

eJournal of Tax Research

Volume 6, Number 1

June 2008

CONTENTS

- 5 Philosophical Paradigms, Inquiry Strategies and Knowledge Claims:
Applying the Principles of Research Design and Conduct to Taxation
Margaret McKerchar
- 23 Avoidance and *abus de droit*: The European Approach in Tax Law
Marco Greggi
- 45 The Economic Benefits of the Use of Guanxi and Business Networks in
a Jurisdiction with Strong Formal Institutions: Minimisation of Taxation
Nolan Cormac Sharkey
- 67 The Effect of Audit Strategy Information on Tax Compliance – An
Empirical Study
Leif Appelgren

eJournal of Tax Research

EDITORS

Associate Professor Binh Tran-Nam, Atax, The University of New South Wales, Australia

Associate Professor Michael Walpole, Atax, The University of New South Wales, Australia

EDITORIAL BOARD

Professor Robin Boadway, Department of Economics, Queen's University

Associate Professor Cynthia Coleman, Faculty of Economics and Business, University of Sydney

Professor Graeme Cooper, Faculty of Law, University of Sydney

Professor Robert Deutsch, Atax, The University of New South Wales

Professor Chris Evans, Atax, The University of New South Wales

Professor Judith Freedman, Faculty of Law, Oxford University

Professor Malcolm Gammie, Chambers of Lord Grabiner QC, London

Professor Jeyapalan Kasipillai, School of Business, Monash University Sunway Campus

Professor Rick Krever, Department of Law and Taxation, Monash University

Professor Charles McLure Jr., Hoover Institution, Stanford University

Professor John Prebble, Faculty of Law, Victoria University of Wellington

Professor Joel Slemrod, University of Michigan Business School

Professor John Tiley, Centre for Tax Law, Cambridge University

Professor Jeffrey Waincymer, Faculty of Law, Monash University

Professor Neil Warren, Atax, The University of New South Wales

Professor Robin Woellner, Faculty of Law, James Cook University

PRODUCTION EDITOR

Cindy Chan, Atax, The University of New South Wales, Australia

PUBLISHER

Atax is the Australian School of Taxation in the Faculty of Law at The University of New South Wales. We are the largest tax school in any university in Australia, bringing together a team of expert academic staff with backgrounds in law, commerce, tax, accounting and economics. At Atax, we're working towards building excellence in the tax profession, looking at tax from both a theoretical and practical perspective.



eJournal of Tax Research

EDITORS' NOTE

The *eJournal of Tax Research* is a refereed journal that publishes original, scholarly works on all aspects of taxation. It aims to promote timely dissemination of research and public discussion of tax-related issues, from both theoretical and practical perspectives. It provides a channel for academics, researchers, practitioners, administrators, judges and policy makers to enhance their understanding and knowledge of taxation. The journal emphasises the interdisciplinary nature of taxation. To ensure the topicality of the journal, submissions will be refereed quickly.

SUBMISSION OF ORIGINAL MATERIAL

Submission of original contributions on any topic of tax interest is welcomed, and should be sent as an email attachment to the Production Editor at <ejtr@unsw.edu.au>. Submission of a manuscript is taken to imply that it is an unpublished work and has not already been submitted for publication elsewhere. Potential authors are requested to follow the "Notes to Authors", which is available from the journal's website.



eJournal of Tax Research

Volume 6, Number 1

June 2008

CONTENTS

- 5 Philosophical Paradigms, Inquiry Strategies and Knowledge Claims:
Applying the Principles of Research Design and Conduct to Taxation
Margaret McKerchar
- 23 Avoidance and *abus de droit*: The European Approach in Tax Law
Marco Greggi
- 45 The Economic Benefits of the Use of Guanxi and Business Networks in
a Jurisdiction with Strong Formal Institutions: Minimisation of Taxation
Nolan Cormac Sharkey
- 67 The Effect of Audit Strategy Information on Tax Compliance – An
Empirical Study
Leif Appelgren

Philosophical Paradigms, Inquiry Strategies and Knowledge Claims: Applying the Principles of Research Design and Conduct to Taxation

Margaret McKerchar*

Abstract

Taxation is a fundamental part of everyday life and it comes as no surprise that it attracts great interest from policymakers, academics, business and the wider community both in Australia and overseas. However, those interested in tax research come from very diverse discipline backgrounds including law, accounting, economics, political science, psychology and philosophy. The prior learning of many tax researchers does tend to be in the study and application of the law and typically they have little training in or exposure to the detail of the theory and practice of research design. This is a limitation for both academics and for the growing body of research students that they are being called upon to supervise. There is capacity to improve the capability of tax researchers by evaluating best practice in closely related disciplines and exploring that of other disciplines that could have relevance to taxation.

The paper explores the fundamental aspects of research design, including a range of philosophical paradigms and strategies of inquiries that could have application to taxation. The framing of research questions (or hypotheses) is considered, as is the need for alignment between research objectives, research questions, philosophical paradigms, strategies of inquiry and knowledge claims. The application of mixed-method designs is also considered. Pitfalls in research design to be avoided are discussed as are suggestions by which the robustness of tax research can be improved.

1. INTRODUCTION

Taxation, with its intrigues and intricacies, is an important part of everyday life. Its shape and form is a reflection of the shared values, goals and aspirations of society, and a means by which its members are bound together. When seeking to understand almost any aspect of taxation, it must be borne in mind that it is much more than the study of the revenue law itself. Taxation is not a discipline in its own right, but a social phenomenon that can be studied through various disciplinary lenses. Commonly, taxation attracts researchers from the disciplines of law, accounting, economics, political science, psychology and philosophy. These disciplinary backgrounds are each understandably narrow and, in spite of researchers being no doubt experts in their respective fields, it can be challenging to apply their skills and knowledge to the complexities of the research problems that emanate from the study

* Associate Professor of Taxation at the Australian School of Taxation (Atax), University of New South Wales, Sydney, Australia. Email: m.mckerchar@unsw.edu.au. An earlier version of this paper was presented at the 20th Annual Conference of the Australasian Tax Teachers Association, University of Tasmania, Hobart, 23-25 January 2008.

of taxation. For example, legal researchers may be well qualified to study the meaning of the letter of the law, but find that they are not quite so well equipped to study how people respond to the law. This in turn may well impinge on any recommendations that researchers may make for law reform. Similarly, an accounting researcher may be experienced in reading data in quantitative form, but not understand the principles of survey design or statistical analysis. This in turn could lead to knowledge claims or conclusions being made or drawn that are not valid.

That is, taxation is complex and researchers in this field are often not fully equipped to grapple with its multidimensional nature, particularly when it involves the study of human behaviour. There are various philosophical research paradigms and strategies of inquiry, the theory and fundamental principles of which a good researcher in taxation (from whatever discipline) needs to understand in order to be able to apply them with confidence. This in turn should allow the researcher to make informed decisions as to their appropriateness in given contexts, and to apply them with rigour. Every researcher is trying to make a meaningful contribution to the body of knowledge and this is best done by producing research of the highest quality. This requires attention to detail when it comes to the design, conduct and writing up of taxation research. The purpose of the paper is to explore these fundamental principles of good research design and conduct in the context of taxation and provide practical examples of best practice. It is hoped that the paper does stimulate further consideration of current practices and also provides some practical guidance for tax researchers.

Following on from the introduction in part 1 of this paper, the philosophical paradigms of research that are considered most relevant to taxation and the fundamental aspects of research design are explored in part 2. The strategies of inquiry used in quantitative research and the knowledge claims that are consistent with this paradigm are considered in part 3. Qualitative research and legal research are discussed in parts 4 and 5 respectively. Other less traditional research approaches including the role of mixed method research are examined in part 6. Part 7 includes a summary of the key concepts covered in the paper and some questions that may help keep tax researchers on track and on the pathway to discovery.

2. PHILOSOPHICAL PARADIGMS AND RESEARCH DESIGN

There is a substantial body of literature on the design and conduct of research from which it can be said that there are traditionally two core research philosophical paradigms. These paradigms, or organising frameworks¹ or disciplinary matrices² that guide researchers, are commonly referred to as positivism and interpretivism. A research paradigm has its own identifying characteristics, its own ways and means (methods and practices) that create expectations about the nature and conduct of research which it embraces. Grix maintains that all research takes place within a paradigm, whether it is explicitly stated or not.³ Paradigm choice is by and large a reflection of how the researcher views the world (ontology) and believes that knowledge is created (epistemology). That is, these implicit beliefs, along with the researcher's disciplinary focus and past experiences, will influence his or her

¹ Grix, J., 2004, *The Foundations of Research*, Palgrave, Basingstoke, p.171.

² Schwandt, T.A., 2001, *Dictionary of Qualitative Inquiry*, 2nd edn, Sage, Thousand Oaks.

³ See note 1, p.98.

philosophical approach to research, even before the topic is chosen.⁴ Whilst every researcher will undoubtedly have a paradigm preference, it is important to keep an open mind to the work of researchers in both tax and other disciplines, particularly where their endeavours can potentially contribute. Further, an effective researcher should be flexible enough to be able to work within the most appropriate paradigm given the nature of the research problem under investigation.

In the context of the social sciences, positivism is adopted by those researchers who seek objectivity in their explanation of social reality. In this paradigm the researcher is viewed as detached from the subjects under study and the explanations derived are based on empirical evidence and tested theories. That is, the knowledge produced by a positivist approach is based on deductive reasoning whereby the researcher follows a precise and structured process leading to the identification of causal relationships, logical conclusions and the making of predictions according to various confidence levels. The positivist approach has been described as being based on a realist, foundationalist ontology which views the world as existing independently of our knowledge of it.⁵

In contrast, interpretivism is based on the assumption that the researcher cannot be detached from the subjects being studied. It is sometimes referred to as anti-positivism. As a research paradigm, interpretivism provides an understanding of social reality that is based on the subjective interpretation of the researcher. It does not provide a hard and fast explanation from which causal relationships can be identified and predictions made. Denscombe⁶ describes the explanations of an interpretivist researcher as likely to be messy and open-ended rather than being nice, neat and complete.

It follows that the researcher whose philosophical paradigm is best described as positivist is more likely to adopt a quantitative methodology⁷, whilst the interpretivist researcher could be expected to employ a qualitative methodology. The parallels are readily apparent. Quantitative research is empirical in nature, relies on deductive reasoning, is commonly used in the sciences, and has been practised as far back as Hippocrates c.450BC.⁸ In contrast, qualitative research began in the 1900s, is more commonly used in the social sciences and is a complex and still evolving paradigm.⁹ Qualitative research requires inductive reasoning to be employed rather than logic, and often calls for more creative and indirect means of collecting data or evidence.¹⁰

However, whilst these two paradigms are quite philosophically opposed, it has been recognised since the 1970s that a continuum exists between them in which other paradigms can and do exist. For example, critical realism and pragmatism both

⁴ See note 1, p.173.

⁵ Guba, E. G. and Y. S. Lincoln, 2005, "Paradigmatic Controversies, Contradictions, and Emerging Confluences" in Denzin, N. K. and Y. S. Lincoln (Eds.), *The Sage Handbook of Qualitative Research*, 3rd edn., Sage, Thousand Oaks, pp.191-215.

⁶ Denscombe, M., 2002, *Ground Rules for Good Research: A Ten-Point Guide for Social Researchers*, Open University Press, Buckingham, pp.21-22.

⁷ In this context a methodology refers to the middle ground between philosophical paradigms and discussions of method. See Schwandt at note 2, p.161.

⁸ Sarantakos, S., 2005, *Social Research*, 3rd edn, Palgrave Macmillan, Basingstoke, p.5.

⁹ Denzin and Lincoln at note 5, p.3. They discuss the 'eight moments' in qualitative research in considerable depth.

¹⁰ Glicken, M.D., 2003, *Social Science: A Simple Guide*, Pearson Education, Boston.

represent philosophical approaches that lie in the middle ground. A critical realist seeks to answer both the 'how' and 'why' questions. In terms of the underlying ontology, a critical realist sees greater complexity in the relationships under study, going beyond the depths of empirical realism. Researchers who subscribe to this paradigm would typically allow the research design to be driven by what was wanted to be learnt, rather than to be pre-ordained.¹¹ A pragmatist has a similar approach and freely chooses the methods, techniques and procedures that best meet the needs and purposes of the research.¹² Typically, the paradigms of critical realism and pragmatism are apparent in mixed method research where the researcher draws from both extremes of the continuum.

Axinn and Pearce¹³ argue that the dichotomous unidimensional distinction between quantitative and qualitative research is unhelpful as it is ultimately only a distinction between whether or not the data is coded into numbers or into text. However, this argument does seem to be too simplistic as it is clear that different methodologies produce different research designs and follow different conventions that are expected by both our peers and our wider audiences. Sarantakos¹⁴ explains that it is the underlying theoretical structure that characterises the research and that this is shaped by the different ontological and epistemological prescriptions. That is, methodology is much more than, and cannot be determined by, the nature of data presentation.

No doubt by now readers from a law background are asking – 'but where do we fit?' It is true, the body of literature on research design is focused heavily on the sciences and the social sciences and there is considerably less said that appears directly relevant to the discipline of law. In terms of paradigms, it is reasonable to posit that legal research could be anywhere on this continuum. It could be positivist and employ a quantitative methodology based on empirical evidence – for example, how often has the law changed this century, how often has a particular section been the subject of a legal dispute, how long are various sections, or how easy is a piece of legislation to read? In contrast, it could be interpretivist and based on social construct – for example, what impact has the introduction of a baby bonus had on fertility rates in Australia, or has the farm management bond scheme helped primary producers better manage their financial self-reliance? Critical realism or pragmatism may well offer a more comfortable paradigm fit for legal researchers than either of the two extremes of positivism and interpretivism. An alternate view is to posit legal research as a different paradigm altogether (for example, socio-legal research¹⁵ and possibly one that lies outside of the standard continuum that has been discussed thus far. This notion of best fit and alternative paradigms in the context of legal research will be considered further in part 5.

Putting aside these philosophical differences, it is readily apparent from the literature that different disciplines do fit more naturally under different paradigms and that disciplines themselves do differ in the emphasis they place on the role and position of theory in research.¹⁶ Whilst challenging the discipline-based principles can bring

¹¹ See note 1, p.85.

¹² Creswell, J., 2003, *Research Design*, 2nd edn, Sage, Thousand Oaks, p.11.

¹³ Axinn, W.G. and L.D. Pearce, 2006, *Mixed Method Data Collection Strategies*, Cambridge University Press, Cambridge, p.3.

¹⁴ See note 8, p.29.

¹⁵ Banakar, R. and M. Travers (Eds), 2005, *Theory and Method in Socio-Legal Research*, Hart, Oxford.

¹⁶ See note 1, p.96.

about change - after all, the rules are not set in concrete - it can be a perilous pathway if the researcher has not based the chosen research methodology on a solid and defensible foundation.

Irrespective of whether or not the researcher is embarking on a ground-breaking methodological pathway or adopting a more traditional paradigm, the research must have a clear purpose at the outset. This may come from the review of the literature in which a gap of knowledge, to be addressed by the research, has been identified. This is a top-down, or funnel, approach whereby from the researcher starts by reading broadly in the area, then drills down to a very specific and narrow aspect that is to be the objective of the research. Alternatively, if there is a broader area of interest, for example, the taxation laws relating to trusts in Australia, it is still worthwhile and infinitely sensible to check the literature early in the piece to identify who else is researching in the area, whether or not there remains scope to make a valid contribution to the field (for example, by adopting a different approach to the topic), or if there is a specific gap of knowledge apparent that offers potential for further research. Further, having a purpose does help the researcher know when the point of completion has been reached and to measure (and rejoice in) progress along the way.

Once identified, the purpose of the research can be expressed as an overarching problem or research question (for example, what are the impacts of employment-related taxes in Australia?). It can then be broken down into smaller defensible issues or questions (for example, consideration could be given to taxes imposed at federal and/or state level; to impacts from different perspectives such as the community, employees, small business employers and/or large business employers; to industry sectors; to impact according to each tax measure (or limited to just one), to short-term and long-term impacts; and/or to alternative models) can help shape the design and give the researcher a clear roadmap. This process of refining the topic should help the researcher place boundaries around the research – for example, the decision may be made that the gap of knowledge exists in respect of state tax laws, and the purpose narrowed accordingly. Going forward, this can help the researcher stay on track and get to the end of the research without unnecessary detours and delays. It goes without saying that these smaller questions must all contribute to resolving the overarching problem. If not, but they are still of interest to the researcher, consider putting them aside for the time being.

With a clear vision of the purpose of the research and its dimensions, the process of designing the research is fairly straightforward. Again, the design of the research should be in such a way that:

- its methodology fits within a paradigm that is (ideally) understood and accepted by others;
- it has a fundamental framework or structure to its conduct;
- it employs appropriate strategies of inquiry or research methods;
- it allows for knowledge claims to be made that are both consistent with the strategy of inquiry; and
- it allows the researcher to answer the research question(s) and meet his or her objectives.

In both traditional and non-traditional research designs the researcher needs to be critically reviewing the chosen methodology looking for its flaws and weaknesses and

seeking to address them until personally satisfied and confident in the appropriateness of the chosen methodology.

Apart from issues of appropriateness, it is fundamental that there is clear alignment between research problem and research design if the research is to have theoretical rigour.¹⁷ It is advisable not to lock in prematurely to a particular design, particularly without considering the alternatives. To achieve the best alignment, it may mean that the researcher needs to develop new skills or else revise the purpose of the research and/or the research questions at the outset. Are they achievable? Time spent at the outset revising and refining all the design details is well worth the investment and can save a lot of headaches down the track. One important design detail is (are) the strategy(ies) of inquiry to be employed and there exists a vast array of possibilities. The next four parts of the paper explore a range of them in more detail according to their underlying philosophical paradigm or methodology – be it quantitative, qualitative, legal research or mixed method.

3. QUANTITATIVE METHODOLOGY

Research designs that embrace quantitative methodology typically use various forms of experiments and surveys as their main strategies of inquiry. These strategies are generally appropriate where the purpose of the research is to relate or compare variables. The purpose generally is to identify if a cause and effect relationship exists, and then to make generalisations about the relationship in the context of a broader population. For example, is the level of taxpayer compliance related to the tax rate? Typically, null hypotheses are developed for each relationship being tested and the findings are written up as empirical evidence (hypotheses either accepted or rejected) with a measure of statistical significance and a specified confidence level.

An important design in both experiments and surveys detail is the correct identification of the variables under consideration and their nature. Variables are usually described as being independent, dependent or mediating. An independent variable is one that (is likely to) cause the outcome under consideration. That is, tax rate could affect compliance behaviour. Dependent variables are the outcome variables and they depend on the independent variables. For example, the dependent variables in this example could be non compliance and compliance; or unintentional non compliance, intentional non compliance and compliance, and so on. Mediating variables are those that mediate the effect of the independent variable on the dependent variable, but are not expected to be the cause. For example, the age of the person being studied or their level of education could have a mediating effect on the outcome. By stating the relationships to be studied (aligned, of course, to the purpose of the research), the variables can be identified and categorised, thus allowing the researcher to design the experiment or the survey accordingly. At this stage it is often quite helpful (for both researchers and readers) to have a diagrammatic representation of the relationships being tested.

Identifying the relationships to be tested does require careful consideration to ensure that ‘causes’ are not confused with ‘mediating factors’ given that no variable is necessarily fixed in terms of its type. The literature may provide some guidance in this respect. For example, in respect of age, level of education and so on (commonly

¹⁷ Liamputtong, P. and D. Ezzy, 2005, *Qualitative Research Methods*, Oxford University Press, Melbourne, p.38.

referred to as the demographics) findings in the tax compliance literature are noticeably mixed and on this basis it is reasonable to conclude that the demographics are mediating variables rather than the 'cause' of a particular type of behaviour. It may be possible to control the independent variable (for example, using a fixed tax rate or set tax brackets) to look more closely at the influence of the mediating variables. An alternate view is that it is difficult to control variables such as taxpayers' attitudes and this may be an inherent design weakness in this methodology.

Selecting populations and then getting access to their data are typically problematic in tax research. How does one identify the population of tax evaders if they are to be surveyed so as to understand what drives their behaviour or the extent of their evasion? Whilst the tax authority may have knowledge of the level of tax evasion, by its very nature it will be imperfect, and regardless, access to this type of data by external researchers is unlikely. What happens in practice in experimental tax research is that university students are typically used as subjects. This in itself may be an inherent weakness and a valid reason for interpreting and generalising experimental results with caution. However, Alm and Jacobson¹⁸ argue that there is no reason to believe that the cognitive processes of students are different to those of 'real' people. Further, they state that there is now much evidence that the experimental responses of students are seldom different than the responses of other subject pools.¹⁹

There are a number of types of experimental designs.²⁰ An experiment means modifying something in a situation, then comparing the outcome with what existed without modification.²¹ A full experimental design (in the real world) is very expensive and not always possible.²² In comparison, a controlled experiment can be a relatively inexpensive means to collect data and has been successfully used in tax research.²³ A controlled experiment needs subjects to be randomly assigned to at least two groups, one being the control group. An alternative approach is to use a quasi experiment where there is no control group nor necessarily random assignment, and commonly pre and post tests are used to measure the effects of a range of treatments (as causes).²⁴ Typically the results of an experiment can only be discussed in relation to the group of subjects and any more broader generalisations made would need to be suitably qualified. It should also be recognised that the attitudes being tested may be so personal that subjects might be unwilling (due to social undesirability²⁵) or unable to disclose them in response to direct questioning.

¹⁸ Alm, J. and S. Jacobson, 2007, "Using Laboratory Experiments in Public Economics", *National Tax Journal*, Vol. LX No. 1, pp.129-152.

¹⁹ See note 18 at p.143.

²⁰ For a comprehensive review of tax experimental design studies see Alm, J., 1991, "A Perspective on the Experimental Analysis of Taxpayer Reporting", *The Accounting Review*, Vol. 66 No. 3, pp.557-593.

²¹ Neuman, W.L., 2003, *Social Research Methods: Qualitative and Quantitative Approaches*, 5th edn, Allyn and Bacon, Boston, p.238.

²² Stouffer, S.A., 1950, "Some observations on study design", *The American Journal of Sociology*, Vol. 55 No. 4, pp.355-361.

²³ For example, see Baldry, J., 1987, "Income Tax Evasion and the Tax Schedule: Some Experimental Results", *Public Finance*, Vol. 42 No. 3, pp.357-383.

²⁴ For example see Loo, E.C., 2006, "Tax Knowledge, Tax Structure and Compliance: A Report on a Quasi-Experiment", *New Zealand Journal of Taxation Law and Policy*, Vol. 12 No. 2, pp.117-140.

²⁵ See note 13, p.51.

In the social sciences generally there is the danger to over-interpret, to ignore the other uncontrolled variables that occur in experiments. Stouffer²⁶ uses an example of comparing the attitudes of two groups of men from the armed services at the same point in time. One group is from the infantry and the other from the air corps. He points out that we cannot know assuredly how much of the difference in attitudes between the two groups can be attributed to the experience in the given branch of service and how much is a function of the attributes of the personnel. True, we can try to rule out various possibilities by matching, comparing men from the two groups on the basis of their age and education for example. But there is all too often a wide-open gate through which other uncontrolled variables can march. 'One lone study, however well designed, can be a very dangerous thing if it exploited beyond its immediate implications.'²⁷ Further, Stouffer expresses a 'central brooding hope that we will have the modesty to recognise the difference between a promising idea and proof.'²⁸ These cautionary comments are of course equally applicable to both experimental and survey designs.

Surveys are generally described as being structured, semi-structured or unstructured and can be conducted via various means including electronic, telephone, in person or by mail. Structured or semi-structured surveys are generally aligned with a quantitative methodology and they are the focus of the discussion in this part whilst unstructured surveys are discussed in more detail in the following part. A structured or semi-structured survey can be used for collecting both empirical and non-empirical data (usually by means of open-ended questions). It can be a useful tool for testing hypotheses and making statistical generalisations to broader populations. However, as a strategy of inquiry, a survey does have inherent weaknesses. Typically there is much attention given to the sample population, its size, sampling rate and means of selection, and to the issues of bias, response rates (i.e. the reliability of the findings) and external validity (i.e. ability to generalise more broadly) as reflections of the objectivity of the research.²⁹

Possibly a more critical issue in survey design is not its objectivity, but the design of the instrument itself. Surveys presume that all the questions and possible answers are known prior to the questions being asked.³⁰ Further, as Axinn and Pearce note, whilst a key feature of surveys is standardized questions, respondents' interpretation of the questions is not standardised.³¹ Further, they assert that this high level of structure makes it difficult to use a survey to uncover completely new hypotheses.

There are clearly means by which the design of survey instruments can be improved. Whilst pilot and cognitive laboratory testing can be used to some extent to identify and correct misunderstandings,³² the importance of developing a survey instrument that has clear, unambiguous and useful questions is a critical first step towards improving

²⁶ See note 22, p.357.

²⁷ See note 22, p.358.

²⁸ See note 22, p.361.

²⁹ For more detailed discussion on these issues and the conduct of surveys see for example Fowler, F., 1993, *Survey Research Methods*, 2nd edn, Sage, Thousand Oaks.

³⁰ See note 17, p.57.

³¹ See note 13, p.4.

³² See for example, McKerchar, M., 2003, *The Impact of Complexity Upon Tax Compliance: A Study of Australian Personal Taxpayers*, Australian Tax Research Foundation, Sydney.

its internal validity³³ (that is, that the findings have not been affected by the instrument). After all, even when a high level of reliability is achieved, a lack of internal validity must cast doubt on any findings. De Vaus³⁴ provides the following useful checklist (and extracted advice) for designing surveys which would be readily applicable in the context of tax research:

1. Is the language simple? (i.e. avoid jargon and technical terms)
2. Can the question be shortened? (i.e. the shorter the question the less confusing and ambiguous it will be)
3. Is the question double-barrelled? (e.g. how often do you visit your parents?)
4. Is the question leading?
5. Is the question negative? (e.g. marijuana should not be legalised. Agree/Disagree)
6. Is the respondent likely to have the necessary knowledge?
7. Will the words have the same meaning for everyone?
8. Is there a prestige bias in the question? (e.g. exaggerate income)
9. Is the question ambiguous?
10. Do you need a direct or indirect question?
11. Is the frame of reference for the question sufficiently clear?
12. Does the question artificially create opinions?
13. Is personal or impersonal wording preferable?
14. Is the question wording unnecessarily detailed or objectionable?

Structured and semi-structured surveys have been described as rather blunt instruments for information gathering. They are powerful in producing statistical generalisations to large populations. They are weak in generating rich understanding of the intricate mechanisms that affect human thought and behaviour. Other methods (typically from the interpretivist paradigm) are preferred for that purposes.³⁵

4. QUALITATIVE METHODOLOGY

The body of literature on research conducted using qualitative methodology, and on the paradigm itself along with its strategies of inquiry, continues to grow and develop. Whilst much of the work on the philosophy of interpretivist research has been undertaken in the United States and the United Kingdom, there are some excellent texts by Australian authors referred to in this section of the paper.

Liamputtong and Ezzy describe writing a guide book for qualitative research as akin to “trying to write a guide for writing poetry. There are certain rules and conventions, pitfalls to be avoided, and good examples to follow. However, in the end, the best qualitative research depends on the creativity and insights of the researcher themselves. The researcher must find the best way of studying how meanings and interpretations are constructed in their particular substantive research area.”³⁶ While

³³ See note 14, p.85.

³⁴ De Vaus, D.A., 1990, *Surveys in Social Research*, 2nd edn, Allen and Unwin, Sydney.

³⁵ Groves, R.M., Fowler Jr, F.J., Couper, M.P., Lepkowski, E., Singer, E. and R. Tourangeau, 2004, *Survey Methodology*, Wiley, New York.

³⁶ See note 17, p.2.

their text is generally written in the field of health, it is primarily social research and thereby of relevance to taxation.

Within the qualitative methodology, a number of theoretical frameworks exist including ethnography, case study, narrative, phenomenology and grounded theory and each of these (and their associated strategies of inquiry) is discussed herein in more detail.³⁷ Other theoretical frameworks do exist, such as symbolic interactionism, feminism, postmodernism and hermeneutics, and no doubt more will emerge over time. Given the transitional state of qualitative methodology, Liamputtong and Ezzy make the important point that a qualitative researcher should not assume that the particular theory and research method used in their project will be understood by all other qualitative researchers. That is, if a piece of research is to be meaningfully understood and assessed by other qualitative researchers, the researcher must explicitly state the theoretical tradition and methodological criteria employed.³⁸ This point is aptly illustrated by the use of 'case study' in the literature as both a theoretical framework and also as a strategy of inquiry in the tradition of Yin.³⁹

An ethnography is the study of an intact cultural group in a natural setting over a prolonged period of time, primarily by collecting observation data, but can also include data collected by focus groups or in-depth interviews. An ethnography attempts to interpret and present its findings from a cultural perspective and is more so associated with anthropological research than with social research. Given its conditions of conduct, its application to tax research is limited.⁴⁰

A case study framework explores in depth a program, an event, or one or more individuals. Creswell describes a case study as being bound by time and activity, with researchers collecting detailed information using a variety of data collections procedures over a sustained period of time.⁴¹ An example of this type of framework being successfully used in the context of tax research in Australia is the study by Wallschutzky and Gibson on small business compliance costs wherein participating business owners diarised their compliance activities over a twelve month period and underwent in-depth interviews on a regular basis.⁴²

Creswell describes narrative research as a theoretical framework in which the researcher studies the lives of individuals and asks one or more of them to provide stories about their lives. Typically, the 'data' would be collected by written account or by an in depth interview. This information is then 'restored' by the researcher into a narrative chronology.⁴³ For example, the lives of former Treasurers or former Taxation Commissioners could be the subject of narrative research. Liamputtong and Ezzy explain that in this approach the researcher typically works with the interview or the person's biography as a whole, making sense of the story, rather than breaking

³⁷ See note 12, pp.9-11.

³⁸ See note 17, p.29.

³⁹ Yin, R., 1989, *Case Study Research*, Sage, Newbury Park CA. For an example of its use in tax research see McKerchar at note 32.

⁴⁰ See note 12, p.14.

⁴¹ See note 12, p.15.

⁴² Wallschutzky, I.G and B. Gibson, 1993, "Small Business Cost of Compliance", *Australian Tax Forum*, Vol. 10 No.4, pp. 511-543.

⁴³ See note 12, p.15.

down the text into discrete parts or collecting information on observations at interview for subsequent analysis.⁴⁴

Phenomenology seeks to identify the ‘essence’ of human experiences concerning a phenomenon, as described by the participants in a study.⁴⁵ The emphasis is on gaining an understanding of the situation from the perspective of those who have experienced it.⁴⁶ For example, the phenomenon of the cash economy could be studied with this framework and an understanding of what drives people to participate could be gleaned. Other examples in a tax context could be the study of new businesses and how they engage with the tax system, or the study of the impact of the baby bonus on taxpayer’s decisions regarding family, or the impact of a tax audit on taxpayer compliance behaviour. Focus groups, in-depth interviews, semi-structured interviews, or Yin’s case study protocol⁴⁷ would all be possible strategies of inquiry in a phenomenology study.

In grounded theory the researcher attempts to derive a general, abstract theory of a process, action or interaction that is ‘grounded’ in the views of participants in a study. Two primary characteristics of this design are the constant comparison of data with emerging categories and theoretical sampling of different groups to maximise the similarities and differences of information.⁴⁸ Strauss and Corbin argue that given the way in which grounded theories are constructed, they are likely to offer insight, enhance understanding and provide a meaningful guide to action.⁴⁹ The challenge for the researcher is to be open-minded, to listen to and hear what is being said and to interpret it as honestly as possible, always checking and rechecking for other possible interpretations. In the context of tax research, a recent scoping study into the drivers of compliance cost for small businesses in Australia successfully used this framework with data being collected primarily by the conduct of in-depth interviews with a range of businesses and accountants using a combination of theoretical and convenience sampling.⁵⁰

Against this background of theoretical frameworks from the qualitative methodology, the strategies of inquiry now deserve attention. Qualitative research is typically about seeking answers to questions (and sub questions) and not about proving or disproving hypothesis. The strategies of inquiry generally seen in qualitative tax research are in-depth interviews and focus groups and their purpose is to collect what is often described as ‘thick’ data. That is, it is the rich information that the researcher is looking for that doesn’t fit into Likert scales, the data that will help the researcher explore the complexity of the research problem and build an understanding or an interpretation. Observation is also a method of data collection that could be

⁴⁴ See note 17, p.125.

⁴⁵ See note 12, p.15.

⁴⁶ See note 17, p.19.

⁴⁷ See note 39.

⁴⁸ See note 12, p.14.

⁴⁹ Strauss, A. and J. Corbin, 1998, *Basics of Qualitative Research*, 2nd edn, Sage, Thousand Oaks CA. p 12.

⁵⁰ McKerchar, M., Hodgson, H. and M. Walpole, 2006, “Scoping Study of Small Business Compliance Costs”, Report to the Board of Taxation, Atax, 31 August.

appropriate in tax research, such as examining subjects' tax returns or record keeping procedures, or the level of stress they exhibit in solving a tax problem.⁵¹

Within this methodology, in-depth interviews can be semi-structured or unstructured and the interviewer can be passive, empathetic, probe or even provocative. It is essential that the interviewer be skilled and well able to build rapport with interviewees. In practice, this strategy is normally conducted by the individual researcher rather than by an assistant. Questions are not normally constructed or standardised in advance (as in the structured survey of the positivist paradigm), but instead the research has a theme list as a guide to establish the topic at interview. The interview is then shaped according to the experiences of the subject. Questions do need to be open ended and to ebb and flow based on the experiences being told and, in the case of grounded theory, the emerging theories under consideration. Subject to ethical considerations, taping of interviews can be helpful to the researcher but may hinder the frankness of the interviewee.

Typically, the interviewer listens actively, takes copious notes, both of what the interviewee has said and also of the interviewer's own observations about what was said. Where there is more than one interviewer, it is important that they regularly debrief and discuss what they believe was being said. Whilst there are many software packages (with variations as to the form of data to be input that ideally should be considered before note taking takes place) to help with text analysis, depending on the number of interviews they may or may not be necessary. Given the time required to conduct in-depth interviews, their numbers may well be limited. If text data is to be analysed manually, it does help strengthen the findings if it is done in a systematic fashion. Liamputtong and Ezzy advocate the use of open coding, axial coding and selective coding as a three-stage methodical and defensible approach to plotting the story.⁵²

Open coding, or 'first run' is used to identify the dimensions (relationships, events, patterns or themes) that could help conceptualise the organisation of the data overall. Axial coding requires closer examination of the details within one dimension, while selective coding takes a bigger picture view again, but with more of the detail. It may be necessary to work both backwards and forwards through the stages, refining the coding and reconfiguring the framework as required. Miles and Huberman⁵³ explain that this conceptual organisation of data can be done in a matrix form and this was the approach adopted in the study by McKerchar, Hodgson and Walpole.⁵⁴ In this study there were seven matrices refined into one meta-matrix which was framed around five emerging themes. A grounded theory was then constructed for each of these themes on the basis of the data analysis.

Focus groups (or a focused interview) came into being after WW2 and have been used extensively since the early 1980s in many fields, predominantly in public health, but

⁵¹ See for example note 32; Woellner, R., Coleman, C., McKerchar, M., Walpole, M. and J. Zetler, 2007, "Can Simplified Legal Drafting Reduce the Psychological Costs of Tax Compliance?: An Australian Perspective", *British Tax Review*, No.6, pp.717-733.

⁵² See note 17 at p.268.

⁵³ Miles, M. and A. Huberman, 1994, *Qualitative Data Analysis*, 2nd edn, Sage, Thousand Oaks.

⁵⁴ See note 50.

their use is also apparent in tax research.⁵⁵ They are regarded as useful when a researcher wishes to explore peoples' knowledge and experiences, focusing on a narrow range of ideas about and acceptances of new programs and evaluating and solving specific program problems and evaluating health programs.⁵⁶ A focus group is not a group interview, but a stimulating group interaction and this can give it a powerful extra dimension as a strategy of inquiry. The setting needs to be relaxed and informal and the participants need to feel uninhibited without any one dominating or withdrawing from the group. Properly managed (in terms of the moderator, the length of time, the nature of the questions, and the recruitment of participants), it can be a very rich source of data. Ideally, participants in a focus group can feel greater confidence in the group setting and can also corroborate or challenge the responses of other members and remind one another of certain phenomena.⁵⁷ However, a focus group does not allow the views of any one individual to be fully explored and the group can be led astray by more vocal participants. Similar processes for data collection as described for in-depth interviews would also be appropriate for a focus group strategy. Further, some tax-related topics (such as tax evasion) may be too sensitive for focus group strategies, or may be beyond the scope of the group's experiences unless purposively chosen (e.g. international taxation).

The qualitative methodology does not lend itself to the making of statistical generalisations consistent with positivism. Instead, its proponents seek to make analytical generalisations or interpretations about a process, rather than its outcomes. They are seeking understanding and explanations rather than a definitive answer about the size and scale of a phenomenon. This is not to say that a researcher in the qualitative methodology should not consider the issues of validity, reliability and sample selection, indeed they should, but the means by which they address these concerns are different to that of the quantitative methodology. In terms of validation, Sarantakos explains that cumulative validation is an appropriate approach, whereby the findings of the study can be validated by various means including the support of other studies. Alternatively, communicative validation can be sought by various means including involving the participants in a subsequent review process, or by employing expert external audits, or by using a Delphi "group of experts" consultative approach.⁵⁸

Reliability (either internal or external) as an issue tends to receive little attention from qualitative researchers. However, it can be addressed to some extent by making greater use of debriefing of interviewers when more than one is used, and/or by peer review. The representativeness of sample selection in the qualitative methodology is generally considered to be irrelevant and unimportant, with Miles and Huberman holding a contrary view.⁵⁹ Unimportance is generally inferred as qualitative studies have no real capacity for their findings to be generalised beyond the boundaries of the sample used. However, it is true that the higher the extent of generalisability, the

⁵⁵ For example, see Coleman, C. and L. Freeman, "Cultural Foundations of Taxpayer Attitudes to Voluntary Compliance", *Australian Tax Forum*, Vol.13 No.3, pp.311-336; Woellner *et al* at note 51.

⁵⁶ See note 17 at p.79.

⁵⁷ See note 13 at p.7.

⁵⁸ See note 8 at p.86. For an example of the use of a Delphi strategy see Evans, C. 2007, "Unravelling the mysteries of the oracle: using the Delphi methodology to inform the personal tax reform debate in Australia", *eJournal of Tax Research*, Vol.5 No.1, pp.105-134.

⁵⁹ See note 53.

higher the value of the study so the issue is one that warrants further attention. Hence the temptation to overgeneralise and to make conclusions that cannot be substantiated.

Sarantakos does discuss strategies by which generalisability of qualitative findings can be improved, such as conducting multi-site research.⁶⁰ Yin's repeat application of a case study protocol and the study of consistency in emerging patterns is another means by which greater generalisability of findings may be achieved.⁶¹ Drawing conclusions that can be substantiated does require researchers to demonstrate methodological rigour. This can be challenging given the fluidity of the methodology. Herein lies some of the inherent weakness of the qualitative methodology and they need to be recognised. On a more positive note it can be a very rewarding experience for the researcher and indeed a privilege, but one that does come with responsibilities to honour the trust placed in the researcher by those who have willingly shared their experiences for the sake of the cause.⁶²

5. LEGAL RESEARCH

Legal research has somewhat lagged behind quantitative and qualitative research when it comes to philosophical paradigms and acceptable conceptual frameworks. Salter and Mason⁶³ explain that this resistance to methodological discussion in the discipline of law can be attributed to a misconception that legal research is about the technical process and the acquisition of skills (such as identifying relevant case and statute law) rather than about the methodology or different approaches by which the objectives and goals of the researcher can be addressed. They posit that, in the case of student dissertations, they can only be improved by reflecting critically upon, and then justifying explicitly, the appropriateness of their methodologies given their research questions and topic.⁶⁴

In the Australian context, Hutchinson⁶⁵ writes that every legal research project is based on an underlying paradigm, but that it is often unarticulated and the research may end up being dysfunctional as a result. She describes the legal research paradigm as a unifying rationale that gives direction and guidance, and notes that success within the discipline tends to be measured within the paradigm boundaries. That is, there is an expectation that legal research is conducted on a methodological basis that is understood by others and can be justified. Clearly the fundamental principles of legal research in respect of making appropriate choices (of questions, methodology and strategies of inquiry) and having alignment throughout the research design are not inconsistent with the expectations of quantitative or qualitative methodologies.

Still, the legal research paradigm is somewhat elusive. The Pearce Committee⁶⁶ identified the two typologies of legal research as being doctrinal or non-doctrinal. Doctrinal research is described as the traditional or 'black letter law' approach and is typified by the systematic process of identifying, analysing, organising and

⁶⁰ See note 8, p.98.

⁶¹ See note 39.

⁶² See note 17, p.72.

⁶³ Salter, M. and J. Mason, 2007, *Writing Law Dissertations*, Pearson Education Ltd, Harlow Essex.

⁶⁴ See note 63, p.31.

⁶⁵ Hutchinson, T., 2005, *Researching and Writing in Law*, 2nd edn, Thomson, Sydney.

⁶⁶ Pearce, D., Campbell, E. and D. Harding, 1987, *Australian Law Schools: A Discipline Assessment for the Commonwealth Tertiary Education Commission*, AGPS, Canberra.

synthesising statutes, judicial decisions and commentary.⁶⁷ It is typically a library-based undertaking, focused on reading and conducting intensive, scholarly analysis.

In contrast, non-doctrinal research is characterised as research ‘about law’ rather than ‘in law’ and employs the methodologies commonly used in other disciplines. It follows that the data used in non-doctrinal research is not limited to the traditional legal sources. The Pearce Committee further divided non-doctrinal research into reform-orientated and theoretical research. Reform-orientated research is designed to accomplish change in the law, and theoretical research is that which fosters a more complete understanding of the conceptual bases of legal principles. It is difficult to identify a philosophical distinction per se in these variations as each appears to have merit and is no doubt more ably suited to a given research purpose. This point is ably demonstrated by Bentley⁶⁸ in his thesis that employs theoretical research in its early chapters to gain an understanding of the conceptual bases of the relevant legal rules and principles, followed by doctrinal research to critically evaluate the legal rules and their interrelationship using both induction and deduction. The underlying thread of the thesis is to propose reform by providing recommendations for change, based on critical examination.⁶⁹

Salter and Mason describe the existence of tension or friction between the traditional black letter academic lawyers⁷⁰ and the concern of many black-letter academics to the apparent threat posed by socio-legal and other alternate approaches appearing as the Trojan horses. It is undoubtedly more productive to not consider the various methodologies as being engaged in a competition of sorts and thereby mutually exclusive, but as providing opportunities for researchers to design and conduct exemplary research that makes a worthwhile contribution to the body of knowledge. It is encouraging to note that in the Australian context, Hutchinson does make the encouraging assertion that the legal paradigm is changing to a more outward looking focus encompassing interdisciplinary approaches to methodology.⁷¹

Whether doctrinal and/or non-doctrinal strategies of enquiry are adopted, there are conventions to be followed, or at the very least, rationales to be provided. It is clear that legal research needs to be systematic, purposive and have a robust framework. For example, the analytical framework employed by Walpole⁷² in his thesis (using a doctrinal strategy) on the reform of the taxation of goodwill was that of the desirable characteristics of a tax structure as identified by the Meade Committee Report (as the most sound and accessible modern attempt to do so).

6. MIXED METHODS

The discussion thus far has focused on the various methodologies as stand-alone approaches, but this is not necessarily the case. Further, there are other paradigms including critical realism, pragmatism, action research (used successfully in education and with potential application in tax research), feminism, queer theory, symbolic

⁶⁷ See note 66, p.309.

⁶⁸ Bentley, D., 2006, “A Model of Taxpayers’ Rights as a Guide to Best Practice in Tax Administration”, PhD thesis, Bond University.

⁶⁹ See note 68, pp.6-7.

⁷⁰ See note 63, p.35.

⁷¹ See note 65, p.21.

⁷² Walpole, M., 2006, “Proposals for the Reform of the Taxation of Goodwill in Australia”, PhD thesis, UNSW.

interactionism, structuralism, postmodernism and methodological anarchism (in which there are no valid methodological rules, hence 'anything goes').⁷³

While each strategy of inquiry or typology can be applied in its own right, there is increasing evidence of researchers drawing from more than one paradigm, or using more than one strategy of inquiry from within the same methodology. Each strategy has its strengths and weaknesses and the drive for mixed method research (i.e. using more than one strategy of inquiry to collect data) is to use one strategy to either inform, validate or compensate for the weaknesses of another. In terms of both research design and conduct, mixed method can offer real scope for researchers who are driven by the purpose of the question and not necessarily bound by any one paradigm. It can help fill in some of the gaps, but may not necessarily be superior than a single strategy of inquiry if it is not well designed and conducted.

Multiple methods of data collection require multiple methods of data analysis, and Creswell explains this makes it more important than ever that there is a rationale for the overall design that will purposively and systematically guide the processes.⁷⁴ There are key decisions to be made about the order of implementation of the strategies, the priority given to the various strategies, and the point of intersection of the strategies.

The implementation of strategies of inquiry may be sequential or concurrent and serve different purposes (such as to explore, to explain or to corroborate). Collecting data sequentially allows the findings from one strategy to inform another. For example, a focus group could be used to discover the key concerns of tax agents and this could then be used to develop a survey to be conducted with the total population of tax agents so that statistical generalisations could be made (i.e. qualitative methodology used to inform the design of a quantitative methodology). Further, in-depth interviews could be conducted with individual small business owners as to their compliance concerns, followed by focus groups with industry groups to gauge the extent to which findings may be applicable more broadly (i.e. a qualitative methodology used to validate another qualitative methodology). Another example could be to first conduct a large-scale survey about taxpayers attitudes followed by a case study protocol to explore how these attitudes might be formed (i.e. a quantitative methodology followed by qualitative methodology). The possibilities are endless and are not limited to just quantitative and qualitative methodologies. Similar designs could include legal research, for example, a focus group to gauge employers' reactions to a proposal to proposed legal reforms arising from doctrinal analysis.

While these examples are of sequential applications, concurrent applications can be readily envisaged. For example, a project recently conducted by the Inspector General of Taxation (IGT) into revenue bias in private binding rulings had at least two concurrent strategies of inquiry – one being an evidence-based technical investigation (conducted by the IGT) and the other being an investigation into the perception of bias held by applicants (conducted by Atax).⁷⁵ This is an example of a problem being able to be divided into components which do not inform each other, but nevertheless do contribute to the overall solution. Alternatively, the independent findings of concurrent strategies directed at addressing the same problem (or the same component

⁷³ See note 8, p.31.

⁷⁴ See note 12 at p.208.

of a problem) can be subsequently compared thereby perhaps achieving a triangulation of findings and improving the validity of the research.

It follows that different strategies may well receive different emphasis within the research, depending on their purpose in the overall design. This is not problematic provided that there is a sound rationale. Similarly, there is a point of intersection of data from the strategies. This can be at the collection phase (such as a survey that includes open-ended questions, or observational and textual data collected at interviews), or at the analysis stage (such as in the thematic approaches of matrix analysis or case study protocols), or in the forming of conclusions and recommendations. Creswell does explain that there does need to be a binding theoretical perspective to mixed method research, whether it be a participatory lens (e.g. small business taxpayer; private binding ruling applicants) or the advocacy lens (e.g. reform of anti-avoidance legislation), as this theoretical perspective helps guide the framework.⁷⁶ That is, the possible combinations for mixed method research are almost unlimited.

The key challenge in mixed method design is to develop and conduct the best and most appropriate combination of strategies. In theory it does offer potential to shed deeper light on perplexing problems, in particular, on the nature of causal relationships, but it still may not necessarily provide all the answers given the complexities of human behaviour.⁷⁷

7. CONCLUSIONS

Throughout this paper there are a number of consistent messages. The overarching message is that there are many approaches to research and no one methodology is necessarily better than another. Each methodology and each strategy of inquiry has its strengths and weaknesses. However, different methodologies and different strategies are more suited to answering different types of research problems. Further, researchers will have personal preferences for different approaches depending on their ontology and epistemology, whether implicit or explicit.

A researcher must adopt a strong conceptual framework for a piece of work and be prepared to justify that framework to others. The process of justification is more straightforward where the research adopts a philosophical approach that is readily understood by others, and a known strategy of inquiry that fits under the paradigm umbrella. That is, it is an easier pathway if the research can be placed in an existing paradigm and to do this the researcher needs to be familiar with the relevant literature. This does not mean that a researcher cannot develop a new systematic and rigorous framework for the design and conduct of research, but if this is done, then the researcher must be prepared to defend it.

Whichever methodology is adopted, it is fundamental that research has a clear purpose and that there is strong alignment between purpose and design. If the design cannot be tweaked, then the purpose may need to be. Ideally the purpose should drive the design, but in practice the researcher may well have to compromise to ensure that the purpose is achievable given existing constraints (for example, time and/or resources). Time spent on developing a strong and robust design is time well spent, as is time

⁷⁶ See note 12, p.213.

⁷⁷ See note 13, p.17.

spent on documenting the design and conduct of the research. Attention to detail is paramount.

Further, there are also other important messages albeit at a more applied level. Knowledge claims must be consistent with the strategy of inquiry adopted (or strategies in the case of mixed method) and with the evidence. The researcher needs to take time to reflect on the analysis and to consider alternate interpretations. Modesty needs to be exercised in making claims and in not overstating the contribution made by the research. Give due recognition to the strengths and weaknesses of the research. No design is perfect, and even if it was as close to perfect as can be, things do not always go according to plan. For example, the response rate to a survey may be lower than expected. The researcher needs to explain what has been done to address the problem (the possibility of which should have been considered in the planning phase) or to minimise its effect. It does not necessarily mean the research is ruined, but the researcher may have to consider alternate strategies and/or suitably qualify the findings. Be prepared for the unexpected, or at least, be prepared to find a solution.

Research is almost never conclusive. There are always some nagging questions or areas that emerge but that are realistically beyond the scope of the research. A good researcher will reflect on his or her contribution and identify these areas that need further investigation, leaving a trail for others. The interest of others needs to be stimulated if they are to prove or disprove the proposed theories and continue to contribute to the discipline, or even to create a new one. A good example of this is the pioneering work of the late Cedric Sandford and the recognition by his peers of his creation of compliance costs as an academic discipline.⁷⁸ While very few tax researchers will ever be honoured in this way, seeking peer review by presenting and publishing works is an important aspect of the process of developing as a researcher and in encouraging and inspiring others.

⁷⁸ Evans, C. Pope J. and J. Hasseldine, (Eds), 2001, *Tax Compliance Costs: A Festschrift for Cedric Sandford*, Prospect, Sydney.

Avoidance and *abus de droit*: The European Approach in Tax Law

Marco Greggi*

Abstract

Defining tax avoidance has been always a nearly impossible quest for tax lawyers. In Continental Europe, however, it could be said that the notion of "avoidance" is strongly embedded to the concept of "abuse" of a right, being an "abuse", according to the Roman law tradition, the exercise of a right inconsistently with the general principles of correctness, good faith or even with the basic rules of ethics. Therefore the definition of avoidance is not purely legal, but it depends also on other disciplines which influence it. In EU law qualifying avoidance is even more complex as the law of the Union is nearly deprived of any influence by other systems of values, the only general principles to rely on are the fundamental freedoms and the nondiscrimination principle both enshrined in the Treaty. That is why when in recent cases the Court had to rule on the "abuse of law" it did its best to find principles or values to build the concept on. The results were different in direct tax and in VAT cases: the notion of abuse can be wider in some of them and narrower in some others as far as it depends on the background to the specific provisions discussed. In case of the VAT the system is more complex and the court has to strike a balance between the need for neutrality of the tax and the coherence. In direct tax cases only the Treaty is applicable, and together with it the freedoms bestowed upon the taxpayer who can make the most of them in any case, unless the purpose of the scheme implemented is only to save taxes.

1. INTRODUCTION, THE ROOTS OF LEGAL AVOIDANCE AND ABUSE OF LAW IN THE CONTINENTAL LEGAL SYSTEMS

The concepts of tax avoidance and tax evasion have been analysed in the past by tax academics and practitioners from all round the world. To a certain extent, those¹ who considered them to be "eternal ones" were right.

Avoidance of tax provisions is an outcome of the infinite struggle between the principles of legal certainty on one side and freedom of business activity on the other:²

* Ph. D., Senior lecturer, University of Ferrara, Faculty of Law, Italy and Research Fellow, Taxation Law and Policy Research Institute, Monash University, Australia. Email: marco.greggi@unife.it.

¹ F. Vanistendael, *Judicial interpretation and the role of anti-abuse provisions in tax law*, in G. S. Cooper (edited by), *Tax Avoidance and the Rule of Law*, Amsterdam, 1997, p. 153; P. Pistone, *L'abuso del diritto nel diritto tributario internazionale*, in (edited by V. Uckmar) *Diritto tributario internazionale*, Padova, 2005, p. 813 ss.; G. Maisto, *Multilingual texts and interpretation of tax Treaties and EC Tax law*, Amsterdam, 2007, *passim*.

² In Europe this feature is evident particularly in VAT cases. The ECJ, for instance, has underlined this conflict in *Halifax plc, Leeds Permanent Development Services ltd, Countrywide Property Investments ltd v. Commissioners of Customs & Excise*, C-255/02 at §§ 71 and 72. For a more detailed analysis, see below § 9.

between the legal form of the commercial operations and the substance of the aims pursued by the taxpayers.

The purpose of this contribution is therefore less far-reaching. The basic assumptions are that (1) the problems avoidance raises are not limited to tax law and (2) it is necessary to analyse avoidance of law in order to understand better the features and the characteristics of it when dealing with tax law. Another assumption is that (3) the EU tax harmonisation via regulations, directives, soft law and ECJ case law progressively compels the legal systems of the different member States to converge on specific characteristics necessarily found in all the tax planning schemes to be qualified as avoidant, even if these conditions might change from tax to tax. However this final result will be difficult to achieve as the fight against tax avoidance is intrinsically national in its roots, depending as it does on factors that do not necessarily belong to the tax systems; coming back in this way to (1).

Last but not least, the linguistic differences³ across Europe also play an important role in the distinction between evasion, avoidance and “abuse of law”.

Some of these barriers are easy to overcome, in so far as they depend only on the language and not on the meaning of the concepts used by the legislator or by the judges. To this extent, it is easy to consider the English “tax avoidance” as the French *évasion fiscale*, the German *Steuerumgehung* or the Italian *Elusione fiscale*. On the other hand, “tax evasion” can be compared to the French *fraude fiscale*, the German *Steuerhinterziehung* and the Italian *Evasione fiscale*, for instance.⁴

Some other differences, however, are of a theoretical nature, although strongly correlated to the different languages used: the concept of “abuse of law” is one of these.

In many cases, especially on the continent, the notion of “abuse” is used instead of that of “avoidance of a legal provision”. In this sense the abuse of law could be considered a synonym of tax avoidance in those countries where the former term is accepted.

Despite the differences, the common ground of all these issues relies on interpretation: dealing with tax avoidance or with “abuse of law” still depends nowadays both on the interpretation of statutory law in the light of the Treaty of Rome and on the Constitution of the State (if any).

³ M. Mossner, *Association of European tax professors: a common language in taxation within Europe*, in EC Tax Review, 1999, p. 158. The author noted that “A first step in developing coordinated tax systems within the Europe Union is therefore to develop a common language within Europe”.

⁴ Some misunderstanding on the continent is, however, always possible: note how the same word “*evasion / evasione*” has two completely different meanings in France and Italy. C. Sacchetto, *Multilingual issues under internal laws with a particular emphasis on tax law: the Italian experience*, in Dir. Prat. trib. int., 2007, p. 881. Another clear example of the difficulties encountered while translating legal concepts is given by P. Pistone, *Centro di attività stabile e stabile organizzazione: l’IVA richiede una evoluzione per il XXI secolo ?*, in Riv. Dir. Trib., 1999, III, p. 31.

Some authors noted that a line should be drawn between avoidance considered as the outcome of loopholes in statutes and avoidance that is the result of different interpretations of statutory law or common law.

While the difference is theoretically correct, it must be remembered that this distinction is in any case based on interpretation of the rule of law, and that understanding the rule is an *a priori* condition to ascertain any loopholes or gaps in the system.

This requirement was well known in the continental (it could be said Roman) experience since the Middle Ages, when the debates on the notion of “abuse” of law and “avoidance” of statutory law arose for the first time, even if they did not immediately involve tax issues.

The historical approach is a good starting point even for current analysis: all in all, the problems faced by the medieval commentators and lawyers in Italy and in other countries were not so different from the current ones.

They had to strike a balance between the statutory laws of the Italian municipalities and the *ius commune* (what was left of the ancient Roman law) considered to be the legal background that was always applicable if not specifically derogated. Nowadays we almost have to do the same thing while managing the relationship between national laws and EU law. The only fundamental distinction relies on the overall trend of the evolution: in the Middle Ages the *ius commune* was progressively being substituted by the *particularisme* of the different municipalities (at least in Italy); in the 21st century it is the national law which has to progressively surrender to communitarian law where necessary.

2. INTERPRETATION OF TAX LAW AND THE ORIGINS OF AVOIDANCE: A LESSON FROM ROMAN AND MEDIEVAL LAW (*NON PUTO DELINQUERE EUM, QUI IN DUBIIS QUAESTIONIBUS CONTRA FISCUM FACILE RESPONDERIT*⁵)

The fragment of the *Corpus Iuris Civilis* compiled and introduced by Emperor Justinian in AD 527⁶ and quoted at the beginning of this article must not be emphasised beyond its limit. Tax law was never considered in Roman law as an autonomous discipline, as it is nowadays. The contributions of academics and lawyers on that are scattered amongst hundreds of references that make it impossible to identify a Roman approach to tax law in terms and conditions we are used to⁷.

⁵ *Corpus Iuris Civilis* 49.14.10.

⁶ The *Corpus* was ended in 533, but it is well known that further amendments were introduced almost until the death of the emperor.

⁷ Some Authors (S. Douma and F. Engelen, *Halifax plc and Others v. Commissioners of Customs and Excises: the ECJ applies the abuse of rights Doctrine in Vat cases*, [2006] BTR 430; see also E. Van der Stok, *General anti avoidance provisions: a Dutch treat* [1998] BTR 151) stress the importance of the Roman background, quoting from the *Corpus iuris* that “In fraudem legis agere” is the behaviour of those “qui salvis verbis legis sententiam ejus circumvenit” [who without infringing the words of the law, deceive the purport thereof] (Dig. 1.3.29). *Fraus legis* is the concept still used in Dutch Tax law to fight back tax avoidance. The reference is still nowadays of extreme interest to tax lawyers, together with another one in the same *Corpus*, but in this case by Ulpianus (1.3.30) “*Fraus legis enim fit, ubi quod fieri noluit, fieri autem non vetuit*” [*Fraus legis* occurs when a certain behaviour is not wanted

However, there's enough to understand the Roman reasoning in respect to interpretation, avoidance and abuse of law. That approach constituted the base on which the medieval commentators in Italy (and first of anywhere in the world, in Bologna) built up medieval law in the form of comments (*glossa*) to the *Corpus Iuris* in the 12th century.

Basically, all the doubtful cases submitted to the lawyers were answered according to one of two rules: "*in dubio contra fiscum*" or "*in dubio pro fisco*". According to the first one, all the cases where tax avoidance⁸ could have taken place had to be resolved in favour of the taxpayer, while under the second one, the solution had to be the opposite. The application of one or the other was a mere consequence not only of the personal position of the interpreter, but also of the source of law which introduced the tax to be applied and the revenue assumed to be avoided. While the ancient custom (*consuetudo*⁹) allowed an extensive interpretation, statutory law could not.

In any case, the reasons that took (Herennius) Modestinus to answer in that way remained incomprehensible to all the medieval commentators. The first and perhaps most influential of them, Accursius, tried to give an answer that is still all but convincing: "*forte est ratio quia (fiscus) dives est*".¹⁰

For all these reasons, and despite the opinion by Modestinus in the *Corpus Iuris*, the interpretation of the tax rules across the Middle Ages was far from clear, there being possible two different approaches to tax avoidance: one relying on the form over substance approach and the other on the opposite. The application of the first or of the second basically depended on the source of laws which introduced the tax in the specific circumstances of the case.

Indeed, if the tax were due according to an ancient tradition (custom), both types of interpretation were possible, but if a statute of a municipality introduced the tax, then the strict interpretation (form over substance) had to be preferred. This was not decided according to specific tax rules, but it was done under a general principle of the *ius commune* that the specific provisions of the municipalities had to be interpreted restrictively,¹¹ considered as they were a derogation of the law.

without being vetoed – by the legislator –]. However. While the definition of *fraus* (circumvention) or abuse is clearly settled, the reader must be aware that Roman lawyers developed their conclusions in Civil law, and not in Tax law. The abuse mentioned in the *Corpus Iuris* dealt with inheritance law (*Lex Cincia de donationibus*, 204 b.C.), and it was the task of the subsequent Academics to extend this reasoning to tax law, using interpretive arguments (such as the analogy or the extensive interpretation).

⁸ Of course the two general rules mentioned above were applicable not only to tax avoidance but in all cases where the wording of the tax provision was not clear enough to produce a straightforward application. It is obvious, however, that the tax avoidance cases were the most relevant ones, as well as those where the rules mentioned played the most relevant role. Of course tax evasion was not involved in the distinction.

⁹ "*Consuetudo servatur in vectigalibus*" [Use must rule in the application of taxes], *Ad. Dig. N. I, dig. XXXIX, 4, 4.*

¹⁰ "Maybe because the Tax Administration is richer than the taxpayer" in *Digestum novum Lugduni*, 1550, col. 1478.

¹¹ See the comments on this topic to the *L. Caesar § expeditis ff. n de publicanis (ad dig. N. II, Dig. L. 4, 18).*

The consequence, as is mentioned above, was that the problem of avoidance was solved in different ways, making the substance over form (or vice versa) prevail, depending (also) on the source of the law to be applied.

This basic rule suffered some exceptions, anyway. Perhaps the most relevant one is given by the same Bartolus (from Saxoferrato) in one of his *responses* (legal advices).¹²

To a certain extent it was a *Ramsay* approach¹³ seven centuries before the decisions by the House of Lords.

3. THE LONG WAY TO A EUROPEAN APPROACH TO TAX AVOIDANCE: FROM THE ROMAN EXPERIENCE TO THE STATUTORY ANTI-AVOIDANCE PROVISIONS. THE CONTINENTAL WAY AND THE “ABUSE OF LAW” (OF *FRAUS LEGIS*)

As has been noted before, Roman law and medieval law (the *ius commune*) weren't directly involved in the phenomenon we call nowadays “tax avoidance”. The legislators and the academics in the first universities across Europe were interested in the interpretation of the law and only in a very broad sense in tax law.

For all these reasons the notion of tax avoidance was never introduced, nor the distinction with tax evasion in the sense mentioned at § 1: while the *ius commune* progressively lost its efficacy and its binding force in most of the emerging national States across the continent (as had progressively happened in the Italian municipalities before), European legal reasoning lost even the possibility of having a common approach to the same issue.

Tax avoidance was named (and defined) in different ways, according to the national experiences, but at least in theory, the distinction made by A. Hensel and E. Blumenstein in their foundational work on tax law is valid nowadays in almost every country.

According to the first author, “... In this latter case [tax fraud] the taxpayer didn't fulfil the obligation to pay the tax already due, while in the case of avoidance the taxpayer prevented the tax from being due, avoiding the legal fact, deed or contract to

¹² Consilium 135 ad L. I. It happened that a municipality introduced a tax to be paid by all those who entered the walls to sell skins and leathers in the market of the town and that the tax was due because of the occupation of the market soil by the goods to be sold. Then the administration of the municipality discovered that many merchants had begun to sell their skins while keeping them on their shoulders and arms, without occupying any portion of the market soil. They also refused to pay the tax, claiming that statutory law was to be interpreted restrictively and that the payment was due only for the occupation of the soil of the market. Obviously that was not the case.

It could be argued that this was one of the first cases of avoidance in the history of tax law. Quite surprisingly Bartolus answered that in cases like these it was necessary to distinguish leathers and skins used by the merchants for their personal purposes from the others. The tax was due for the latter; despite their being sold without occupying any portion of the market. Arguably, Bartolus in this case applied the principle of substance over form to overrule the strict interpretation of the statute introducing the tax (named *Gabella*).

¹³ *W. T. Ramsay Ltd. v. Internal Revenue Commissioner* [1981] 1 All E.R. 865.

whom tax law linked the payment of a tax ...”¹⁴ From Blumenstein’s point of view,¹⁵ however, tax avoidance occurs when “... According to a carefully planned and intentional procedure the taxpayer enters into a contract or sets up a specific operation which is capable of reducing the amount of the tax otherwise due, or to prevent it; thus far the avoidance is different from the tax evasion. In this latter case the tax is due to all the effects but the determination of its amount by the Tax Administration is prevented by an unlawful behaviour of the taxpayer ...”.

Only in the latter case, therefore, (tax evasion) do the authors qualify the behaviour of the taxpayer as unlawful, while in the former it is consistent with law although in conflict with the normality of business activity or the exercise of the same rights.

Arguably it is also here where the legal traditions of the common law countries and the civil law ones began to diverge sensibly from this specific point of view.

The need to levy taxes on the business effectively carried on was achieved by the first ones through the “business purpose test” as developed later on in the case law by English, Australian and American courts¹⁶, and by the second ones through the notion of “abuse of law”.

Even if the effect of the application of the two anti-avoidance provisions should be the same (limiting the taxpayer in the exercise of his business economic freedoms when no other goal seems to be pursued other than an improper tax saving), in common law countries the courts seek to understand what the normal business practice should include, while the continental ones seem to stress the importance of what is not reasonably included in the specific right attributed to the taxpayer.

The outcomes of the two interpretive approaches do not necessarily coincide.

As was mentioned before, Roman law vastly influences the continental interpretations summarised by the German and Swiss authors.

Under the Roman civil law, at first stage, the current notion of “abuse of law” was nonsensical (as it is probably nowadays in common law). In the case of property rights, for example, no abusive enjoyment was conceivable.

The owner of any goods either enjoyed them under law or in an illegal way.

In the second case, civil law had remedies for any person who claimed to be injured by the improper exercise of the right; in the first one, no limit could be imposed on the owner of any goods. The rule of law was clearly expressed in “*Nullus videtur dolo facere qui suo iure utitur*”¹⁷ or, even better in our perspective, Ulpianus added: “... *is qui iure publico utitur non videtur iniuriae faciendae causa hoc facere: iuris enim*

¹⁴ A. Hensel, *Steuerrecht*, Berlin, 1924. Translation by the Author from the Italian version by D. Jarach, *Diritto tributario*, Milan, 1956.

¹⁵ E. Blumenstein, *System des Steuerrechts*, Zurich, 1951. Translation by the author from the Italian version by F. Forte, *Sistema di diritto delle imposte*, Milan, 1954.

¹⁶ See for instance Y. Keinan, The many faces of the economic substance’s two-prong test: time for reconciliation?, N.Y.U. J. L. & Bus., 2005, 371 ss.; J. Rector, A review of the economic substance doctrine, 10 Stan. J.L. Bus. & Fin., 2004, p. 173.

¹⁷ Dig. 50, 17, 55. “It is never in bad faith for one to exercise his right”.

executio non habet iniuriam,”¹⁸ which is not completely new to the UK legal tradition, where “... if it was a lawful act, however ill the motive might be, he had a right to do it ...”¹⁹

The situation was clear, for instance, in the case of slavery. The owner of a slave had any right over him (or better, *it*, according to the Roman law) including the right to kill him at will (*ius vitae necisque*): a right that was accorded to the *pater familias* over all his family in the early stage of Roman law.

After centuries, however, during the later Empire era, this absolute right was somehow clarified without being explicitly limited. In other words, the property right had to be exercised over the slave according to some guidelines, so that the *pater familias* still had an absolute right “... *sed dominorum interest ne auxilium contra saevitiam vel famem vel intolerabilem iniuriam denegetur his qui iuste deprecantur*”;²⁰ soon after the principle was extended by Gaius in the rule “... *male enim nostro iure uti non debemus ...*” or also “... *expedit enim rei publicae ne qui re sua male utatur ...*”.

The turning point of the Roman legal approach towards the abuse is evident.

For the first time the Emperor clearly ruled that the exercise of an absolute right must meet specific conditions. The most important aspect of the *revirement* relies on the fact that neither Antoninus Pius nor Gaius seem to reduce the content of the property right, which still remains absolute, but rather they appear to introduce external and objective limits to its exercise, because a public interest may require it. This aspect is particularly evident in the passage quoted by the *Institutiones*, where the author clearly refers to the “common interest” as a limit to the improper exercise of rights. The step from “improper exercise” to “abuse” of right is particularly short.

It could be interesting to question what might have been the reason for such a drastic change (which, although adopted, at the same time respected the traditional content of the right).

Italian authors who faced this issue argued that a reason might have been the advent of the Christian religion²¹ or, in the specific circumstances, the personality of the Emperor, who was particularly keen on these issues. In a Christian perspective, slavery was more than inconceivable and any legal rules that allowed it were unacceptable.

¹⁸ Dig. 47, 10, 13, 1. “The one who exercises his own rights never damages others: indeed the exercise of a right is never an offence”. To a certain extent, the step-transaction doctrine could be seen as grounded on this basic assumption: where every step of a more complex transaction is taken according to the law of contracts, no abuse is conceivable and therefore no avoidance occurs.

¹⁹ Lord Halsbury in *Bradford v. Pickles* [1895 A.C. 587].

²⁰ Inst. 1, 8, 2; Dig. 1, 6, 2 “It is in the same interest of the owner not to refuse help in case a slave is maintained in a condition of starvation, famine or subject to disproportionate work”. It is worthwhile remembering that this excerpt is taken from a ruling of the Emperor Antoninus Pius.

²¹ G. Grosso, *Abuso del diritto*, Enc. dir., Milan, 1958, I, p. 161.

So here began a sort of war of attrition between two sets of rules: legal on one side, moral or religious on the other, with the latter influencing the former, or at least their interpretation.

In this early stage of the evolution of the *ius commune* the abuse of law or any avoidance of the law (including, arguably, tax avoidance) were behaviours consistent with the first set of rules, but not with the latter. The conflict has been a serious issue for all lawyers throughout the ages, and therefore not only for tax practitioners and academics.

The solution the Romans tried to introduce intended to strike a balance between the fundamental provisions of their civil law and the need for a more evolutionary approach. That's why in the Roman literature abuse is never a factor that limits the right in its exercise, but is rather an external condition that must be observed.

In other words, the Roman legal literature in the aforementioned case never rejected the argument that the owner of a thing has a complete right over it, but rather that it is in the interest of somebody else (the *Rei publicae* in the Roman literature – we could say the State nowadays) to exercise that right in a specific way and within specific boundaries, and the latter interest prevails.

The common law approach is probably different. Here the “abuse of right” is probably treated in the law of tort, and the “abuse” is associated in most cases with nuisance and the remedies provided by law for the latter.

It's also worthwhile mentioning that nuisance²² is relevant and allows compensation to the plaintiff despite the absence of a deliberate intention to offend in the counterpart (i.e., of the Roman *animus nocendi*) as ruled in *Bradford*, unless the intention also played a fundamental role in the causation of the event.

The ECJ in its most recent case law seems to adopt a similar approach to the notion of abuse of law in a communitarian context, as is noted below: in other words, the European notion of abuse seems to be founded on objective rather than subjective aspects²³.

4. AVOIDANCE AND ABUSE IN THE MIDDLE AGES: “LEX SAECULI MERITO COMPRIMERETUR, JUSTITIA DEI CONSERVARETUR”²⁴

The offspring given birth by the Roman law in its latest stage of evolution flourished across the continent during the Middle Ages.

The basic idea was that (1) the correct exercise of any right should not damage others without determining at the same time an advantage to the owner and that (2) any exercise of the right should be consistent with the good faith of the user, who should

²² A. Gambaro, *Abuso del diritto II, Diritto comparato e straniero*, Enc. giur. Treccani, Roma, 1988, I, p. 6 and 7; the author also quotes from *Grandel v Mason*, in 3 D.L.R., 1953, 65 and *Hollywood Silver Fox Farm v Emmett*, in [1936] 2 K.B. See also U. Magnus, *Tort law in general*, in (edited by J. M. Smits) *Elgar encyclopedia of comparative law*, Cheltenham, 2006, 725.

²³ See also Poiars Maduro's opinion in *Halifax* at §§ 70 and 87

²⁴ “[Sometimes] it is necessary to modify the application of the law to let God's justice prevail”.

try to achieve the scope the right was intended to protect or the legislator intended when attributing the right in question.

On the contrary, should the first condition not be met, that would be a case of objective abuse of right, while in the second case, the abuse would be of a subjective nature, according to the medieval commentators.

It is evident that the notion of abuse in tax law is similar to the one depicted above in (2). As was mentioned in § 3, the work of the commentators was aimed at defining the nature of the limits that could have been imposed to the exercise of the right in question, and some Roman commentators discovered them in the moral principles of the Christian religion.

In this sense, the *ius naturalis* introduced a rule of equity that was capable of derogating the rules of the old Roman law, and even more so the rules introduced in later centuries via statutory law by kings, municipalities, etc. The equity criteria in the Middle Ages soon became the most efficient anti-abuse provision, but at the same time also showed all the limits that instruments like this continue to have nowadays, especially in terms of lack of certainty of the rule of law.

5. THE EVOLUTION OF AVOIDANCE IN THE MODERN ERA AND THE CONTRIBUTION OF FRENCH LEGISLATION TO THE NOTION OF “ABUSE”

The boundaries of this article prevent a more detailed analysis of the notion of abuse in law and particularly in tax law in subsequent years, but we must at least mention the rise of the issue again after the *Code Napoleon* was implemented in France. The legal framework obviously had changed: the sovereign State assumed the monopoly of the sources of law, leaving but a narrow place to the interpreter.

In this context, the notion of abuse was, as it could seem today, nonsensical. The *Code* implemented straightforwardly a sort of “binary” approach to the law: any taxpayer (or any citizen) could either behave respecting the law, and therefore no abuse was conceivable, or act outside of the law, and then the concept of abuse was insignificant to that respect.

That’s why in the French tradition, and the traditions of those countries which followed the French example, such as Italy, the abuse of law was substantially unknown to academics and practitioners.

Different solutions were not legally possible. As was noted above, the abuse can arise when two different rules, belonging to different formal systems, are in conflict: it could be the case of moral or religious rules against the rule of law (Roman law versus Christian principles) or principles of law against the specific regulations of the case (*ius commune* versus statutes of the municipalities); but where a public body, such as the State, assumes the monopoly of the sources of law and excludes any other subject, then no sort of abuse is possible, and to a certain extent, the principle decided in *Bradford* above is confirmed.

More to the point, the concept of abuse as intended in the medieval tradition was no longer acceptable by national legislators across the European countries, while at the same time and to a certain extent, a different notion of abuse began to be implemented and recognised by national courts only in the last century.

The new concept of abuse was grounded on the necessity for the Courts (especially in Italy) to control the use of civil rights, such as those with regards to property, in a way consistent with the social function the Constitution attributed to them.²⁵ To this extent an abuse could be considered to take place whenever the exercise of any right was not consistent with the social function recognised and protected by the Constitution²⁶.

It could be argued that the place once occupied by the moral values of the Christian tradition was now taken by a more secular approach, closer to the socialist philosophy²⁷.

The constitutional approach to tax avoidance, once more, at least in the Italian experience, remains upheld nowadays as far as the duty to pay taxes in Italy is imposed under two basic articles: art. 53 and art. 3²⁸.

But while the first one is more accurate, introducing the “ability to pay principle” to our legal system, the second one deals with the notion of “economic solidarity” among citizens, leaving significant leeway for the interpreter to bend the notion in different ways.

That means, for example, that in the Italian legal system (and in all those on the continent which have a similar inspiration) the social function of taxes has been used many times as the overarching principle to oppose tax avoidance²⁹.

Namely, every time a taxpayer avoids the payment of taxes without evading them, he infringes the general principle of solidarity that the Constitution recognises. It is not enough, of course, to qualify art. 3 of our Constitution as an anti-tax avoidance provision, or that it was (and still is) considered a self-executing principle (the most influential authors state that rather, it constitutes a guideline to be followed by the parliament while passing new tax acts³⁰), but it is the foundation of the current anti-tax avoidance approach: an interpretive instrument.

The economic solidarity enshrined in the Constitution and, to a certain extent, the social function that must assist the exercise of many rights under our fundamental law,

²⁵ In the Italian Constitution, the right of property is recognised and protected only in so far as it is exercised also in a social interest. Even if the wording of the article didn't have a significant impact on the subsequent legislation that is still in force nowadays, it was the heritage of the Communist Party in the Constituent Assembly after the Second World War. So far, an abuse of right has been considered every time the right has been exercised without taking into due account the social function attached to the right.

²⁶ The constitutional background of the “*abus de droit*” theory is clearly explained by P. Harris, ‘Abus de droit’ in the field of value added taxation [2003] BTR, p. 136.

²⁷ C. Salvi, *Abuso del diritto*, in Enc. Giur. Treccani, Roma, I, 1988, p.3. See also C. Sacchetto, *Ethics and Taxation*, Dir. e prat. trib. int., 2007, I, p.9 ss.

²⁸ A. Di Pietro, The principle of equality in European taxation – Italy, G. Meussen (ed. by) *The principle of Equality in European taxation*, The Hague, 1999, p. 118; W. B. Barker, *The three faces of equality: constitutional requirements in taxation*, Dir. Prat. trib. int., 2007, p.405.

²⁹ The situation is therefore significantly different from the American one, see for instance Y. Edrey, *Constitutional review and tax law: an analytical framework*, in *American University law review*, 2007, 118; B. Ackerman, *Taxation and the constitution*, Colum. L. Rev., 1999, p. 39.

³⁰ See A. Di Pietro, *cit.*, p. 119.

today take the place of what once was the morality of some religion or the general principles of the *ius commune*.

This historical journey has eventually come to an end.

It was not useless. The most important result of this brief research is that in this perspective, the notion of avoidance is not absolute but might vary through the centuries, and, most important of all, according to the legal framework of tax law (but also of constitutional and civil law), it has to be implemented within.

In other words, it could change from State to State.

The more the system is inspired to social solidarity or accepts a wider notion of welfare, the more it is likely to find an aggressive approach to tax avoidance as an infringement of the duty (and not only of the obligation) to pay taxes.

This is not only a descriptive outcome but is at the same time the starting point for the study of tax avoidance in Europe and, more to the point, in ECJ case law.

The communitarian law, based on freedoms and not on a Constitution³¹, had to develop a different approach to the issue: in other words, it had to find elsewhere the reasons and the legal principles to oppose tax avoidance, if it wished to do that.

Eventually, it still has to be seen whether, and how, an EU concept of tax avoidance would be suitable for implementation by the different Member States³², as the potential conflict could cover fundamental rights and duties of some Member States. The opposite is also true.

It happened in the past that the need to battle against tax avoidance was used by Member States to justify limitations to EU fundamental freedoms, with various outcomes before the ECJ.

Generally, the need for anti-tax avoidance provisions is not a *conditio per quam*; a derogation to the EU freedoms is allowed: the court always asked for something more or tested the presence of tax avoidance under its standards.

This could potentially constitute another field of conflict between the national experience in combating tax avoidance and the different necessities of the European Union. In other words, some anti-avoidance provisions introduced to enforce the economic solidarity among citizen taxpayers could not work in an EU framework: they could be considered as disproportionate or unreasonable in the light of the communitarian approach.

³¹ It is worth while mentioning that the draft scheme of an European constitution, refused by French and Dutch citizens after national referenda contains no specific provisions in this respect. Nonetheless the existence of a European constitution, enshrining social and political rights could have grounded better any extensive approach to the “Abuse of law” theory.

³² This aspect is specifically clear where EU law directly legitimates an anti abuse approach, such as for instance in the Mergers Directive (1990/434/EEC), or in the Parent – Subsidiary one (1990/435/EEC): see B. J. M. Terra and P. J. Wattel, *European Tax Law*, The Hague, 2005, p. 563.

The Rule of reason, therefore, should succeed where morality or religious or philosophical obligation failed in the past.

6. RULES AND BOUNDARIES: REASONABLENESS AND FUNDAMENTAL FREEDOMS

The fundamental freedoms enshrined in the EU Treaty are not different from the rights analyzed above, starting with Roman law. They are not introduced mainly for tax law purposes but have a relevant impact in this context.

When a national taxpayer adopts an approach to the national tax rule that leans towards tax avoidance, he could try and justify his behaviour under the aegis of a fundamental freedom.

It could be said, therefore, that the abuses of the Treaty rights (freedom of movement of the citizens of the Union, the free movement of goods, the free movement of workers, the right of establishment, the freedom to provide services and the free movement of capital³³) raise the same problem discussed above while analysing the notion of abuse in different fields.

The solution differs: here it is the rule of reason³⁴ that distinguishes between proper exercise of the right and abuse (and therefore improper avoidance of tax).³⁵

According to this approach, the national anti-avoidance provisions must pass several tests in order to qualify as consistent with EU law.

The first condition is that the specific circumstances of the case might fall into the exercise of one of the fundamental freedoms mentioned above, as far as the conflict (between national anti-avoidance rules and EU law) arises only where the second is theoretically applicable³⁶.

In this respect, it was noted that the real purpose of the taxpayer (namely the subjective approach to avoidance) is now absolutely irrelevant.³⁷ The ECJ always judged the application of a fundamental freedom without taking into account the purpose followed by the taxpayer in the specific case, at least in this first stage of the inquiry.

While the subjective aspect is underestimated by the ECJ, the objective aspect plays a more relevant role. The economic activity for which the taxpayer asks protection must be effective and genuine, and it must not constitute a mere device by which to qualify for the protection of the Treaty.

In other words, the business activity must be actual and not artificial: the exercise of the freedom must be consistent with an economic interest of the taxpayer.

It is self-evident that while at first glance the “rule of reason” seems to constitute an ideal concept capable of distinguishing the use from the abuse of a right, and therefore

³³ Respectively articles 23, 39, 43, 46 and 49 of the Treaty.

³⁴ See *Cassis de Dijon*, C-120/79.

³⁵ D. Weber, *Tax avoidance and the EC Treaty freedoms*, The Hague, 2005, p. 253.

³⁶ In *Leur Bloem* (C-28/95) however communitarian tax law was considered applicable even to a purely National case.

³⁷ D. Weber, *cit.*, p. 254.

to identify the avoidance, on the other hand, under a more practical approach, it is difficult for everybody, and for the ECJ in particular, to assess in some cases the existence of reasonableness in a business activity³⁸ carried on by a communitarian taxpayer in the exercise of a fundamental right.

The second feature of the “rule of reason” is so evident that even in recent papers and seminars some influential authors stressed the fact that the ECJ seems to be losing the route followed in more recent years, deciding tax avoidance cases (both in direct taxation and in VAT) without a clear general picture.

The situation may be different from the one represented in the past: the ECJ approach has not changed through the years, but rather the objects of judgement of many recent cases are taking the ECJ into the maelstrom of harmonising its tax approach to an issue that may not be clear also within a national perspective.

The supranational approach can't but emphasise the divergences, and the decisions reached are, of course, harder to accept for the Member States, as the most recent case law clearly shows.

7. USE, ABUSE AND *CADBURY SCHWEPPE*S: THE FINGERPRINTS OF THE EU APPROACH TO TAX AVOIDANCE

*Cadbury Schweppes*³⁹ is not the first case, from a chronological perspective, in which the ECJ dealt with abuse of law: it is, however, one amongst those wherein the need to apply a clear concept of abuse has been considered with a particular focus, and where the court had to use once more the rule of reason (and the principle of proportionality to properly apply the Treaty) to provide reliable answers to the involved parties.

It is not important here to describe the facts and the circumstances of the case, which are well known and of lesser relevance to the extent of this research. It is significant, however, to focus attention on the fundamental freedoms allegedly infringed (arts. 43 and 48 of the Treaty: freedom of establishment) and on the anti-avoidance rule challenged before the court: the British CFC regulations.

In this case, the *Cadbury* group set up an Irish-controlled company to take advantage of the lower tax rates of that country, when compared to the UK rates. The tax saving was then reached through a leasing and financing operation from the latter to the former: the interests were tax deductible in the UK and, of course, taxed at a lower rate on the Irish subsidiary.

In order to avoid this tax deferral, the UK, just like many other European countries, introduced CFC regulations, attributing to the resident parent company the profits (or part of them) realised by the non-resident subsidiary, without any necessity to wait for the distribution of dividends by the foreign subsidiary.

³⁸ Or also in a specific commercial operation, such as in *Part Service* (C- 425/06) still pending in front of the ECJ. In this case the taxpayer (unreasonably) unbundled a single financial leasing contract into a loan and a leasing. The overall commercial operation didn't change significantly, but the tax saving was considerable.

³⁹ *Cadbury Schweppes plc, Cadbury Schweppes Overseas Ltd v. Commissioners of Inland Revenue*, C-196/04.

At first glance, a provision like this could impair the freedom of establishment⁴⁰ because a resident company could be prevented from establishing a subsidiary abroad if this decision could then immediately and directly increase its taxable base.

The UK provision, considered alone, is reasonable because it is aimed at preventing the foreign allocation of profit where this decision is inspired mainly by tax considerations: the abuse in this case could consist of a business activity developed abroad (and not in the homeland) only for tax purposes.

The answers that Advocate General Léger and the ECJ gave on this issue are quite surprising, but only to a certain extent.

The first statement is that tax planning across Europe is a legitimate way to spare taxes and that the delocalisation of subsidiaries or branches of a company only for tax reasons cannot be challenged by the home State Tax Administration. It could be said, then, that a certain level of tax competition amongst Member States is not only tolerated but reflects the need of a common market. In other words, tax motives are, in themselves, legal motives.⁴¹

The Advocate General clearly pointed out in his final remarks that when a parent company chooses another Member State in which to establish a company only because in the latter the taxation is significantly lower, it isn't abusing the freedom of establishment enshrined in the Treaty. This conclusion was, all in all, already upheld in the *Centros*⁴² case, even if the latter was not exactly a tax case.

So far, no iron curtain runs between use and abuse of law, but rather a thin red line that can shape a different border when the abuse is tested under commercial law versus tax law.

For instance, in *Centros* the freedom of establishment was defended by the ECJ even if in the specific case, the seat of the company set up in the UK was merely artificial, as the main business was developed in Denmark. The same goes for a case in which the company was set up according to a Member State commercial law and then was transferred to the UK or elsewhere. In these circumstances, the company tried to circumvent only commercial law provisions and no abuse was assessed even if the main seat abroad was fictitious or not effectively involved in any business at all.

⁴⁰ In the specific case art. 53 was invoked by the taxpayer as well. The Court refused to go into details as far as art. 43 was considered sufficient to assess the infringement of the Treaty. According to a settled case law, when an infringement of art. 43 is recorded, it is not necessary to examine the same issue under another point of view.

This conclusion was possible in the specific circumstances of the case because the host State of the subsidiary (Ireland) is an EU member country. At the moment, therefore, it is not possible to confirm beyond any reasonable doubt that the decision in *Cadbury* would have been confirmed where that country would have been a non EU one.

It could be argued, in this latter case, that art. 56 (applicable also to non EU taxpayers) would have taken the place of art. 43 in the reasoning of the Court.

⁴¹ N. Vinther and E. Werlauff, *Tax motives are legal motives – The borderline between the use and abuse of the freedom of establishment with reference to the Cadbury Schweppes case*, European taxation, 2006, p. 384.

⁴² *Centros ltd. v. Erhvervs-og Selskabsstyrelsen*, C-212/97. In this case the UK was chosen because of the lesser conditions to be met in order to set up a limited company, in comparison to those in Denmark.

In tax cases, such as *Cadbury*, the distinction is not so straightforward, and the artificial nature of the subsidiary could constitute a reasonable limit to the fundamental freedoms of the Treaty.

Of course *Centros* and *Cadbury* have a significant common ground. In both cases the court ruled that the nationals of any Member States cannot use EU law to “improperly circumvent”⁴³ the national legislation or, in another point of the decision, that they are not allowed to “improperly or fraudulently”⁴⁴ take advantage of the community law.⁴⁵

When such conditions are not met, then no restriction to the freedom of establishment is allowed in a EU perspective, unless, and this is specifically a case relevant to tax law, other overriding reasons of public interest⁴⁶ do not emerge. And even in this case the derogations and the limits to the fundamental freedoms must be appropriate, proportional and consistent with the aim pursued.

It is self-evident that the need of every Member State for specific countermeasures to tax avoidance falls into one of the overriding reasons of public interest, but those provisions must be proportionate to the aims pursued and to the actual dangers. If this condition is not met, then it is the Member State that abuses this possibility.

That’s why the Court plays a fundamental role in defining the notion and the condition of abuse: defining tax avoidance, it also contributes to defining the abuse of right by the Member State.

In EU law this is particularly true: the abuse of the taxpayer is counterweighted by the abuse of the Member State; both subjects are, to a certain extent, in an equal position before the court⁴⁷.

Cadbury is paradigmatic to this extent. From the UK point of view, the taxpayer is abusing his right to set up an economic activity abroad, as far as his decision is basically inspired by tax deferral and tax reduction, and from the taxpayer point of view, the UK is abusing its right to override the Treaty freedoms. The concept of tax avoidance (and the closely related “abuse of law”) therefore is at a crossroads the court had to regulate, potentially catching two pigeons with one seed.

8. RESOLVING A FEARFUL SYMMETRY? THE ECJ ATTEMPTS TO TACKLE TAX AVOIDANCE

In the previous paragraph it is noted that the EU perspective is a privileged one from which to look at the phenomenon called “tax avoidance” or “abuse of tax law”.

The reasons are quite evident: while sovereign States deal with tax avoidance with a one-way approach, that is, seeking to limit rights and freedoms of the taxpayer in a convenient way, the EU approach compels Member States to do this with respect to some of them, unless other very specific circumstances are not met.

⁴³ See *Cadbury, cit.*, at § 35.

⁴⁴ Once more *Cadbury, cit.*, at § 35. It is important to note how the court uses words and attributions such as “improper”, “fraudulent”, “genuine” (at § 54) that recall the Roman experience, where the abuse of law was discovered by using moral or ethical concepts.

⁴⁵ Before *Centros* see also *Knoors*, C-115/78 and *Bouchoucha*, C-61/89.

⁴⁶ *Cadbury*, § 47.

⁴⁷ This situation resembles to a certain extent the relation between “*Abus de droit*” and “*Abus de pouvoir*” in the French doctrine, as Harris mentioned in his article quoted above.

So far, the EU approach to tax avoidance is “two way”: the ECJ in cases like *Cadbury* judges the limits imposed by the Member States according to limits imposed on those States by the Treaty.

Cadbury is once more helpful to this extent, because the Court set out, albeit not for the first time, the circumstances in which a business operation falls outside the scope of the EU law protection. This situation occurs, just like in the specific case, when together with the “subjective element of obtaining a tax advantage” objective circumstances exist and show that “despite formal observance of the conditions laid down by community law, the objective pursued by freedom of establishment has not been achieved”⁴⁸.

It is a helpful benchmark for future rulings even if the conditions are deeply influenced by the circumstances of the case, by the rules challenged (CFC legislation) and by the specific ways and means to set up an improper tax-saving scheme.

In the text of the sentence it is possible to underline three specific aspects which played an important role in the final decision: three conditions, two positive and one negative, that help to define better the notion of abuse, in both the mentioned ways.

The first one, the negative, is for the Member States.

From the Court’s point of view, “anti-tax avoidance” provisions are matters for the different Member States: a need for harmonising them arises neither from the Treaty nor from any specific directive, excluding some incidental steps in directives whose primary target was different.⁴⁹

As a matter of principle, every Member State is free to implement any of them as far as it deems necessary to safeguard the general principles of the system and the general duty to pay taxes.

However, once those provisions limit or, broadly speaking, interfere with the exercise of any fundamental freedom, their application can be allowed only if targeted at contrasting “wholly artificial arrangements”⁵⁰ concluded by the taxpayer aimed at “circumventing the application of the legislation of the Member State concerned”.⁵¹

This is not the first case⁵² in which the court has ruled on this, but it could be considered one of the most clear, taking into account the impact on typical anti-avoidance provisions (CFC regulations) whose use is quite widespread across Europe.

The second ones, the positives, are for the taxpayer.

⁴⁸ See *Cadbury* at § 64.

⁴⁹ There are references to tax avoidance and abuse of law in virtually any directive involving direct taxation in the EU. The same could be said also for VAT, even if the situation in that case is much more complex and will be analysed in a subsequent paragraph of this research. In the case of the Sixth Directive, for example, the notion of avoidance and abuse were mentioned in art.13A, 13B, 14, 15, 22.8, 27.1 (see again P. Harris, *cit.*, p. 147)

⁵⁰ See *Cadbury*, § 51.

⁵¹ See *Cadbury*, § 51.

⁵² See also *ICI*, § 26; *Lankhorst – Hohorst*, C-324/00, § 37 where the court challenged the “thin capitalisation rules” in the light of the same principle (freedom of establishment); *De Lasteyrie du Saillant*, C-9/02, § 50; *Marks and Spencer*, C-446/03, § 57.

Freedom of establishment means that (1) the Treaty protects not every kind of business, but only those that are aimed at actually pursuing an economic goal through a fixed establishment for an indefinite period. Moreover, (2) the business must underpin a genuine economic activity: this means that it must reflect, as the Court says, a “real economic activity”.

Even if closely related, the two sets of conditions do not necessarily coincide all the time.

First of all, time matters for the positive conditions and not for the negative one. “Establishment” in the meaning of the Treaty means a place of business qualified by stability in time and an intrinsic attitude to be developed (according to the specific case) for an indeterminate period (depending on the nature of the business developed). The Treaty of Rome seems to guarantee protection where a business is carried on in another EU country only through a set of investments that will develop during the time.

The first conditions don’t seem to require such a qualified duration. So, for instance, an anti-avoidance provision aimed at contrasting a business operation in another Member State that is not going to create in that other State a fixed establishment under the meaning of the word in *Cadbury* could not find an adequate protection under art. 43.

This will not mean that the Treaty doesn’t cover the mentioned operation, as it could fall into the protection of another article or another principle of communitarian law (for instance, the free movement of capital⁵³).

Timing is not the only difference between the two sets of conditions, then.

It could be argued that a second and a third difference rely on the subjective and objective elements, because later on in *Cadbury* the court realises that even if, theoretically speaking, the distinction between genuine activity and artificial arrangement could suffice to introduce an acceptable difference, from a more practical point of view, the aim is still far from being reached.

That’s why the Court stresses the importance of the subjective element⁵⁴ in the abuse of law: underlying this requirement is the intention to obtain a tax advantage. Only if this condition is met, and the motive is the exclusive reason which pushed the taxpayer to set up the entire operation, is the anti-avoidance provision consistent with EU law.

But this is not enough: the Tax Administration of the home State has to demonstrate⁵⁵ that the taxpayer didn’t achieve, in the specific circumstances of the case, the objective pursued by the freedom of establishment.⁵⁶

⁵³ And in this case even non EU companies could qualify for the protection of the Treaty.

⁵⁴ It is not the first time we meet such a condition in the definition of “abuse of (tax) law”; see § 2 above. Even if the subjective element is not decisive *per se*, it still shows a certain importance in the reasoning of the Court.

⁵⁵ The burden of proof is neither clearly established in this case nor in the others preceding it (see, for instance, *Emsland – Stärke*, C-110/99 at §§ 52 and 53; *Halifax and others*, C-255/02, at §§ 74 and 75). However from § 64 of the decision in *Cadbury* emerges the need for the

In other words, that there has been a *detournement*, that is, a deviation, of the *ratio* of the rule in question. The taxpayer invoked the freedom of establishment, but the operation he set out is not intended to establish anything in the other Member State.⁵⁷

In this way the Court sought to strike a balance between the two abuses underlined before: on one hand, the abuse of the taxpayer exercising his freedoms and rights in a improper way (inconsistent with the duty to pay taxes and the solidarity principle that are introduced in almost every European Constitution, if any, or fundamental law of the State) and on the other hand of the State introducing an anti-tax avoidance that could constitute an overkill, in some cases, of the business operation.

A clear example of this were the British CFC regulations: unable to distinguish true tax avoidance from the exercise of the freedom of establishment.

The Court tried to reach the best solution available, considering that the playing field was all but level, in the circumstances of the case.

In other words, tax avoidance is perceived differently (and answered differently) in light of the general principle of any tax system⁵⁸. If, like in many Member States, a duty of economic solidarity is present, then a certain answer is possible; if, as in EU law, this principle is not present (the Treaty is not a Constitution and arguably it will never become one), the Court can't but use overarching rules of reason in a desperate attempt to find an acceptable solution for all the parties involved.

This answer is, however, valid not for all EU law, but only for those aspects that are not harmonised and where the tax avoidance and the abuse of law can be tested only in the light of the fundamental freedoms enshrined in the Treaty.

Where the EU legislation is more detailed and contains principles and rules applicable to all the Member States it is possible for the Court to implement a notion of abuse and of tax avoidance that is more compatible with the system it has to work within, without any need to refer to anything but the rule of reason.

In the case of the VAT, the Court has a complete and detailed legislation built upon several directives, later consolidated into one text⁵⁹, and a principle, the neutrality of the tax, that must necessarily inspire the different legislations of the Member States without entering into any conflict with other issues at a constitutional level (such as solidarity or the ability to pay taxes...).

The notion of abuse (and of avoidance) shall, therefore, be flavoured differently.

Member State to demonstrate something more of the subjective element; that's why it could be argued that it is the burden of the Tax Administration to demonstrate that the business operation is not sound at all.

⁵⁶ See *Cadbury*, § 64.

⁵⁷ See *Cadbury*, § 64.

⁵⁸ P. Harris, *cit.*, p. 136.

⁵⁹ Directive 2006/112/EC.

9. ABUSING NEUTRALITY? PLAYING WITH THE VAT DEDUCTION IN *HALIFAX*

The most striking difference between *Cadbury* and *Halifax*⁶⁰ is clearly summarised by the court in §§ 57 and 62 of the latter judgement.

According to the first paragraph of the sentence, the notion of abuse (if any) in the VAT mechanism is not qualified by any subjective element.⁶¹ In other words, the VAT application must be necessarily inspired by two general principles, facility in the application and certainty of the legal relationship between businesses.

From this point of view, any possibility of tax reassessment which is grounded on subjective factors, such as the deliberate will to avoid the payment of taxes by setting up a complex commercial operation, could constitute, *per se*, an unacceptable impairment of that goal, introducing an element of uncertainty and unpredictability to the reassessment which could interfere with the application of the VAT mechanism.

At first glance, it seems therefore that the taxpayer could enjoy a greater leeway in the abuse of the VAT provision. The impression is arguably wrong, and all in all it would be surprising that in a harmonised field of tax law the possibilities for tax avoidance could be greater than those in non-harmonised sectors⁶².

The court clarifies this point later on in the judgement, stating that clearly the sixth directive (and not the 2006/112/EC directive) is aimed also at preventing tax evasion, avoidance and abuse of law,⁶³ but that this objective must be achieved with respect to the legal certainty of the business relation between entrepreneurs localised in different Member States of the Union.

In other words, priority must be given to the need for trust in business relations or, more to the point, to tax consequences of the business relations. Every entrepreneur across Europe must know in advance the amount of the VAT (if any) applicable to the sale of goods or purchase of services in a cross-border dimension.⁶⁴

It could be argued at this point that this second necessity is peculiar to the VAT system and no issues of this kind were raised by the Court (or at least were raised

⁶⁰ It is not the intent to summarise here the facts of the judgement. Anyway, *Halifax* was submitted to the ECJ after the British Revenue discovered a VAT deduction mechanism set up by the Halifax company and other subsidiaries aimed at allowing them to deduct the VAT even in cases and circumstances where the core business activity of the parent would not allow it. Basically, running a VAT-exempt business (see *Halifax*, at § 12) it would have been impossible for Halifax to deduct the VAT paid for repairs on real estates. That's why the parent company involved in the operation of a considerable number of companies of the same group (see *Halifax*, § 19), allowing them to recover the VAT. The Revenue considered the entire operation a sort of abuse of the step-transaction approach that still inspires the VAT mechanism, and therefore the commissioners refused any claims of deductions. See also H. L. McCarthy, *Abuse of rights: the effect of the Doctrine on VAT planning*, in BTR, 2007, p. 160.

⁶¹ A similar conclusion was upheld also in *BLP Group*, C-4/94.

⁶² The context of the two cases is different. In *Cadbury* the taxpayer challenged the compatibility of a national anti abuse provision with EU law; in *Halifax* the Commissioners sought to implement an anti-abuse rule relying only on EU law despite the fact that the national provisions didn't say anything about that.

⁶³ In this sense, the case *Gemente Leusden and Holin Groep*, C-487/01 and C-7/02, at § 76.

⁶⁴ See *Netherlands v. Commission*, C-326/85, § 24 and *Sudholz*, C-17/01, § 34.

without such a priority) in cases involving direct taxation and fundamental freedoms. The first and foremost effect, as mentioned above, is that the anti-avoidance provisions have to pass a tougher control of their compatibility with EU law.

It is not a matter of abuse here, considered as a balance between fundamental rights and the need of every Member State to raise money for welfare, but also a matter of predictability of the tax consequences⁶⁵. The dimensions of the overall judgement are therefore different, and we could argue that the concept of avoidance and abuse is no longer the same.

Looking to this issue from the point of view of the taxpayer, then, the need for neutrality of the tax and the priority of legal certainty become the two fundamental benchmarks with which to assess the compatibility of the anti-avoidance provision to EU law.

That's why it is not entirely accurate to state that the step-transaction approach followed by the ECJ in all the cases involving avoidance of VAT allows more loopholes to be abusively exploited by the fraudulent taxpayer. Rather, it is the mechanism of the tax that imposes this approach.

The Court implicitly clarifies this point later on in the *Halifax* judgement, when it notes that even if the step-transaction approach has to be favoured in VAT cases, and even if it's up to the national Courts to assess the eventual abuse of the neutrality (in most of the cases, of the VAT deduction input by the taxpayer), taking into account the specific characteristics of the operation.

From the ECJ approach, the "purely artificial nature"⁶⁶ of the operation from an objective point of view and the "legal, economic or personal relations"⁶⁷ between the parties (that is, the seller and the purchaser of the goods, for instance) are significant. If these conditions are met,⁶⁸ the operation can be disregarded for VAT purposes and the deduction of the tax denied by the Tax Administration.

The Court is walking once more on a minefield⁶⁹ and seems quite aware of that. It is contradictory to stress the importance of legal certainty and soon after to introduce elements of judgement recalling a sort of "substance over form" approach that is quite unfit to give *ex ante* a clear solution, depending as it does on the specific circumstances of the case.

That is probably why soon after, the Court clarifies that, anyway, the abuse must be assessed under "a number of objective factors" and that only after that positive

⁶⁵ This aspect was not debated in *Cadbury* as far as it involved the existence of an anti avoidance provision. Even if its compatibility with EU law was discussed, every taxpayer was aware of its existence and applicability to all qualifying cases.

⁶⁶ *Halifax*, at § 81.

⁶⁷ *Halifax*, at § 81.

⁶⁸ Even if the Court is not clear on this point, it could be argued that the subjective conditions (special relationships between the involved parties) could be considered a factor emphasising the artificial nature of the operations.

⁶⁹ This could constitute a partial answer to the provocative thesis by H. Rasmussen in *On law and Policy in the European Court of Justice. A Comparative study in judicial policymaking*, Dordrecht, 1986, *passim*.

assessment of the operation it is possible to say that the essential (not only the main) reason for the existence of the specific operation is to obtain a tax advantage.

All in all, the approach seems to recall the one analysed before, where the Court stressed the importance of the effective exercise of a fundamental right or freedom enshrined in the Treaty (freedom of establishment, of movement of capital, etc., ...), but here the soundness of the business is the condition to be met and the neutrality of the tax the feature to be preserved as far as possible.

10. CONCLUDING REMARKS: WHERE DO WE GO FROM HERE?

The first part of this brief research began with a quotation from an author⁷⁰ indicating that the issue relative to tax evasion and, first of all, avoidance, is an eternal one. After ten years the same author implicitly confirmed his opinion some months ago at another academic meeting⁷¹. Moreover, other academics recently presented a different and original approach⁷² to the notion of tax avoidance, which seems to focus not only on legal issues but also on different factors derived from the business experience, such as risk management and the need for an overall different approach to the taxpayer by the HMRC, at least in the UK. In a relationship based on trust, there is no space for avoidance⁷³.

Both the approaches are absolutely correct, and where the first underlines a problem, the second tries to offer a new solution with different means.

However this is intended to be an investigation made from a legal point of view and therefore in a quite traditional way. The purpose has not been to provide another approach to tax avoidance, tax evasion and abuse of law. Such a purpose would be too far-reaching and impossible to achieve. The goal has been, to a certain extent, to try to understand the notion of avoidance and abuse of law from a different perspective.

The basic assumption of this research was that the notions of tax avoidance and abuse of tax law must be understood and interpreted according to what happens in the other fields of the national legal system. The historical evolution of each of them, and of the Roman law, is determinant in clarifying when abuse takes place.

This perspective has supported the idea that avoidance is not an “eternal problem” in tax law, but an issue that arises when legal provision collides with other sets of rules

⁷⁰ F. Vanistendael, at footnote 1.

⁷¹ F. Vanistendael, *Halifax and Cadbury Schweppes: one single European theory of abuse in tax law ?*, in EC Tax Review, 2006, p. 192; see also the report by J. Freedman, G. Loomer and J. Vella, *Moving beyond avoidance? Tax risk and the relationship between large business and HMRC*, pp.18 and 42.

⁷² J. Freedman, G. Loomer, J. Vella, *Moving beyond tax avoidance? Tax risk and relationship between large business and HMRC*, Oxford, 2007.

⁷³ Commenting on the Varney Delivery Plan and the relationship between specific taxpayers and the HMRC, J. Freedman, G. Loomer and J. Vella noted that a vital element “... is the desire from both sides for a relationship based on mutual trust” (see J. Freedman, G. Loomer and J. Vella, *cit.*, p. 5).

of a non-legal nature: of course this conflict depends on the national background and on the strength (real or perceived) of the non-legal rules⁷⁴.

When the ECJ tries to harmonise the different tax systems as far as this is allowed by the Treaty, it progressively has to solve the same problems faced by the different national States across the centuries. However, while in every Member State the abuse is always a judgement of comparison between an individual right (not necessarily of the taxpayer) and other principles enshrined in the legal system (ability to pay, solidarity, etc., ...), in Europe these principles are missing, in so far as something more of a Treaty between States will not be implemented.

This is helpful towards explaining why the Court, in some cases, seems to wander in obscurity⁷⁵, recalling concepts such as “genuine business” or even “non-artificial business” that although sufficient to solve the problem from a theoretical point of view are of little use from a practical point of view. The practitioners need something more, as the same ECJ implicitly admits in all cases involving VAT.

The answers now provided, even if workable in direct taxation, are not sufficient. Assessing avoidance only in cases of letterbox companies (a case that in the Italian point of view could be qualified, to a certain extent, as tax evasion and not as avoidance) is of little help for all those States that are striving to defend constitutional principles of economic solidarity, while it could be considered enough by others which made a different choice, providing limited welfare and at the same time requesting a lower level of taxes.

All in all, tax avoidance and abuse of tax law are not “eternal problems” in tax law, but issues that are only partially related to tax law and depend first and foremost on the culture and the historical evolution of every State, issues about which the European tax lawyer has little to say, especially without a constitutional background.

⁷⁴ Avoidance varies according to the legal tradition of each State and to the specific tax it refers to.

⁷⁵ M. Cappelletti, *Is the European Court of Justice “running wild”?*, in *European law Review*, 1987, p. 3.

The Economic Benefits of the Use of Guanxi and Business Networks in a Jurisdiction with Strong Formal Institutions: Minimisation of Taxation

Nolan Cormac Sharkey *

Abstract

Why rely on informal institutions when conducting business? Because it leads to a greater economic return

The current academic status quo on the use of guanxi and business networks by Chinese entrepreneurs in both China, South East Asia and beyond is that it is a response to the weak formal institutional environment¹. It is argued that weak institutional environments create an economic rationale for the use of guanxi and business networks. From this base, it is further asserted that the use of guanxi and business networks will reduce and disappear with the strengthening of formal institutions in China and other jurisdictions. This paper will challenge this latter assertion by arguing that there are strong economic benefits to those that are able to operate through guanxi and business networks in a jurisdiction with strong formal institutions. The primary benefit that will be examined is the reduction or elimination of business taxation which is a very significant issue in most jurisdictions with strong institutional environments. This paper will demonstrate how guanxi and business networks can be used to avoid taxation in the strong formal institutional environment such as that represented by Australia.

THE USE OF GUANXI IN BUSINESS

The discussion of the use of networks and guanxi in Chinese business is not free from controversy. Most scholars do not question its existence (at least in part). That is, there is clear evidence to show the instrumental and extensive use of strong reciprocal informally constituted inter-personal bonds (guanxi) by Chinese entrepreneurs in the conduct of business in China and beyond. It has also been shown that people can be bound by these bonds into networks that are used to facilitate economic activity (guanxi networks). Thus the extensive use of guanxi and networks can be viewed as a Chinese business modality.

There is, however, critical debate as to the causes, rationale and therefore future of this business modality. Early work attributed the phenomenon to the fact that Chinese business people's economic activities were socially embedded² and Chinese social

* Senior lecturer at the Australian School of Taxation (Atax), University of New South Wales, Sydney, Australia. Email: nolan.sharkey@unsw.edu.au.

¹ Chan, K. B. (2002). Preface. Chinese Entrepreneurship and Asian Business Networks. T. Menkhoff, Solvay. London, RoutledgeCurzon: xii - xiv.

² Fei, X. (1992). From the Soil: The Foundations of Chinese Society. Berkley, University of California Press.

values dictated that business be pursued through guanxi and networks. The behaviour was therefore not pursued with a rational objective but was simply the result of social embeddedness.

A conclusion that may be drawn from this is that network business activity leads to economic inefficiency. The argument here is quite clear. That is, that the most efficient economy is one where market forces and prices determine economic decisions and therefore allocation. Given that in the case of Chinese entrepreneurs, it is being asserted that economic decisions are being highly influenced by inter-personal relationships, economic inefficiency must accordingly result. Therefore, as long as Chinese business people hold “Chinese” social values, the behaviour will continue to the detriment of economic performance. For this reason Chinese capitalism is seen as a significant contributing factor to problems such as the economic meltdown in Asia beginning in 1997³.

Other scholars have however asserted that the social embeddedness of Chinese business and the resulting use of guanxi and networks is a major contributing factor to the success of Chinese entrepreneurs in China and South-East Asia⁴. They point to the phenomenal economic growth of China and other economies dominated by Chinese business as evidence of this. In addition, they highlight the success of Chinese business persons relative to other groups in countries such as Malaysia, Singapore and Thailand. The theoretical argument here is that the most efficient economy is not one where the market dictates all prices and economic allocation. Rather there is a completely different structure and logic to Asian markets and the use of guanxi and business networks is a more efficient allocation mechanism as evidenced by the success of the tiger economies⁵.

More recent academic works have challenged the fundamental idea that the use of guanxi and networks results from social embeddedness. That is, it is contested that Chinese business persons let personal relationships dictate their economic decision making simply because they are Chinese. Rather, it is asserted that the core business values that Chinese entrepreneurs hold are not materially different from any other group. Instead, the phenomena of guanxi and network use where they occur are seen as a rational response to the circumstances in which Chinese entrepreneurs have had to do business⁶.

Fei never uses the term “social embeddedness” but see the following for why I argue that Fei is referring to social embeddedness: Granovetter, M. (1985). “Economic Action and Social Structure: The Problem of Embeddedness.” *American Journal of Sociology* **91**: 481 - 510.

³ Backman, M. (1999). *Asian eclipse : exposing the dark side of business in Asia*. Singapore, J. Wiley.

⁴ Lever-Tracy, C., D. F.-K. Ip, et al. (1996). *The Chinese diaspora and mainland China : an emerging economic synergy*. Houndmills, Basingstoke, Hampshire

New York, Macmillan Press ;

St. Martins Press.

Biggart, N. W. and G. G. Hamilton (1997). On the Limits of Firm-Based Theory to Explain Business Networks: The Western Bias of Neoclassical Economics. *The economic organization of East Asian capitalism*. Thousand Oaks, Calif., Sage Publications.: 33 - 54.

⁵ See generally: Redding, S. G. (1990). *The spirit of Chinese capitalism*. Berlin, W. de Gruyter.

⁶ Krug, B. (2004). *China's rational entrepreneurs : the development of the new private business sector*. London ; New York, RoutledgeCurzon.

Hendrischke, H. (2002). *The Role of Social Capital, Networks and Property Rights in China's Privatization Process*. Chinese Enterprise Models, University of New South Wales, Sydney, UNSW Centre for Chinese Studies.

The above assertions are supported by demonstrating that Chinese entrepreneurs operating in both transformation China and many parts of South-East Asia have had to contend with formal institutional environments that are very poorly developed and weak, aggressively negative towards them or both⁷.

Developed formal institutions for doing business were largely non-existent in pre-colonial South-East Asia when Chinese business first took root in the region⁸. The colonial period bore witness to formal institutions largely geared towards European colonial interests while the post-colonial period has seen the slow and incomplete development of formal institutions in many of the newly independent states. In China, formal institutions for private economic activity were obliterated by communism and have only begun to reemerge since 1979. They still remain weak⁹.

Formal institutional environments that are or have been negative towards Chinese entrepreneurs have also been common in the region. The source of this negativity being that many states in the region have been anti-entrepreneur, anti-Chinese or both. At many points in history, the imperial Chinese state has been at the very least unsupportive of entrepreneurial activity. The origins of this attitude being a belief in a state order from peasants to emperor that does not value entrepreneurs or merchants¹⁰. The period of staunch communism in China was clearly against entrepreneurs and lingering socialist values have ensured that entrepreneurs during much of the transition period have tended to be discrete about their affairs.

In certain other parts of the region similar state values based upon anti-merchant traditions and communist experience have had a similar effect in many eras. However, the situation has been made significantly more severe by anti-Chinese sentiment in both colonial and post-colonial states in the region. Fear and resentment of Chinese influence and potential economic domination in most post-colonial states in South-East Asia have led to both direct and indirect measures to suppress Chinese economic activities or prevent them¹¹.

Colonial states on the other hand simply attempted to ensure that the European colonists were the primary beneficiaries of all economic activities. These situations have therefore been anything but strong formal institutional environments in which Chinese entrepreneurs could operate. A particularly difficult situation when many of the Chinese pursuing business activities have been doing so due to their exclusion from other sectors of the economy such as primary production. Even in the case of imperial China itself, it has been argued that those that have turned to commerce have done so due to their in-substance exclusion from other sectors due to the impossibility

⁷ Chan, K. B. (2000). Chinese business networks : state, economy and culture. Singapore Copenhagen S, Demark, Prentice Hall Pearson Education Asia Pte. Ltd. ; Nordic Institute of Asian Studies. Pp 1- 13

⁸ For a review of early Chinese business in South East Asia see: Wang, T. P. (1994). The Origins of the Chinese Kongsj. Selangor, Pelanduk Publications.

⁹ Wu, W.-p. (2000). Transaction Cost, Cultural Values and Chinese Business Networks. Chinese Business Networks. C. K. Bun. Singapore, Prentice Hall.

¹⁰ For a detailed study of Chinese commercial practices and their origins see: Gates, H. (1996). China's Motor: A Thousand Years of Petty Capitalism. Ithaca, NY, Cornell University Press.

¹¹ Chan, K. B. (2000). Chinese business networks : state, economy and culture. Singapore Copenhagen S, Demark, Prentice Hall Pearson Education Asia Pte. Ltd. ; Nordic Institute of Asian Studies.

of economic success¹². In this case, however, the lack of opportunity has been caused by the inability of these people to sustain themselves through primary production rather than formal exclusion.

Trade and business activity involve shifts of wealth between individuals and groups. It is therefore a general prerequisite for the existence of such activities at any meaningful scale that institutions are available to protect rights to wealth and reinforce adjustments to them. At the very least formalized property rights should exist. Thus in a developed capitalist state property rights are upheld by the state and courts while trade is reinforced by contracts and the legal institutions that support these. Commerce is therefore supported by formal state institutions.

As has been shown above, formal institutional support has often not been available to Chinese entrepreneurs in China and the region. The answer as to how Chinese entrepreneurs have been able to engage in extensive commercial activities without strong formal institutional support is that institutional support has been obtained through informal or popular institutions¹³. Chinese entrepreneurs have relied upon the use of trust obtained through appropriate interpersonal relationships (guanxi) and networks to provide the institutional support for their business activities. The substitutability of these informal institutions for formal institutions is borne out by the general success of these entrepreneurs in not only getting business done but in often prospering.

The rationality of the use of guanxi and business networks in undeveloped or exclusionary formal institutional environments is therefore clear. Chinese business methods are highly suited to prospering in these environments due to their lack of reliance upon formal institutions to engage in economic activity. Chinese entrepreneurs have therefore been able to succeed without legal support and with extremely constrictive bureaucratic impediments.

The argument that Chinese business modalities are not the result of social embeddedness but a rational response to a particular environment is therefore a strong one attested to by historical circumstance. Further credence is added to the argument by studies that show that in China, South East Asia and other areas, when the formal institutional environment has provided entrepreneurs with the appropriate support, they have often made use of the formal institutions instead of informal institutions¹⁴. There are studies that show that large firms in China seek employees through the market and not through connections. In the case of Singapore¹⁵, Hong Kong, Taiwan

¹² See generally: Gates, H. (1996). China's Motor: A Thousand Years of Petty Capitalism. Ithaca, NY, Cornell University Press.

¹³ Wu, W.-p. (2000). Transaction Cost, Cultural Values and Chinese Business Networks. Chinese Business Networks. C. K. Bun. Singapore, Prentice Hall.

Hendrischke, H. (2002). The Role of Social Capital, Networks and Property Rights in China's Privatization Process. Chinese Enterprise Models, University of New South Wales, Sydney, UNSW Centre for Chinese Studies.

Wank, D. L. (1999). Commodifying communism : business, trust, and politics in a Chinese city. Cambridge England ; New York, Cambridge University Press.

¹⁴ Guthrie, D. (2002). Information Asymmetries and the Problem of Perception: The Significance of Position in Assessing the Importance of Gunaxi in China. Social Connections in China. T. Gold, Guthrie, Doug and Wank, David. Cambridge, Cambridge University Press: 37 -56.

¹⁵ Chan, K. B. a. N. B. K. (2000). Myths and Misperceptions of Ethnic Chinese Cpitalism. Chinese Business Networks. K. B. Chan. Singapore, Prentice Hall. pg 292

and western countries, it may be argued that Chinese entrepreneurs base their decisions on market factors due to the strong institutional support¹⁶. Finally, reinforcing the argument in the other direction, there are studies that show that certain non-Chinese groups also make use of similar informal institutions when faced with a lack of formal institutional support¹⁷.

Based on the above, it is becoming increasingly common to assert that Chinese business modalities that make use of networks and guanxi are not the product of Chinese culture and society at all. Rather they are simply a rational response to the formal institutional environment. The conclusion that is drawn from this is that as the formal institutional environment in China and the region improves, the use of informal institutions will disappear as reliance upon formal institutions replaces them¹⁸. Thus the demand for guanxi and network institutions disappears as formal institutions appear. This, after all, is why informal institutions are not a feature of modern developed economies.

There is little doubt that the use of guanxi and networks is a rational response to poor formal institutions. However, studies that see it as nothing more than such a response do not question whether there is a demand for these informal institutions even when there are fully developed formal institutions to support business. If such a demand can be shown to exist then it is unlikely that the use of guanxi and business networks will simply disappear with the strengthening of formal institutions. Rather it is likely to be perpetuated.

In addition to the above, if a significant demand for such informal institutions can be shown to exist in economies with strong formal institutions but these informal institutions do not generally exist, it reopens the question as to whether Chinese culture, society and values are a critical part of the phenomena of guanxi and network use after all.

Mayfair Yang¹⁹ has already argued that the use of guanxi and networks is a tool in the toolbox of cultural assets available to Chinese entrepreneurs when needed. Her argument being an attempt to bring culture back into a debate that increasingly attempts to argue that essentially the use of guanxi and networks is the rational course that will be taken by any person in a poor formal institutional environment. Yang does view the use of guanxi as a rational response. However, she argues that it is a tool available to those with a Chinese culture and not necessarily to those without.

The fact that the use of guanxi and networks is not an available tool to many cultures is already attested to by the failure of many non-Chinese groups to succeed in weak

¹⁶ Li, P. S. (2000). Overseas Chinese Networks: A Reassessment. Chinese Business Networks: State Economy and Culture. C. K. Bun. Singapore.

¹⁷ Rutten, M. Individualism and Collective Forms of Business Organisation: Rural capitalists in India, Malaysia and Indonesia. Chinese Entrepreneurship and Asian Business Networks. T. Menkhoff, Solvay. London, Routledge.

¹⁸ Gomez, E. T. and X. Xiao (2001). Chinese business in Southeast Asia : contesting cultural explanations, researching entrepreneurship. Richmond, Surrey, Curzon, Guthrie, D. (2002). Information Asymmetries and the Problem of Perception: The Significance of Position in Assessing the Importance of Guanxi in China. Social Connections in China. T. Gold, Guthrie, Doug and Wank, David. Cambridge, Cambridge University Press: 37 -56.

¹⁹ Yang, M. M.-h. (2002). "The Resilience of Guanxi and Its New Deployments." China Quarterly 170: 459 - 476.

formal institutional environments due to the lack of institutional support. This fact being something which is too often ignored by those that seek to show that Chinese culture is irrelevant by demonstrating that some other group uses similar informal institutions to overcome problematic formal institutions.

The fact that Chinese culture and society does not have a monopoly on guanxi/network type institutions does not mean that these very institutions are not a particular characteristic of Chinese society. It simply means that some other societies or cultures have similar characteristics but not all of them and seemingly not modern western secular cultures with Christian roots.

This paper seeks to add strength to the type of argument put forward by Yang that the use of guanxi and networks are tools available to those with a Chinese culture to be used when it is rational to do so. It is generally accepted that there is a demand for guanxi and network use in poor formal institutional environments. The fact that many non-Chinese groups do not respond with the use of guanxi in these environments is ignored in favour of showing certain groups that do. In this paper, it is intended to show a demand for guanxi and networks in strong formal institutional environments. The fact that this additional demonstrated demand goes unmet in many societies without Chinese culture will further strengthen the argument that Chinese culture is a critical part of the use of guanxi and networks.

Demand for guanxi and networks in strong institutional environments will be demonstrated in the context of income tax evasion and arbitrage in such environments. As with other studies of guanxi, entrepreneurs will be the key focus. First the very high value of reduced tax burdens in most strong institutional environments will be demonstrated. Secondly, the key methods of reducing tax burdens through arbitrage and evasion as commonly encountered will be examined to see how tax reduction may be achieved and consider how the government may react to constrain these methods. Finally, the utility of informal bonds such as guanxi and social networks (where they exist) in facilitating tax arbitrage and evasion and circumventing government constraints will be demonstrated. Through these steps it will ultimately be asserted that guanxi and social networks significantly increase entrepreneurs' ability to engage in tax evasion and arbitrage. Given the high value associated with reducing tax burdens in most strong institutional environments, there is a continued demand for guanxi and social networks in such environments.

THE VALUE OF REDUCED BUSINESS TAX BURDENS TO ENTREPRENEURS

Income tax is a highly significant cost associated with business activities in most developed countries. It can in fact be the largest cost that any entrepreneur will face in relation to their business activities. This is because it is not unusual for the ultimate income tax burden to approach 50% of business profits. In addition to this other taxes such as value added tax may add further significant costs to a business activity.

The significance of the cost of taxation to entrepreneurs in developed countries is amply demonstrated in the Australian case. Here, the top marginal income tax rate including associated levies has just been reduced to 46.5% of net income. While the company tax rate is only 30% (already a highly significant amount), company business profits will ultimately be taxed at marginal individual rates as company profits are distributed as dividends to shareholders. This means that net income earned through companies is still potentially taxed at 46.5%.

In addition to income tax, Australia's goods and services tax (GST) will be imposed at a rate of 10% on all sales of goods and services, an amount that can be highly significant in relation to net profit. It would constitute 30% of net profit if net profit was 30% of sales. While GST is arguably a cost on consumers and not business, this is not entirely the case for all businesses and there is certainly a significant commercial advantage to those that can reduce their GST burden. If the GST is evaded the business can either reduce the cost of their product thereby increasing sales or keep the evaded GST and increase the return on their activities.

The rates of taxation imposed in countries such as Australia mean that there is indisputably significant value attached to the reduction of taxation. This will be the case whether this reduction is achieved through planning, avoidance or evasion. For example, a restaurant owner that sells a meal for \$22 may only make \$5.35 after costs (say \$10), income tax (\$4.65) and GST (\$2). This amounts to 24.3% of sales value. If this person were to successfully evade taxation by not declaring the sale for tax purposes the return becomes \$12 or 54.5%. This amounts to a more than doubling of the return on the activity.

The value associated with reducing tax burdens through more legitimate means in Australia is further attested to by the significance of the tax profession and the value of the tax planning industry. The potential cost of taxation is of such significance that a tax reduction technique, legitimate or otherwise, can determine the actual viability of a business venture or even an industry. It is indisputable that the ability to successfully reduce your tax burden is highly desirable and valuable.

REDUCING BUSINESS TAX BURDENS THROUGH ARBITRAGE

High general-rate income and other taxes inevitably allow for a lower tax burden in the case of certain taxpayers and situations. They may also impose a higher burden or in some manner differential tax treatment on certain other taxpayers or situations. There are numerous reasons for this including equity, social policy, simplicity, economic policy, enforceability and jurisdictional limitation. Thus taxpayers on lower incomes may attract a lower rate of taxation due to their inability to reasonably pay the general rates or to encourage their businesses to grow. On the other hand undesirable activities may attract heavier tax burdens while those that are desirable may attract a lighter burden. Finally, it may be impossible to collect a high rate of tax in relation to certain situations and the Government may then impose an enforceable lower amount of tax.

All the above issues inevitably result in a tax where highly differential tax burdens will arise depending on characterisation of taxpayer, business, both of these or some other factor. This in turn results in significant arbitrage possibilities as financial benefit accrues to those that can alter the relevant factors. These financial benefits are very valuable due to the high general rate of taxation imposed and the fact that one of the aforementioned policies may apply a nil tax rate or even a negative tax rate²⁰. The exploitation of these arbitrage opportunities is at the heart of tax planning and tax avoidance.

²⁰ This may be the case when a welfare payment is made to those on low incomes.

NON-MARKET BASED TRANSACTIONS AND ARBITRAGE

Non-market based transactions are integral to tax arbitrage through planning and avoidance. They can be used to shift income from one taxpayer to another who is granted differential treatment. They can also be used to transfer value from one situation to another. Finally, non-market based transactions can be used to alter the actual nature of a taxpayer or situation.

Many well known tax arbitrage strategies ranging from income splitting by professionals and small business to international transfer pricing by multinationals rest upon essentially simple non-market based transactions that shift income from one taxpayer to another. It may be a self-employed doctor paying a 'wage' to their spouse in the case of the former or the payment of royalties and management fees to offshore associates in the latter. In terms of altering the actual nature of a situation through non-market based transactions, an example may be inflating investment in research and development through an associate to qualify for preferential treatment.

EVASION OF TAXATION AND THE HIDING OF EVIDENCE

Evasion of taxation differs somewhat from arbitrage through planning and avoidance. In the case of evasion the taxpayer may seek to simply hide a transaction or activity entirely in order that they do not pay tax in relation to it. This may be readily done by simply not declaring an amount of cash income. However, to be done successfully at a significant scale, the taxpayer will need to be able to hide outward signs of business activity and income accumulation. This is because tax authorities can turn to these outward signs to assess income in cases of evasion. Successful large scale evasion therefore requires the hiding of asset ownership, the disguising of assets' origins, the hiding or disguising of business input transactions (For example employees and purchases) as well as the hiding of sales transactions. Tax authorities will otherwise be able to determine that revenue has been earned on the basis of business inputs or wealth accumulation alone.

From the above discussion it should be clear that the prevention of tax arbitrage by tax law and tax authorities depends upon the ability to identify and adjust non-market transactions. The prevention of evasion on the other hand depends upon the identification of any signs of business activity or wealth accumulation. Conversely, successful tax arbitrage rests upon disguising non-market transactions as market transactions while successful tax evasion rests upon hiding or disguising transactions and assets.

COUNTERING ARBITRAGE AND EVASION

Developed tax laws contain a range of measures that are aimed at dealing with both unacceptable arbitrage and evasion. These are complemented by policies and measures adopted by taxation administrators to prevent the same. In relation to arbitrage, the tax law will adjust or nullify or allow for the adjustment or nullification of non-market transactions. It may also deem certain transactions to be non-market transactions. Tax administration activity will in its turn be focused upon the identification of non-market transactions so that they can be adjusted or nullified as appropriate.

In relation to evasion, the tax law will contain measures that allow for the determination of a taxpayer's tax liability from available information when the usual complete information is unavailable. More critically, tax administration activity will

focus on identifying tax evasion and collecting any available information that will allow the law to operate correctly.

IDENTIFYING NON-MARKET TRANSACTIONS – THE ‘ASSOCIATE’ CONNECTION

A key difficulty that arises in relation to nullifying or adjusting non-market transactions is identifying them. Without such identification, adjustment is impossible. Once they are positively identified, adjustment or nullification of non-market transactions may be readily achieved. In carrying out the task of identifying non-market transactions both the tax law and administration make use of the separate concept of an associate taxpayer or related party (hereafter associate).

The concept of the associate is essential to the identification of non-market transactions and arrangements. To this point the concept of a non-market or non-arm's length transaction is often used synonymously with the concept of an associate or related party transaction. The concepts are, however, quite distinct. A non-market or non-arm's length transaction is a transaction that is not priced through market interactions such as bargaining. An associate or related party transaction, on the other hand, is a transaction between taxpayers that are related to one another. The link between the two concepts is that it is presumed that non-market transactions will only occur between associates.

If the presumption that association is a prerequisite to non-market transactions holds true, then identification of the latter can commence with the identification of the former. This makes the task much simpler as associates will generally be far more readily identifiable than non-market transactions due to the fact that characterizing entities and persons is easier than characterizing transactions. In some cases the law or administration may simply deem transactions between associates to be the relevant non-market value transactions. Doing this simplifies the process even further but may lead to inequities as it cannot be asserted that all transactions between associates are non-market based.

The presumption that association of some sort is a prerequisite for non-market transactions would appear to be self-evidently true. A taxpayer would be expected to maximize their individual utility through a market based bargaining process in all transactions unless the transaction is with another taxpayer associated with them. When the taxpayers are associated, however, they may be willing to depart from market bargaining as they may not ultimately suffer genuine economic loss by doing so.

This may be because they ultimately share wealth in some way or because they have a myriad of interactions through which they can correct for the prima-facie economic loss. Thus, a tax administrator may question the wage paid to a close family member in a small business but assume that the wage paid to an unrelated taxpayer is based upon the market. Likewise a purchase of goods from a subsidiary company overseas may be scrutinized but not one from an independent supplier.

There is therefore a strong link between the extent of arbitrage opportunity and non-market transactions. There is also a strong link between non-market transactions and associates. These two links together mean that the greater and more diverse a taxpayer's group of associates is, the greater that taxpayer's scope for tax arbitrage is. In addition the greater the group of associates, the larger the job of identifying them and scrutinizing their interactions is for the tax administration.

WHO ARE ASSOCIATES?

In considering the link between associates and non-market transactions, it is instructive to consider which taxpayers are usually considered to be associates. Section 318 of the Australian *Income Tax Assessment Act (1936)* (ITAA 1936) contains the most extensive definition of associate used in Australian income tax law. It is used in a number of key areas along with other less extensive definitions for certain purposes²¹. It has been reproduced in Appendix A along with an extract from s995-1 (*Income Tax Assessment Act 1997*) that defines what a relative is. The definition of a relative is essential to the definition of an associate as a relative will be an associate. Relatives are also “related entities” being a term that is used in the stead of associate for some purposes²².

A review of taxpayers that are included as associates in the definition indicates that they fall into several broad categories:

- Individuals that are relatives (including spouses) are associated
- Entities that have a control connection based upon shareholdings and other legal factors are associated
- Those in a legal partnership relationship are associated
- Trusts and their beneficiaries are associated
- Companies that customarily act in accordance with the wishes of others are associates of the others

A review of the above categories, in turn, indicates why in most cases associates will be readily identifiable. It should not be too difficult to identify if two natural persons are relatives as defined in s995-1. This can be done through family records. Identifying and proving control relationships between entities may in some cases be time consuming but share register records and legal documentation make the task accomplishable. Likewise with partnerships and trusts, contracts and establishment documentation will in most cases allow for ready identification of associates. It is acknowledged that this may not be so in all cases due to the concealment of contracts and other documentation. However, the need to reveal deeds and agreements in many formal processes ensures that concealment is not as common as it could be.

The final category may prove more difficult to determine and the extent of its application is far less clear than with the other categories. However, its inclusion is necessitated by the fact that companies can be controlled or influenced in a number of situations when the formal control tests are not otherwise satisfied. This results from the multiple powers that are available to both officers and shareholders in relation to a company that can be used to direct company actions. Thus while it is impossible to specifically list in the law all the possible situations when sufficient influence will exist, it will be possible to determine its existence in many circumstances.

²¹ The definition was originally intended for use in the Controlled Foreign Company regime but its use has been expanded to that of a general definition of “associate” in many respects. A key example of this is its use in Division 7A of the ITAA1936 dealing with deemed dividends. There was a separate definition of associate contained in s26AAB for s108 and other issues but this was largely similar to the s318 definition and has in fact been replaced by it.

²² Primarily s26-35 which limits deductions when they are paid to related entities.

The success of legal and administrative measures that attempt to contain tax arbitrage through non-market transactions is brought about through the scrutiny and adjustment of interactions between associated taxpayers. Such scrutiny and adjustment is feasible as associated taxpayers are readily defined and identifiable as being relatives and entities with defined formal legal relationships that are recorded. The underpinning logic of this mechanism is that taxpayers will only engage in non-market transactions and arrangements with relatives or entities with the type of recorded formal relationship specified above.

ASSOCIATES AND THE HIDING OF TRANSACTIONS, ASSETS AND ACTIVITIES

As with arbitrage through planning and avoidance, the process of identifying wealth and signs of business activity in countering tax evasion needs to pay particular attention to the associates of the taxpayer being scrutinised. In addition associates need to be considered in the process of reconstructing income figures during the deemed assessment process that is the remedy for tax evasion.

Wealth can be hidden by placing it formally in the names of associates. However, most taxpayers would be loath to put their assets formally into the hands of non-associates for fear of losing the assets that are no longer legally theirs. The origins of wealth can also be attributed to associates rather than the earning of income. When questioned on how a taxpayer afforded their house, for example, they may state that they received a large cash gift or loan from a relative in the past knowing that the relative will back them up. This strategy is unlikely to be feasible with non-associates who will be unlikely to support the taxpayer due to the perceived consequences for themselves.

Signs of business activity or factors that may be used in the process of generating a deemed assessment may also be hidden through the use of associates. Business inputs are a key basis for estimating gross income as they often have a strong statistical relationship to outputs. However if major business inputs such as purchases and labour are provided by associates they will be far more readily disguised than if non-associates are the relevant providers. This is for the same reason as stated above in the context of non-arm's length transactions. That is, non-associated employees will insist on formal records being kept in relation to their employment and wages for their own protection. These records are then available to the tax administration. However, associate employees may be willing to not record their activities or record their activities in a misleading fashion. This is because either the protection does not matter to them as they are essentially supported by the taxpayer or because they can ensure that they are treated correctly through non-formal mechanisms.

The availability of non-formal institutions for protection and trust in business activities and transactions comes hand in hand with dealing with associates. This alternative institutional support means that formal institutional support is not required if it is undesirable for a particular reason. A simple example would be that a parent that lends money to their child may be able to rely on family pressure to ensure that the money is repaid. There is then no need to draw up a loan contract, register a mortgage or ask for detailed business accounts in relation to the loan. These tasks will then be avoided if they are undesirable or even if they are simply not felt to add to the situation.

The need for formal institutional support is the driving factor behind the creation and availability of most documents, registers, records, accounts and contracts that record business activity, transactions and assets. These records are however then available to tax administrators and constitute their primary source of information to work with. This is clearly an undesirable outcome for a would-be tax evader. Returning to the parent above, the recording of the loan may lead to income tax on interest and stamp duty on contracts. It would also highlight the activity of the child for which the money is being used.

There is therefore a clear benefit and rationale for dealing with associates and relying on informal institutions for protection. If you sell to associates you may not need to issue receipts, if you buy from associates you may not need contracts and invoices, if you use associates as sources of credit and equity, you may not need to produce detailed and accurate financial statements or record ownership through a corporate register or partnership contract. Without these various records, the ability of the tax administration to determine the extent of a taxpayer's business activities is seriously hindered.

ASSOCIATE SCOPE

The previous sections of this paper have made several key assertions. The first is that the ability to successfully reduce or eliminate tax burdens through either arbitrage or evasion is highly valuable due to the significance of the legal tax burden in relation to private business in a developed economy.

The second key assertion is that the existence, extent and diversity of a taxpayer's associates have a direct bearing on the taxpayer's ability to successfully reduce or eliminate their income tax burden. This is the case regardless of whether this reduction or elimination is done through tax planning, tax avoidance or tax evasion. A person with a small group of associates may be able to divert a minor amount of income to take advantage of lower tax rates but is unlikely to be able to run their business through associate connections. The pool of associates would simply be inadequate to seek major debt or equity or locate supplies or custom.

The third key assertion is that the limiting of associates to specific close relatives such as those defined in s995-1 of Australia's Income Tax Act (1997) and entities with a formal specified relationship as covered in s318 (ITAA 1936) allows income tax law and administration to counter arbitrage and evasion by providing mechanisms to scrutinise, adjust or nullify interactions between what is a relatively discrete group of taxpayers. At the very least, these definitions mean that when deciding whether the interactions between two taxpayers should be investigated, the administration will be able to determine whether they are associated and therefore whether arbitrage and evasion may be an issue. They will also be within their rights in making any adjustments.

ASSOCIATES IN A WESTERN SOCIETY

Thus much arbitrage and evasion depends on associates while counter arbitrage and evasion measures target associates. The efficacy and appropriateness of these counter measures does however rest on the accuracy of the underlining assumption that non-market transactions and arrangements or those arrangements that are supported by informal institutions are generally limited to associates. That is, close relatives as

defined or taxpayers with a specified formal relationship. Put another way it depends upon associates only being those close relatives and specified taxpayers.

The accuracy of this assumption in societies such as modern Australian society and other western societies where such rules of association have been developed seems reasonable. If any criticism were to be leveled at the definition of associate in the ITAA, it would be that it is overly harsh. This is because the assumption that dealings between relatives such as adult siblings or individuals and their uncles and aunts would be off-market or informal is just as often false as it is true. Even many parents and their adult children in modern Australia can be said to deal with one another at arm's length. However, the inclusion of relatives as defined is necessary as it is clear that in western societies such relatives may deal with one another informally and off market.

The ideal of individual independence is deeply rooted and strong in western society²³. It is also growing stronger as witnessed by the growing propensity of even the closest familial relationships such as spouses to rely upon formal institutional specification of individual property rights rather than operate in a unitary fashion or trust informal institutions. This independence is reinforced by formal institutions and dependent on formal institutions. It in turn in itself reinforces these formal institutions. It leads to the belief that reality is that which is represented by formal institutions.

That is, an individual's property can only be that which is recorded as his or hers in titles, registers or contracts. To alter these records is to alter reality and to not have these records of ownership, say, is to not own. There are no strong informal institutions to create an alternative reality other than in the case, sometimes, of close relatives. Otherwise there must be a formally constituted relationship to offer appropriate formal institutional protection.

The result of the above situation is that associates (as defined in the ITAA) does largely capture the extent of possible non-market transactions and probably goes far beyond this extent. The fact that the associate definition used in the ITAA is not generally thought to be too narrow by tax administrators, policy makers or academics in Australia is testament to the adequacy of the definition in capturing those who may interact through informal institutions.

INCREASED SCOPE FOR ARBITRAGE IN A CHINESE SOCIETY

Focusing on Chinese society, a very different outcome emerges. The assumption that the group of taxpayers captured by the definition of relative and associate in the ITAA defines the extent of possible off-market or informal institutional interactions in relation to Chinese society is highly inaccurate. The extent of the group of persons who should be considered associates in Chinese society goes far beyond this group. It follows that in accordance with the second key assertion above, members of Chinese society have far more scope to engage in successful taxation arbitrage or evasion. In turn, in accordance with the first key assertion above, members of Chinese society possess a highly valuable tool if they operate as entrepreneurs in a state, such as Australia, that imposes income tax on business income at a high rate. If the entrepreneurs operating in that state come from various different social backgrounds,

²³ Fei, X. (1992). *From the Soil: The Foundations of Chinese Society*. Berkley, University of California Press. See Chapter 4 in particular

then the Chinese entrepreneurs have a significant competitive advantage through their extensive group of associates. As explained earlier, this competitive advantage may be so valuable in a competitive industry that it leads to domination of that industry by Chinese entrepreneurs.

THE PERPETUATION OF CHINESE BUSINESS MODALITIES IN STRONG FORMAL ENVIRONMENTS

Thus, it is a fallacy to assert that the strengthening of formal institutions will lead to the demise of the use of informal institutions in the manner that Chinese entrepreneurs have used them in the past. This is because their remains a significant benefit associated with these business modalities in the form of reduced tax burdens. Of course, there are also competitive disadvantages associated with the use of the modalities as an entrepreneur may be limited in the scope of their business operations due to the need to deal with associates. However, the cost associated with this often highlighted disadvantage needs to be compared to the value associated with reduced tax burdens. The advantage may indeed outweigh the disadvantage.

It is also worth noting that tax benefits may not be the only benefits associated with the use of informal institutions. In a highly regulated formal economy informal institutions can be used to reduce or eliminate other obligations in the same manner as taxation. For example liability and risk arising from customers and creditors can be reduced. In Australia, entrepreneurs go to great lengths to protect their assets from bankruptcy. This is attempted through formal methods such as trusts. However, the use of formal methods again opens the doors to authorities to counter such avoidance of liability. The widened scope of associates in the Chinese context is again an advantage here. As developed jurisdictions with strong formal institutions generally go hand in hand with high tax burdens and more regulation, the value of informal institutions is actually increased. It is interesting to note that the developed Chinese jurisdictions of Hong Kong and Singapore impose very low tax burdens.

COUNTERING ARBITRAGE IN CHINESE SOCIETY

Finally, to round off this discussion, the third assertion above needs to be returned to. It was asserted that authorities and laws can counter tax arbitrage and evasion also by focusing on associates. Here it must therefore be considered whether the increased scope of associates in Chinese society can be effectively countered by a wider definition of associate to be used in tax law and administration. If this is the case, the advantage of the informal institutions will be diminished or eliminated and there will be no value in them and no demand for them.

ASSOCIATES IN CHINESE SOCIETY

The literature on Chinese social connections, guanxi and networks is extensive. The key focus of this literature is the strong informal bonds (guanxi) that individuals in Chinese society have with certain other individuals that are mutually beneficial. The concept of the network refers to groups of individuals linked by these bonds for common benefit. The extent both in terms of numbers of individuals and in geographical terms is also a noted feature of networks. The strength of these informal bonds is what has allowed Chinese entrepreneurs to succeed in the face of a poor or antagonistic formal institutional environment. It would follow from the very nature of guanxi and networks that anybody that is linked to another person through a guanxi network should be considered to be an associate for taxation purposes if such a definition is to be effective in a Chinese society.

The question that arises from the above is whether persons that are likely to have the appropriate relationship with one another can be characterized in a way that allows the law to identify them. This is what is done in the current Australian law where close relatives (as defined) are included as associates. Referring to seminal work on Chinese society such as that done by Fei Xiaotong²⁴ as well as to some studies of 'patricorps' in South East Asia and beyond²⁵, the impression is that membership of the patrilineage-extended family would be the appropriate characteristic.

FAMILIES AND PATRILINEAGES

Fei's work argues that economic activity is organized through the patrilineage in Chinese society and that relatives with the appropriate relationship will almost inevitably collude for mutual benefit. Very specific detail is often given on the nature and strength of the relationship within the family as well. For example married daughters would not be expected to have a strong collaborative relationship with their parents and brothers but sons would. If these assertions hold true for Chinese society, it means that it is possible to define associates in a manner more suited to Chinese society. The members of the patrilineage as defined by Fei would all be defined as associates creating a far wider net than that provided by the definition of relative in s995-1 of the ITAA.

The definition would also differ as arguably (for reasons just mentioned) married daughters, say, would not be appropriately branded associates. The strength of the informal bonds between patrilineage members as portrayed in Fei's work would merit a case for the patrilineage to constitute a tax unit let alone a group of associates.

FLEXIBILITY AND STRENGTH IN RELATIONSHIPS

Having seen a possible solution to defining associates in a Chinese society, a problem arises when further evidence on Chinese social connections is reviewed. That is that Fei's descriptions are not an accurate portrayal of Chinese social relationships at least in contemporary Chinese society. Arguably, they were not accurate in the past either. Rather, they represent an ideal for Chinese social relationships, a core of ideas which does underpin Chinese society but does not portray it accurately. Indeed, a deeper reading of Fei's work indicates his awareness of this²⁶.

There is clear evidence that guanxi relationships or network membership in contemporary Chinese society is not limited to blood relations. There is also evidence that two closely related (by blood) persons may not be involved in collusive economic activity even when the opportunity exists. This is particularly so in China itself. Rather strong informal relationships are developed between individuals through a commonality that cannot be typified in any concrete manner. It may be that the individuals are old school friends or army friends. It may be that they are blood relatives. It may be that they originate from the same village or that their ancestors are from the same village. The commonality may in itself be quite superficial, however, a strong bond may nevertheless be formed.

²⁴ Ibid. pp 83 -85

²⁵ See amongst many: Gates, H. (1996). *China's Motor: A Thousand Years of Petty Capitalism*. Ithaca, NY, Cornell University Press.

EAAU (1995). *Overseas Chinese business networks in Asia*. Parkes Australia, East Asia Analytical Unit Dept. of Foreign Affairs and Trade.

²⁶ (Fei, 1992) Pp 63 -64

The strength and success of these bonds rests at least partly on the invocation of the institutions of trust, loyalty and connectivity as described by Fei. However, the very flexibility of whom these bonds can be formed between is in itself a key feature of Chinese social connections. The most important factor is not the commonality seized upon but the talent of the individuals involved at building relationships. This flexibility is recognized even in the earliest works such as Fei's where after indicating that the familial bonds are so strong carries his analysis further to show that despite the strength of these bonds they can actually be discarded when they are not useful and be generated with others who are not members of the family. He indicates that a successful person is one that is good at "extending" himself²⁷. That is, at extending his strong informal bonds.

ALL THAT MATTERS IS WHETHER THERE IS A RELATIONSHIP

Ultimately, in Chinese society what counts is whether two persons have the "guanxi" relationship and not whether they are family or school friends²⁸. This means that it is not possible to describe a category of persons such as "relatives" in the law and hope to capture the majority of those that have strong informal bonds. The only way to describe them would be that associates are those that have a (say) "guanxi" bond but this does absolutely nothing to help identify those that do.

If a wide definition was drawn up to include the extensive blood relative network, old school and army associates, those from a common village or dialect then many may be included. However such a definition would be far too extensive to be workable and would include too many people that do not have any such relationship. Thus it is impossible to combat arbitrage and avoidance through the use of associates by focusing on associates when dealing with Chinese society. Associates while many and diverse cannot be identified.

THE VALUE OF GUANXI IN A STRONG FORMAL INSTITUTIONAL ENVIRONMENT

There is ultimately a great deal of value associated with doing business through informal institutions such as guanxi networks in a strong formal institutional environment. This value does not arise from the protection that informal institutions offer when there are no strong formal institutions as has been the case historically in China and South East Asia. Rather, it arises from the ability to avoid or reduce costs and obligations associated with strong institutional environments. The case here has been made in relation to taxation but it is clear that issues of liability to creditors and others are also relevant.

The fact that this demand for guanxi and networks in doing business in strong formal institutional environments such as Australia goes unmet on the part of the local "western" society is evidence in favour of the hypothesis that the use of guanxi is a cultural tool available to those with Chinese culture²⁹. Otherwise why are strongly bounded networks of individuals not a feature of Australian business large or small? There are clear advantages to their use and, at least in the privately owned sector, not necessarily significant disadvantages at the level of the entrepreneurs themselves³⁰. However, the inability to trust others without first having a formal relationship and

²⁷ (Fei, 1992) Pg. pp 60 -70

²⁸ (Hendrischke, 2006; M.-h. M. Yang, 1994)

²⁹ (M. M.-h. Yang, 2002)

³⁰ This is not to say there are not disadvantages to the economy as a whole. There are.

record of specific rights is more a feature of these Australian entrepreneurs and their relationships with others.

The advantages of the use of informal institutions are actually maximized when other groups operating in a particular jurisdiction do not rely upon informal institutions as the reduced costs of taxation and potential liability give those operating through informal institutions a competitive advantage that is extremely significant. Of course, they have competitive disadvantages as well but these need to be valued against the massive advantages.

Thus the use of informal institutions such as guanxi networks should be expected to continue regardless of attempts made to strengthen formal institutions. This is because there is significant value associated with this modality of operation. This is regardless of whether the future development of China or Malaysia is being considered or whether the exposure of, say, Australia to Chinese business modalities is being considered. In the global economy this exposure of states such as Australia is not a case of immigrant society. Without any local Chinese community, the Australian tax system is fully exposed to Chinese business modalities through international trade and investment.

Finally, it is worth considering whether strong formal institutions can be developed or maintained in the face of the significant use of informal institutions. When a significant group of persons makes use of informal institutions, the formal institutions are ultimately weakened for those that depend on them. This weakening increases the value of the use of informal institutions for the same reason they were used in China in the first place. That is to counteract a weak formal institutional environment. Thus the use of informal institutions weakens the formal institutional environment while the weakening of the formal institutional environment increases the use of informal institutions. This is a sobering thought in a globalizing world that calls for a single institutional environment as this global institutional environment may not be dominated by strong formal institutions.

APPENDIX A – DEFINITION OF AN ASSOCIATE

Section 318 Associates

318(1) [Entity]

For the purposes of this Part, the following are associates of an entity (in this subsection called the "primary entity") that is a natural person (otherwise than in the capacity of trustee):

- (a) a relative of the primary entity;
- (b) a partner of the primary entity or a partnership in which the primary entity is a partner;
- (c) if a partner of the primary entity is a natural person otherwise than in the capacity of trustee - the spouse or a child of that partner;
- (d) a trustee of a trust where the primary entity, or another entity that is an associate of the primary entity because of another paragraph of this subsection, benefits under the trust;

- (e) a company where:
 - (i) the company is sufficiently influenced by:
 - (A) the primary entity; or
 - (B) another entity that is an associate of the primary entity because of another paragraph of this subsection; or
 - (C) another company that is an associate of the primary entity because of another application of this paragraph; or
 - (D) 2 or more entities covered by the preceding sub-subparagraphs; or
 - (ii) a majority voting interest in the company is held by:
 - (A) the primary entity; or
 - (B) the entities that are associates of the primary entity because of subparagraph (i) of this paragraph and the preceding paragraphs of this subsection; or
 - (C) the primary entity and the entities that are associates of the primary entity because of subparagraph (i) of this paragraph and because of the preceding paragraphs of this subsection.

318(2) [Company]

For the purposes of this Part, the following are associates of a company (in this subsection called the "primary entity"):

- (a) a partner of the primary entity or a partnership in which the primary entity is a partner;
- (b) if a partner of the primary entity is a natural person otherwise than in the capacity of trustee - the spouse or a child of that partner;
- (c) a trustee of a trust where the primary entity, or another entity that is an associate of the primary entity because of another paragraph of this subsection, benefits under the trust;
- (d) another entity (in this paragraph called the "controlling entity") where:
 - (i) the primary entity is sufficiently influenced by:
 - (A) the controlling entity; or
 - (B) the controlling entity and another entity or entities; or
 - (ii) a majority voting interest in the primary entity is held by:
 - (A) the controlling entity; or
 - (B) the controlling entity and the entities that, if the controlling entity were the primary entity, would be associates of the controlling entity because of subsection (1), because of subparagraph (i) of this

paragraph, because of another paragraph of this subsection or because of subsection (3);

(e) another company (in this paragraph called the "controlled company") where:

(i) the controlled company is sufficiently influenced by:

(A) the primary entity; or

(B) another entity that is an associate of the primary entity because of another paragraph of this subsection; or

(C) a company that is an associate of the primary entity because of another application of this paragraph; or

(D) 2 or more entities covered by the preceding sub-subparagraphs; or

(ii) a majority voting interest in the controlled company is held by:

(A) the primary entity; or

(B) the entities that are associates of the primary entity because of subparagraph (i) of this paragraph and the other paragraphs of this subsection; or

(C) the primary entity and the entities that are associates of the primary entity because of subparagraph (i) of this paragraph and the other paragraphs of this subsection;

(f) any other entity that, if a third entity that is an associate of the primary entity because of paragraph (d) of this subsection were the primary entity, would be an associate of that third entity because of subsection (1), because of another paragraph of this subsection or because of subsection (3).

[Subsection (2) amended by Act No 41 of 1998 s 3 and Sch 6, Item 18, with effect effect from 4 June 1998, by deleting 'application of this' from sub-subparagraph (e)(i)(B).]

318(3) [Trustee]

For the purposes of this Part, the following are associates of a trustee (in this subsection called the "primary entity"):

(a) any entity that benefits under the trust;

(b) if a natural person benefits under the trust - any entity that, if the natural person were the primary entity, would be an associate of that natural person because of subsection (1) or because of this subsection;

(c) if a company is an associate of the primary entity because of paragraph (a) or (b) of this subsection - any entity that, if the company were the primary entity, would be an associate of the company because of subsection (2) or because of this subsection.

318(4) [Partnership]

For the purposes of this Part, the following are associates of a partnership (in this subsection called the "primary entity"):

- (a) a partner in the partnership;
- (b) if a partner in the partnership is a natural person - any entity that, if that natural person were the primary entity, would be an associate of that natural person because of subsection (1) or (3);
- (c) if a partner in the partnership is a company - any entity that, if the company were the primary entity, would be an associate of the company because of subsection (2) or (3).

318(5) [Time of association]

In determining, for the purposes of this section, whether an entity is an associate of another entity at a particular time (in this subsection called the "test time"):

- (a) an entity (in this subsection called the "public unit trust entity") that, apart from this subsection, is the trustee of a public unit trust at the test time is to be treated as if it were a company instead of a trustee; and
- (b) the public unit trust entity is taken to be sufficiently influenced by another entity or other entities if the public unit trust entity is accustomed or under an obligation (whether formal or informal), or might reasonably be expected, to act in accordance with the directions, instructions or wishes of the other entity or other entities (whether those directions, instructions or wishes are, or might reasonably be expected to be, communicated directly or through interposed companies, partnerships or trusts); and
- (c) another entity or other entities are taken to hold a majority voting interest in the public unit trust entity if either of the following percentages is not less than 50%:
 - (i) the percentage of the income of the trust represented by the share of the income to which the other entity or other entities are entitled, or that the other entity or other entities are entitled to acquire;
 - (ii) the percentage of the corpus of the trust represented by the share of the corpus to which the other entity or other entities are entitled, or that the other entity or other entities are entitled to acquire.

318(6) [Terms explained]

For the purposes of this section:

- (a) a reference to an entity benefiting under a trust is a reference to the entity benefiting, or being capable (whether by the exercise of a power of appointment or otherwise) of benefiting, under the trust, either directly or through any interposed companies, partnerships or trusts; and
- (b) a company is sufficiently influenced by an entity or entities if the company, or its directors, are accustomed or under an obligation (whether formal or informal), or might reasonably be expected, to act in accordance with the directions, instructions or wishes of the entity or entities (whether those directions, instructions or wishes are, or

might reasonably be expected to be, communicated directly or through interposed companies, partnerships or trusts); and

(c) an entity or entities hold a majority voting interest in a company if the entity or entities are in a position to cast, or control the casting of, more than 50% of the maximum number of votes that might be cast at a general meeting of the company.

318(7) ["Spouse"]

In this section, and in any other provision of this Act insofar as that provision has effect for the purposes of this section, a reference to the spouse of a person (in this subsection called the "first person") does not include a reference to a person who is legally married to the first person but is living separately and apart from the first person on a permanent basis.

Section 995-1 (ITAA 1997)

relative of a person means:

- (a) the person's *spouse; or
- (b) the parent, grandparent, brother, sister, uncle, aunt, nephew, niece, lineal descendent or *adopted child of that person, or of that person's spouse; or
- (c) the spouse of a person referred to in paragraph (b).

BIBLIOGRAPHY

- Backman, M. (1999). Asian eclipse : exposing the dark side of business in Asia. Singapore, J. Wiley.
- Biggart, N. W. and G. G. Hamilton (1997). On the Limits of Firm-Based Theory to Explain Business Networks: The Western Bias of Neoclassical Economics. The economic organization of East Asian capitalism. Thousand Oaks, Calif., Sage Publications.: 33 - 54.
- Chan, K. B. (2000). Chinese business networks : state, economy and culture. Singapore
Copenhagen S, Demark, Prentice Hall Pearson Education Asia Pte. Ltd.; Nordic Institute of Asian Studies.
- Chan, K. B. (2002). Preface. Chinese Entrepreneurship and Asian Business Networks. T. Menkhoff, Solvay. London, RoutledgeCurzon: xii - xiv.
- Chan, K. B. a. N. B. K. (2000). Myths and Misperceptions of Ethnic Chinese Capitalism. Chinese Business Networks. K. B. Chan. Singapore, Prentice Hall.
- EAAU (1995). Overseas Chinese business networks in Asia. Parkes Australia, East Asia Analytical Unit Dept. of Foreign Affairs and Trade.
- Fei, X. (1992). From the Soil: The Foundations of Chinese Society. Berkley, University of California Press.
- Gates, H. (1996). China's Motor: A Thousand Years of Petty Capitalism. Ithaca, NY, Cornell University Press.
- Gomez, E. T. and X. Xiao (2001). Chinese business in Southeast Asia : contesting cultural explanations, researching entrepreneurship. Richmond, Surrey, Curzon.

- Granovetter, M. (1985). "Economic Action and Social Structure: The Problem of Embeddedness." American Journal of Sociology **91**: 481 - 510.
- Guthrie, D. (2002). Information Asymmetries and the Problem of Perception: The Significance of Position in Assessing the Importance of Guanxi in China. Social Connections in China. T. Gold, Guthrie, Doug and Wank, David. Cambridge, Cambridge University Press: 37 -56.
- Hendrischke, H. (2002). The Role of Social Capital, Networks and Property Rights in China's Privatization Process. Chinese Enterprise Models, University of New South Wales, Sydney, UNSW Centre for Chinese Studies.
- Hendrischke, H. (2006). "Some Thoughts on New Chinese Business Networks." Forthcoming.
- Krug, B. (2004). China's rational entrepreneurs : the development of the new private business sector. London ; New York, RoutledgeCurzon.
- Lever-Tracy, C., D. F.-K. Ip, et al. (1996). The Chinese diaspora and mainland China : an emerging economic synergy. Houndmills, Basingstoke, Hampshire, New York, Macmillan Press; St. Martins Press.
- Li, P. S. (2000). Overseas Chinese Networks: A Reassessment. Chinese Business Networks: State Economy and Culture. C. K. Bun. Singapore.
- Redding, S. G. (1990). The spirit of Chinese capitalism. Berlin, W. de Gruyter.
- Rutten, M. Individualism and Collective Forms of Business Organisation: Rural capitalists in India, Malaysia and Indonesia. Chinese Entrepreneurship and Asian Business Networks. T. a. G. Menkhoff, Solvay. London, Routledge.
- Wang, T. P. (1994). The Origins of the Chinese Kongs. Selangor, Pelanduk Publications.
- Wank, D. L. (1999). Commodifying communism : business, trust, and politics in a Chinese city. Cambridge England ; New York, Cambridge University Press.
- Wu, W.-p. (2000). Transaction Cost, Cultural Values and Chinese Business Networks. Chinese Business Networks. C. K. Bun. Singapore, Prentice Hall.
- Yang, M.-h. M. (1994). Gifts, favors, and banquets : the art of social relationships in China. Ithaca, N.Y., Cornell University Press.
- Yang, M. M.-h. (2002). "The Resilience of Guanxi and Its New Deployments." China Quarterly **170**: 459 - 476.

The Effect of Audit Strategy Information on Tax Compliance – An Empirical Study

Leif Appelgren*

Abstract

This paper deals with an experiment by the Swedish Tax Agency to test the effect of information to taxpayers regarding different audit strategies¹. The experiment involved approximately 900 sole proprietors, divided into three groups, where one was informed that audits would focus on taxpayers declaring the lowest income, i.e. according to a rational audit strategy. Another group was told that audits would be made at random whereas the third was a control group. The effect of strategy information was measured as the change in declared income between years. The principal finding was that declared income increased significantly more in the rational-audit-strategy group than in the control group.

INTRODUCTION AND THEORY

Tax audit theory prescribes that the audit risk should decrease with declared income, thus giving the taxpayer an incentive to reduce the fraud amount in order to reduce the risk of detection and sanctions. The objective of this paper is to study whether taxpayers in reality behave as predicted by theory.

Tax audits have a direct effect, i.e. that fraud is detected, resulting in the collection of tax and penalties. The audits also have an indirect deterrent effect, i.e. rational taxpayers are made aware that they may be audited and therefore adapt their behaviour to the expected degree of auditing.

One specific difficulty in the design of audits is in modelling how the deterrent effect depends on the actions of the auditor. A natural starting point is to assume that the taxpayer behaves rationally, in other words, maximizes his expected utility. This approach leads to the model introduced by Gary Becker (1968), which is based on the assumption that a crime is committed if the expected utility of the crime exceeds the expected cost of the sanction.

* Economic Information Systems, Department of Management and Engineering, Linköpings universitet, Linköping, Sweden. Email: leif.appelgren@lac.se. The author would like to thank Lennart Wittberg at the Swedish Tax Agency for the opportunity to conduct this study and for valuable assistance in the analysis. The author also wishes to express his appreciation to Birger Rapp, Professor of Economic Information Systems, Linköpings universitet, for financial support. Finally, the author is grateful to Stig Danielsson, Associate Professor of Statistics, Department of Mathematics, Linköpings universitet, for advice and assistance concerning the statistical analysis.

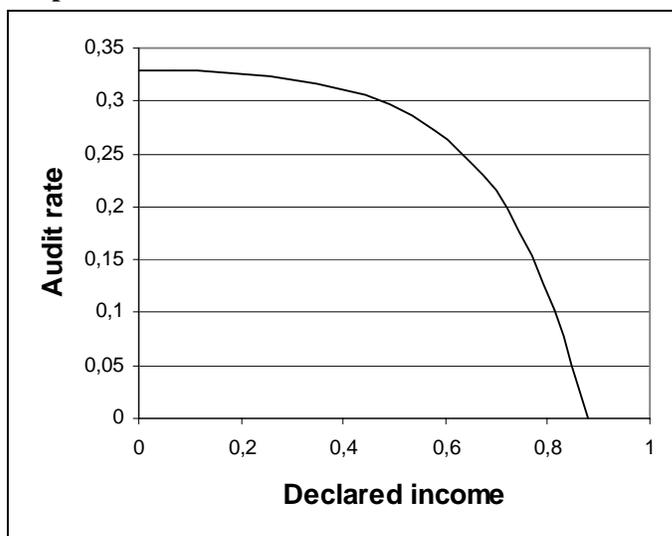
¹ The results have been briefly described in Swedish in Appelgren (2006)

The Becker model was first applied to tax fraud by Allingham & Sandmo (1973), who used a concave utility function, i.e. one with decreasing marginal utility, in order to determine the optimal amount of fraud when the audit rate is constant and known to the taxpayer. Other applications, which also included the behaviour of the auditor, have been studied by Reinganum & Wilde (1986) and Erard & Feinstein (1994). The theoretical tax fraud models are well described in a survey by Andreoni, Erard & Feinstein (1998).

Reinganum & Wilde (1986) studied the optimal allocation of audit resources to a homogenous group of taxpayers when the cost per audit is given. Such a homogenous group may consist of craftsmen in one-man enterprises or taxi companies with one car. An important assumption is that the only information available to the auditor regarding the individual taxpayer is the declared income. In addition, the auditor knows the distribution of true income, for instance from earlier random audits.

In the model developed by Erard & Feinstein (1994), the cost of an audit is replaced by a constraint that the number of audits is given. The authors also improved the model by introducing the concept of a known fraction of honest taxpayers, i.e. taxpayers who always declare their true income. The remaining taxpayers are assumed to behave rationally. The model has been further developed by Appelgren (2003).

Figure 1: A typical optimal audit-rate function



The basis of the models developed by Reinganum & Wilde and Erard & Feinstein is that the auditor observes declared income only and bases his audit decision on this observation. The models lead to an optimal audit-rate function which decreases with declared income, as illustrated in Figure 1.

PREVIOUS EMPIRICAL RESEARCH

The effects of information about audit activity have been studied previously in field experiments in Minnesota, USA, the United Kingdom and Australia. The limited number of experiments is probably due to high costs and confidentiality issues involved in the use of actual taxpayer data. In the Minnesota and UK experiments, the

effect of tax advisors (paid preparers) is studied since this may have influence on the effect of audit information.

In the Minnesota experiment (Blumenthal et al, 2001), taxpayers were informed that their tax returns would be audited; this led to significantly higher declared incomes for high-opportunity groups (taxpayers with business or farm income) with low and medium incomes. The same effect was not noted for high income earners. A possible explanation for this surprising result is that high-income earners increased their use of tax advisors under the threat of audit, and that those advisors were able to identify legal means for tax evasion, leading to lower declared income.

In the UK experiment (Hasseldine et al, 2007), more than 7,300 small enterprises were studied. They were considered to belong to a high-risk group, with a turnover just below the limit above which a more detailed tax reporting would be required. The companies were divided into six groups, one control group and five groups which received different letters characterised by the terms “*Enabling*”, “*Citizenship*”, “*Increased audit*”, “*Audit/penalties*” and “*Preselected audit*”. The principal result was that significantly higher turnover and net profit were reported in the three audit-related groups than in the control group. The monetary profit effects amounted to GBP 176 – 770 in the three groups compared to the control group.

The Australian experiment (Wenzel & Taylor, 2004) was carried out on 9,000 taxpayers in order to measure the effect of a specific form (Rental Property Schedule, RPS) for itemizing deductions made in conjunction with rental property income. The main result of the study was that when the RPS was used to account towards the tax office for the deduction claims, it reduced deductions with 5-7.5%. A mere warning letter or a schedule for personal use only had no effect on cost deductions. The tone of letter was either “soft” or “hard”, where the latter included an audit threat. Overall, the tone of letter had no effect, but it had a positive effect (smaller deductions) on taxpayers who received the RPS for the first time, whereas it had a negative effect on taxpayers who had received the schedule before.

The results in the studies above concerning the effect of audit information are mixed. The effect is clearly positive in the UK study, mainly positive in the Minnesota study and small in the Australian study. A reason for the mixed results may be that the taxpayer has an ex ante assessment concerning the audit risk, and if the audit letter merely confirms this assessment, the effect will be small.

RESEARCH QUESTION

All the models described above assume that the taxpayer is rational and tries to select the amount of tax fraud in order to maximize his expected utility. With an audit rate decreasing with increasing declared income, the taxpayer has an incentive to increase his declared income, i.e. decrease the amount of fraud, in order to reduce the audit risk. In order to obtain this effect, the auditor must inform the taxpayers about the rational audit strategy with decreasing audit rate.

It is not obvious that taxpayers behave as rationally as the theoretical models assume. The scope of this paper is to study whether real-world taxpayers adapt to information about a decreasing audit rate by reducing the fraud amount. More specifically, the main research question is:

Does information about a rational audit strategy with a decreasing audit rate reduce tax fraud as compared to information about a random audit strategy?

TEST DESIGN

This paper concerns an empirical test of the effect of information to taxpayers concerning different audit strategies. The test was carried out by the Linköping Regional Office of the Swedish Tax Agency in 2003-2004 on approximately 900 sole proprietors. The primary objective was to investigate whether information to taxpayers about a near-optimal audit strategy reduces tax fraud compared to information about a more conventional audit strategy, i.e. pure random audits. Information concerning the use of tax advisors/paid preparers was not collected.

The opportunities for tax evasion for individuals with income from employment are limited in Sweden as employers supply the tax authorities with statements on employee remuneration. It is therefore natural to perform an experiment on a group of enterprises. In order to obtain a large homogenous group, sole proprietors mainly in craft trades were selected.

The test was conducted on sole proprietors without employees and with little or no income from employment (maximum SEK 10,000 in the year 2002, where 1 SEK is approximately equal to 0.1 Euro). These owners were supposed to support themselves with their business. Further, the sample was limited to men below the age of 55 in order to concentrate on a high risk group (younger men are more fraudulent than women and older men). The trades included were craftsmen in the building industry, auto-repair craftsmen and hairdressers. Those trades were selected by the Tax Agency as they are the largest groups of sole proprietorships.

According to the theoretical work referred to above, the optimal audit strategy for a homogeneous group of taxpayers is to concentrate audits on those who declare the lowest income. In the experiment, however, the total net cash flow of the household was used instead as the basis for audit selection. Net cash flow is defined as declared income after tax, adjusted for non-cash items like depreciation and allocation to tax allocation reserves, as well as for cash items not included in income such as amortisation and new borrowing.

Three groups were studied, each with around 300 firms.

- | | |
|-------------------|--|
| A. Rational Group | The members were informed by mail that audits would concentrate on taxpayers who declare the lowest net cash flow (Appendix 1) |
| B. Random Group | The members were informed by mail that taxpayers to be audited would be selected at random (Appendix 2) |
| C. Control Group | The members received no information |

The three groups were geographically separated within the region, which encompasses the counties of Östergötland and Jönköping. Without such separation, there might be confusion if two colleagues were to find out that they had received different information regarding the upcoming tax audit.

Information was sent out during the second half of 2003 in order for it to affect the accounting for the remainder of the year and the tax return in May 2004. The Tax

Agency expected a possible negative reaction to the audit letters, especially from the Rational Group. Therefore, a service phone number was provided in the audit letters. The Tax Agency registered a total of only 11 phone calls, none of which with negative or critical content.

The effect of the strategy information in the audit letters was measured by comparing declared income for 2003 with declared income for 2002. The hypothesis was that the Rational Group would show a larger increase compared to the Control Group, and that the Random Group would fall between the other two since all information to taxpayers regarding audits is assumed to have a certain deterrent effect.

In the analysis, it was evident that additional delimitations should have been made in the selection of taxpayers. First, firms with income from employment in the year 2003 should have been excluded, in consistency with the exclusion of such firms in 2002. Moreover, a number of firms had used subcontractors extensively, and it can be argued that they should have been excluded like firms with employees.

As Sweden has a loss-carry-forward system, firms showing a loss in 2002 may have had losses in previous years which are included in the 2002 loss. In addition, the result in 2003 includes a carry-forward of the loss shown in 2002. Therefore, firms showing negative income in 2002 should be excluded.

The results reported below refer to data where firms with employment income exceeding SEK 10,000 in 2003 as well as firms with negative income in 2002 have been excluded. Regarding subcontractors, we present results from two sets of data, one excluding firms using subcontractors and one including such firms.

The data set included the Swedish Industry Classification Code (SNI) for each firm. It is evident from Tables 1 and 2 that the construction industry is predominant in the data.

An assumption behind the experiment is that all taxpayers belong to a homogenous group with random variations in income change between years and with randomly varying response to the audit letters. However, an analysis of the distribution of income change (Appendix 3) indicates that the population may consist of two distinct sub-groups, one “honest” group with a smaller standard deviation in income change and a smaller response to the audit letters, and one “fraudulent” group with a larger standard deviation in income change and a stronger response to the audit letters. This sub-group hypothesis has not been taken into account in the analysis below but would be of great interest as a subject for further research.

Table 1. Industry classification of participating firms, excluding firms with subcontractors

| Industry | SNI code | Rational Group | Random Group | Control Group |
|--|----------|----------------|--------------|---------------|
| Manufacture of metal products, machinery and equipment | 28, 29 | 0 | 0 | 0 |
| Demolition of buildings, earth moving | 451 | 6 | 13 | 19 |
| Construction of buildings etc | 452 | 40 | 54 | 41 |
| Installation (electric, plumbing etc) | 453 | 16 | 20 | 13 |
| Painting, plastering, floor and wall covering, glazing etc | 454 | 39 | 48 | 28 |
| Renting of construction/demolition equipment with operator | 455 | 4 | 6 | 18 |
| Auto repair | 502 | 6 | 5 | 8 |
| Hairdressers | 93021 | 7 | 12 | 15 |
| Total number of taxpayers | | 118 | 158 | 142 |

Table 2. Industry classification of participating firms, including firms with subcontractors

| Industry | SNI code | Rational Group | Random Group | Control Group |
|--|----------|----------------|--------------|---------------|
| Manufacture of metal products, machinery and equipment | 28, 29 | 1 | 0 | 2 |
| Demolition of buildings, earth moving | 451 | 16 | 25 | 24 |
| Construction of buildings etc | 452 | 71 | 91 | 71 |
| Installation (electric, plumbing etc) | 453 | 47 | 35 | 30 |
| Painting, plastering, floor and wall covering, glazing etc | 454 | 79 | 76 | 65 |
| Renting of construction/demolition equipment with operator | 455 | 9 | 8 | 25 |
| Auto repair | 502 | 12 | 12 | 12 |
| Hairdressers | 93021 | 13 | 22 | 22 |
| Total number of taxpayers | | 248 | 269 | 251 |

STATISTICAL ANALYSIS

The standard deviation of income change

In the preparations for the experiment, we expected that the standard deviation in income change between the two years would not exceed 25% of the average income. If two groups are compared, the standard deviation in the difference increases by the square root of 2, i.e. up to 35%. With 300 firms in each group, the standard deviation of the average income is reduced by the square root of 300, i.e. to approximately

2.0%. If a change in audit strategy would result in a change in declared income by 4%, the change would be statistically significant at the 5% level.

Actual data for the three groups are shown in Tables 3 and 4. The expectation regarding the standard deviation was apparently wrong, as the standard deviation was 35-50% of the average income instead of 25%. Furthermore, the size of the groups was reduced due to the additional limitations made above. Therefore, the difference in income change between groups had to be between 6 and 10% in order to be statistically significant.

Table 3. Average income, average income change and standard deviation in income change, excluding firms with subcontractors

| Amounts in SEK 1000 | Group size | Average income 2002 | Average increase 2002-2003 | Standard deviation | Relative standard deviation |
|---------------------|------------|---------------------|----------------------------|--------------------|-----------------------------|
| Rational Group | 118 | 134 | 27.2 | 50 | 37% |
| Random Group | 158 | 150 | 17.0 | 58 | 39% |
| Control Group | 142 | 149 | 13.4 | 56 | 38% |

Table 4. Average income, average income change and standard deviation in income change, including firms with subcontractors

| Amounts in SEK 1000 | Group size | Average income 2002 | Average increase 2002-2003 | Standard deviation | Relative standard deviation |
|---------------------|------------|---------------------|----------------------------|--------------------|-----------------------------|
| Rational Group | 248 | 154 | 19.5 | 77 | 50% |
| Random Group | 269 | 164 | 14.0 | 55 | 34% |
| Control Group | 251 | 157 | 11.6 | 54 | 34% |

Testing

The results of a simple statistical test are shown in Tables 5 and 6, where the hypothesis tested is that two groups have the same mean. The distribution of the difference in average income is approximately normal with an estimated standard deviation s calculated from

$$s^2 = s_1^2/n_1 + s_2^2/n_2$$

where s_i and n_i are the sample size and the estimated standard deviation of the compared groups, respectively.

Table 5. Data for testing the difference in average income change between groups, excluding firms with subcontractors

| Amounts in SEK 1000 | Difference in average income change Δm | $\Delta m/s$ | Significance level |
|---------------------------------|--|--------------|--------------------|
| Rational Group vs Control Group | 13.8 | 2.10 | 3.6% |
| Random Group vs Control Group | 3.6 | 0.55 | 58% |
| Rational Group vs Random Group | 10.2 | 1.56 | 12% |

In the *Subcontractors Excluded* case, shown in Table 5, the Rational Group shows a larger income change compared to the Control Group, significant on the 4% level. This indicates that information concerning the “near-optimal” audit strategy has had the intended effect of reducing tax fraud and thereby increasing declared income.

In regard to the income change, the Random Group falls between the two other groups. However, the Random Group does not show a significantly higher income change than the Control Group, nor does it show a significantly lower income change than the Rational Group. The groups have thus been too small to permit any clear conclusions as to whether the results are due to the effect of information in general or to the effect of information on the near-optimal audit strategy.

For the case of *Subcontractors Included* (Table 6), no significant income changes have been obtained, probably because those groups are less homogenous and therefore show larger standard deviations.

Table 6. Data for testing the difference in average income change between groups, including firms with subcontractors

| Amounts in SEK 1000 | Difference in average income change Δm | $\Delta m/s$ | Significance level |
|---------------------------------|--|--------------|--------------------|
| Rational Group vs Control Group | 7.9 | 1.33 | 18% |
| Random Group vs Control Group | 2.4 | 0.50 | 62% |
| Rational Group vs Random Group | 5.5 | 0.93 | 35% |

From Tables 5 and 6, we find that the monetary effect of the rational strategy compared to the control group is SEK 7,900 – 13,800. This can be compared to the results of the UK study (Hasseldine et al, 2007), where the effect of the three audit letters, converted to Swedish currency, amounted to SEK 2,600 – 11,500. Thus the effect is of the same order of magnitude.

Profit before tax allocations

The data supplied by the Tax Agency also included profit before depreciation and other tax allocations. This measure of income should be closer to the “true” result of the business since it is not affected by the adjustments that the taxpayer can make in order to reduce his tax burden or show a smoother income pattern over time. This measure is also closer to the net-cash-flow measure used for audit selection in the Rational case.

The Swedish tax system allows for two main instruments for income tax management. First, tax depreciation is very liberal for machinery and equipment with up to 20% depreciation on cost or 30% depreciation on depreciated value. Second, 25% of income may be allocated to a tax allocation reserve and retained there for a six-year period.

The most striking difference compared with the income measure used in Tables 3 and 4 is much larger standard deviations in profit change between years, SEK 100,000 – 180,000 as compared to SEK 50,000 – 77,000 in Tables 3 and 4. This indicates that tax allocations are actively used to equalize declared income between years.

The highest standard deviations are observed for the Rational Group. This is due to a few extreme outliers, with negative income changes exceeding SEK 500,000. Excluding the extreme outliers, statistically significant differences between the Rational Group and the Control Group are found at the 5% level for subcontractors excluded as well as included. When subcontractors are included, the Rational Group is even found to be significantly better than the Random Group.

Median tests

In addition to the tests above, using the average income/profit increase between years, we have also studied the median of the income/profit increase since this parameter is independent of extreme outliers. In three of the four cases, the Rational Group is significantly better than the Control Group, thus the results are quite similar to those obtained from using the increase in average income/profit.

Test summary

Comparisons have been drawn for the three groups in eight combinations (average/median, income/profit, with/without subcontractors). In addition, the average profit case has been studied with and without extreme outliers. Thus, ten cases in all have been evaluated. In nine of these, the ordering of the three groups was as expected, i.e. with the greatest changes between years for the Rational Group and the smallest changes for the Control Group.

Table 7. Significance level for eight tests

| Significance level | Rational Group vs Control Group | Rational Group vs Random Group |
|---|------------------------------------|-----------------------------------|
| Subcontr. Excluded | | |
| Average income change | 3.6% | 12% |
| Average profit change, extreme outliers excluded | 3% | 45% |
| Median income change | 12% | 25% |
| Median profit change | 5.0% | 91% |
| Subcontr. Included | | |
| Average income change | 18% | 35% |
| Average profit change, extreme outliers excluded | 0.4% | 0.7% |
| Median income change | 1.9% | 13% |
| Median profit change | 1.0% | 19% |

In six of the eight cases, the income/profit increase was significantly larger for the Rational Group than for the Control Group. In view of the limited volume of data in the study, however, it cannot be determined with statistical significance whether this result is due to the quality of the rational audit strategy or to the audit information in general.

DISCUSSION AND CONCLUSIONS**Method of measurement**

The selected method of measuring changes in tax-fraud behaviour by changes in declared income or profit between years, is low in cost but has several drawbacks. Income and profit changes may have other causes, such as changes in business volume, changes in profitability, investment, sale of assets etc. The data include several firms with zero sales in one or both years, those firms were not excluded as the low declared sales volume may be due to large-scale fraud.

A better measure of fraud would possibly be obtained with random audits but this method is much more costly. As shown below, random audits were made in the Random Group, but unfortunately not in the other groups. It should be remarked that

audits do not discover all fraud, especially not hidden income which is kept out of the accounts.

The quality of the study may have been affected by the use of net household cash flow as a parameter for the audit strategy when the effect of the strategy is measured as the change in declared income. A taxpayer with a high net household cash flow, perhaps due to employment income from his spouse, has no incentive to reduce any fraudulent behaviour as he does not expect to become audited. On the other hand, the incentive works as intended for taxpayers with a low net household cash flow.

Main conclusions

The statistical tests indicate strongly that information concerning the use of rational, “near-optimal” audit strategies is superior to information concerning random audits and that audit information is general is superior to no information.

It can be stated with statistical significance that information concerning the rational audit strategy reduces tax fraud compared to no information.

The results are well in line with the Minnesota and UK experiments. Similar results should be expected if the study would be repeated in other countries where sole proprietors make self-assessments for income tax purposes. .

Other results

According to Appendix 3, it seems possible that a separation into two sub-groups provides a realistic model, where one sub-group is less sensitive to audit-strategy information (thus less fraudulent) whereas the second sub-group is more sensitive to audit-strategy information (thus more fraudulent).

In Appendix 4, we find that the indirect effect of changing from random to rational audits is smaller than the direct effect, contrary to the experience of the author. A possible reason is that the audit rate in the study is higher than normal, resulting in larger direct effects. As the taxpayer is not informed of the high audit rate, the behaviour does not change as much as if the taxpayer had been aware of this fact.

Further research

It would be desirable to conduct new experiments on a larger scale in order to obtain statistically significant differences between the rational and random groups. It would also be helpful to use more specific information regarding actual audit rates. Disclosing such information is against the policy of the Swedish Tax Agency, however.

The hypothesis that the groups consist of two distinct sub-groups, one stable and less fraudulent and one volatile and more fraudulent, would be very interesting to follow up on a larger set of empirical data.

BIBLIOGRAPHY

- Allingham, M.G., & Sandmo, A. (1972). “Income Tax Evasion: A Theoretical Analysis”, *Journal of Public Economics*, Vol.1, pp. 323-338.
- Andreoni, J., Erard, B. & Feinstein, J. (1998). “Tax Compliance”, *Journal of Economic Literature*, Vol. 36, pp. 818-860.

Appelgren, L. (2003), "Audit Strategies for Tax Fraud Detection and Prevention", Paper presented at the *Stockholm International Symposium on Economic Crime*, 1-3 December 2003.

Appelgren, L. (2006), "Optimal revision – användning av ekonomiska modeller för ekobrottsprevention". In: Nilsson, F. & Olve, N-G., (eds), 2006. *Ekonomiska informationssystem*", Studentlitteratur, Lund, pp. 175-195.

Becker, G. S. (1968), "Crime and Punishment: An Economic Approach", *Journal of Political Economy*, Vol. 76, March-April, pp. 169-217.

Blumenthal, M., Christian, C. & Slemrod, J. (2001), "Do normative appeals affect tax compliance: Evidence from a controlled experiment in Minnesota", *National Tax Journal*, Vol. 54, pp. 125-138.

Eide, E. (2000), "Oversikt over litteratur om svart arbeid og skatteunndragelser", Rapport 6/2000, *Ragnar Frisch Centre for Economic Research*, Oslo.

Erard, B., & Feinstein, J. S. (1994), "Honesty and Evasion in the Tax Compliance Game", *RAND Journal of Economics*, Vol. 25, pp. 1-19.

Hasseldine, J., Hite, P., James, S. & Tuomi, M. (2007), "Persuasive Communications: Tax Compliance Enforcement Strategies for Sole Proprietors", *Contemporary Accounting Research*, Vol. 24, pp. 171-94.

Reinganum, J. F., & Wilde, L. L. (1986), "Equilibrium Verification and Reporting Policies in a Model of Tax Compliance", *International Economic Review*, Vol. 27, pp. 739-760.

Wenzel, M. & Taylor, N. (2004), "An experimental evaluation of tax-reporting schedules: a case of evidence-based tax administration", *Journal of Public Economics*, Vol. 88, pp. 2785-2799.

APPENDIX 1: LETTER TO MEMBERS OF THE RATIONAL GROUP

Riksskatteverket (Swedish Tax Agency)

NN

Each year the Linköping Regional Office of the Swedish Tax Agency conducts various activities for purposes of information and control so that the tax assessment will be as correct as possible. In some cases we provide advance notice that a certain type of control will be carried out. You are part of a group of randomly selected business proprietors who are being informed at this early stage that their income-tax returns to be submitted in 2004 may be audited.

Which returns will be audited?

After the tax returns have been filed, we will select the returns to be audited. Your return is one of those subject to a possible special audit, where we will select the returns of taxpayers with the lowest net cash flow.

In order to determine your net cash flow, we examine the data that you have provided in your tax return and the remuneration statements received by the Tax Agency. We then calculate how much money you have received and how much you have paid out. The difference between what you have received and what you have paid out is your

net cash flow. In the enclosure to this letter, you can see a sample calculation of net cash flow.

If most of the taxpayers in the group have a lower net cash flow than you, your tax return will not be audited. On the other hand, if your tax return is one of those with the lowest net cash flow, it will be selected for audit.

There need not be any error

There need not be any error in your tax return just because you have a low net cash flow. But a low net cash flow may be an indication of unreported income.

This audit concerns your business income. If you also have income from employment or capital, your tax return may be audited for other reasons – in that case there would be no difference between your tax return and all others.

Advance notice

Normally an audit comes as a surprise. We now want to test what happens when let taxpayers know before they file their tax returns how we will select which returns will be audited. We hope that as a result more taxpayers will file correct returns in the first place.

If you have questions regarding this letter, please call NN at XX.

Best regards
Bertil Olofson
Director, Linköping Regional Office, Swedish Tax Agency

APPENDIX 2: LETTER TO MEMBERS OF THE RANDOM GROUP

Riksskatteverket (Swedish Tax Agency)

NN

Each year the Linköping Regional Office of the Swedish Tax Agency conducts various activities for purposes of information and control so that the tax assessment will be as correct as possible. In some cases we provide advance notice that a certain type of control will be carried out. You are part of a group of randomly selected business proprietors who are being informed at this early stage that their income-tax returns to be submitted in 2004 may be audited.

Which returns will be audited?

After the tax returns have been submitted, we will select the ones to be audited. Your return is one of those subject to a possible special audit, where we will select a number of returns for closer examination on a totally random basis.

This audit concerns your business income. If you also have income from employment or capital, your tax return may be audited for other reasons – in that case there would be no difference between your tax return and all others.

Advance notice

Normally an audit comes as a surprise. We now want to test what happens when let taxpayers know before they file their tax returns how we will select which returns will be audited. We hope that as a result more taxpayers will file correct returns in the first place.

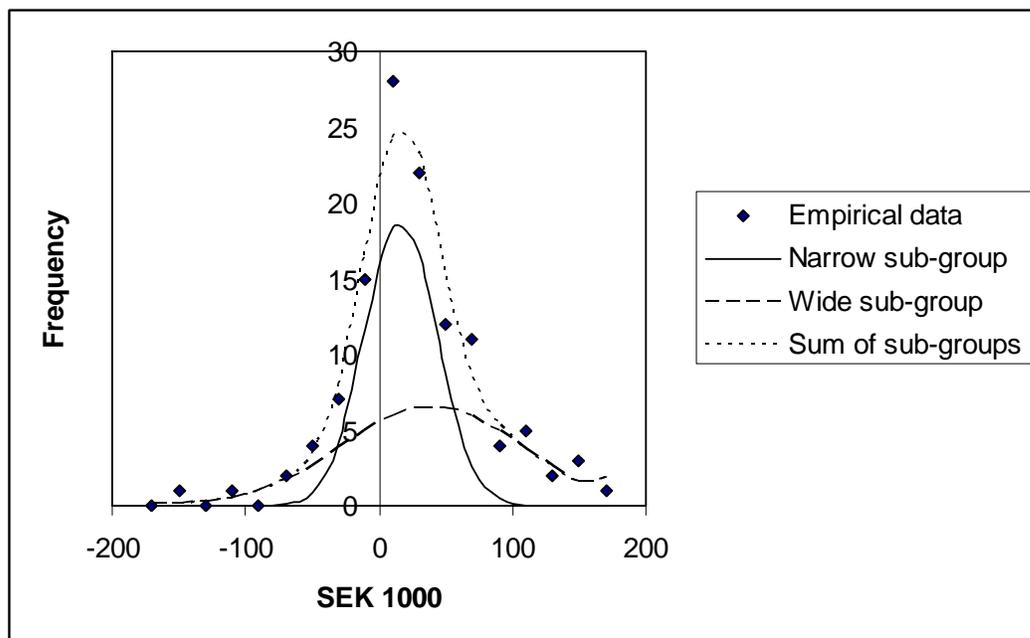
If you have questions regarding this letter, please call NN at XX.

Best regards
Bertil Olofson
Director, Linköping Regional Office, Swedish Tax Agency

APPENDIX 3: DISTRIBUTION OF INCOME CHANGE BETWEEN YEARS

The distributions of income change for the *Rational* and *Control* groups in the case *Subcontractors Excluded* are shown in Figures A3:1 and A3:2. Is the income change between 2002 and 2003 normally distributed?

Figure A3:1. Distribution of income change for the Rational Group, Subcontractors Excluded



In a normal distribution, the ratio between standard deviation and the quartile distance, i.e. the distance between the third and the first quartile, is 0.74. In Table A3:1, we compute this ratio for the six cases. It is apparent that in all cases except *Subcontractors Excluded/Rational*, the distribution is far from normal, with extreme outliers causing an abnormally high standard deviation. A χ^2 test for the three cases *Subcontractors Excluded* confirms that the distribution is not normal for the *Random* and *Control* cases.

Figure A3:2. Distribution of income change for the Control Group, Subcontractors Excluded

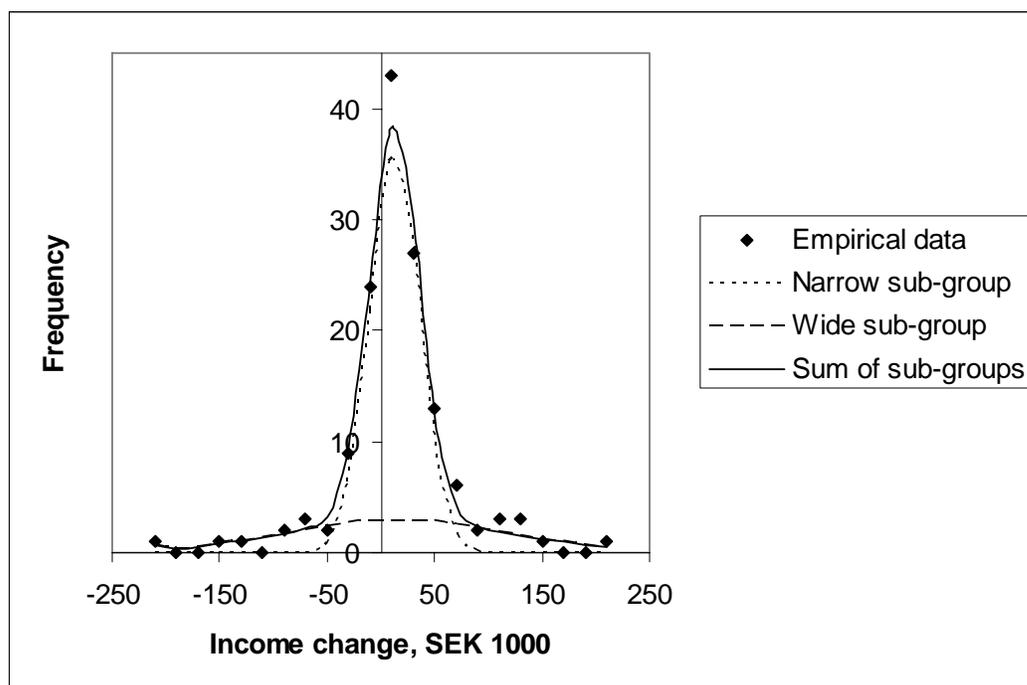


Table A3:1. Comparison between standard deviation and quartile distance

| Amounts in SEK 1000 | Group | Standard dev. | Quartile distance | Ratio |
|-------------------------|----------|---------------|-------------------|-------|
| Subcontractors Excluded | Rational | 50 | 56 | 0.89 |
| | Random | 58 | 48 | 1.21 |
| | Control | 56 | 36 | 1.56 |
| Subcontractors Included | Rational | 77 | 56 | 1.37 |
| | Random | 55 | 53 | 1.04 |
| | Control | 54 | 42 | 1.29 |

A possible approximation of the income change distribution is that each group consists of two normally distributed sub-groups, one with a small standard deviation (narrow group) and one with a large standard deviation (wide group). The best fit for two normal distributions has been determined with a maximum-likelihood model with five parameters (two mean values, two standard deviations and one relative weight factor). The results for the case of *Subcontractors Excluded* are presented in Table A3:2 below.

Table A3:2. Maximum likelihood estimates for two normally distributed sub-groups, subcontractors excluded

| Amounts in SEK 1000 | Narrow sub-group | | | Wide sub-group | | |
|------------------------|------------------|-----------------------|------------------|----------------|-----------------------|---------------|
| | Mean | Standard deviation | Weight factor | Mean | Standard deviation | Weight factor |
| Rational Group | 16.6 | 26.2 | 0.539 | 39.5 | 66.1 | 0.461 |
| Random Group | 12.9 | 21.2 | 0.678 | 20.7 | 93.5 | 0.322 |
| Control Group | 12.9 | 22.4 | 0.734 | 15.7 | 101.4 | 0.266 |

It is striking that the mean values and standard deviations for the three narrow sub-groups are so similar. This finding gives rise to a hypothesis that the populations consist of two distinct groups, one with stable income from year to year and one with volatile income. When information regarding future audits is supplied, the members of the volatile group respond with an increase in declared income, i.e. there is a reduction in fraud.

It must be emphasised that the above results are quite uncertain because of the limited size of the groups. The same results were not obtained for the case *Subcontractors Included*, possibly because those groups are less homogenous.

APPENDIX 4: DIRECT AND INDIRECT EFFECTS

The Swedish Tax Agency has carried out audits according to its announced strategies, i.e. on taxpayers with the lowest net household cash flow in the Rational Group and randomly in the Random Group. No audits were conducted in the Control Group. An equal number of audits were made in the Rational and Random groups. They resulted in SEK 846,000 and SEK 260,000 respectively in increased taxes and tax penalties. Thus the direct effect of a transition from random to rational audits is SEK 586,000, a strong indicator that the latter strategy is considerably more efficient than random auditing.

The direct effect should be compared to the indirect, deterrent effect, which for the case of *Subcontractors Excluded* is an average income increase amounting to SEK 10,200 according to Table 7, i.e. SEK 1,204,000 for 118 taxpayers. With the Swedish local tax rate around 30%, the indirect effect on public revenues would be about SEK 360,000. The corresponding numbers for the case of *Subcontractors Included* are SEK 5,500 for 248 taxpayers, with a revenue effect of roughly SEK 410,000.

Since the Tax Agency did not exclude taxpayers with subcontractors in the selection of audit targets, the comparison should be made with the case *Subcontractors Included*. The direct effect of switching from random to rational audits, SEK 586,000, should thus be compared to the indirect effect of SEK 410,000.