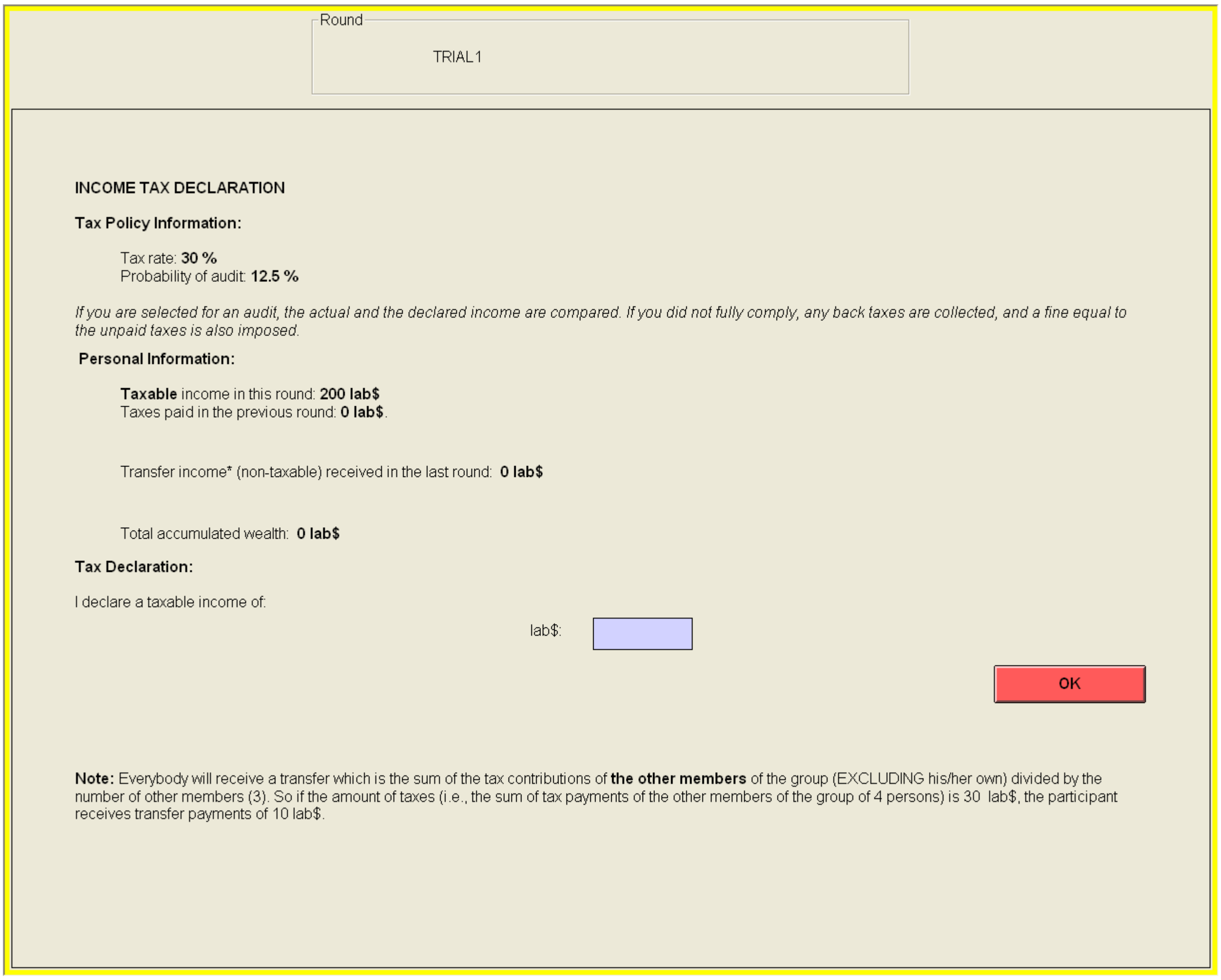
Your decisions are anonymous. You are identified solely with your participant number. Thus the payments are strictly confidential as well. You will receive them in a sealed envelope with your number. Nobody, neither the experimenter nor the other participants, will be able to attribute a decision to a person.

There are several stages in the experiment, you will receive instructions for each stage later on. You can also find the instructions for the main part of the experiment on the sheet on your desk. You will be randomly grouped together with 4 other participants and this group will stay the same over the whole experiment. You will not know the identities of the other members of your group at any time.

Thank you for your participation. We hope you enjoy this experiment. Do you have any questions?

Please enter your participation number:

Figure A4: Income Tax Declaration Screen



Round

TRIAL 1

INCOME TAX DECLARATION

Tax Policy Information:

Tax rate: **30 %**
Probability of audit: **12.5 %**

If you are selected for an audit, the actual and the declared income are compared. If you did not fully comply, any back taxes are collected, and a fine equal to the unpaid taxes is also imposed.

Personal Information:

Taxable income in this round: **200 lab$**
Taxes paid in the previous round: **0 lab$**.

Transfer income* (non-taxable) received in the last round: **0 lab$**

Total accumulated wealth: **0 lab$**

Tax Declaration:

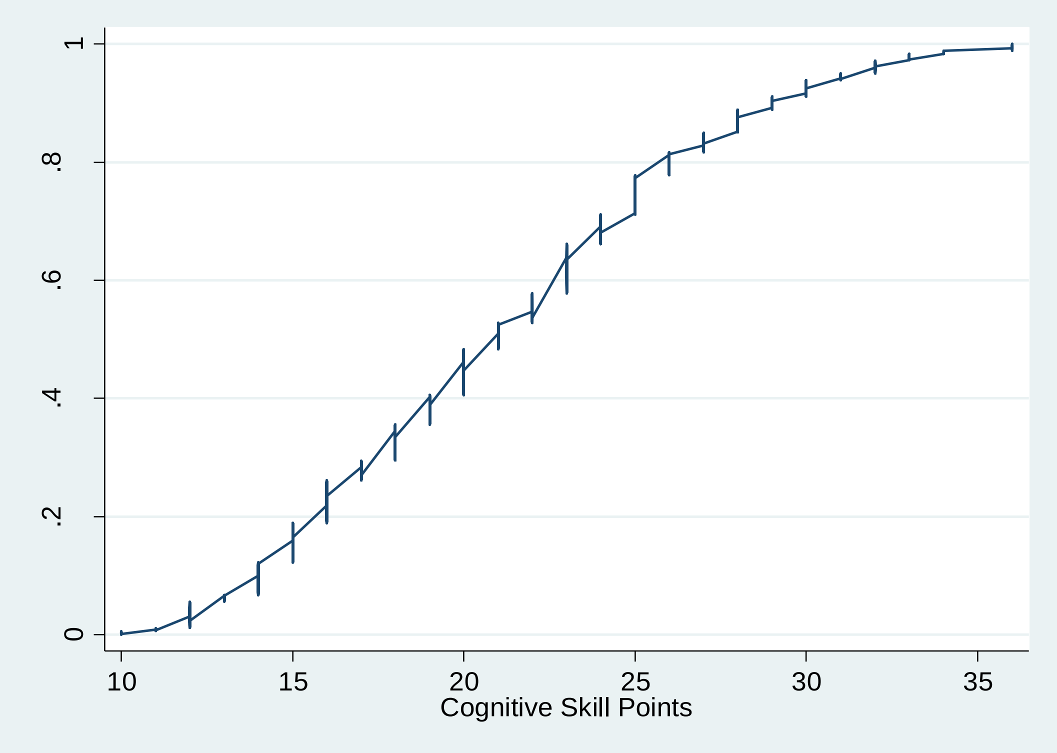
I declare a taxable income of:

lab$:

Note: Everybody will receive a transfer which is the sum of the tax contributions of **the other members** of the group (EXCLUDING his/her own) divided by the number of other members (3). So if the amount of taxes (i.e., the sum of tax payments of the other members of the group of 4 persons) is 30 lab$, the participant receives transfer payments of 10 lab$.

Notes: Income tax declaration screen. Because the time data on events in the computer-based experiment can be prone to variation caused by network traffic issues, the time recorded gives only an indication of when the event actually happened (i.e., when the screen with the information was actually visible to the participant). We therefore isolated the laboratory network from the university network and assigned it to its own subnet, which substantially reduced the discrepancy in recording (<6 miliseconds).

Figure A5: Cognitive Skills



Notes: Cumulative distribution function of participants' cognitive skills with a time restriction of 12 minutes (between 0 and 50 possible points). Participants with inconsistent risk aversion (see Figure A6) also exhibited lower cognitive skills (P < 0.0150). Rational thinking was assessed using questions characterized by concrete calculation and evaluations related to numerical, verbal and spatial reasoning. Example questions are given below:

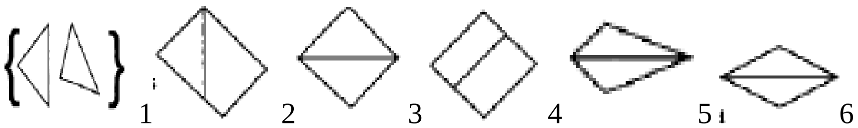
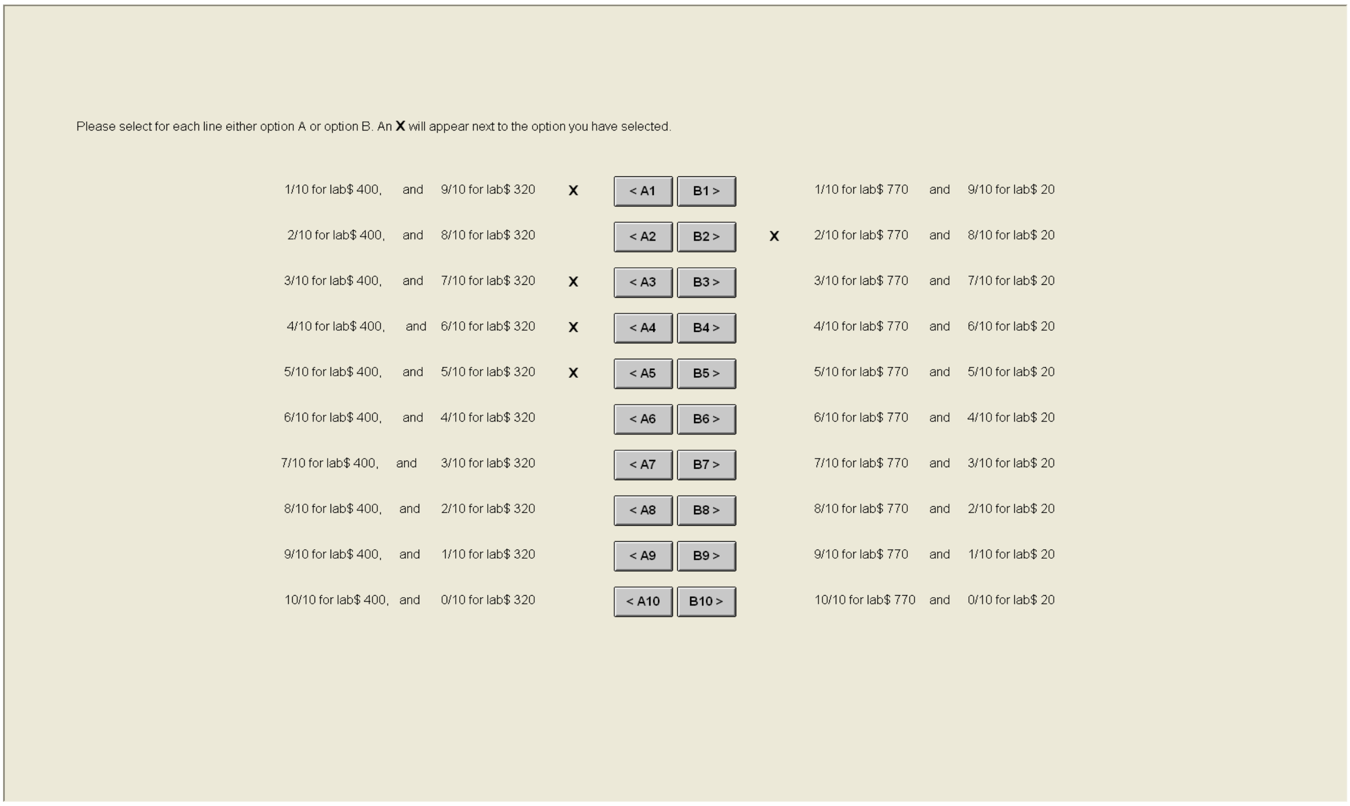
Numerical reasoning	Look at the row of numbers. What number should come next? 27 9 3 1 1/3 1/9 1/?
Verbal reasoning	IMPRISON is the opposite of 1. capture, 2. endanger, 3. free, 4. discover, 5. heal
Spatial reasoning	Which figure can be made from the two figures in the brackets? 

Figure A6: Risk Aversion



Please select for each line either option A or option B. An **X** will appear next to the option you have selected.

1/10 for lab$ 400, and 9/10 for lab$ 320	X	< A1	B1 >		1/10 for lab$ 770 and 9/10 for lab$ 20
2/10 for lab$ 400, and 8/10 for lab$ 320		< A2	B2 >	X	2/10 for lab$ 770 and 8/10 for lab$ 20
3/10 for lab$ 400, and 7/10 for lab$ 320	X	< A3	B3 >		3/10 for lab$ 770 and 7/10 for lab$ 20
4/10 for lab$ 400, and 6/10 for lab$ 320	X	< A4	B4 >		4/10 for lab$ 770 and 6/10 for lab$ 20
5/10 for lab$ 400, and 5/10 for lab$ 320	X	< A5	B5 >		5/10 for lab$ 770 and 5/10 for lab$ 20
6/10 for lab$ 400, and 4/10 for lab$ 320		< A6	B6 >		6/10 for lab$ 770 and 4/10 for lab$ 20
7/10 for lab$ 400, and 3/10 for lab$ 320		< A7	B7 >		7/10 for lab$ 770 and 3/10 for lab$ 20
8/10 for lab$ 400, and 2/10 for lab$ 320		< A8	B8 >		8/10 for lab$ 770 and 2/10 for lab$ 20
9/10 for lab$ 400, and 1/10 for lab$ 320		< A9	B9 >		9/10 for lab$ 770 and 1/10 for lab$ 20
10/10 for lab$ 400, and 0/10 for lab$ 320		< A10	B10 >		10/10 for lab$ 770 and 0/10 for lab$ 20

Notes: Prior to the tax declaration process, in a simple choice task to measure degree of risk aversion, participants were presented with a menu of 10 choices. Risk aversion was then measured by the sum of the number of safe (left-hand side) choices made.

Text A1: Data Processing

Estimation of the power spectral density (PSD)

All the following estimation methods employ a number of parameters. In addition, because of the trade-off between time resolution and frequency resolution, a decision must be made on the resolution of each. The lowest feasible time resolution is 27 seconds, which corresponds to the lower end of the LF, for which a reasonable frequency resolution would be 0.001 Hz. The result is a total of 512 frequency bins. A second time resolution must also be estimated, however, that determines the distance between two time points. Since estimations are overall values within a certain window, the time window is adjusted in all three methods using a Gaussian or Hamming scaling window, which places more emphasis on the signal at the center of the time window. In the case of wavelet transformation, the same procedure is also used to amplify the frequency windows.

Short-term fast Fourier transformation (SFFT)

The oldest and simplest estimation method is the Fourier transformation (FT), most commonly computed using the fast Fourier transformation (FFT) algorithm, which is based on the principle that any finite signal s can be transformed into a series of parameterized sine and cosine functions (Eq. 1). This transformation allows estimation of the strength k of the frequency ω

$$s(t) = \alpha_0 + \sum_{k=1}^{\infty}(\alpha_k \cos(k\omega t) + \beta_k \sin(k\omega t)) \qquad (1)$$

because the parameters α_k and β_k also enable derivation of the desired PSD, making FFT very well suited to overlapping periodic signals. Unfortunately, however, to a large extent, the heart rate variation is not periodic and time variant, so the common solution is to take short time intervals (that correspond to the period of the longest frequency likely to be observed), detrend this signal, and apply the FFT after detrending (short-term FFT, SFFT). However, as this method is not efficient for detecting isolated, short, nonperiodic changes in heart rate, the SFFT is used primarily as a reference point for other methods.

Autoregressive spectral estimation (AR)

Another widely used method for obtaining the PSD of a signal is to fit an autoregressive model of the form

$$X_t + \alpha_1 X_{t-1} + ... + \alpha_M X_{t-M} = \epsilon_t \qquad (2)$$

where ϵ_t is a white noise process with a constant variance σ^2. The PSD estimate $\hat{h}(.)$ for different frequencies ω can then be determined as follows:

$$\hat{h}(\omega) = \frac{\sigma_\epsilon^2}{2\pi} \frac{1}{|1 + \sum_{k=1}^{M} \hat{\alpha}_k e^{-i\omega k}|^2} \qquad (3)$$

where $\hat{\alpha}_1, ..., \hat{\alpha}_M$ are the estimated model coefficients. Estimation with any order M larger than 50 has been shown to yield a good “finite order” approximation of the heart rate signal. The specific estimation technique used here is based on the “maximum entropy” method, whose application to physiological data using both the SFFT and AR estimation techniques is reviewed by Spyers-Ashby, Bain and Roberts (1998).

Wavelet analysis using the smoothed pseudo Winger-Ville distribution (SPWVD)

The most recently proposed estimation method is a wavelet analysis based on the intuitive assumption that a signal does not simply fall into one specific time window but into a whole series of different windows. To obtain a power estimate, each of these windows is matched to a set of short periodic functions with particular frequencies ω. The resulting time-scale representation is thus a collection of time-frequency distributions with different time and frequency resolutions.

One particularly suitable technique for HRV estimation is to build a smoother pseudo Wigner-Ville distribution (SPWVD) (Seong et al. 2004), a method that condenses the time-scale representation of the wavelet transformation into a single time-frequency distribution with a high time-frequency resolution (Eq. 4).

$$PSD_{WV}(t,\omega) = \int W_\omega(\tau) \times \left[\int W_t(u-t) x(u+\frac{\tau}{2}) x^*(u+\frac{\tau}{2}) du \right] e^{-i\omega t} d\tau \qquad (4)$$

Comparison and evaluation of the estimation approaches in the context of HRV measurement

Although the SFFT and the AR methods are well-established in the literature and provide reliable results, because both techniques must use a predefined time window, they suffer quite heavily from the trade-off between time and frequency resolution. This problem stems from the Heisenberg uncertainty principle: it is impossible to obtain the exact frequency and exact time of frequency occurrence in the heart rate signal. The SPWVD technique, on the other hand, allows for much finer gradations, although it does produce results that contain overlapping time intervals for the same data. This problem, however, can be mitigated by using data averages. Nevertheless, because the use of SPWVD is still infrequent in practice, it seems advisable to check the results for consistency using the other two methods to ensure that no major discrepancies have arisen.